MTH 235 - Differential Equations

Gabriel Nagy

Mathematics Department,
Michigan State University,
East Lansing, MI, 48824.
gnagy@msu.edu

November 1, 2024

Contents

Preface

These notes are an introduction to ordinary differential equations. We study a variety of models in physics and engineering, which use differential equations, we learn how to construct differential equations, corresponding to different ecological or physical systems. In addition to introducing various analytical techniques for solving basic types of ODEs, we also study qualitative techniques for understanding the behavior of solutions. We describe a collection of methods and techniques used to find solutions to several types of differential equations, including first order scalar equations, second order linear equations, and systems of linear equations. We introduce Laplace transform methods to solve constant coefficients equations with generalized source functions. We also provide a brief introduction to boundary value problems, Sturm-Liouville problems, and Fourier Series. Near the end of the course we combine previous ideas to solve an initial boundary value problem for a particular partial differential equation, the heat propagation equation.

CHAPTER 1

First Order Equations

This first chapter is an introduction to differential equations. We focus on particular techniques—developed in the eighteenth and nineteenth centuries—to solve certain first order differential equations. These equations include separable equations, Euler homogeneous equations, and linear equations. Soon this way of studying differential equations reached a dead end. Most of the differential equations cannot be solved by any of the techniques presented in the first sections of this chapter. Then, people tried something different. Instead of solving the equations they tried to show whether an equation has solutions or not, and what properties such solution may have. This is less information than obtaining the solution, but it is better than giving up. The results of these efforts are shown in the last sections of this chapter. We present theorems describing the existence and uniqueness of solutions to a wide class of first order differential equations.

1.1. Separable Equations

A differential equation is an equation, where the unknown is a function and both the function and its derivatives appear in the equation. When the function depends on only one variable, the equation is called an ordinary differential equation. In the first part of this section we give a few examples of ordinary differential equations, then we focus on a particular type of equations that are specially simple to solve—separable equations. The simplest idea to solve a differential equation actually works well with separable equations—integrate on both sides of the equation. In the second part of this section we introduce Euler homogeneous equations. These equations are not separable but we can transform them into separable equations for a different variable. Then, we can solve them for the new variable and transform the solution back to the original variable appearing in the Euler equation.

1.1.1. Definitions and Examples. A differential equation is an equation for a function and its derivatives. The function in a differential equation is called the *unknown function*. An *Ordinary Differential Equation* (ODE) is an equation containing derivatives with respect to only one variable. A *Partial Differential Equation* (PDE) is an equation containing derivatives with respect to more than one variable. The *order* of a differential equation is the number of the highest derivative in the equation. For example, a first order ODE is a differential equation containing derivatives with respect to only one variable and the highest derivative of the unknown function is the first derivative. We can write all this in a more concise way.

Definition 1.1.1. *A* ***first order ODE*** *on the unknown function y is*

$$y'(t) = f(t, y(t)), \tag{1.1.1}$$

where f is a given function of two variables and $y' = \dfrac{dy}{dt}$.

The function y above is usually called the *dependent variable* while t above is called the *independent variable*. The function $y(t)$ is also called the *unknown function*, since in most problems involving a differential equation we try to find the solutions, $y(t)$.

Example 1.1.1. Here are a few examples of first order ordinary differential equations.

(a)

$$y' = 2y,$$

which is called the exponential growth equation and will be aolved later on in this section.

(b)

$$y' = a\,y + b,$$

where a, b are arbitrary constants. The previous example is the case $a = 2$ and $b = 0$.

(c)

$$y' = \frac{2t}{y(3-y)},$$

which is an equation with explicit t dependence on the right side, unlike the previous two examples.

(d)

$$2t\,y' - \sin(y) = 0,$$

which is not written as in (1.1.1), but simple algebraic transformations give us

$$y' = \frac{\sin(y)}{2t}.$$

These two forms of the equation share most of their solutions, so we consider them as the same equation.

(e)

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

which is an interesting equation, called scale invariant, and we will study them later.

A differential equation can be written in many different ways. We say that a differential equation is written in *normal* form if the equation is written as in (1.1.1), that is, y' is alone on the left-hand side and the right-hand side contains only functions of t and y but no derivatives, y'. For example, the differential equation

$$3y' + 6y = 9$$

is not written in normal form, but we can rewrite it in normal form after a few simple algebraic transformations,

$$y' = -2y + 3.$$

We define *simple algebraic transformations* of a differential equation as adding the same function on both sides of the equation and multiplying by non-zero functions on both sides of the equation. Simple algebraic transformations change how the equation looks like but they do not change its solutions.

A function g is a *solution* of a differential equation $y' = f(t, y)$, if the function $g'(t)$ is the same as the function composition $f(t, g(t))$ for all t. For example, the function

$$g(t) = e^{2t} - \frac{3}{2}$$

is a solution of the differential equation

$$y' = 2y + 3.$$

Because its derivative is

$$g'(t) = 2\,e^{2t}$$

while the algebraic calculation below give us

$$2\,g(t) + 3 = 2\Big(e^{2t} - \frac{3}{2}\Big) + 3 = 2\,e^{2t}.$$

Therefore, this function g satisfies the differential equation

$$g' = 2g + 3.$$

Differential equations may have infinitely many solutions. For example, the functions

$$h(t) = k\,e^{2t} - \frac{3}{2},$$

where k is an arbitrary constant are also solutions of the differential equation

$$y' = 2y + 3.$$

Indeed, if we compute the derivative,

$$h'(t) = 2k\,e^{2t},$$

while the algebraic calculation below gives us

$$2\,h(t) + 3 = 2\Big(k\,e^{2t} - \frac{3}{2}\Big) + 3 = 2k\,e^{2t}.$$

Therefore, this function h satisfies the differential equation

$$h' = 2h + 3$$

for any value of the constant k. So we have infinitely many solutions.

Definition 1.1.2. *An **initial value problem** (IVP) is to find a solution $y(t)$ to an equation*

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

*that also satisfies the **initial condition***

$$y(t_0) = y_0,$$

where t_0 and y_0 are arbitrary constants.

Example 1.1.2. Find a solution to the initial value problem

$$y' = 2\,y + 3, \qquad y(0) = 5.$$

Solution: We have seen in the previous paragraph that the differential equation in this example has infinitely many solutions given by

$$y(t) = k\,e^{2t} - \frac{3}{2}.$$

From these solutions we select the one that satisfies the initial condition. That is, the initial condition determines the value of the constant k, as we can see below.

$$5 = y(0) = k\,e^{0} - \frac{3}{2} \quad \Rightarrow \quad k = 5 + \frac{3}{2} \quad \Rightarrow \quad k = \frac{13}{2}.$$

Therefore, the solution of the initial value problem above is

$$y(t) = \frac{13}{2}\,e^{2t} - \frac{3}{2}.$$

$\triangleleft$

Differential equations are essential to describe change in nature. Newton's equation for the motion of a point particle is $ma = f$, mass times acceleration equals the sum of all forces acting on the particle. This equation is a differential equation. In fact, it is the differential equation that started the whole field of differential equations. The acceleration is the second time derivative of the position function. If we call the position function by $y(t)$, then $a = y''$. The force term may depend on time t, on the position y, and sometimes on the velocity y', therefore, Newton's equation is a second order differential equation of the form

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

If the particle is a projectile and the force is the Earth's gravity near its surface, then Newton's equations predict the parabolic motion of the projectile. If the particle is the Moon and the force is Earth's gravitational attraction, then Newton's equation describe the orbit of the Moon around the Earth.

Electricity and magnetism are described by Maxwell's equations, which are a set of differential equations for the electric and magnetic fields, partially discovered by Ampère, Faraday, and with the last touches given by Maxwell. These equations not only describe how electricity and magnetism behave, but they also say that light is an electromagnetic phenomena having a fixed speed of propagation in vacuum. That observation is the corner stone used by Einstein to build his Special Theory of Relativity.

Quantum mechanics rests on the Schrödinger equation, which is a differential equation describing how objects behave at scales of electrons orbiting the nucleus of atoms. The variable in Schrödinger equation is the probability of finding these microscopic objects, say electrons, at a certain time and position. Quantum mechanics explain how atoms interact with other atoms, converting Chemistry into a part of Physics.

Einstein equations of General Relativity are differential equations that improve Newton's theory of gravitation in the case of large masses, such as near big stars or at the early times of the universe. General Relativity predicts both the existence of black holes, which produce extreme effects in the curvature of space and the passing of time, and the big bang at the beginning of the universe.

As we see above, most natural phenomena are described by differential equations. The job of scientists is to try to find these equations and figure out the behavior of their solutions.

1.1.2. Overview of Modeling. Modeling is to create mathematical representations or natural porcesses, called models. These models usually involve differential equations because natural processes change in time. To create a mathematical model of a natural system usually involves following a few basic steps:

(1) To state all the *assumptions* needed to isolate the feature we are interested to describe.
(2) To introduce the *independent variables* and the *dependent variables*.
(3) To introduce all the *parameters* needed to specify the model.
(4) To use these assumptions to derive equations relating the variables and parameters.
(5) To analyze the predictions of the model—do they make physical sense, do they agree with your data? If not, you might need to revise your assumptions, going back to part (1).

In the following examples we show a few models that accurately describe different situations, including population dynamics with infinite food resources, the radioactive decay equation, and Newton's cooling law describing the temperature of an object in an environment with fixed temperature.

Example 1.1.3 (Exponential Growth Equation)**.** The population as function of time, $y(t)$, of a living organism that has access to infinite food resources, can be described by the solutions of the differential equation

$$y' = r\,y,$$

where the parameter $r > 0$ is a constant called the growth rate, that has units of 1/time. Solutions of this equation are exponentials of the form

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

where y_0 is the initial population, $y(0) = y_0$. These functions are solutions of the differential equation because

$$y'(t) = r\, y_0\, e^{rt} = r\, y(t).$$

This equation is called the exponential growth equation because its solutions are functions that grow exponentially.

Example 1.1.4 (Radioactive Decay)**.** The amount of a radioactive material as function of time, $y(t)$, when the material is left alone, can be described by the solutions of the differential equation

$$y' = -r\,y.$$

In this case the parameter $r > 0$ is called the decay constant, which also has units of 1/time. Solutions of this equation are exponentials of the form

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

where y_0 is the initial amount or radioactive material, $y(0) = y_0$. These functions are solutions of the differential equation because

$$y'(t) = -r\, y_0\, e^{-rt} = -r\, y(t).$$

This equation is called the exponential decay equation because its solutions are functions that decay exponentially.

Example 1.1.5 (Newton's Cooling Law)**.** The temperature as function of time, $T(t)$, of a material placed in a surrounding medium kept at a constant temperature T_s can be described by the solutions of the differential equation

$$T' = -k\,T + kT_s,$$

where the parameter $k > 0$ is a constant characterizing the material thermal properties. Since the medium temperature is constant, we can rewrite this differential equation as

$$(\Delta T)' = -k\,(\Delta T),$$

where we introduced the temperature difference

$$\Delta T(t) = T(t) - T_s,$$

Notice that this is the exponential decay equation for the temperature difference ΔT. Although Newton's law is called a "Cooling Law", it also describes objects that warm up. When the initial temperature difference, $(\Delta T)(0) = T(0) - T_s$ is positive the object cools down, but when $(\Delta T)(0)$ is negative the object is warms up. The solution of Newton's cooling law equation is

$$(\Delta T)(t) = (\Delta T)(0)\,e^{-kt},$$

for some initial temperature difference $(\Delta T)(0) = T(0) - T_s$. Since

$$(\Delta T)(t) = T(t) - T_s,$$

then we can write the solution of the differential equation as

$$T(t) = (T(0) - T_s)\,e^{-kt} + T_s.$$

All the equations in the models above are particular cases of equations we call separable differential equations. These are equations where the dependent and independent variables can be separated on different sides of the equation, which allow us to find exact formulas for their solutions. We introduce these equations in our next subsection.

1.1.3. Separable Equations. We now introduce a particular type of differential equations called separable equations. They include the exponential growth equation, exponential decay equation, and Newton's cooling law equation. In these equations the variables can be separated on each side of the equation, which makes the equation simple to solve.

Definition 1.1.3. *A differential equation for a function $y(t)$ is **separable** iff after simple algebraic transformations the equation can be written in the form*

$$h(y)\,y' = g(t), \tag{1.1.2}$$

where h, g are given functions.

Remarks:

- There is no y-dependence on the right-hand side of (1.1.2), it depends *only* on t.
- The left-hand side depends explicitly *only* on y; any t dependence is through y.
- The left-hand side is of the form (something on y) multiplied by y'.
- The normal form of a separable equation is

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

Example 1.1.6.

(a) The exponential grow-decay equations are separable, since the equation

$$y' = a\,y$$

for a constant, can be rewritten as

$$\frac{y'}{y} = a \quad \Rightarrow \quad \begin{cases} g(t) = a, \\ h(y) = \dfrac{1}{y}. \end{cases}$$

(b) The differential equation $y' = \dfrac{t^2}{1 - y^2}$ is separable, since it can be easily transformed as

$$\left(1 - y^2\right) y' = t^2 \quad \Rightarrow \quad \begin{cases} g(t) = t^2, \\ h(y) = 1 - y^2. \end{cases}$$

(c) The differential equation $y' + y^2 \cos(2t) = 0$ is separable, since it can be transformed into (1.1.2) as follows,

$$\frac{1}{y^2}\, y' = -\cos(2t) \quad \Rightarrow \quad \begin{cases} g(t) = -\cos(2t), \\ h(y) = \dfrac{1}{y^2}. \end{cases}$$

The functions g and h are not uniquely defined; another choice in this example is:

$$g(t) = \cos(2t), \quad h(y) = -\frac{1}{y^2}.$$

(d) The equation $y' = e^y + \cos(t)$ is not separable because there are no simple algebraic transformations that can changed it into the form in (1.1.2).

(e) The exponential grow-decay with migration equation is separable, since

$$y' = a\,y + b$$

for a and b constants, can be rewritten as

$$\frac{1}{(a\,y + b)}\, y' = 1 \quad \Rightarrow \quad \begin{cases} g(t) = 1, \\ h(y) = \dfrac{1}{(a\,y + b)}. \end{cases}$$

(f) The variable coefficient equation

$$y' = a(t)\,y + b(t),$$

with $a \neq 0$ and b/a nonconstant, is not separable because there are no simple algebraic transformations that can changed it into the form in (1.1.2).

◁

Separable differential equations are simple to solve. We just write them as in (1.1.2) and integrate on both sides of the equation with respect to the independent variable t. We show this idea in the following example.

Example 1.1.7. Find all solutions y to the differential equation

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

Solution: The differential equation above is separable, with

$$g(t) = \cos(2t), \qquad h(y) = -\frac{1}{y^2}.$$

Therefore, it can be integrated as follows:

$$-\frac{y'}{y^2} = \cos(2t) \quad \Leftrightarrow \quad \int -\frac{y'(t)}{y^2(t)}\,dt = \int \cos(2t)\,dt + c.$$

The integral on the right-hand side can be computed explicitly. The integral on the left-hand side can be done by substitution. The substitution is

$$u = y(t), \qquad du = y'(t)\,dt \quad \Rightarrow \quad \int -\frac{du}{u^2} = \int \cos(2t)\,dt + c.$$

This notation makes clear that u is the new integration variable, while $y(t)$ are the unknown function values we look for. However it is common in the literature to use the same name for the variable and the unknown function. We follow that convention and we write the substitution as

$$y = y(t), \qquad dy = y'(t)\,dt \quad \Rightarrow \quad \int -\frac{dy}{y^2} = \int \cos(2t)\,dt + c.$$

We hope this is not too confusing. Integrating on both sides above we get

$$\frac{1}{y} = \frac{1}{2}\sin(2t) + c.$$

So, we get the implicit and explicit form of the solution,

$$\frac{1}{y(t)} = \frac{1}{2}\sin(2t) + c \quad \Leftrightarrow \quad y(t) = \frac{2}{\sin(2t) + 2c}.$$

◁

Remark: Notice the following about the equation and its implicit solution:

$$-\frac{1}{y^2}\,y' = \cos(2t) \quad \Leftrightarrow \quad h(y)\,y' = g(t), \qquad h(y) = -\frac{1}{y^2}, \qquad g(t) = \cos(2t),$$

$$\frac{1}{y} = \frac{1}{2}\sin(2t) \quad \Leftrightarrow \quad H(y) = G(t), \qquad H(y) = \frac{1}{y}, \qquad G(t) = \frac{1}{2}\sin(2t).$$

- H is an antiderivative of h, that is, $H(y) = \int h(y)\,dy$.
- G is an antiderivative of g, that is, $G(t) = \int g(t)\,dt$.

This remark help us summarize the calculation done in the previous example as follows.

Theorem 1.1.4 (Separable Equations)**.** *If h, g are continuous, with $h \neq 0$, then*

$$h(y)\,y' = g(t) \tag{1.1.3}$$

has infinitely many solutions y satisfying the algebraic equation

$$H(y(t)) = G(t) + c, \tag{1.1.4}$$

where $c \in \mathbb{R}$ is arbitrary, $H = \int h(y)\,dy$ and $G = \int g(t)\,dt$ are antiderivatives of h and g.

Proof of Theorem 1.1.4: Integrate with respect to t on both sides in Eq. (1.1.3),

$$h(y)\,y' = g(t) \quad \Rightarrow \quad \int h(y(t))\,y'(t)\,dt = \int g(t)\,dt + c,$$

where c is an arbitrary constant. Introduce on the left-hand side of the second equation above the substitution

$$y = y(t), \qquad dy = y'(t)\,dt.$$

The result of the substitution is

$$\int h(y(t))\,y'(t)\,dt = \int h(y)dy \quad \Rightarrow \quad \int h(y)\,dy = \int g(t)\,dt + c.$$

To integrate on each side of this equation means to find a function H, primitive of h, and a function G, primitive of g. Using this notation we write

$$H(y) = \int h(y)\,dy, \qquad G(t) = \int g(t)\,dt.$$

Then the equation above can be written as follows,

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

which implicitly defines a function y, which depends on t. This establishes the Theorem. □

Example 1.1.8. Find all solutions y to the differential equation

$$y' = \frac{t^2}{1 - y^2}. \tag{1.1.5}$$

Solution: We write the differential equation in (1.1.5) in the form $h(y)\,y' = g(t)$,

$$\big(1 - y^2\big)\,y' = t^2.$$

In this example the functions h and g defined in Theorem 1.1.4 are given by

$$h(y) = (1 - y^2), \qquad g(t) = t^2.$$

We now integrate with respect to t on both sides of the differential equation,

$$\int \big(1 - y^2(t)\big)\,y'(t)\,dt = \int t^2\,dt + c,$$

where c is any constant. The integral on the right-hand side can be computed explicitly. The integral on the left-hand side can be done by substitution. The substitution is

$$y = y(t), \qquad dy = y'(t)\,dt \quad \Rightarrow \quad \int \big(1 - y^2(t)\big)\,y'(t)\,dt = \int (1 - y^2)\,dy.$$

This substitution on the left-hand side integral above gives,

$$\int (1 - y^2)\,dy = \int t^2\,dt + c \quad \Leftrightarrow \quad y - \frac{y^3}{3} = \frac{t^3}{3} + c.$$

The equation above defines a function y, which depends on t. We can write it as

$$y(t) - \frac{y^3(t)}{3} = \frac{t^3}{3} + c.$$

We have solved the differential equation, since there are no derivatives in the last equation. When the solution is given in terms of an algebraic equation, we say that the solution y is given in *implicit form*. ◁

Definition 1.1.5. *A function y is a solution in **implicit form** of the equation*

$$h(y)\,y' = g(t)$$

iff the function y is solution of the algebraic equation

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

where H and G are any antiderivatives of h and g. In the case that function H is invertible, the solution y above is given in ***explicit form*** *iff is written as*

$$y(t) = H^{-1}\big(G(t) + c\big).$$

In the case that H is not invertible or H^{-1} is difficult to compute (as in Example 1.1.8), we leave the solution y in implicit form. Now we solve the problem in Example 1.1.8, but now we just use the result of Theorem 1.1.4.

Example 1.1.9. Use the formula in Theorem 1.1.4 to find all solutions y to the equation

$$y' = \frac{t^2}{1 - y^2}. \tag{1.1.6}$$

Solution: Theorem 1.1.4 tell us how to obtain the solution y. Writing Eq. (1.1.6) as

$$\big(1 - y^2\big)\, y' = t^2,$$

we see that the functions h, g are given by

$$h(y) = 1 - y^2, \qquad g(t) = t^2.$$

Their primitive functions, H and G, respectively, are simple to compute,

$$h(y) = 1 - y^2 \quad \Rightarrow \quad H(y) = y - \frac{y^3}{3},$$

$$g(t) = t^2 \quad \Rightarrow \quad G(t) = \frac{t^3}{3}.$$

Then, Theorem 1.1.4 implies that the solution y satisfies the algebraic equation

$$y(t) - \frac{y^3(t)}{3} = \frac{t^3}{3} + c, \tag{1.1.7}$$

where $c \in \mathbb{R}$ is arbitrary. $\lhd$

Remark: In general it is simpler to remember ideas than formulas, specially when the number of formulas to remember is large. Usually it is better to solve a separable equation as we did in Example 1.1.8 instead of using the solution formulas, as in Example 1.1.9. (Although in the case of separable equations both methods are very close.)

In the next Example we show that an initial value problem can be solved even when the solutions of the differential equation are given in implicit form.

Example 1.1.10. Find the solution of the initial value problem

$$y' = \frac{t^2}{1 - y^2}, \qquad y(0) = \frac{1}{2}. \tag{1.1.8}$$

Solution: From Example 1.1.8 we know that all solutions to the differential equation in (1.1.8) are given by

$$y(t) - \frac{y^3(t)}{3} = \frac{t^3}{3} + c,$$

where $c \in \mathbb{R}$ is arbitrary. This constant is now fixed with the initial condition in Eq. (1.1.8)

$$y(0) - \frac{y^3(0)}{3} = \frac{0}{3} + c \quad \Rightarrow \quad \frac{1}{2} - \frac{1}{3}\frac{1}{2^3} = c \quad \Leftrightarrow \quad c = \frac{11}{24} \quad \Rightarrow \quad y(t) - \frac{y^3(t)}{3} = \frac{t^3}{3} + \frac{11}{24}.$$

So we can rewrite the algebraic equation defining the solution functions y as the (time dependent) roots of a cubic (in y) polynomial,

$$y^3(t) - 3y(t) + t^3 + \frac{33}{24} = 0.$$

$\triangleleft$

Example 1.1.11. Find the solution of the initial value problem

$$y' + y^2 \cos(2t) = 0, \qquad y(0) = 1. \tag{1.1.9}$$

Solution: The differential equation above can be written as

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

We know, from Example 1.1.7, that the solutions of the differential equation are

$$y(t) = \frac{2}{\sin(2t) + 2c}, \qquad c \in \mathbb{R}.$$

The initial condition implies that

$$1 = y(0) = \frac{2}{0 + 2c} \quad \Leftrightarrow \quad c = 1.$$

So, the solution to the IVP is given in explicit form by

$$y(t) = \frac{2}{\sin(2t) + 2}.$$

$\triangleleft$

Example 1.1.12. Follow the proof in Theorem 1.1.4 to find all solutions y of the equation

$$y' = \frac{4t - t^3}{4 + y^3}.$$

Solution: The differential equation above is separable, with

$$h(y) = 4 + y^3, \qquad g(t) = 4t - t^3.$$

Therefore, it can be integrated as follows:

$$\big(4 + y^3\big)\, y' = 4t - t^3 \quad \Leftrightarrow \quad \int \big(4 + y^3(t)\big)\, y'(t)\, dt = \int (4t - t^3)\, dt + c.$$

Again the substitution

$$y = y(t), \qquad dy = y'(t)\, dt$$

implies that

$$\int (4 + y^3)\, dy = \int (4t - t^3)\, dt + c_0. \quad \Leftrightarrow \quad 4y + \frac{y^4}{4} = 2t^2 - \frac{t^4}{4} + c_0.$$

Therefore, the solution $y(t)$ can be given in implicit form as

$$y^4(t) + 16y(t) - 8t^2 + t^4 = c_1,$$

where $c_1 = 4c_0$. $\triangleleft$

Example 1.1.13. Find the solution of the initial value problem below in explicit form,

$$y' = \frac{2-t}{1+y}, \qquad y(0) = 1. \tag{1.1.10}$$

Solution: The differential equation above is separable with

$$h(y) = 1 + y, \qquad g(t) = 2 - t.$$

Their primitives are respectively given by,

$$h(y) = 1 + y \quad \Rightarrow \quad H(y) = y + \frac{y^2}{2},$$
$$g(t) = 2 - t \quad \Rightarrow \quad G(t) = 2t - \frac{t^2}{2}.$$

Therefore, the implicit form of all solutions y to the ODE above are given by

$$y(t) + \frac{y^2(t)}{2} = 2t - \frac{t^2}{2} + c,$$

with $c \in \mathbb{R}$. The initial condition in Eq. (1.1.10) fixes the value of constant c, as follows,

$$y(0) + \frac{y^2(0)}{2} = 0 + c \quad \Rightarrow \quad 1 + \frac{1}{2} = c \quad \Rightarrow \quad c = \frac{3}{2}.$$

We conclude that the implicit form of the solution y is given by

$$y(t) + \frac{y^2(t)}{2} = 2t - \frac{t^2}{2} + \frac{3}{2}, \quad \Leftrightarrow \quad y^2(t) + 2y(t) + (t^2 - 4t - 3) = 0.$$

The explicit form of the solution can be obtained realizing that $y(t)$ is a root in the quadratic polynomial above. The two roots of that polynomial are given by

$$y_{\pm}(t) = \frac{1}{2}\big[-2 \pm \sqrt{4 - 4(t^2 - 4t - 3)}\big] \quad \Leftrightarrow \quad y_{\pm}(t) = -1 \pm \sqrt{-t^2 + 4t + 4}.$$

We have obtained two functions y_+ and y_-. However, we know that there is only one solution to the initial value problem. We can decide which one is the solution by evaluating them at the value $t = 0$ given in the initial condition. We obtain

$$y_+(0) = -1 + \sqrt{4} = 1,$$
$$y_-(0) = -1 - \sqrt{4} = -3.$$

Therefore, the solution is y_+, that is, the explicit form of the solution is

$$y(t) = -1 + \sqrt{-t^2 + 4t + 4}.$$

$\lhd$

1.1.4. The Logistic Equation.

The population as function of time, $y(t)$, of a living organism that has access to finite food resources, can be described by the solutions of the differential equation, called the logistic equation

Definition 1.1.6. *the **logistic equation** for the function $y(t)$ is given by*

$$y' = r\,y\Big(1 - \frac{y}{K}\Big), \tag{1.1.11}$$

where $r > 0$ is a constant called the growth rate and $K > 0$ is a constant called the carrying capacity of the environment.

The logistic equation was introduced in a series of articles from 1838 to 1844 by Pierre Francois Verhulst, who used to describe population growth when subject to limited resources. The name logistic is an adaptation of the french *logistique* used by Verhulst. He studied population growth with both unlimited and limited resources. Verhulst called the former logarithmic growth and latter logistic growth. In more recent times we changed the first name to exponential growth and kept the second name.

We can get a qualitative idea of the behavior of the solutions of this equation if we study the behavior of the equation itself in some limiting cases.

(a) If the population is small enough so that the food seems unlimited for this population, which happens when $y(t) \ll K$, then the rate of growth of the population is close to be proportional to its size, just as in the exponential growth equation.

$$\text{If } \; y(t) \; \text{ is small enough, then } \quad y'(t) \simeq r\, y(t), \; \text{ then } \quad y(t) \simeq y(0)\, e^{rt}.$$

(b) If the population is $y(t) < K$ and near K, that is $y(t) \lessapprox K$, then the rate of growth, $y'(t)$, decreases and becomes smaller than the exponential growth given by $r\, y(t)$.

$$\text{If } \; y(t) \lessapprox K \; \text{ then } \quad y'(t) \gtrapprox 0.$$

We interpret this result saying that the population is getting close to the maximum population sustained by the food available.

(c) If the population is larger than the critical population K, that is $y(t) > K$, then the population actually decreases in time.

$$\text{If } \; y(t) > K \; \text{ then } \quad y'(t) < 0.$$

We interpret this result saying that the population becomes too large to be supported by the finite food in the environment and decreases.

(d) If the population is exactly equal to K, then $y' = 0$ and that population number does not change in time. So, one solution to the logistic equation is the constant $y(t) = K$.

Therefore, the solutions of the logistic equation for positive but small enough data should behave like growing exponentials in time, but then the increase should slow down and approach the constant solution $y(t) = K$, giving the solution curve an s-shape. We now find a formula for all the solutions of the logistic equation.

Theorem 1.1.7 (Logistic Equation)**.** *The initial value problem for the logistic equation*

$$y' = r\, y \Big(1 - \frac{y}{K}\Big), \qquad y(0) = y_0,$$

where r and K are positive constants and y_0 is an arbitrary constant, are given by

$$y(t) = \frac{K\, y_0}{y_0 + (K - y_0)\, e^{-rt}}.$$

Proof of Theorem 1.1.7: We start writing the logistic equation given in (1.1.11) as follows,

$$y' = \frac{r}{K}\, y\, (K - y),$$

and then we rewrite this equation in the separable form

$$\frac{y'}{y\, (K - y)} = \frac{r}{K}.$$

To solve this equation we only need to integrate on both sides of the equation on the independent variable,

$$\int \frac{y'(t)\, dt}{y(t)\, \big(K - y(t)\big)} = \int \frac{r}{K}\, dt.$$

The integral on the right side is simple, while on the left side we do the usual substitution $y = y(t)$ gives $dy = y'(t)\,dt$. Then, we get

$$\int \frac{dy}{y\,(K-y)} = \frac{r}{K}t + c_0, \tag{1.1.12}$$

where c_0 is an arbitrary integration constant. The integral on the left side can be done using partial fractions on the integrand,

$$\frac{1}{y\,(K-y)} = \frac{a}{y} + \frac{b}{(K-y)},$$

where a and b are specific constants that we can compute as follows. First add up the right-side above,

$$\frac{a}{y} + \frac{b}{(K-y)} = \frac{a(K-y)+by}{y\,(K-y)}$$

which means we got the equation

$$\frac{1}{y\,(K-y)} = \frac{a(K-y)+by}{y\,(K-y)}.$$

This last equation says the numerators must be equal, meaning

$$1 = a(K-y) + by \qquad \text{for all } y.$$

This equation can be rewritten as

$$(b-a)y + (aK-1) = 0 \qquad \text{for all } y.$$

This can only happen for

$$b-a=0, \quad aK-1=0\} \quad \Rightarrow \quad a = \frac{1}{K}, \quad b = \frac{1}{K}.$$

Therefore, we have shown the following partial fractions decomposition,

$$\frac{1}{y\,(K-y)} = \frac{1}{K}\,\frac{1}{y} + \frac{1}{K}\,\frac{1}{(K-y)}$$

Now we can come back to the integral in (1.1.12),

$$\frac{1}{K}\int \frac{dy}{y} + \frac{1}{K}\int \frac{dy}{(K-y)} = \frac{r}{K}t + c_0.$$

We multiply by K the whole equation and integrate each term to get

$$\ln|y| - \ln|K-y| = rt + c_1,$$

where $c_1 = Kc_0$. We have solved the logistic differential equation and the expression above gives all the solutions in implicit form. We can work a bit on the left-side of the equation to get a nicer expression, for example,

$$\ln\Big|\frac{y}{K-y}\Big| = rt + c_1 \quad \Rightarrow \quad \Big|\frac{y}{K-y}\Big| = e^{rt+c_1} = e^{rt}\,e^{c_1},$$

which gives us

$$\frac{y}{K-y} = c_2\,e^{rt},$$

where $c_2 = \pm e^{c_1}$. The expression above is still an implicit expression of all solutions to the logistic equation. But now it is simpler to get an explicit expression, because

$$y = c_2\,e^{rt}\,(K-y) = c_2K\,e^{rt} - c_2\,e^{rt}\,y \quad \Rightarrow \quad (1 + c_2\,e^{rt})\,y = c_2K\,e^{rt},$$

which leads us to the explicit expression

$$y(t) = \frac{c_2 K\, e^{rt}}{(1 + c_2\, e^{rt})},$$

or, equivalently

$$y(t) = \frac{c_2 K}{(c_2 + e^{-rt})}.$$

We have obtained an explicit expression of all solutions of the logistic equation. Now we find the solution that satisfies the initial condition, that is,

$$y_0 = y(0) = \frac{c_2 K}{(c_2 + 1)}.$$

Some simple algebraic transformations give us

$$(1 + c_2) y_0 = c_2 K \quad \Rightarrow \quad c_2 (y_0 - K) = -y_0 \quad \Rightarrow \quad c_2 = \frac{y_0}{K - y_0}.$$

If we put this expression for c_2 into the formula for the solution $y(t)$ we get

$$y(t) = \frac{y_0 K}{(K - y_0)\big(\frac{y_0}{(K - y_0)} + e^{-rt}\big)},$$

and we arrive to the final expression

$$y(t) = \frac{y_0 K}{y_0 + (K - y_0)\, e^{-rt}}.$$

This establishes the Theorem. □

Example 1.1.14. Find the solution of the initial value problem

$$y' = 2y\left(1 - \frac{y}{3}\right), \qquad y(0) = 1.$$

Solution: We follow the calculations in the proof of Theorem 1.1.7. First we write

$$y' = \frac{2}{3} y\,(3 - y),$$

then we rewrite this in separable form

$$\frac{y'}{y\,(3 - y)} = \frac{2}{3}$$

and we integrate on both sides with respect to the variable t,

$$\int \frac{y'(t)\, dt}{y(t)\,(3 - y(t))} = \int \frac{2}{3}\, dt.$$

We integrate on the right side and we substitute $y = y(t)$, with $dy = y'(t)\, dt$, on the left side,

$$\int \frac{dy}{y\,(3 - y)} = \frac{2}{3}\, t + c_0,$$

where c_0 is an arbitrary integration constant. Now we use partial fractions to simplify the integrand on the left side. That is, we find constants a and b such that

$$\frac{1}{y(3 - y)} = \frac{a}{y} + \frac{b}{(3 - y)} \quad \Rightarrow \quad \frac{1}{y(3 - y)} = \frac{a(3 - y) + by}{y\,(3 - y)} \quad \Rightarrow \quad 1 = a(3 - y) + by.$$

From this last equation we get the values of the constants a and b, since

$$(b - a)y + (3a - 1) = 0 \quad \text{for all } y \quad \Rightarrow \quad a = \frac{1}{3}, \quad b = \frac{1}{3}.$$

Therefore, the partial fraction decomposition of the integrand above is

$$\frac{1}{y\,(3-y)} = \frac{1}{3}\,\frac{1}{y} + \frac{1}{3}\,\frac{1}{3-y},$$

which means we now have to integrate

$$\frac{1}{3}\int \frac{dy}{y} + \frac{1}{3}\int \frac{dy}{3-y} = \frac{2}{3}\,t + c_0.$$

We multiply by 3 the whole equation, introduce the new constant $c_1 = 3c_0$, and then we integrate the left side, which gives us

$$\ln|y| - \ln|3-y| = 2t + c_1.$$

This is an implicit expression of all solutions of the logistic equation. WE now simplify this expression a bit using the properties of the log function,

$$\ln\Big|\frac{y}{(3-y)}\Big| = 2t + c_1 \quad\Rightarrow\quad \Big|\frac{y}{(3-y)}\Big| = e^{2t+c_1} = e^{2t}\,e^{c_1} \quad\Rightarrow\quad \frac{y}{(3-y)} = c_2\,e^{2t},$$

where $c_2 = \pm e^{c_1}$. This is still an implicit expression for all solutions of the logistic equation, but now we are closer to an explicit expression, because

$$y = c_2\,e^{2t}(3-y) = 3\,c_2\,e^{2t} - c_2\,e^{2t}\,y \quad\Rightarrow\quad (1 + c_2\,e^{2t})\,y = 3\,c_2\,e^{2t},$$

which gives us our first explicit expression for all solutions of this logistic equation,

$$y(t) = \frac{3\,c_2\,e^{2t}}{(1 + c_2\,e^{2t})}.$$

A nicer explicit expression is

$$y(t) = \frac{3\,c_2}{(c_2 + e^{-2t})}.$$

To find the solution of the initial value problem we need to use the initial condition,

$$1 = y(0) = \frac{3\,c_2}{(c_2+1)} \quad\Rightarrow\quad c_2 + 1 = 3\,c_2 \quad\Rightarrow\quad c_2 = \frac{1}{2}.$$

Which gives us the solution of the initial value problem,

$$y(t) = \frac{3}{2(\frac{1}{2} + e^{-2t})} \quad\Rightarrow\quad y(t) = \frac{3}{(1 + 2\,e^{-2t})}.$$

◁

1.1.5. Euler Homogeneous Equations. We have seen several examples of differential equations that are not separable but they can be transformed into a separable equation by simple changes in the equation. For example the equation

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

can be written as a separable equation as follows,

$$t^2y^2 + y^2 + y' = 0 \quad\Rightarrow\quad y' = -y^2 - y^2t^2 \quad\Rightarrow\quad \frac{y'}{y^2} = -(1+t^2)$$

Sometimes a differential equation is not separable but it can be transformed into a separable equation by a more complicated transformation that involves to **change the unknown function**. This is the case for differential equations known as Euler homogenous equations.

Definition 1.1.8. *An **Euler homogeneous** differential equation has the form*

$$y'(t) = F\Big(\frac{y(t)}{t}\Big).$$

Remark:

(a) Any function F of t, y that depends only on the quotient y/t is *scale invariant*. This means that F does not change when we do the transformation $y \to cy$, $t \to ct$,

$$F\Big(\frac{(cy)}{(ct)}\Big) = F\Big(\frac{y}{t}\Big).$$

The the differential equations above are also called *scale invariant* equations.

(b) Scale invariant functions are a particular case of *homogeneous functions of degree n*, which are functions f satisfying

$$f(ct, cy) = c^n\, f(y, t).$$

Scale invariant functions are the case $n = 0$.

(c) An example of an homogeneous function is the energy of a thermodynamical system, such as a gas in a bottle. The energy, E, of a fixed amount of gas is a function of the gas entropy, S, and the gas volume, V. Such energy is an homogeneous function of degree one,

$$E(cS, cV) = c\, E(S, V), \qquad \text{for all } c \in \mathbb{R}.$$

Example 1.1.15. Show that the functions f_1, f_2 are homogeneous and find their degree,

$$f_1(t, y) = t^4 y^2 + t y^5 + t^3 y^3, \qquad f_2(t, y) = t^2 y^2 + t y^3.$$

Solution: The function f_1 is homogeneous of degree 6, since

$$f_1(ct, cy) = c^4 t^4\, c^2 y^2 + ct\, c^5 y^5 + c^3 t^3\, c^3 y^3 = c^6\, (t^4 y^2 + t y^5 + t^3 y^3) = c^6\, f(t, y).$$

Notice that the sum of the powers of t and y on every term is 6. Analogously, function f_2 is homogeneous degree 4, since

$$f_2(ct, cy) = c^2 t^2\, c^2 y^2 + ct\, c^3 y^3 = c^4\, (t^2 y^2 + t y^3) = c^4\, f_2(t, y).$$

And the sum of the powers of t and y on every term is 4. ◁

Example 1.1.16. Show that the functions below are scale invariant functions,

$$f_1(t, y) = \frac{y}{t}, \qquad f_2(t, y) = \frac{t^3 + t^2\, y + t\, y^2 + y^3}{t^3 + t\, y^2}.$$

Solution: Function f_1 is scale invariant since

$$f_1(ct, cy) = \frac{cy}{ct} = \frac{y}{t} = f_1(t, y).$$

The function f_2 is scale invariant as well, since

$$f_2(ct, cy) = \frac{c^3 t^3 + c^2 t^2\, cy + ct\, c^2 y^2 + c^3 y^3}{c^3 t^3 + ct\, c^2 y^2} = \frac{c^3 (t^3 + t^2\, y + t\, y^2 + y^3)}{c^3 (t^3 + t\, y^2)} = f_2(t, y).$$

◁

More often than not, Euler homogeneous differential equations come from a differential equation $N\,y' + M = 0$ where both N and M are homogeneous functions of the same degree.

Theorem 1.1.9. *If the functions N, M, of t, y, are homogeneous of the same degree, then the differential equation*

$$N(t,y)\,y'(t) + M(t,y) = 0$$

is can be transformed into the Euler homogeneous equation

$$y' = F\Big(\frac{y}{t}\Big), \quad \textit{with} \quad F\Big(\frac{y}{t}\Big) = -\frac{M(1,(y/t))}{N(1,(y/t))}.$$

Proof of Theorem 1.1.9: Rewrite the equation as

$$y'(t) = -\frac{M(t,y)}{N(t,y)},$$

The function $f(y,y) = -\dfrac{M(t,y)}{N(t,y)}$ is scale invariant, because

$$f(ct,cy) = -\frac{M(ct,cy)}{N(ct,cy)} = -\frac{c^n\,M(t,y)}{c^n\,N(t,y)} = -\frac{M(t,y)}{N(t,y)} = f(t,y),$$

where we used that M and N are homogeneous of the same degree n. We now find a function F such that the differential equation can be written as

$$y' = F\Big(\frac{y}{t}\Big).$$

Since M and N are homogeneous degree n, we multiply the differential equation by "1" in the form $(1/t)^n/(1/t)^n$, as follows,

$$y'(t) = -\frac{M(t,y)}{N(t,y)}\,\frac{(1/t^n)}{(1/t^n)} = -\frac{M((t/t),(y/t))}{N((t/t),(y/t))} = -\frac{M(1,(y/t))}{N(1,(y/t))} \quad \Rightarrow \quad y' = F\Big(\frac{y}{t}\Big),$$

where

$$F\Big(\frac{y}{t}\Big) = -\frac{M(1,(y/t))}{N(1,(y/t))}.$$

This establishes the Theorem. □

Example 1.1.17. Show that $(t-y)\,y' - 2y + 3t + \dfrac{y^2}{t} = 0$ is an Euler homogeneous equation.

Solution: Rewrite the equation in the normal form

$$(t-y)\,y' = 2y - 3t - \frac{y^2}{t} \quad \Rightarrow \quad y' = \frac{\Big(2y - 3t - \dfrac{y^2}{t}\Big)}{(t-y)}.$$

So the function f in this case is given by

$$f(t,y) = \frac{\Big(2y - 3t - \dfrac{y^2}{t}\Big)}{(t-y)}.$$

This function is scale invariant, since numerator and denominator are homogeneous of the same degree, $n = 1$ in this case,

$$f(ct,cy) = \frac{\Big(2cy - 3ct - \dfrac{c^2y^2}{ct}\Big)}{(ct-cy)} = \frac{c\Big(2y - 3t - \dfrac{y^2}{t}\Big)}{c(t-y)} = f(t,y).$$

So, the differential equation is Euler homogeneous. We now write the equation in the form $y' = F(y/t)$. Since the numerator and denominator are homogeneous of degree $n = 1$, we multiply them by "1" in the form $(1/t)^1/(1/t)^1$, that is

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

Distribute the factors $(1/t)$ in numerator and denominator, and we get

$$y' = \frac{\big(2(y/t) - 3 - (y/t)^2\big)}{(1 - (y/t))} \quad \Rightarrow \quad y' = F\Big(\frac{y}{t}\Big),$$

where

$$F\Big(\frac{y}{t}\Big) = \frac{\big(2(y/t) - 3 - (y/t)^2\big)}{(1 - (y/t))}.$$

So, the equation is Euler homogeneous and it is written in the standard form. ◁

Example 1.1.18. Determine whether the equation $(1 - y^3)\, y' = t^2$ is Euler homogeneous.

Solution: If we write the differential equation in the normal form, $y' = f(t, y)$, then we get $f(t, y) = \dfrac{t^2}{1 - y^3}$. But

$$f(ct, cy) = \frac{c^2 t^2}{1 - c^3 y^3} \neq f(t, y),$$

hence the equation is not Euler homogeneous. ◁

1.1.6. Solving Euler Homogeneous Equations.

Theorem 1.1.10 transforms an Euler homogeneous equation into a separable equation, which we know how to solve.

Theorem 1.1.10. *The Euler homogeneous equation*

$$y' = F\Big(\frac{y}{t}\Big)$$

for the function y determines a separable equation for $v = y/t$, given by

$$\frac{v'}{\big(F(v) - v\big)} = \frac{1}{t}.$$

Remark: The original homogeneous equation for the function y is transformed into a separable equation for the unknown function $v = y/t$. One solves for v, in implicit or explicit form, and then transforms back to $y = t\, v$.

Proof of Theorem 1.1.10: Introduce the function $v = y/t$ into the differential equation,

$$y' = F(v).$$

We still need to replace y' in terms of v. This is done as follows,

$$y(t) = t\, v(t) \quad \Rightarrow \quad y'(t) = \big(t\, v(t)\big)' = v(t) + t\, v'(t).$$

Introducing these expressions into the differential equation for y we get

$$v + t\, v' = F(v) \quad \Rightarrow \quad v' = \frac{\big(F(v) - v\big)}{t} \quad \Rightarrow \quad \frac{v'}{\big(F(v) - v\big)} = \frac{1}{t}.$$

The equation on the far right is separable. This establishes the Theorem. □

Example 1.1.19. Find all solutions y of the differential equation

$$y' = \frac{t^2 + 3y^2}{2ty}, \qquad t > 0.$$

Solution: The equation is Euler homogeneous, since

$$f(ct, cy) = \frac{c^2t^2 + 3c^2y^2}{2(ct)(cy)} = \frac{c^2(t^2 + 3y^2)}{c^2(2ty)} = \frac{t^2 + 3y^2}{2ty} = f(t, y).$$

Next we compute the function F. Since the numerator and denominator are homogeneous degree "2" we multiply the right-hand side of the equation by "1" in the form $(1/t^2)/(1/t^2)$,

$$y' = \frac{(t^2 + 3y^2)}{2ty} \frac{\left(\frac{1}{t^2}\right)}{\left(\frac{1}{t^2}\right)} \quad \Rightarrow \quad y' = \frac{1 + 3\left(\frac{y}{t}\right)^2}{2\left(\frac{y}{t}\right)}.$$

Now we introduce the change of functions $v = y/t$,

$$y' = \frac{1 + 3v^2}{2v}.$$

Since $y = t\,v$, then $y' = v + t\,v'$, which implies

$$v + t\,v' = \frac{1 + 3v^2}{2v} \quad \Rightarrow \quad t\,v' = \frac{1 + 3v^2}{2v} - v = \frac{1 + 3v^2 - 2v^2}{2v} = \frac{1 + v^2}{2v}.$$

We obtained the separable equation

$$v' = \frac{1}{t}\left(\frac{1 + v^2}{2v}\right).$$

We rewrite and integrate it,

$$\frac{2v}{1 + v^2}\, v' = \frac{1}{t} \quad \Rightarrow \quad \int \frac{2v}{1 + v^2}\, v'\, dt = \int \frac{1}{t}\, dt + c_0.$$

The substitution $u = 1 + v^2(t)$ implies $du = 2v(t)\, v'(t)\, dt$, so

$$\int \frac{du}{u} = \int \frac{dt}{t} + c_0 \quad \Rightarrow \quad \ln(u) = \ln(t) + c_0 \quad \Rightarrow \quad u = e^{\ln(t) + c_0} = e^{\ln(t)} e^{c_0}.$$

We denote $c_1 = e^{c_0}$, then $u = c_1 t$. So, we get

$$1 + v^2 = c_1 t \quad \Rightarrow \quad 1 + \left(\frac{y}{t}\right)^2 = c_1 t \quad \Rightarrow \quad y(t) = \pm t\,\sqrt{c_1 t - 1}.$$

◁

Example 1.1.20. Find all solutions y of the differential equation

$$y' = \frac{t(y + 1) + (y + 1)^2}{t^2}, \qquad t > 0.$$

Solution: This equation is Euler homogeneous when written in terms of the unknown $u(t) = y(t) + 1$ and the variable t. Indeed, $u' = y'$, thus we obtain

$$y' = \frac{t(y + 1) + (y + 1)^2}{t^2} \quad \Leftrightarrow \quad u' = \frac{tu + u^2}{t^2} \quad \Leftrightarrow \quad u' = \frac{u}{t} + \left(\frac{u}{t}\right)^2.$$

Therefore, we introduce the new variable $v = u/t$, which satisfies $u = t\,v$ and $u' = v + t\,v'$. The differential equation for v is

$$v + t\,v' = v + v^2 \quad \Leftrightarrow \quad t\,v' = v^2 \quad \Leftrightarrow \quad \int \frac{v'}{v^2}\, dt = \int \frac{1}{t}\, dt + c,$$

with $c \in \mathbb{R}$. The substitution $w = v(t)$ implies $dw = v'\,dt$, so

$$\int w^{-2}\,dw = \int \frac{1}{t}\,dt + c \quad \Leftrightarrow \quad -w^{-1} = \ln(t) + c \quad \Leftrightarrow \quad w = -\frac{1}{\ln(t) + c}.$$

Substituting back v, u and y, we obtain $w = v(t) = u(t)/t = [y(t) + 1]/t$, so

$$\frac{y+1}{t} = -\frac{1}{\ln(t) + c} \quad \Leftrightarrow \quad y(t) = -\frac{t}{\ln(t) + c} - 1.$$

$\lhd$

Notes. This section corresponds to Boyce-DiPrima [4] Section 2.2. Zill and Wright study separable equations in [13] Section 2.2, and Euler homogeneous equations in Section 2.5. Zill and Wright organize the material in a nice way, they present first separable equations, then linear equations, and then they group Euler homogeneous and Bernoulli equations in a section called Solutions by Substitution. Once again, a one page description is given by Simmons in [8] in Chapter 2, Section 7.

1.1.7. Exercises.

1.1.1.- Find an explicit expression of all solutions y to the equation
$$y' = \frac{t^2}{y}.$$

1.1.2.- Find an implicit expression of all solutions y to the equation
$$3t^2 + 4y^3 y' - 1 + y' = 0.$$

1.1.3.- Find the solution y to the initial value problem
$$y' = t^2 y^2, \quad y(0) = 1.$$

1.1.4.- Find all solutions y of the equation
$$ty + \sqrt{1+t^2}\, y' = 0.$$

1.1.5.- Find every solution y of the Euler homogeneous equation
$$y' = \frac{y+t}{t}.$$

1.1.6.- Find all solutions y to the equation
$$y' = \frac{t^2 + y^2}{ty}.$$

1.1.7.- Find the explicit solution to the initial value problem
$$(t^2 + 2ty)\, y' = y^2, \quad y(1) = 1.$$

1.1.8.- Prove that if $y' = f(t, y)$ is an Euler homogeneous equation and $y_1(t)$ is a solution, then
$$y(t) = \frac{1}{k}\, y_1(kt)$$
is also a solution for every non-zero $k \in \mathbb{R}$.

1.2. Linear Equations

Linear differential equations look simpler than separable equations, but they could be more complicated to solve. There are linear equations that cannot be transformed into separable equations, meaning their solutions cannot be obtained by simply integrating on both sides of the equation with respect to the independent variable. It turns out we need a new idea to solve all linear equations. In the first part of this section we study one of such new ideas called the integrating factor method.

In the second part of this section we show a simple physical system that can be described by a linear differential equation. The physical system is a tank containing salty water and where fresh or salty water can come in and out at the same or different rates. This type of physical systems are called Mixing Problems.

In the last part of this section we turn our attention to a particular *nonlinear* differential equation—the Bernoulli equation. This nonlinear equation has a particular property: it can be transformed into a linear equation by an appropriate *change of the unknown function*. Then, one solves the linear equation for the changed function using the integrating factor method. The last step is to transform the changed function back into the original function.

1.2.1. Constant Coefficients.

We start with a precise definition of linear differential equations, and then we focus on a particular type of linear equations—those with constant coefficients.

Definition 1.2.1. *A* ***linear non-homogenous*** *differential equation on the function y is*

$$y' = a(t)\, y + b(t) \tag{1.2.1}$$

The equation is called ***linear****, also linear homogeneous, if $b(t) = 0$ for all t. The equation has* ***constant coefficients*** *if both a and b are constants, otherwise we say that the equation has* ***variable coefficients****.*

Examples of linear equations with constant coefficients are population models with unlimited food resources.

Example 1.2.1 (Exponential Population Model with Immigration)**.** Consider a population of rabbits in a region with *unlimited* food resources. Suppose that the growth rate is proportional to the population with a proportionality factor of 2 per month and there is an immigration rate of 25 rabbits per month. Find the differential equation that describes this population of rabbits.

Solution: The population of rabbits has unlimited food resources, meaning that the rate of change of the rabbits population, $y'(t)$, must be proportional to the actual rabbit population, $y(t)$. We also have immigration, so we must add a constant term into $y'(t)$ with the immigration information. Therefore, the differential equation for the rabbits population is

$$y'(t) = a\, y(t) + b,$$

where $y(t)$ is the rabbit population at the time t, a is the growth rate per capita coefficient, and b is the immigration rate coefficient. So in this case we have

$$a = 2, \qquad b = 25.$$

So the equation describing the rabbit population under the assumptions of this example is

$$y'(t) = 2\, y(t) + 25.$$

◁

Our next result is a formula for the solutions of all linear differential equations with constant coefficients.

Theorem 1.2.2 (Constant Coefficients)**.** *The linear non-homogeneous equation*

$$y' = a\,y + b \tag{1.2.2}$$

with $a \neq 0$, b constants, has infinitely many solutions,

$$y(t) = c\,e^{at} - \frac{b}{a}, \qquad c \in \mathbb{R}. \tag{1.2.3}$$

Furthermore, the initial value problem

$$y' = a\,y + b, \qquad y(0) = y_0, \tag{1.2.4}$$

has a unique solution given by

$$y(t) = \Big(y_0 + \frac{b}{a}\Big)\,e^{at} - \frac{b}{a}. \tag{1.2.5}$$

Remarks:

(a) The expression in Eq. (1.2.3) is called the *general solution* of the differential equation.

(b) We prove the Theorem *transforming the linear equation into a separable equation* and integrating on both sides with respect to t. If we do not transform the equation into separable form and we just integrate with respect to t the linear equation itself, Eq. (1.2.2), then we cannot solve the problem. Indeed,

$$\int y'(t)\,dt = a \int y(t)\,dt + bt + c, \quad c \in \mathbb{R}.$$

The Fundamental Theorem of Calculus implies $y(t) = \int y'(t)\,dt$, so we get

$$y(t) = a \int y(t)\,dt + bt + c.$$

The equation above is not a solution given in implicit form, because we still need to find a primitive of y. We have only rewritten the original differential equation as an integral equation. Simply integrating both sides of a linear equation does not solve the equation.

Proof of Theorem 1.2.2: Rewrite the differential equation as a separable equation,

$$\frac{y'(t)}{a\,y(t) + b} = 1 \quad \Rightarrow \quad \frac{1}{a}\int \frac{y'(t)\,dt}{y(t) + (b/a)} = \int dt.$$

In the left-hand side introduce the substitution

$$\left.\begin{cases} u = y + (b/a) \\ du = y'\,dt. \end{cases}\right\} \Rightarrow \frac{1}{a}\int \frac{du}{u} = \int dt.$$

The integral is is now simple to find,

$$\frac{1}{a}\,\ln(|u|) = t + c_0 \quad \Rightarrow \quad \ln|y + (b/a)| = at + c_1, \qquad c_1 = a\,c_0.$$

Compute exponentials on both sides above, and recall that $e^{(a_1 + a_2)} = e^{a_1}\,e^{a_2}$,

$$|y + (b/a)| = e^{at + c_1} = e^{at}\,e^{c_1}.$$

We take out the absolute value,

$$y(t) + (b/a) = (\pm e^{c_1})\, e^{at} \quad \Rightarrow \quad y(t) = c_2\, e^{at} - \frac{b}{a}.$$

where $c_2 = (\pm e^{c_1})$. We know that at $t = 0$ we have $y(0) = y_0$, so

$$y_0 = y(0) = c_2 - \frac{b}{a} \quad \Rightarrow \quad c_2 = y_0 + \frac{b}{a}.$$

So the solution to the initial value problem is

$$y(t) = \Big(y_0 + \frac{b}{a}\Big)\, e^{at} - \frac{b}{a}.$$

This establishes the Theorem. □

Now we solve the exponential growth equation with immigration from Example 1.2.1.

Example 1.2.2 (Exponential Population Model with Immigration)**.** Solve the differential equation that describes the rabbit population in Example 1.2.1 knowing that the initial rabbit population is 100 rabbits.

Solution: We follow the calculation in the proof of Theorem 1.2.2. The population of rabbits has unlimited food resources and immigration, so the differential equation for such system is

$$y'(t) = a\, y(t) + b,$$

where $y(t)$ is the rabbit population at the time t, a is the growth rate per capita coefficient, and b is the immigration rate coefficient. So in this case we have

$$a = 2, \qquad b = 25.$$

We have seen in previous sections how to solve this differential equation, but we do it again here. Rewrite the differential equation as a separable equation,

$$\frac{y'(t)}{2\, y(t) + 25} = 1 \quad \Rightarrow \quad \frac{1}{2} \int \frac{y'(t)\, dt}{y(t) + (25/2)} = \int dt.$$

In the left-hand side introduce the substitution

$$\left\{ \begin{aligned} u &= y + (25/2) \\ du &= y'\, dt. \end{aligned} \right\} \Rightarrow \frac{1}{2} \int \frac{du}{u} = \int dt.$$

The integral is is now simple to find,

$$\frac{1}{2} \ln(|u|) = t + c_0 \quad \Rightarrow \quad \ln|y + (25/2)| = 2t + c_1, \qquad c_1 = 2\, c_0.$$

Compute exponentials on both sides above, and recall that $e^{(a_1 + a_2)} = e^{a_1}\, e^{a_2}$,

$$|y + (25/2)| = e^{2t + c_1} = e^{2t}\, e^{c_1}.$$

We take out the absolute value,

$$y(t) + (25/2) = (\pm e^{c_1})\, e^{2t} \quad \Rightarrow \quad y(t) = c_2\, e^{2t} - \frac{25}{2}.$$

where $c_2 = (\pm e^{c_1})$. We know that at $t = 0$ we have 100 rabbits, so,

$$100 = y(0) = c_2 - \frac{25}{2} \quad \Rightarrow \quad c_2 = 100 + \frac{25}{2} = \frac{225}{2}.$$

So the solution to our problem is

$$y(t) = \frac{225}{2}\, e^{2t} - \frac{25}{2}.$$

◁

1.2.2. Newton's Cooling Law. Our next example is about the *Newton's cooling law*, which says that the temperature T at a time t of a material placed in a surrounding medium kept at a constant temperature T_s satisfies the linear non-homogeneous differential equation

$$(\Delta T)' = -k\,(\Delta T),$$

where $\Delta T(t) = T(t) - T_s$ is the temperature difference between the material's temperature and the surroundings, while $k > 0$ is a constant that characterizes the material's thermal properties.

Remarks:

(a) Although Newton's law is called a "Cooling Law", it also describes objects that warm up. When the initial temperature difference, $(\Delta T)(0) = T(0) - T_s$ is positive the object cools down, but when $(\Delta T)(0)$ is negative the object is warms up.

(b) Newton's cooling law for ΔT is the linear differential equation we called the radioactive decay equation. But now $(\Delta T)(0)$ can be either positive or negative.

(c) The solution of Newton's cooling law equation is

$$(\Delta T)(t) = (\Delta T)(0)\, e^{-kt},$$

for some initial temperature difference $(\Delta T)(0) = T(0) - T_s$. Since $(\Delta T)(t) = T(t) - T_s$,

$$T(t) = (T(0) - T_s)\, e^{-kt} + T_s.$$

Example 1.2.3 (Newton's Cooling Law)**.** A cup with water at 45 C is placed in the cooler held at 5 C. If after 2 minutes the water temperature is 25 C, when will the water temperature be 15 C?

Solution: The differential equation satisfied by the temperature difference is

$$\Delta T' = -k\,\Delta T.$$

The solution to the Newton cooling law equation is

$$(\Delta T)(t) = (\Delta T)(0)\, e^{-kt} \quad \Rightarrow \quad T(t) = (T(0) - T_s)\, e^{-kt} + T_s.$$

We know that in this case we have

$$T(0) = 45, \qquad T_s = 5, \qquad T(2) = 25.$$

Notice that we do not have the constant k yet. The problem is to find the time t_1 such that $T(t_1) = 15$. In order to find that t_1 we first need to find the constant k,

$$T(t) = (45 - 5)\, e^{-kt} + 5 \quad \Rightarrow \quad T(t) = 40\, e^{-kt} + 5.$$

Now use the fact that $T(2) = 25$ C, that is,

$$20 = T(2) = 40\, e^{-2k} \quad \Rightarrow \quad \ln(1/2) = -2k \quad \Rightarrow \quad k = \frac{1}{2}\ln(2).$$

Having the constant k we can now go on and find the time t_1 such that $T(t_1) = 15$ C.

$$T(t) = 40\, e^{-t\ln(\sqrt{2})} + 5 \quad \Rightarrow \quad 10 = 40\, e^{-t_1\ln(\sqrt{2})} \quad \Rightarrow \quad t_1 = 4.$$

◁

1.2.3. Variable Coefficients. We have solved constant coefficient linear equations by transforming them into separable equations, and then integrating with respect to the independent variable. Linear equations with *variable coefficients* are also separable in the case that $b = 0$ (homogenous equations) or b/a is constant (a and b are proportional).

Example 1.2.4 (Variable Coefficient Case $b = 0$)**.** Find all solutions of the equation

$$y' = a(t)\, y.$$

Solution: We transform the linear equation into a separable equation and we integrate,

$$\frac{y'}{y} = a(t) \quad \Rightarrow \quad \ln(|y|)' = a(t) \quad \Rightarrow \quad \ln(|y(t)|) = A(t) + c_0,$$

where $A = \int a\, dt$, is a primitive or antiderivative of a. Therefore,

$$y(t) = \pm e^{A(t)+c_0} = \pm e^{A(t)}\, e^{c_0},$$

so we get the solution $y(t) = c\, e^{A(t)}$, where $c = \pm e^{c_0}$. ◁

Example 1.2.5. Find all solutions of

$$y' = t^2\, y.$$

Solution: We use the result of the previous example with $a(t) = t^2$, which gives

$$A(t) = \frac{t^3}{3}$$

then the solutions of the differential equations are

$$y(t) = c\, e^{t^3/3}, \qquad c \in \mathbb{R}.$$

◁

The linear equation $y' = a(t)\, y + b(t)$ is not separable in the case that $b(t)/a(t)$ is not constant. We need a new idea to solve these equations. We now introduce the *integrating factor method*, which is an idea that works well to solve all linear equations, including the linear equations with variable coefficients.

Theorem 1.2.3 (Variable Coefficients)**.** *If a, b are continuous on (t_1, t_2), then*

$$y' = a(t)\, y + b(t), \tag{1.2.6}$$

has infinitely many solutions on (t_1, t_2) given by

$$y(t) = c\, e^{A(t)} + e^{A(t)} \int e^{-A(t)}\, b(t)\, dt, \tag{1.2.7}$$

where $A(t) = \int a(t)\, dt$ is any antiderivative of the function a and $c \in \mathbb{R}$. Furthermore, for any $t_0 \in (t_1, t_2)$ and $y_0 \in \mathbb{R}$ the initial value problem (IVP)

$$y' = a(t)\, y + b(t), \qquad y(t_0) = y_0, \tag{1.2.8}$$

has the unique solution y on the same domain (t_1, t_2), given by

$$y(t) = y_0\, e^{\hat{A}(t)} + e^{\hat{A}(t)} \int_{t_0}^{t} e^{-\hat{A}(s)}\, b(s)\, ds, \tag{1.2.9}$$

where the function $\hat{A}(t) = \int_{t_0}^{t} a(s)\, ds$ is a particular antiderivative of function a.

Remarks:

(a) The expression in Eq. (1.2.7) is called the *general solution* of the differential equation.
(b) The function $\mu(t) = e^{-A(t)}$ is called an *integrating factor* of the equation.

Example 1.2.6 (Consistency with Constant Coefficients General Solution)**.** Show that for constant coefficient equations the solution1 given in Eq. (1.2.7) reduces to Eq. (1.2.3).

Solution: In the particular case of constant coefficient equations, a primitive, or antiderivative, for the constant function a is $A(t) = at$, so

$$y(t) = c\, e^{at} + e^{at} \int e^{-at}\, b\, dt.$$

Since b is constant, the integral in the second term above can be computed explicitly,

$$e^{at} \int b\, e^{-at}\, dt = e^{at}\Big(-\frac{b}{a}\, e^{-at}\Big) = -\frac{b}{a}.$$

Therefore, in the case of a, b constants we obtain $y(t) = c\, e^{at} - \dfrac{b}{a}$ given in Eq. (1.2.3). ◁

Example 1.2.7 (Consistency with Constant Coefficients IVP)**.** Show that for constant coefficient equations the solution formula given in Eq. (1.2.9) reduces to Eq. (1.2.5).

Solution: In the particular case of a constant coefficient equation, where $a, b \in \mathbb{R}$, the solution given in Eq. (1.2.9) reduces to the one given in Eq. (1.2.5). Indeed,

$$\hat{A}(t) = \int_{t_0}^{t} a\, ds = a(t - t_0), \qquad \int_{t_0}^{t} e^{-a(s-t_0)}\, b\, ds = -\frac{b}{a}\, e^{-a(t-t_0)} + \frac{b}{a}.$$

Therefore, the solution y can be written as

$$y(t) = y_0\, e^{a(t-t_0)} + e^{a(t-t_0)}\Big(-\frac{b}{a}\, e^{-a(t-t_0)} + \frac{b}{a}\Big) = \Big(y_0 + \frac{b}{a}\Big)\, e^{a(t-t_0)} - \frac{b}{a}.$$

◁

We now prove Theorem 1.2.3. The proof is based on the integrating factor method. We find a special function $\mu(t)$ such that the linear differential equation multiplied by this function is a total derivative, which are simple to integrate.

Proof of Theorem 1.2.3: Write the differential equation with y on one side only,

$$y' - a\, y = b,$$

and then multiply the differential equation by a function μ, called an integrating factor,

$$\mu\, y' - a\, \mu\, y = \mu\, b. \tag{1.2.10}$$

The critical step is to choose a function μ such that

$$-\, a\, \mu = \mu'. \tag{1.2.11}$$

If μ is solution of Eq. (1.2.11), then the differential equation in (1.2.10) has the form

$$\mu\, y' + \mu'\, y = \mu\, b.$$

But the left-hand side is a total derivative of a product of two functions,

$$\big(\mu\, y\big)' = \mu\, b. \tag{1.2.12}$$

This is the property we want in an integrating factor, μ. We want to find a function μ such that the left-hand side of the differential equation for y can be written as a total derivative, just as in Eq. (1.2.12). We need to find just one of such functions μ. So we go back to Eq. (1.2.11), the differential equation for μ, which is separable, so we know how to solve it,

$$\mu' = -a\,\mu \quad \Rightarrow \quad \frac{\mu'}{\mu} = -a \quad \Rightarrow \quad \ln(|\mu|)' = -a \quad \Rightarrow \quad \ln(|\mu|) = -A + c_0,$$

where $A = \int a\,dt$, a primitive or antiderivative of a, and c_0 is an arbitrary constant. Computing the exponential of both sides we get

$$\mu = \pm e^{c_0}\, e^{-A} \quad \Rightarrow \quad \mu = c_1\, e^{-A}, \qquad c_1 = \pm e^{c_0}.$$

Since c_1 is a constant which will cancel out from Eq. (1.2.10) anyway, we choose the integration constant $c_0 = 0$, hence $c_1 = 1$. The integrating factor is then

$$\mu(t) = e^{-A(t)}.$$

This function is an integrating factor, because if we start again at Eq. (1.2.10), we get

$$e^{-A}\,y' - a\,e^{-A}\,y = e^{-A}\,b \quad \Rightarrow \quad e^{-A}\,y' + \big(e^{-A}\big)'\,y = e^{-A}\,b,$$

where we used the main property of the integrating factor, $-a\,e^{-A} = \big(e^{-A}\big)'$. Now the product rule for derivatives implies that the left-hand side above is a total derivative,

$$\big(e^{-A}\,y\big)' = e^{-A}\,b.$$

Integrating on both sides we get

$$\big(e^{-A}\,y\big) = \int e^{-A}\,b\,dt + c \quad \Rightarrow \quad \big(e^{-A}\,y\big) - \int e^{-A}\,b\,dt = c.$$

The function $\psi(t,y) = \big(e^{-A}\,y\big) - \int e^{-A}\,b\,dt$ is called a *potential function* of the differential equation. The solution of the differential equation can be computed from the second equation above, $\psi = c$, and the result is

$$y(t) = c\,e^{A(t)} + e^{A(t)} \int e^{-A(t)}\,b(t)\,dt.$$

This establishes the first part of the Theorem. For the furthermore part, let us use the notation $K(t) = \displaystyle\int e^{-A(t)}\,b(t)\,dt$, and then introduce the initial condition in (1.2.8), which fixes the constant c in the general solution,

$$y_0 = y(t_0) = c\,e^{A(t_0)} + e^{A(t_0)} K(t_0).$$

So we get the constant c,

$$c = y_0\,e^{-A(t_0)} - K(t_0).$$

Using this expression in the general solution above,

$$y(t) = \Big(y_0\,e^{-A(t_0)} - K(t_0)\Big)\,e^{A(t)} + e^{A(t)}\,K(t) = y_0\,e^{A(t)-A(t_0)} + e^{A(t)}\big(K(t) - K(t_0)\big).$$

Let us introduce the particular primitives $\hat{A}(t) = A(t) - A(t_0)$ and $\hat{K}(t) = K(t) - K(t_0)$, which vanish at t_0, that is,

$$\hat{A}(t) = \int_{t_0}^{t} a(s)\,ds, \qquad \hat{K}(t) = \int_{t_0}^{t} e^{-A(s)}\,b(s)\,ds.$$

Then the solution y of the IVP has the form

$$y(t) = y_0\,e^{\hat{A}(t)} + e^{A(t)} \int_{t_0}^{t} e^{-A(s)}\,b(s)\,ds$$

which is equivalent to

$$y(t) = y_0\, e^{\hat{A}(t)} + e^{A(t)-A(t_0)} \int_{t_0}^{t} e^{-(A(s)-A(t_0))}\, b(s)\, ds,$$

so we conclude that

$$y(t) = y_0\, e^{\hat{A}(t)} + e^{\hat{A}(t)} \int_{t_0}^{t} e^{-\hat{A}(s)}\, b(s)\, ds.$$

This establishes the IVP part of the Theorem. □

In the following examples we show the integrating factor method used in the proof of the Theorem 1.2.3 above to find solutions to linear equations with variable coefficients.

Example 1.2.8 (The Integrating Factor Method)**.** Find all solutions y of the equation

$$y' = \frac{3}{t}\, y + t^5, \qquad t > 0.$$

Solution: Rewrite the equation with y on only one side,

$$y' - \frac{3}{t}\, y = t^5.$$

Multiply the differential equation by a function μ, which we determine later,

$$\mu(t)\Big(y' - \frac{3}{t}\, y\Big) = t^5\, \mu(t) \quad \Rightarrow \quad \mu(t)\, y' - \frac{3}{t}\, \mu(t)\, y = t^5\, \mu(t).$$

We need to choose a positive function μ having the following property,

$$-\frac{3}{t}\, \mu(t) = \mu'(t) \quad \Rightarrow \quad -\frac{3}{t} = \frac{\mu'(t)}{\mu(t)} \quad \Rightarrow \quad -\frac{3}{t} = \big(\ln(|\mu|)\big)'$$

Integrating,

$$\ln(|\mu|) = -\int \frac{3}{t}\, dt = -3\, \ln(|t|) + c_0 = \ln(|t|^{-3}) + c_0 \quad \Rightarrow \quad \mu = (\pm e^{c_0})\, e^{\ln(|t|^{-3})},$$

so, $\mu = (\pm e^{c_0})\, |t|^{-3}$. Rename $c_1 = \pm e^{c_0}$,

$$\mu(t) = c_1\, |t|^{-3}.$$

The absolute value is not needed since $t > 0$, and we need only one integrating factor, so we choose $c_1 = 1$ and we get

$$\mu = t^{-3}.$$

Back to the differential equation for y, we multiply it by the integrating factor above,

$$t^{-3}\Big(y' - \frac{3}{t}\, y\Big) = t^{-3}\, t^5 \quad \Rightarrow \quad t^{-3}\, y' - 3\, t^{-4}\, y = t^2.$$

Using that $-3\, t^{-4} = (t^{-3})'$ and $t^2 = \Big(\dfrac{t^3}{3}\Big)'$, we get

$$t^{-3}\, y' + (t^{-3})'\, y = \Big(\frac{t^3}{3}\Big)' \quad \Rightarrow \quad \big(t^{-3}\, y\big)' = \Big(\frac{t^3}{3}\Big)' \quad \Rightarrow \quad \Big(t^{-3}\, y - \frac{t^3}{3}\Big)' = 0.$$

This last equation is a total derivative of a potential function $\psi(t,y) = t^{-3}\, y - \dfrac{t^3}{3}$. Since the equation is a total derivative, this confirms that we got a correct integrating factor. Now we integrate the total derivative, which is simple to do,

$$t^{-3}\, y - \frac{t^3}{3} = c \quad \Rightarrow \quad t^{-3}\, y = c + \frac{t^3}{3} \quad \Rightarrow \quad y(t) = c\, t^3 + \frac{t^6}{3},$$

where c is an arbitrary constant. ◁

Example 1.2.9 (The Integrating Factor Method)**.** Find all solutions y of the equation

$$ty' + 2y = 4t^2, \qquad \text{with} \quad t > 0.$$

Solution: Rewrite the equation in normal form,

$$y' = -\frac{2}{t}\,y + 4t \quad \Leftrightarrow \quad a(t) = -\frac{2}{t}, \quad b(t) = 4t. \tag{1.2.13}$$

Rewrite again,

$$y' + \frac{2}{t}\,y = 4t.$$

Multiply by a function μ,

$$\mu\, y' + \frac{2}{t}\,\mu\, y = \mu\, 4t.$$

Choose the function μ to be solution of

$$\frac{2}{t}\,\mu = \mu' \quad \Rightarrow \quad \ln(|\mu|)' = \frac{2}{t} \quad \Rightarrow \quad \ln(|\mu|) = 2\ln(|t|) + c_0 = \ln(t^2) + c_0.$$

From here we get

$$\mu(t) = (\pm e^{c_0})\, t^2, \qquad c_1 = \pm e^{c_0} \quad \Rightarrow \quad \mu(t) = c_1\, t^2.$$

We choose the integrating factor $\mu = t^2$. Multiply the differential equation by this μ,

$$t^2\, y' + 2t\, y = 4t\, t^2 \quad \Rightarrow \quad (t^2\, y)' = 4t^3.$$

If we write the right-hand side also as a derivative,

$$\big(t^2\, y\big)' = \big(t^4\big)' \quad \Rightarrow \quad \big(t^2\, y - t^4\big)' = 0.$$

So a potential function is $\psi(t, y(t)) = t^2\, y(t) - t^4$. Integrating on both sides we obtain

$$t^2\, y - t^4 = c \quad \Rightarrow \quad t^2\, y = c + t^4 \quad \Rightarrow \quad y(t) = \frac{c}{t^2} + t^2.$$

◁

In the next example we show how to solve an initial value problem for a linear variable coefficient equation.

Example 1.2.10 (IVP)**.** Find the function y solution of the initial value problem

$$ty' + 2y = 4t^2, \qquad t > 0, \qquad y(1) = 2.$$

Solution: In Example 1.2.9 we computed the general solution of the differential equation,

$$y(t) = \frac{c}{t^2} + t^2, \qquad c \in \mathbb{R}.$$

The initial condition implies that

$$2 = y(1) = c + 1 \quad \Rightarrow \quad c = 1 \quad \Rightarrow \quad y(t) = \frac{1}{t^2} + t^2.$$

◁

In our next example we solve again the problem in Example 1.2.10 but this time we use Eq. (1.2.9).

Example 1.2.11 (IVP)**.** Find the solution of the problem given in Example 1.2.10, but this time using Eq. (1.2.9) in Theorem 1.2.3.

Solution: We find the solution simply by using Eq. (1.2.9). First, find the integrating factor function μ as follows:

$$A(t) = -\int_1^t \frac{2}{s}\,ds = -2\big[\ln(t) - \ln(1)\big] = -2\ln(t),$$

that is,

$$A(t) = \ln(t^{-2}).$$

The integrating factor is $\mu(t) = e^{-A(t)}$, that is,

$$\mu(t) = e^{-\ln(t^{-2})} = e^{\ln(t^2)}.$$

From here we get the integrating factor

$$\mu(t) = t^2.$$

Note that Eq. (1.2.9) contains $e^{A(t)} = 1/\mu(t)$. Then, compute the solution as follows,

$$\begin{aligned} y(t) &= \frac{1}{t^2}\Big(2 + \int_1^t s^2\,4s\,ds\Big) \\ &= \frac{2}{t^2} + \frac{1}{t^2}\int_1^t 4s^3 ds \\ &= \frac{2}{t^2} + \frac{1}{t^2}(t^4 - 1) \\ &= \frac{2}{t^2} + t^2 - \frac{1}{t^2}. \end{aligned}$$

Therefore, the solution of the initial value problem is

$$y(t) = \frac{1}{t^2} + t^2.$$

◁

1.2.4. Mixing Problems. Consider a tank containing salty water as pictured in Fig. 1, where salty water comes in and goes out of the tank. The amount of water in the tank at a time t is proportional to the water volume, $V(t)$, while the amount of salt dissolved in the water at that time is given by $Q(t)$. Water is pouring into the tank at a rate $r_i(t)$ with a salt concentration $q_i(t)$. Water is also leaving the tank at a rate $r_o(t)$ with a salt concentration $q_o(t)$. Recall that a water rate, r, means water volume per unit time, and a salt concentration, q, means salt mass per unit volume. If we denote by $[r_i]$ the units of the quantity r_i, then we have

$$[V] = \text{Volume}, \quad [Q] = \text{Mass}, \quad [r_i] = [r_o] = \frac{\text{Volume}}{\text{Time}}, \quad [q_i] = [q_o] = \frac{\text{Mass}}{\text{Volume}}.$$

We want to write a mathematical model to describe how the water volume and salt mass change in time. We make one important assumption that will simplify such mathematical model. We assume that the salt inside the tank gets *instantaneously mixed*, which means that—at every time—the salt concentration in one part of the tank is the same as in any other part of the tank. When that happens we say that the salt concentration inside the tank is constant in space and changes only in time.

We now introduce our mathematical model of this physical situation, which we call it a *mixing problem* in the following definition.

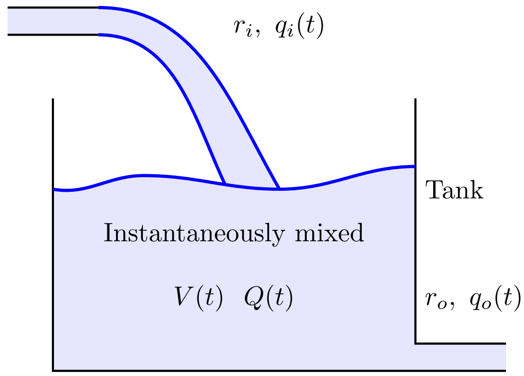


FIGURE 1. Description of a water tank problem.

Definition 1.2.4. *A **Mixing Problem** refers to water coming into a tank at a rate r_i with salt concentration q_i, and going out the tank at a rate r_o and salt concentration q_o, so that the water volume V and the total amount of salt Q, which is instantaneously mixed, in the tank satisfy the following equations,*

$$V'(t) = r_i(t) - r_o(t), \tag{1.2.14}$$

$$Q'(t) = r_i(t)\, q_i(t) - r_o(t)\, q_o(t), \tag{1.2.15}$$

$$q_o(t) = \frac{Q(t)}{V(t)}, \tag{1.2.16}$$

$$r_i'(t) = r_o'(t) = 0. \tag{1.2.17}$$

Remarks:

(a) The first equation says that the variation in time of the water volume inside the tank is the difference of volume rates coming in and going out of the tank.

(b) The second equation above says that the variation in time of the amount of salt in the tank is the difference of the amount of salt rates coming in and going out of the tank. These salt rates are the product of a water rate r times a salt concentration q. Notice that this product has units of mass per time, which are the units of salt rates.

(c) Eq. (1.2.16) is the consequence of the instantaneous mixing mechanism in the tank. Since the salt in the tank is well-mixed, the salt concentration is homogeneous in the tank, with value $Q(t)/V(t)$.

(d) Finally the equations in (1.2.17) say that both rates in and out are time independent, hence constants. We include this assumption to get simple mathematical models.

Theorem 1.2.5 (Mixing Problem)**.** *The amount of salt in the mixing problem above satisfies the equation*

$$Q'(t) = a(t)\, Q(t) + b(t), \tag{1.2.18}$$

where the coefficients in the equation are given by

$$a(t) = -\frac{r_o}{(r_i - r_o)\, t + V_0}, \qquad b(t) = r_i\, q_i(t). \tag{1.2.19}$$

Proof of Theorem 1.2.5: The equation for the salt in the tank given in (1.2.18) comes from Eqs. (1.2.14)-(1.2.17). We start noting that Eq. (1.2.17) says that the water rates are constant. We denote them as r_i and r_o. This information in Eq. (1.2.14) implies that V' is constant. Then we can easily integrate this equation to obtain

$$V(t) = (r_i - r_o)\,t + V_0, \tag{1.2.20}$$

where $V_0 = V(0)$ is the water volume in the tank at the initial time $t = 0$. On the other hand, Eqs.(1.2.15) and (1.2.16) imply that

$$Q'(t) = r_i\,q_i(t) - \frac{r_o}{V(t)}\,Q(t).$$

Since $V(t)$ is known from Eq. (1.2.20), we get that the function Q must be solution of the differential equation

$$Q'(t) = r_i\,q_i(t) - \frac{r_o}{(r_i - r_o)\,t + V_0}\,Q(t).$$

This is a linear ODE for the function Q. Indeed, introducing the functions

$$a(t) = -\frac{r_o}{(r_i - r_o)\,t + V_0}, \qquad b(t) = r_i\,q_i(t),$$

the differential equation for Q has the form

$$Q'(t) = a(t)\,Q(t) + b(t).$$

This establishes the Theorem. □

Example 1.2.12 (General Case for $V(t) = V_0$)**.** Consider a mixing problem with equal constant water rates $r_i = r_o = r$, with constant incoming concentration q_i, and with a given initial water volume in the tank V_0. Find the solution to the initial value problem

$$Q'(t) = a(t)\,Q(t) + b(t), \qquad Q(0) = Q_0,$$

where function a and b are given in Eq. (1.2.19). Graph the solution function Q for different values of the initial condition Q_0.

Solution: The assumption $r_i = r_o = r$ implies that the function a is constant, while the assumption that q_i is constant implies that the function b is also constant too,

$$a(t) = -\frac{r_o}{(r_i - r_o)\,t + V_0} \quad \Rightarrow \quad a(t) = -\frac{r}{V_0} = a_0,$$
$$b(t) = r_i\,q_i(t) \quad \Rightarrow \quad b(t) = r_i\,q_i = b_0.$$

Then, we must solve the initial value problem for a constant coefficients linear equation,

$$Q'(t) = a_0\,Q(t) + b_0, \qquad Q(0) = Q_0,$$

The integrating factor method can be used to find the solution of the initial value problem above. The formula for the solution is given in Theorem 1.2.3 and Example 1.2.7,

$$Q(t) = \Big(Q_0 + \frac{b_0}{a_0}\Big)\,e^{a_0 t} - \frac{b_0}{a_0}.$$

In our case the we can evaluate the constant b_0/a_0, and the result is

$$\frac{b_0}{a_0} = (rq_i)\Big(-\frac{V_0}{r}\Big) \quad \Rightarrow \quad -\frac{b_0}{a_0} = q_i V_0.$$

Then, the solution Q has the form,

$$Q(t) = \big(Q_0 - q_i V_0\big)\,e^{-rt/V_0} + q_i V_0. \tag{1.2.21}$$

The initial amount of salt Q_0 in the tank can be any non-negative real number. The solution behaves differently for different values of Q_0. We classify these values in three classes:

(a) If $Q_0 = q_i V_0$, the initial amount of salt in the tank is the critical value, then the solution Q remains constant equal to this critical value, that is, $Q(t) = q_i V_0$.

(b) If $Q_0 > q_i V_0$, the initial amount of salt in the tank is larger than the critical value, then the salt in the tank Q decreases exponentially towards the critical value.
(c) If $Q_0 < q_i V_0$, the initial amount of salt in the tank is smaller than the critical value, then the salt in the tank Q increases exponentially towards the critical value.

The graphs of a few solutions in these three classes are plotted in Fig. 2.

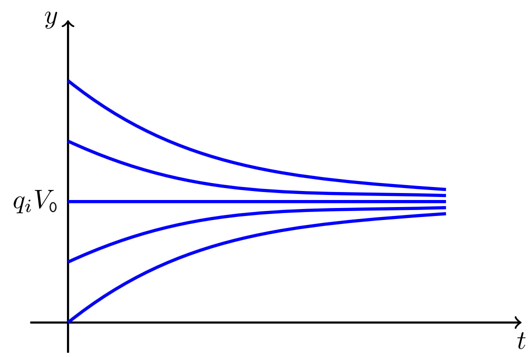


FIGURE 2. The function Q in (1.2.21) for a few values of the initial condition Q_0.

◁

Example 1.2.13 (Finding a particular time, case $V(t) = V_0$)**.** Consider a mixing problem with equal constant water rates $r_i = r_o = r$ and with fresh water coming into the tank, hence $q_i = 0$. Find the time t_1 such that the salt concentration in the tank $Q(t)/V(t)$ is 1% the initial value. Write that time t_1 in terms of the rate r and initial water volume V_0.

Solution: The first step to find the time t_1 is to solve the initial value problem for Q,

$$Q'(t) = a(t)\,Q(t) + b(t), \qquad Q(0) = Q_0,$$

where function a and b are given in Eq. (1.2.19). In this case they are

$$a(t) = -\frac{r_o}{(r_i - r_o)\,t + V_0} \qquad \Rightarrow \qquad a(t) = -\frac{r}{V_0},$$
$$b(t) = r_i\, q_i(t) \qquad \Rightarrow \qquad b(t) = 0.$$

Therefore, the initial value problem we need to solve is

$$Q'(t) = -\frac{r}{V_0}\,Q(t), \qquad Q(0) = Q_0.$$

We know from Theorem 1.2.3 that the solution is given by

$$Q(t) = Q_0\, e^{-rt/V_0}.$$

The second step to find t_1 is to find the concentration inside the tank,

$$q(t) = \frac{Q(t)}{V(t)}.$$

We already have $Q(t)$ and we know that $V(t) = V_0$, since $r_i = r_o$. Therefore,

$$q(t) = \frac{Q(t)}{V_0} \quad \Rightarrow \quad q(t) = \frac{Q_0}{V_0}\, e^{-rt/V_0}.$$

We can now find t_1. The condition that defines t_1 is

$$q(t_1) = \frac{1}{100}\,\frac{Q_0}{V_0}.$$

Therefore, we get

$$\frac{1}{100}\,\frac{Q_0}{V_0} = \frac{Q_0}{V_0}\,e^{-rt_1/V_0} \quad \Rightarrow \quad \frac{1}{100} = e^{-rt_1/V_0}.$$

Computing Logs on both sides

$$\ln\Big(\frac{1}{100}\Big) = -\frac{rt_1}{V_0} \quad \Rightarrow \quad \ln(100) = \frac{rt_1}{V_0},$$

and then we get the final result,

$$t_1 = \frac{V_0}{r}\,\ln(100).$$

◁

Example 1.2.14 (Variable Water Volume)**.** A tank with a maximum capacity for 100 liters originally contains 20 liters of water with 50 grams of salt in solution. Fresh water is poured in the tank at a rate of 5 liters per minute. The well-stirred water is allowed to pour out the tank at a rate of 3 liters per minute. Find the amount of salt in the tank at the time t_c when the tank is about to overflow.

Solution: We start with the water volume conservation,

$$V'(t) = r_i - r_o = 5 - 3 = 2 \quad \Rightarrow \quad V(t) = 2t + V_0.$$

Since $V(0) = 20$ we get the function volume of water in the tank,

$$V(t) = 2t + 20.$$

At this point we can compute the time when the tank overflows,

$$100 = V(t_c) = 2\,t_c + 20 \quad \Rightarrow \quad t_c = \frac{100 - 20}{2} \quad \Rightarrow \quad t_c = 40\,\text{min}.$$

We now need to find the equation for the salt in the tank. We start with the mass salt conservation

$$Q'(t) = r_i\,q_i - r_o\,q_o(t), \qquad q_i = 0 \quad \Rightarrow \quad Q'(t) = -r_o\,q_o(t).$$

Since the water in the tank is well-stirred, $q_o(t) = Q(t)/V(t)$, so

$$Q'(t) = -\frac{r_o}{V(t)}\,Q(t) \quad \Rightarrow \quad Q'(t) = -\frac{3}{(2t + 20)}\,Q(t).$$

This is a linear, homogeneous, differential equation with variable coefficients, so it can be converted into a separable equation,

$$\frac{Q'(t)}{Q(t)} = -\frac{3}{2t + 20} \quad \Rightarrow \quad \int \frac{dQ}{Q} = -\int \frac{3\,dt}{2t + 20}.$$

Integrating we get

$$\ln|Q(t)| = -\frac{3}{2}\int \frac{dt}{t + 10} = -\frac{3}{2}\,\ln(t + 10) + c_0 \quad \Rightarrow \quad \ln|Q(t)| = \ln((t + 10)^{-3/2}) + c_0.$$

We can now compute exponentials on both sides,

$$|Q(t)| = (t + 10)^{-3/2}\,e^{c_0} \quad \Rightarrow \quad Q(t) = c\,(t + 10)^{-3/2}, \qquad c = (\pm e^{c_0}).$$

The initial condition $Q(0) = 50$ fixes the constant c,

$$50 = Q(0) = c\,(0 + 10)^{-3/2} = c\,\frac{1}{(10)^{3/2}} \quad \Rightarrow \quad c = 50\,(10)^{3/2}.$$

Therefore,

$$Q(t) = 50\,(10)^{3/2}\,\frac{1}{(t+10)^{3/2}} \quad\Rightarrow\quad Q(t) = \frac{50}{\big((t/10)+1\big)^{3/2}}.$$

This result is reasonable, the amount of salt decreases as function of time, since we are adding fresh water to the tank. The amount of salt at the time the tank overflows is $Q(t_c)$,

$$Q(40) = \frac{50}{\big((40/10)+1\big)^{3/2}} = \frac{50}{(5)^{3/2}} \quad\Rightarrow\quad Q(40) = 4.47 \text{ grams}.$$

◁

Example 1.2.15 (Nonzero q_i, for $V(t) = V_0$)**.** Consider a mixing problem with equal constant water rates $r_i = r_o = r$, with only fresh water in the tank at the initial time, hence $Q_0 = 0$, and with a given initial volume of water in the tank V_0. Find the amount of salt in the tank, Q, in the case that the incoming salt concentration is given by the function

$$q_i(t) = 2 + \sin(2t).$$

Solution: We need to find the function Q solution of the initial value problem

$$Q'(t) = a(t)\,Q(t) + b(t), \qquad Q(0) = 0,$$

where the functions a and b are given in Eq. (1.2.19). In this case we have

$$a(t) = -\frac{r_o}{(r_i - r_o)\,t + V_0} \quad\Rightarrow\quad a(t) = -\frac{r}{V_0} = -a_0,$$

$$b(t) = r_i\,q_i(t) \quad\Rightarrow\quad b(t) = r\,\big[2 + \sin(2t)\big].$$

Notice that $a_0 > 0$. The initial value problem we need to solve is

$$Q'(t) = -a_0\,Q(t) + b(t), \qquad Q(0) = 0.$$

The solution is computed using the integrating factor method and the result is

$$Q(t) = e^{-a_0 t}\int_0^t e^{a_0 s}\,b(s)\,ds,$$

where we used that the initial condition is $Q_0 = 0$. Recall the definition of the function b,

$$Q(t) = e^{-a_0 t}\int_0^t e^{a_0 s}\big[2 + \sin(2s)\big]\,ds.$$

This is the formula for the solution of the problem, we only need to compute the integral given in the equation above. This is not straightforward though. Notice that two integrations by parts gives us the formula

$$\int e^{ks}\sin(ls)\,ds = \frac{e^{ks}}{k^2 + l^2}\,\big[k\sin(ls) - l\cos(ls)\big],$$

where k and l are constants. Therefore,

$$\int_0^t e^{a_0 s}\big[2 + \sin(2s)\big]\,ds = \Big[\frac{2}{a_0}e^{a_0 s}\Big]\Big|_0^t + \Big[\frac{e^{a_0 s}}{a_0^2 + 2^2}\,\big[a_0\sin(2s) - 2\cos(2s)\big]\Big]\Big|_0^t,$$

$$= \frac{2}{a_0}q\big(e^{a_0 t} - 1\big) + \frac{e^{a_0 t}}{a_0^2 + 2^2}\,\big[a_0\sin(2t) - 2\cos(2t)\big] + \frac{2}{a_0^2 + 2^2}.$$

With the integral above we can compute the solution Q as follows,

$$Q(t) = e^{-a_0 t}\Big[\frac{2}{a_0}\big(e^{a_0 t} - 1\big) + \frac{e^{a_0 t}}{a_0^2 + 2^2}\,\big[a_0\sin(2t) - 2\cos(2t)\big] + \frac{2}{a_0^2 + 2^2}\Big],$$

recalling that $a_0 = r/V_0$. We rewrite expression above as follows,

$$Q(t) = \frac{2}{a_0} + \Big[\frac{2}{a_0^2 + 2^2} - \frac{2}{a_0}\Big]\, e^{-a_0 t} + \frac{1}{a_0^2 + 2^2}\,\big[a_0 \sin(2t) - 2\cos(2t)\big]. \tag{1.2.22}$$

◁

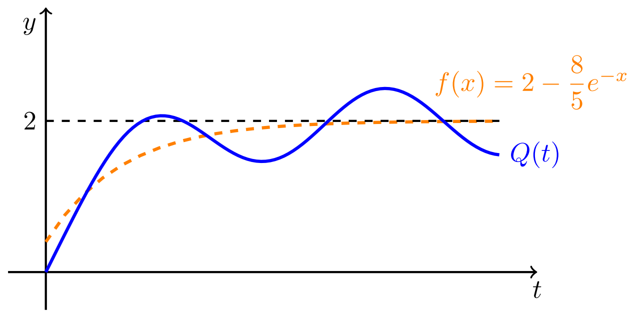


FIGURE 3. The graph of the function Q given in Eq. (1.2.22) for $a_0 = 1$.

1.2.5. The Bernoulli Equation. In 1696 Jacob Bernoulli solved a first order *nonlinear* differential equation, which is now known as the Bernoulli differential equation. This is not the Bernoulli equation from fluid dynamics, though. The following year Leibniz solved this equation by transforming it into a linear equation. This is a fruitful idea in mathematics: You have an equation you do not know how to solve; you transform that equation into an equation you do know how to solve; you solve it; you transform back the solution. We now explain Leibniz's idea in more detail.

Definition 1.2.6. *The **Bernoulli equation** is*

$$y' = p(t)\, y + q(t)\, y^n. \tag{1.2.23}$$

where p, q are given functions and $n \in \mathbb{R}$.

Remarks:

(a) For $n \neq 0, 1$ the equation is nonlinear.
(b) If $n = 2$ we get the *logistic equation*, which we have studied in § 1.1 and § 1.3,

$$y' = ry\Big(1 - \frac{y}{k}\Big).$$

(c) This is not the Bernoulli equation from fluid dynamics.

The Bernoulli equation is special in the following sense: it is a nonlinear equation that can be transformed into a linear equation.

Theorem 1.2.7 (Bernoulli)**.** *The function y is a solution of the Bernoulli equation*

$$y' = p(t)\, y + q(t)\, y^n, \qquad n \neq 1,$$

iff the function $v = 1/y^{(n-1)}$ is solution of the linear differential equation

$$v' = -(n-1)p(t)\, v - (n-1)q(t).$$

Remark: This result summarizes Laplace's idea to solve the Bernoulli equation. To transform the Bernoulli equation for y, which is nonlinear, into a linear equation for $v = 1/y^{(n-1)}$. One then solves the linear equation for v using the integrating factor method. The last step is to transform back to $y = (1/v)^{1/(n-1)}$.

Proof of Theorem 1.2.7: Divide the Bernoulli equation by y^n,

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

Introduce the new unknown $v = y^{-(n-1)}$ and compute its derivative,

$$v' = \big[y^{-(n-1)}\big]' = -(n-1)y^{-n}\, y' \quad \Rightarrow \quad -\frac{v'(t)}{(n-1)} = \frac{y'(t)}{y^n(t)}.$$

If we substitute v and this last equation into the Bernoulli equation we get

$$-\frac{v'}{(n-1)} = p(t)\, v + q(t) \quad \Rightarrow \quad v' = -(n-1)p(t)\, v - (n-1)q(t).$$

This establishes the Theorem. □

Example 1.2.16. Find every nonzero solution of the differential equation

$$y' = y + 2\, y^5.$$

Solution: This is a Bernoulli equation for $n = 5$. Divide the equation by y^5,

$$\frac{y'}{y^5} = \frac{1}{y^4} + 2.$$

Introduce the function $v = 1/y^4$ and its derivative $v' = -4(y'/y^5)$, into the differential equation above,

$$-\frac{v'}{4} = v + 2 \quad \Rightarrow \quad v' = -4\, v - 8 \quad \Rightarrow \quad v' + 4\, v = -8.$$

The last equation is a linear differential equation for the function v. This equation can be solved using the integrating factor method. Multiply the equation by $\mu(t) = e^{4t}$, then

$$\big(e^{4t} v\big)' = -8\, e^{4t} \quad \Rightarrow \quad e^{4t} v = -\frac{8}{4}\, e^{4t} + c.$$

We obtain that $v = c\, e^{-4t} - 2$. Since $v = 1/y^4$,

$$\frac{1}{y^4} = c\, e^{-4t} - 2 \quad \Rightarrow \quad y(t) = \pm \frac{1}{\big(c\, e^{-4t} - 2\big)^{1/4}}.$$

◁

Example 1.2.17. Given constants a_0, b_0, find all solutions y of the differential equation

$$y' = a_0\, y + b_0\, y^3.$$

Solution: This is a Bernoulli equation with $n = 3$. Divide the equation by y^3,

$$\frac{y'}{y^3} = \frac{a_0}{y^2} + b_0.$$

Introduce the function $v = 1/y^2$ and its derivative $v' = -2(y'/y^3)$, into the differential equation above,

$$-\frac{v'}{2} = a_0 v + b_0 \quad \Rightarrow \quad v' = -2a_0 v - 2b_0 \quad \Rightarrow \quad v' + 2a_0 v = -2b_0.$$

The last equation is a linear differential equation for v. This equation can be solved using the integrating factor method. Multiply the equation by $\mu(t) = e^{2a_0 t}$,

$$\big(e^{2a_0 t} v\big)' = -2b_0\, e^{2a_0 t} \quad \Rightarrow \quad e^{2a_0 t} v = -\frac{b_0}{a_0}\, e^{2a_0 t} + c$$

We obtain that $v = c\, e^{-2a_0 t} - \dfrac{b_0}{a_0}$. Since $v = 1/y^2$,

$$\frac{1}{y^2} = c\, e^{-2a_0 t} - \frac{b_0}{a_0} \quad \Rightarrow \quad y(t) = \pm \frac{1}{\big(c\, e^{-2a_0 t} - \frac{b_0}{a_0}\big)^{1/2}}.$$

◁

Example 1.2.18. Find every solution of the equation

$$t\, y' = 3y + t^5\, y^{1/3}.$$

Solution: Rewrite the differential equation as

$$y' = \frac{3}{t}\, y + t^4\, y^{1/3}.$$

This is a Bernoulli equation for $n = 1/3$. Divide the equation by $y^{1/3}$,

$$\frac{y'}{y^{1/3}} = \frac{3}{t}\, y^{2/3} + t^4.$$

Define the new unknown function $v = 1/y^{(n-1)}$, that is, $v = y^{2/3}$, compute is derivative, $v' = \dfrac{2}{3} \dfrac{y'}{y^{1/3}}$, and introduce them in the differential equation,

$$\frac{3}{2}\, v' = \frac{3}{t}\, v + t^4 \quad \Rightarrow \quad v' - \frac{2}{t}\, v = \frac{2}{3}\, t^4.$$

This is a linear equation for v. Integrate this equation using the integrating factor method. To compute an integrating factor we need to find

$$A(t) = \int \frac{2}{t}\, dt = 2\ln(t) = \ln(t^2).$$

Then, the integrating factor is $\mu(t) = e^{-A(t)}$. In this case we get

$$\mu(t) = e^{-\ln(t^2)} = e^{\ln(t^{-2})} \quad \Rightarrow \quad \mu(t) = \frac{1}{t^2}.$$

Therefore, the equation for v can be written as a total derivative,

$$\frac{1}{t^2}\Big(v' - \frac{2}{t}\, v\Big) = \frac{2}{3}\, t^2 \quad \Rightarrow \quad \Big(\frac{v}{t^2} - \frac{2}{9}\, t^3\Big)' = 0.$$

The potential function is

$$\psi(t, v) = \frac{v}{t^2} - \frac{2}{9}\, t^3$$

and the solution of the differential equation is $\psi(t, v(t)) = c$, that is,

$$\frac{v}{t^2} - \frac{2}{9}\, t^3 = c \quad \Rightarrow \quad v(t) = t^2\Big(c + \frac{2}{9}\, t^3\Big) \quad \Rightarrow \quad v(t) = c\, t^2 + \frac{2}{9}\, t^5.$$

Once v is known we compute the original unknown $y = \pm v^{3/2}$, where the double sign is related to taking the square root. We finally obtain

$$y(t) = \pm\Big(c\,t^2 + \frac{2}{9}\,t^5\Big)^{3/2}.$$

$\lhd$

Notes. This section corresponds to Boyce-DiPrima [4] Section 2.1, and Simmons [8] Section 2.10. The Bernoulli equation is solved in the exercises of section 2.4 in Boyce-Diprima, and in the exercises of section 2.10 in Simmons.

1.2.6. Exercises.

1.2.1.- Find all solutions of the following differential equations:

(a) $y' = 4t\,y$.

(b) $y' = -y + e^{-2t}$.

(c) $\dfrac{y'}{(t^2+1)y} = 4t$.

(d) $ty' + n\,y = t^2, \qquad n > 0$.

(e) $y' = y - 2\sin(t)$.

(f) $y' + t\,y = t\,y^2$.

(g) $y' = -x\,y + 6x\,\sqrt{y}$.

1.2.2.- Find the solution y of the initial value problems

(a) $y' = y + 2te^{2t}, \quad y(0) = 0$.

(b) $t\,y' + 2\,y = \dfrac{\sin(t)}{t}, \quad y\Big(\dfrac{\pi}{2}\Big) = \dfrac{2}{\pi}, \quad$ for $t > 0$.

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

(d) $t\,y' = 2\,y + 4t^3\cos(4t), \quad y\big(\frac{\pi}{8}\big) = 0$.

(e) $y' = y + \dfrac{3}{y^2}, \quad y(0) = 1$.

1.2.3.-

(a) A cup with some liquid is placed in a fridge held at 3 C. If k is the (positive) liquid cooling constant, find the *differential equation* satisfied by the temperature of the liquid.

(b) Find the liquid temperature T as function of time knowing that the liquid initial temperature when it was placed in the fridge was 18 C.

(c) After 3 minutes the liquid temperature inside the fridge is 13 C. Find the liquid cooling constant k.

1.2.4.-

(a) A pizza is placed in a oven held at 100 C. If k is the (positive) pizza cooling constant, find the *differential equation* satisfied by the temperature of the pizza.

(b) Find the pizza temperature T as function of time knowing that the pizza initial temperature when it was placed in the oven was 20 C.

(c) After 2 minutes the liquid temperature inside the fridge is 80 C. Find the liquid cooling constant k.

1.2.5.- A tank initially contains $V_0 = 100$ liters of water with $Q_0 = 25$ grams of salt. The tank is rinsed with fresh water flowing in at a rate of $r_i = 5$ liters per minute and leaving the tank at the same rate. The water in the tank is well-stirred. Find the time such that the amount the salt in the tank is $Q_1 = 5$ grams.

1.2.6.- A tank initially contains $V_0 = 100$ liters of pure water. Water enters the tank at a rate of $r_i = 2$ liters per minute with a salt concentration of $q_1 = 3$ grams per liter. The instantaneously mixed mixture leaves the tank at the same rate it enters the tank. Find the salt concentration in the tank at any time $t \geqslant 0$. Also find the limiting amount of salt in the tank in the limit $t \to \infty$.

1.2.7.- A tank with a capacity of $V_m = 500$ liters originally contains $V_0 = 200$ liters of water with $Q_0 = 100$ grams of salt in solution. Water containing salt with concentration of $q_i = 1$ gram per liter is poured in at a rate of $r_i = 3$ liters per minute. The well-stirred water is allowed to pour out the tank at a rate of $r_o = 2$ liters per minute. Find the salt concentration in the tank at the time when the tank is about to overflow. Compare this concentration with the limiting concentration at infinity time if the tank had infinity capacity.

1.3. Graphical Analysis

Right after the invention of differential calculus, by Newton in 1671 and independently by Leibniz in 1684, the first differential equations were solved using ideas similar to those we studied in § 1.1 and § 1.2. It didn't take long before people realized that most of the differential equations could not be solved in this way—meaning there was no clear way to find formulas for solutions of all differential equations in terms of known functions. Instead, two other ideas emerged.

The first idea was to study whether a differential equation has solutions or not, without finding a formula for the solutions. If we know that a differential equation has solutions, we can define new functions by saying they are the solutions of that differential equation. In this section we state Theorem 1.3.1–which we will prove in § 1.4—saying that a wide class of first order differential equations have solutions. Then we use this result to prove that solutions of a differential equation with different initial conditions cannot intersect.

The second idea was to developed methods to graph the qualitative behavior of solutions to differential equations without actually computing the explicit expression of these solutions. In this section we study three of these methods. The first method works with a particular type of differential equations called autonomous equations, where the independent variable does not show explicitly in the equation. In this method we use the differential equation to determine regions of the solution values where these solutions are increasing or decreasing functions of the independent variable. This information is enough to sketch an approximate graph of these solutions. The second method works with any differential equation, autonomous or not. In this method we use a computer to find the slope (or direction) field of the equation. These slope fields are line segments determined by the differential equation, which are tangent to a solution of the differential equation at the point they are computed. So, the slope fields also give us a way to draw approximate graphs of solutions to a differential equation. The third method is an application of the slope field method. We use the slope field to construct a collection of segments that approximate solutions to a differential equation. This third method is called the Euler method.

1.3.1. Existence-Uniqueness Result.

Solutions of differential equations are not unique. Solving a differential equation involves integration and for each integration we get an integration constant. These integration constants can take any values, hence differential equations have infinitely many solutions, one for each value of the integration constants.

An *initial value problem* is to find *particular* solutions to a differential equation, solutions satisfying extra conditions, called *initial conditions*. Usually, the initial conditions are introduced so that they determine all the integration constants. That is why solutions to initial value problems are usually unique. For example, in the case of Newton's second law of motion for a point particle—force equals mass times acceleration—one could be interested only in solutions such that the particle is at a specific position and having a specific velocity at the initial time. These conditions determine only one possible motion, that is, only one solution of Newton's second law of motion.

We now present the Picard-Lindelöf Theorem, which shows that a large class of initial value problems have solutions uniquely determined by appropriate initial data.

Theorem 1.3.1 (Picard-Lindelöf). *Consider the initial value problem*

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0. \tag{1.3.1}$$

If the function f is continuous in t and differentiable in y on some rectangle on the ty-plane containing the point (t_0, y_0) in its interior, then there is a unique solution y of the initial value problem in (1.3.1) defined on an open interval (t_1, t_2) containing the point t_0.

Remarks:

(a) We do not have an explicit formula for the solutions of the differential equation.
(b) We do not specify how large is the domain of the solution, (t_1, t_2).
(c) The domain (t_1, t_2) could even change when we change the initial data y_0.
(d) The result still holds if assumption that the function f being differentiable in the variable y is relaxed to Lipschitz continuous in y, which is something more than continuity but less than differentiability in y.

Theorem 1.3.2. *Two functions y_1, y_2 solutions with different initial data of a differential equation satisfying the hypotheses of Theorem 1.3.1 cannot intersect as in Fig. 4.*

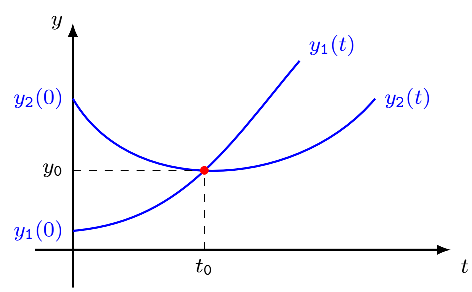


FIGURE 4. Intersections like in this picture *cannot happen* to solutions of differential equations satisfying the hypotheses in Theorem 1.3.1.

Proof of Theorem 1.3.2: Suppose the functions y_1 and y_2 are solutions of the same differential equation,

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

but with different initial conditions, $y_1(0) \neq y_2(0)$. Therefore, in a neighborhood of the initial conditions these solutions are different. If these solutions intersect at a point (t_0, y_0), as pictured in Fig. 4, then we can use the intersection point as the initial condition for the same differential equation. That is, we get the initial value problem

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0.$$

On the one hand, this initial value problem has two different solutions, as we see in Fig. 4. On the other hand, this initial value problem satisfies the hypotheses in Theorem 1.3.1, so it must have a unique solution in a neighborhood of (t_0, y_0). This is a contradiction, hence intersections such as in Fig. 4 cannot happen. This establishes the Theorem. □

1.3.2. Autonomous Equations.

We introduce an idea to find qualitative properties of solutions to differential equations without having to solve the equation. But this idea only works on a particular type of differential equations—autonomous differential equations.

Definition 1.3.3. *A first order **autonomous** differential equation is*

$$y' = f(y), \tag{1.3.2}$$

where $y' = \dfrac{dy}{dt}$ and the function f does not depend explictly on t.

Autonomous equations are differential equations where the independent variable does not appear explicitly in the equation. In symbols we can write

$$y'(t) = f(\not{t}, y(t)).$$

It is simple to see that autonomous equations are a particular type of the separable equations we studied in § 1.1,

$$h(y)\,y' = g(t), \quad \text{with} \quad h(y) = \frac{1}{f(y)}, \quad g(t) = 1.$$

Here we show a few examples of autonomous and nonautonomous equations.

Example 1.3.1. The following first order equations are autonomous:

(a) $y' = 2\,y + 3$.

(b) $y' = \sin(y)$.

(c) $y' = r\,y\left(1 - \dfrac{y}{k}\right)$.

The independent variable t does not appear explicitly in these equations. The following equations are not autonomous.

(a) $y' = 2\,y + 3t$.

(b) $y' = t^2 \sin(y)$.

(c) $y' = t\,y\left(1 - \dfrac{y}{k}\right)$. ◁

Remark: Since the autonomous equation in (1.3.2) is a particular case of the equations in the Picard-Lindelöf Theorem 1.3.1. Then, the initial value problem

$$y' = f(y), \qquad y(0) = y_0,$$

with f differentiable, always has a unique solution in the neighborhood of $t = 0$ for every value of the initial data y_0.

1.3.3. Qualitative Solution Curves. In the following example we explain how we can obtain qualitative information about solutions of an autonomous equation by using the equation itself without solving it.

Example 1.3.2. Sketch a qualitative graph of solutions to the initial value problem

$$y' = \sin(y), \qquad y(0) = \hat{y}_0,$$

for different initial data conditions y_0.

Solution: The differential equation has the form $y' = f(y)$, where $f(y) = \sin(y)$. The first step in this method is to graph f as function of y.

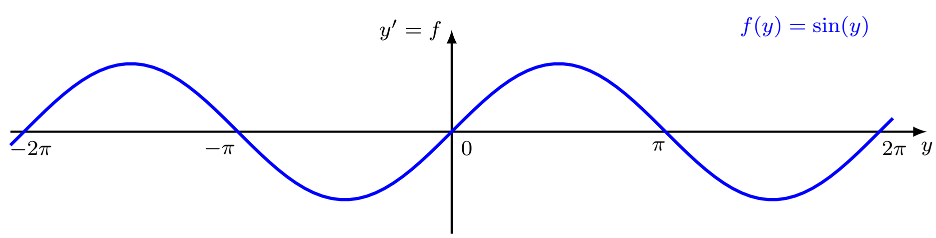


FIGURE 5. Graph of the function $f(y) = \sin(y)$.

The second step is to identify all the zeros of the function f. In this case,

$$f(y) = \sin(y) = 0 \quad \Rightarrow \quad y_n = n\pi, \quad \text{where} \quad n = \cdots, -2, -1, 0, 1, 2, \cdots.$$

These constants, y_n are called *equilibrium solutions* because they are t-independent solutions of the differential equation. Indeed, they satisfy $y_n' = 0$ and they are the zeros of f, hence $f(y_n) = 0$. From here we get that the y_n satisfy

$$0 = y_n' = f(y_n) = 0.$$

The third step is to identify the regions on the y-line where f is positive and where f is negative. These regions are important because of the following argument. Let $y(t)$ be any solution of the differential equation

$$y' = f(y).$$

Now, fix any time t_1 and evaluate the solution $y(t)$ at that time, let's call it $y_1 = y(t_1)$.

(a) If $y_1 \in (0, \pi)$, then $f(y_1) > 0$, and therefore, this solution satisfies that

$$0 < f(y(t_1)) = y'(t_1) \quad \Rightarrow \quad y'(t_1) > 0.$$

We see that this solution is increasing at t_1. Then, for a time t_2 close to t_1 but with $t_2 > t_1$ we have that

$$y_2 = y(t_2) > y(t_1) = y_1.$$

Therefore, the point y_2 is on the *right* of the point y_1 on the horizontal y-axis. We represent this behavior by a green *arrow pointing to the right* on the interval $(0, \pi)$ in Fig. 6. The same behavior occurs on every interval where $f > 0$.

(b) If $y_1 \in (-\pi, 0)$, then $f(y_1) < 0$, and therefore, this solution satisfies that

$$0 > f(y(t_1)) = y'(t_1) \quad \Rightarrow \quad y'(t_1) < 0.$$

We see that this solution is increasing at t_1. Then, for a time t_2 close to t_1 but with $t_2 > t_1$ we have that

$$y_2 = y(t_2) < y(t_1) = y_1.$$

Therefore, the point y_2 is on the *left* of the point y_1 on the horizontal y-axis. We represent this behavior by a green *arrow pointing to the left* on the interval $(-\pi, 0)$ in Fig. 6. The same behavior occurs on every interval where $f < 0$.

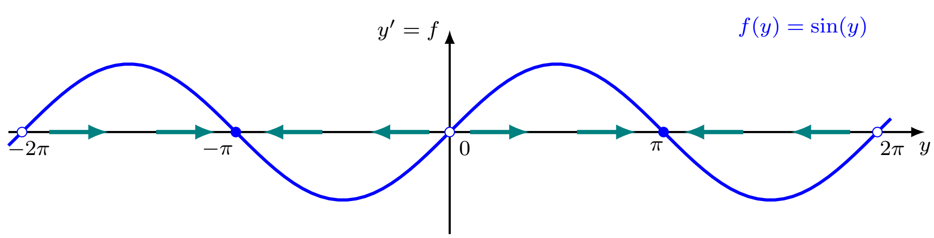


FIGURE 6. Critical points and increase/decrease information added to Fig. 5.

There are two types of equilibrium solutions in Fig. 6.

(a) Points such as $y_{-1} = -\pi$ and $y_1 = \pi$ have arrows on both sides pointing to them

$$\rightarrow \bullet \leftarrow .$$

They are called *stable* equilibrium solutions or *attractors*, and they are pictured with solid blue dots in Fig. 5.

(b) Points such as $y_{-2} = -2\pi$, $y_0 = 0$, and $y_2 = 2\pi$ have arrows on both sides pointing away from them

$$\leftarrow \bullet \rightarrow .$$

They are called *unstable* equilibrium solutions or *repellers*, and they are pictured with white dots in Fig. 5.

(c) In other equations there could be equilibrium solutions that have increasing solutions on either side or decreasing solutions on either side,

$$\rightarrow \bullet \rightarrow \qquad \leftarrow \bullet \leftarrow .$$

They are also called *unstable* equilibrium solutions or *mixed* points. We do not have these type of equilibrium solutions in Fig. 5.

The fourth step is to find the regions where the curvature of a solution is *concave up* or *concave down*. That information is given by the second derivative of the solution function y, which can be computed by taking one more derivative of the differential equation,

$$y'' = (y')' = \big(f(y)\big)' = f'(y)\, y' = f'(y)\, f(y),$$

that is,

$$y'' = f'(y)\, f(y).$$

Now we repeat the analysis done in the third step above: fix any time t_1 and evaluate a solution $y(t)$ at that time, $y_1 = y(t_1)$.

(a) If y_1 is in any region where $f(y_1)\, f'(y_1) > 0$, then the second derivative of this solution satisfies that $y''(t_1) > 0$, hence this solution is concave up (CU) at that time.

(b) If y_1 is in any region where $f(y_1)\, f'(y_1) < 0$, then the second derivative of this solution satisfies that $y''(t_1) < 0$, hence this solution is concave down (CD) at that time.

In Fig. 7 we graph both f and f', so it is simple to see the sign of their product $f'\, f$. From this product we get the intervals where solutions are concave up or concave down.

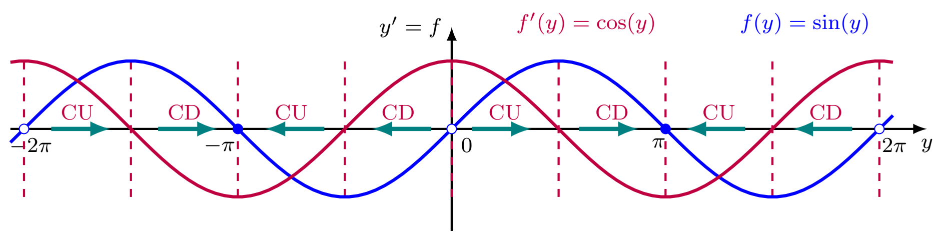


FIGURE 7. Concavity information on the solution y added to Fig. 6.

The fifth step is to sketch a qualitative graph of solutions to the differential equation on a ty-plane. All the information we collected on the horizontal axis in Fig. 7 is now displayed in the vertical axis on Fig. 8. In the horizontal axis in Fig. 8 we plot the independent variable t.

Fig. 8 contains the graph of several solutions y for different choices of initial data $y(0)$. Equilibrium solutions are in blue and t-dependent solutions in green. The equilibrium solutions are separated in two types. The stable equilibrium solutions

$$y_{-1} = -\pi, \qquad y_1 = \pi,$$

are pictured with solid blue lines. The unstable equilibrium solutions

$$y_{-2} = -2\pi, \qquad y_0 = 0, \qquad y_2 = 2\pi,$$

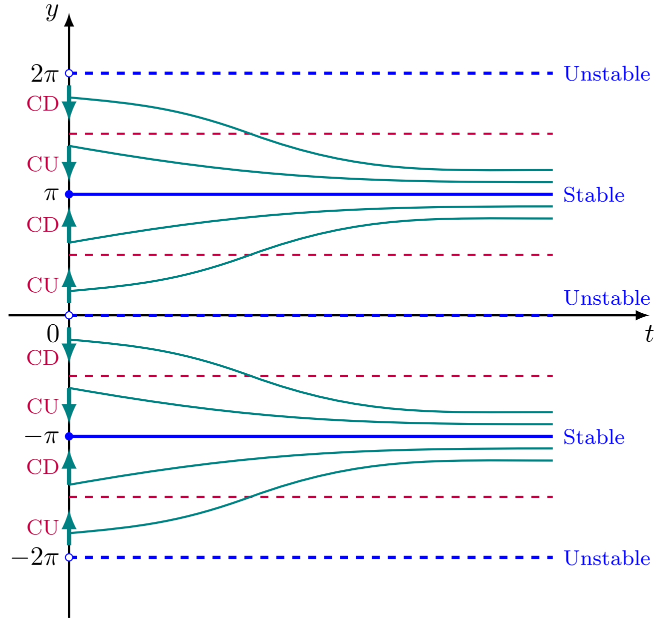


FIGURE 8. Qualitative graphs of solutions y for different initial conditions.

are pictured with dashed blue lines. The graph in time of the non-equilibrium solutions is done with the intervals in y where $y(t)$ is increasing or decreasing and with the concavity at those intervals. ◁

Remark: A qualitative graph of the solutions does not provide all the possible information about the solution. For example, we know from the graph above that for some initial conditions the corresponding solutions have inflection points at some $t > 0$. But we cannot know the exact value of t where the inflection point occurs. Such information could be useful to have, since $|y'|$ has its maximum value at those points.

In the Example 1.3.2 above we found the concavity of solutions from the sign of the second derivative of these solutions. The second derivative of solutions is related to f and f'. We remark this result in its own statement.

Theorem 1.3.4. *If y is a solution of the autonomous system $y' = f(y)$, then*

$$y'' = f'(y)\, f(y).$$

Proof: The differential equation relates y'' to $f(y)$ and $f'(y)$, because of the chain rule,

$$y'' = \frac{d}{dt}\Big(\frac{dy}{dt}\Big) = \frac{d}{dt} f(y(t)) = \frac{df}{dy}\,\frac{dy}{dt} \quad \Rightarrow \quad y'' = f'(y)\, f(y).$$

□

Example 1.3.3. Sketch a qualitative graph of solutions of the logistic equation

$$y' = ry\Big(1 - \frac{y}{k}\Big), \qquad y(0) = y_0,$$

for different values of the initial condition y_0, where r and k are given positive constants.

Solution: The logistic equation for population growth can be written $y' = f(y)$, where function f is the polynomial

$$f(y) = ry\left(1 - \frac{y}{k}\right).$$

The first step is to graph f as function of y. The result is in Fig. 9.

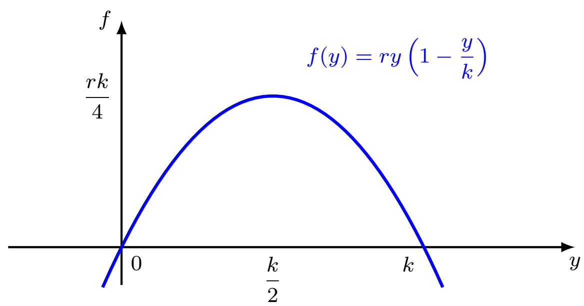


FIGURE 9. The graph of $f = ry\left(1 - \frac{y}{k}\right)$.

The second step is to identify all the equilibrium solutions of the equation, which are the zeros of the function f. In this case, $f(y) = 0$ implies

$$y_0 = 0, \quad y_1 = k.$$

The third step is to identify the regions on the y-line where f is positive and where f is negative. We repeat the argument from the previous example. Let $y(t)$ be any solution of the differential equation

$$y' = f(y).$$

Now, fix any time t_1 and evaluate the solution $y(t)$ at that time, let's call it $y_1 = y(t_1)$.

(a) If $y_1 \in (0, k)$, then $f(y_1) > 0$, and therefore, this solution satisfies that

$$0 < f(y(t_1)) = y'(t_1) \quad \Rightarrow \quad y'(t_1) > 0.$$

We see that this solution is increasing at t_1. Then, for a time t_2 close to t_1 but with $t_2 > t_1$ we have that

$$y_2 = y(t_2) > y(t_1) = y_1.$$

Therefore, the point y_2 is on the *right* of the point y_1 on the horizontal y-axis. We represent this behavior by green *arrows pointing to the right* on the interval $(0, k)$ in Fig. 10.

(b) If $y_1 \in (-\infty, 0)$ or $y_1 \in (k, \infty)$, then $f(y_1) < 0$, and then this solution satisfies that

$$0 > f(y(t_1)) = y'(t_1) \quad \Rightarrow \quad y'(t_1) < 0.$$

We see that this solution is increasing at t_1. Then, for a time t_2 close to t_1 but with $t_2 > t_1$ we have that

$$y_2 = y(t_2) < y(t_1) = y_1.$$

Therefore, the point y_2 is on the *left* of the point y_1 on the horizontal y-axis. We represent this behavior by a green *arrow pointing to the left* on the interval $(-\infty, 0)$ and on the interval (k, ∞) in Fig. 10.

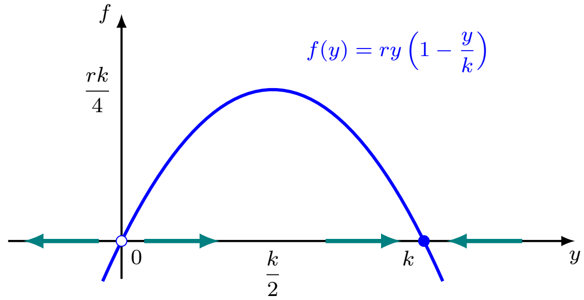


FIGURE 10. Critical points added.

The fourth step is to find the regions where the curvature of a solution is *concave up* or *concave down*. That information is given by the second derivative of the solution function y, which can be computed by taking one more derivative of the differential equation,

$$y'' = (y')' = \big(f(y)\big)' = f'(y)\,y' = f'(y)\,f(y),$$

that is,

$$y'' = f'(y)\,f(y).$$

Now we repeat the analysis done in the third step above: fix any time t_1 and evaluate a solution $y(t)$ at that time, $y_1 = y(t_1)$.

(a) If y_1 is in any region where $f(y_1)\,f'(y_1) > 0$, then the second derivative of this solution satisfies that $y''(t_1) > 0$, hence this solution is concave up (CU) at that time.

(b) If y_1 is in any region where $f(y_1)\,f'(y_1) < 0$, then the second derivative of this solution satisfies that $y''(t_1) < 0$, hence this solution is concave down (CD) at that time.

In Fig. 11 we graph both f and f', so it is simple to see the sign of their product $f'\,f$. From this product we get the intervals where solutions are concave up or concave down.

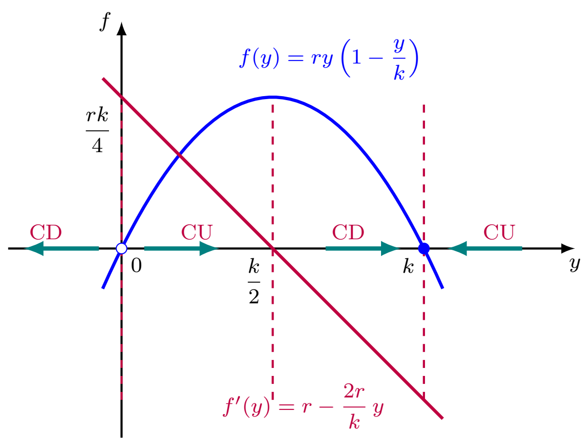


FIGURE 11. Concavity information added.

The fifth step is to sketch a qualitative graph of solutions to the differential equation on a ty-plane. All the information we collected on the horizontal axis in Fig. 11 is now displayed in the vertical axis on Fig. 12. In the horizontal axis in Fig. 12 we plot the independent variable t.

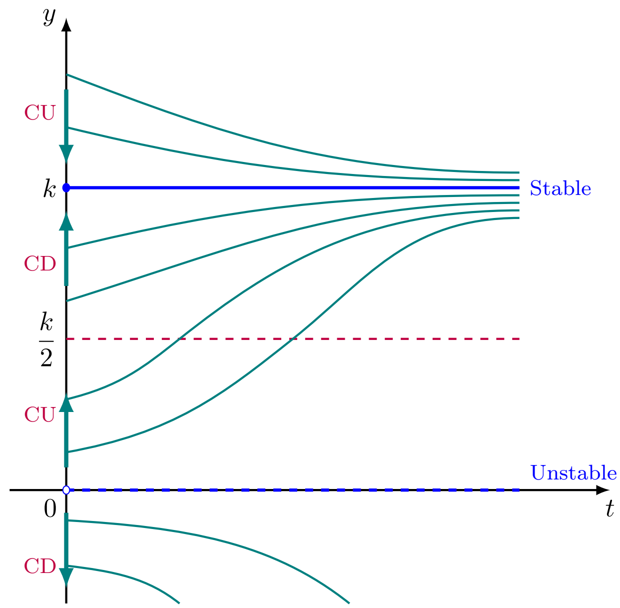


FIGURE 12. Qualitative graphs of solutions y for different initial conditions.

Fig. 12 contains the graph of several solutions y for different choices of initial data $y(0)$. Equilibrium solutions are in blue and t-dependent solutions in green. The equilibrium solutions are separated in two types: the stable equilibrium solution

$$y_1 = k,$$

which is graphed with a solid blue line; and the unstable equilibrium solution,

$$y_0 = 0,$$

which is graphed with dashed blue line. The graph in time of the non-equilibrium solutions is done with the intervals in y where $y(t)$ is increasing or decreasing and with the concavity at those intervals. ◁

1.3.4. Slope Fields. We now introduce a second idea to find qualitative properties of solutions to differential equations without having to solve the equation. This idea only works on any first order differential equation, whether autonomous or not. However, this second idea requires the use of a computer. Consider a differential equation

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

We want to display in a graph all the information given by the right-hand side of the equation. This information can be graphed in at least two different ways.

(a) In the *usual way*, the graph of $f(t,y)$ is a surface in the tyz-space, where $z = f(t,y)$. The values of $f(t,y)$ are points pictured in the z-axis, perpendicular to the domain plane, the ty-plane.
(b) In the *new way*, we graph the values of $f(t,y)$ as the slopes of segments at each point (t,y) on the ty-plane.

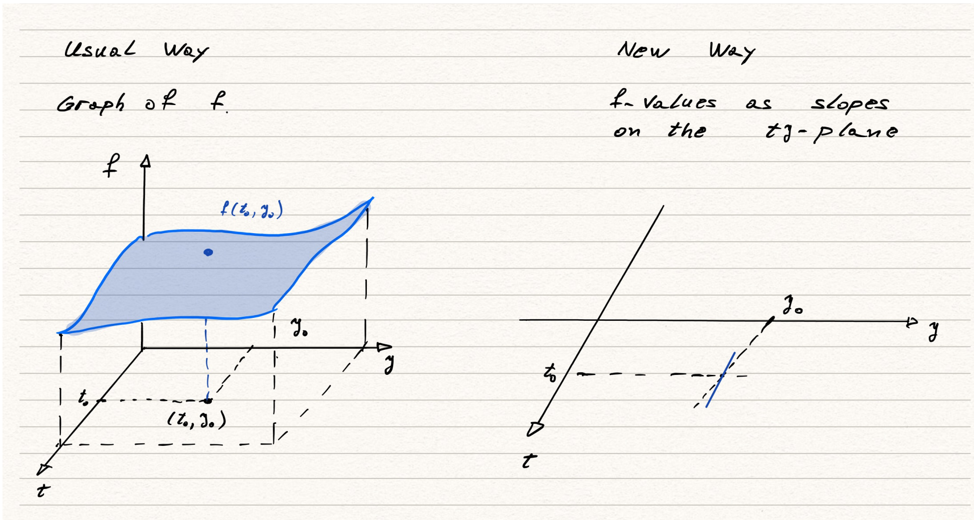


FIGURE 13. The function f graphed as a point in the z-axis and as a slope of a segment in the ty-plane.

The new way pictured above comes from using the differential equation

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

to interpret the value of $f(t, y)$. Given a solution, $y(t)$, the value of $f(t, y(t))$ is the value of the derivative of that solution, $y'(t)$. The latter can be represented graphically on the ty-plane as the slope of a segment tangent to the graph of the solution $y(t)$ at t. The ideas above suggest the following definition.

Definition 1.3.5. *A **slope field** (or direction field) for the differential equation*

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

is the graph on the ty-plane of the values $f(t, y)$ as slopes of a small segments.

Example 1.3.4. Find the slope field of the equation $y' = y$, and sketch a few solutions to the differential equation for different initial conditions.

Solution: We use a computer to find the slope field, which is shown in Fig. 14. We have also plotted solution curves corresponding to four solutions.

Notice that the solution curves are *tangent* at every point to the slope field. This is no accident. *Since each curve is a solution os the differential equation, these curves must be tangent to every segment in the slope field.* Also notice that the solution curves in this example agree with the uniqueness property of solutions to initial value problems showed in Theorem 1.3.1, which implies that the solution curves corresponding to different initial conditions *do not intersect*.

Recall that in this example the solutions are functions of the form $y(t) = y_0\, e^t$. $\lhd$

Example 1.3.5. Use a computer to find the slope field of the equation

$$y' = \sin(y).$$

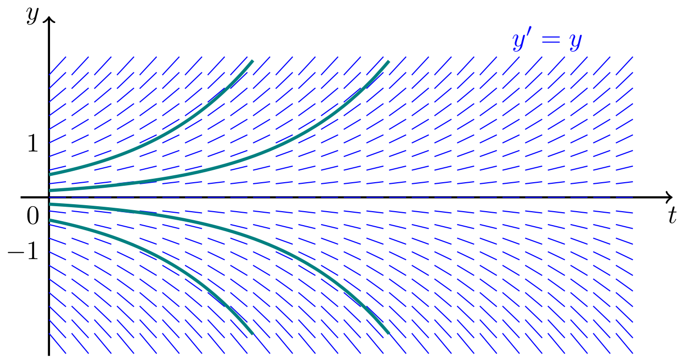


FIGURE 14. Slolpe field for the equation $y' = y$ and a few solutions.

Solution: We first mention that the equation above can be solved exactly, in implicit form, and the solutions are

$$\frac{\sin(y)}{\big(1+\cos(y)\big)} = \frac{\sin(y_0)}{\big(1+\cos(y_0)\big)}\, e^t.$$

for any $y_0 \in \mathbb{R}$. This is an equation that defines the solution function y. There are no derivatives in the equation, so this is not a differential equation; We call it an algebraic equation. However, the graphs of these solutions are not simple to do. But the direction field is simple to plot and it can be seen in Fig. 15. From that direction field one can see what the graph of the solutions should look like. ◁

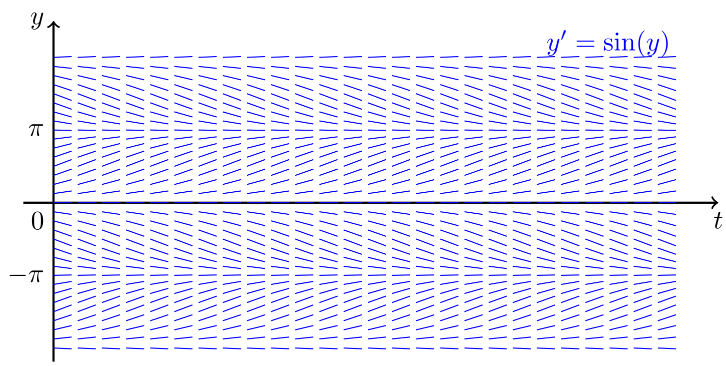


FIGURE 15. Slope field for the equation $y' = \sin(y)$.

Example 1.3.6. Use a computer to find the slope field of the equation

$$y' = 2\cos(t)\cos(y).$$

Solution: We do not need to compute the explicit solution of $y' = 2\cos(t)\cos(y)$ to have a qualitative idea of its solutions. The slope field can be seen in Fig. 16. ◁

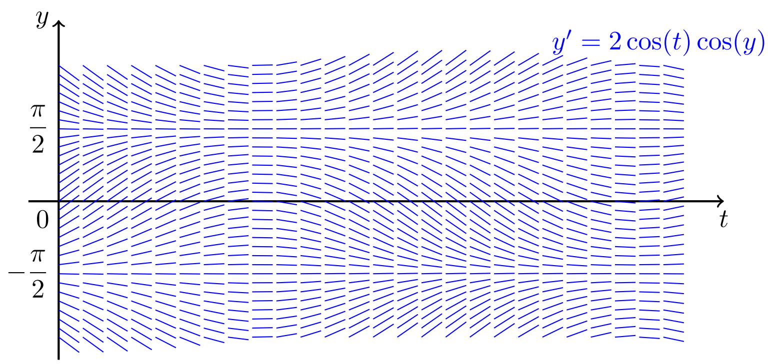


FIGURE 16. Slope field for the equation $y' = 2\cos(t)\cos(y)$.

1.3.5. The Euler Method. The Euler method, also called the tangent line method, is a simple way to obtain an approximation to a solution of an initial value problem

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0. \tag{1.3.3}$$

The idea is to use the information given by the equation, provided by the function values $f(t, y)$, to construct a linear spline—a collection of segments where the end of one segment is the beginning of the next—that is close to the solution $y(t)$ of (1.3.3).

The function values $f(t, y)$ provide all possible slopes for all possible solutions $y(t)$ at all possible times t. In Figure 19 we picture this meaning in the case of the differential equation $y' = y$, which means that $f(t, y) = y$. On the left we have the geometrical meaning of $f(t, y)$ at a point (t_0, y_0); on the right we have plotted $f(t, y)$, as the slope of small segments, at several points on the ty-plane.

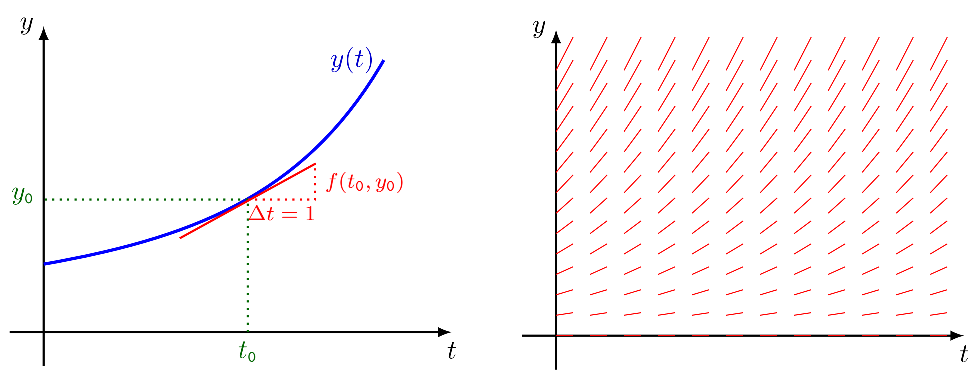


FIGURE 17. On the left we have the geometrical meaning of the function value $f(t_0, y_0)$ at the point (t_0, y_0). On the right we have a plot of the values of $f(t, y)$ at several points on the ty-plane. These pictures are made in the case $f(t, y) = y$.

The Euler method is simple to understand with a graphical representation. Suppose the solution of the initial value problem in (1.3.3), for some arbitrary $f(t, y)$, is given in picture on the left of Figure 20. In that picture we also plot the intial condition $y(t_0) = y_0$, and the value of $f(t_0, y_0)$ as the slope of the red segment. Fix a time step $\Delta t > 0$, and introduce the partition

$$\{t_0, t_1 = t_0 + \Delta t, t_2 = t_1 + \Delta t, \cdots, t_N = t_{N-1} + \Delta t\},$$

where in the picture on the right of Figure 20 we chose $N = 9$. To construct the Spline $s_N(t)$ we proceed as follows:

- Find the equation of the line $L_0(t)$ that passes through the point (t_0, y_0) and has slope given by $f(t_0, y_0)$.
- Use L_0 to find the next point in the spline, y_1, by evaluating L_0 at $t_1 = t_0 + \Delta t$, that is,

$$y_1 = L_0(t_1), \qquad t_1 = t_0 + \Delta t.$$

- Find the equation of the line $L_1(t)$ that passes through the point (t_1, y_1) and has slope given by $f(t_1, y_1)$.
- Use L_1 to find the next point in the spline, y_2, by evaluating L_1 at $t_2 = t_1 + \Delta t$, that is,

$$y_2 = L_1(t_2), \qquad t_2 = t_1 + \Delta t.$$

- Repeat until reaching t_N.

This procedure would give a spline $s_n(t)$ like the one shown on the right in Figure 20.

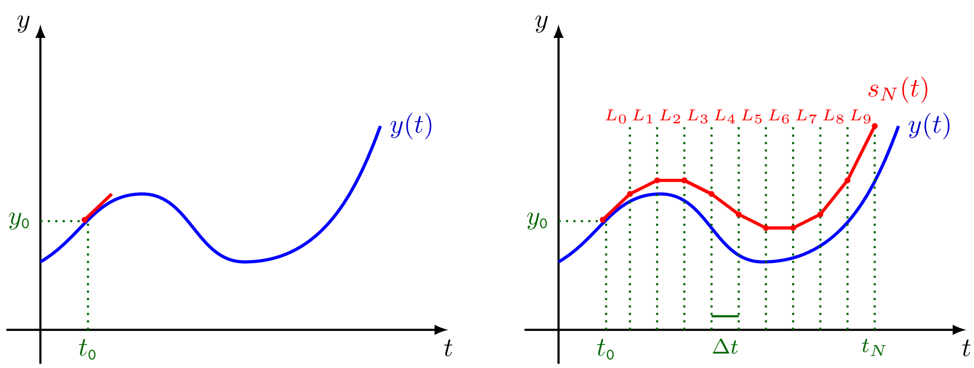


FIGURE 18. This is what a linear spline $s_N(t)$, that approximates a solution $y(t)$, constructed with the Euler method could look like. Here we used a number of steps $N = 9$ and a time step Δt. The linear functions $L_0, \cdots, L_9$ are the line segments that form the spline s_N.

We are now ready to find an analytic expression for the lines $L_i(t)$, for $i = 0, 1, 2, \cdots, N$, that form the spline $s_N(t)$.

Step 0: We need to find the equation of $L_0(t)$ and then use that equation to find the next y-value, $y_1 = L_0(t_1)$, for $t_1 = t_0 + \Delta t$. Since $L_0(t)$ is a line, it can be written as

$$L_0(t) = m_0 t + b_0.$$

We know that the line contains the point (t_0, y_0), that is,

$$y_0 = m_0 t_0 + b_0 \quad \Rightarrow \quad b_0 = y_0 - m_0 t_0.$$

We also know that the slope of the line is given by $f(t_0, y_0)$, that is,

$$m_0 = f(t_0, y_0).$$

Therefore, the equation of the line L_0 is

$$L_0(t) = f(t_0, y_0)\,(t - t_0) + y_0.$$

Since $t_1 = t_0 + \Delta t$, we get that $y_1 = L_0(t_1)$ is given by

$$y_1 = f(t_0, y_0)\,\Delta t + y_0.$$

Step 1: Now we need to find the equation of $L_1(t)$ and then use that equation to find the next y-value, $y_2 = L_1(t_2)$, for $t_2 = t_1 + \Delta t$. Since $L_1(t)$ is a line, it can be written as

$$L_1(t) = m_1 t + b_1.$$

We know that the line contains the point (t_1, y_1), that is,

$$y_1 = m_1 t_1 + b_1 \quad \Rightarrow \quad b_1 = y_1 - m_1 t_1.$$

We also know that the slope of the line is given by $f(t_1, y_1)$, that is,

$$m_1 = f(t_1, y_1).$$

Therefore, the equation of the line L_1 is

$$L_1(t) = f(t_1, y_1)\,(t - t_1) + y_1.$$

Since $t_2 = t_1 + \Delta t$, we get that $y_2 = L_1(t_2)$ is given by

$$y_2 = f(t_1, y_1)\,\Delta t + y_1.$$

Step n: If we continue this process, for the n-term we get

$$L_n(t) = f(t_n, y_n)\,(t - t_n) + y_n, \qquad y_{n+1} = f(t_n, y_n)\,\Delta t + y_n.$$

Then the spline $s_N(t)$ is determined by the $N + 1$ points

$$(t_0, y_0),\ (t_1, y_1),\ (t_2, y_2), \cdots,\ (t_N, y_N).$$

The discussion above can be summarized in the following result.

Theorem 1.3.6 (Euler Method)**.** *The Euler approximation of $y(t)$ solution of the initial value problem*

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0,$$

with $N > 0$ terms and time step $\Delta t > 0$, is the linear spline $s_N(t)$ through the points

$$(t_0, y_0),\ (t_1, y_1), \cdots, (t_N, y_N),$$

where

$$t_{n+1} = t_n + \Delta t, \qquad y_{n+1} = f(t_n, y_n)\,\Delta t + y_n, \qquad n = 0, 1, 2, \cdots, N - 1.$$

Example 1.3.7. Compute the Euler approximation of the solution of

$$y' = 2y + 3, \qquad y(0) = 1, \tag{1.3.4}$$

on the interval $[0, 3]$ with time step $\Delta t = 1$.

Solution: We need to construct the approximation on the interval $[0, 3]$, with time-step $\Delta t = 1$, which means we have $N = (3 - 0)/1$, so $N = 3$. In this case, the Euler approximation of the initial value problem in (1.3.4) is a table of numbers of the form

t_n	y_n
t_0	y_0
t_1	y_1
t_2	y_2
t_3	y_3

We know the values of $t_{n+1} = t_n + \Delta t$, since $\Delta t = 1$ and $t_0 = 0$ give us

$$t_0 = 0, \quad t_1 = 1, \quad t_2 = 2, \quad t_3 = 3.$$

Now we can compute the values of $y_{n+1} = f(t_n, y_n)\,\Delta t + y_n$, knowing $y_0 = 1$ from the initial condition and $\Delta t = 1$, which give us

$$\begin{aligned} y_1 &= f(t_0, y_0)\,\Delta t + y_0 = (2(1) + 3)(1) + 1 \quad \Rightarrow \quad y_1 = 6 \\ y_2 &= f(t_1, y_1)\,\Delta t + y_1 = (2(6) + 3)(1) + 6 \quad \Rightarrow \quad y_2 = 21 \\ y_3 &= f(t_2, y_2)\,\Delta t + y_2 = (2(21) + 3)(1) + 21 \quad \Rightarrow \quad y_3 = 66. \end{aligned}$$

Therefore, the spline is given by the points (t, y) as follows,

$$(0,1),\ (1,6),\ (2,21),\ (3,66). \tag{1.3.5}$$

Equivalently, the spline is given by the table

t_n	y_n
0	1
1	6
2	21
3	66

◁

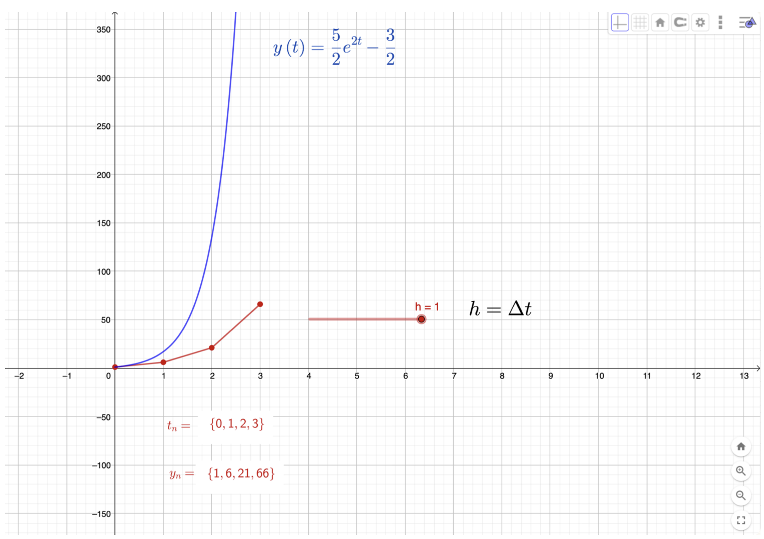


Figure 19. The exact solution of the initial value problem in (1.3.4) is graphed in blue while the spline $s_N(t)$ computed in (1.3.5) is plotted in red.

We now find the Euler approximation to the solution of the initial value problem in Example 1.3.7, but this time with a time-step half of the one used in that Example.

Example 1.3.8. Compute the Euler approximation of the solution of

$$y' = 2y + 3, \qquad y(0) = 1, \tag{1.3.6}$$

on the interval $[0, 3]$ with time step $\Delta t = 0.5$.

Solution: We need to construct the approximation on the interval $[0, 3]$, with time-step $\Delta t = 1$, which means we have $N = (3 - 0)/0.5$, so $N = 6$. In this case, the Euler approximation of the initial value problem in (1.3.6) is a table of numbers of the form

t_n	y_n
t_0	y_0
t_1	y_1
t_2	y_2
t_3	y_3
t_4	y_4
t_5	y_5
t_6	y_6

We know the values of $t_{n+1} = t_n + \Delta t$, since $\Delta t = 0.5$ and $t_0 = 0$ give us

$$t_0 = 0, \quad t_1 = 0.5, \quad t_2 = 1, \quad t_3 = 1.5, \quad t_4 = 2, \quad t_5 = 2.5, \quad t_6 = 3.$$

Now we can compute the values of $y_{n+1} = f(t_n, y_n)\,\Delta t + y_n$, knowing $y_0 = 1$ from the initial condition and $\Delta t = 0.5$, which give us

$$\begin{aligned}
y_1 &= f(t_0, y_0)\,\Delta t + y_0 = (2(1) + 3)(0.5) + 1 \quad \Rightarrow \quad y_1 = 3.5\\
y_2 &= f(t_1, y_1)\,\Delta t + y_1 = (2(3.5) + 3)(0.5) + 3.5 \quad \Rightarrow \quad y_2 = 8.5\\
y_3 &= f(t_2, y_2)\,\Delta t + y_2 = (2(8.5) + 3)(0.5) + 8.5 \quad \Rightarrow \quad y_3 = 18.5,\\
y_4 &= f(t_3, y_3)\,\Delta t + y_3 = (2(18.5) + 3)(0.5) + 18.5 \quad \Rightarrow \quad y_4 = 38.5\\
y_5 &= f(t_4, y_4)\,\Delta t + y_4 = (2(38.5) + 3)(0.5) + 38.5 \quad \Rightarrow \quad y_5 = 78.5\\
y_6 &= f(t_5, y_5)\,\Delta t + y_5 = (2(78.5) + 3)(0.5) + 78.5 \quad \Rightarrow \quad y_6 = 158.5.
\end{aligned}$$

Therefore, the spline is given by the points (t, y) as follows,

$$(0, 1),\ (0.5, 3.5),\ (1, 8.5),\ (1.5, 18.5), (2, 38.5),\ (2.5, 78.5),\ (3, 158.5). \tag{1.3.7}$$

Equivalently, the spline is given by the table

t_n	y_n
0	1
0.5	3.5
1	8.5
1.5	18.5
2	38.5
2.5	78.5
3	158.5

$\triangleleft$

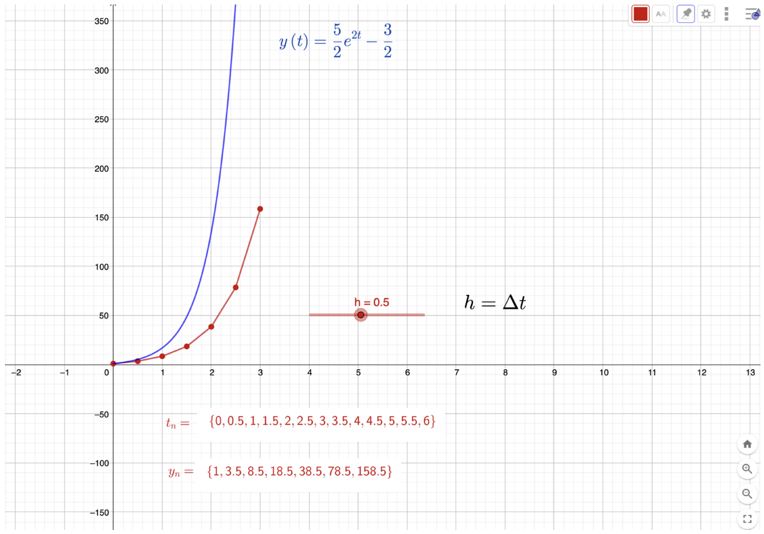


FIGURE 20. The exact solution of the initial value problem in (1.3.6) is graphed in blue while the spline $s_N(t)$ computed in (1.3.7) is plotted in red.

1.3.6. Exercises.

1.3.1.- For the differential equations below do the following:

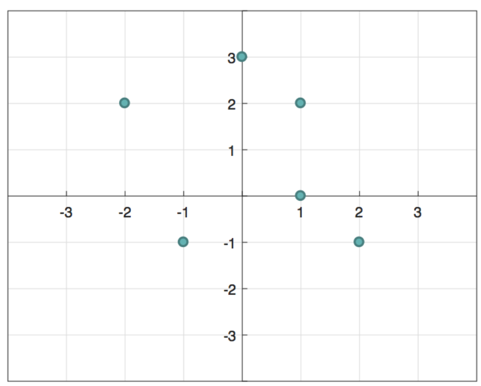
(a) Find all the equilibrium solutions of the differential equation.
(b) Find the open intervals where solutions are increasing.
(c) Find the open intervals where solutions are decreasing.
(d) Find the open intervals where solutions are concave up.
(e) Find the open intervals where solutions are concave up.
(f) With the information above sketch a qualitative graph of the several solutions for each differential equation.

$$\text{(i)}\ y' = \sin(3y), \quad \text{(ii)}\ y' = \cos(3y), \quad \text{(iii)}\ y' = y^2 - 9, \quad \text{(iv)}\ y' = 9 - y^2, \quad \text{(v)}\ y' = y\Big(1 - \frac{y}{2}\Big).$$

1.3.2.- Consider the differential equation

$$\frac{dy}{dx} = xy.$$

Sketch vectors from the corresponding slope field of this differential equation, at the points indicated on the figure below.



1.3.3.- Match the slope fields of the following three differential equations to the three figures below. Provide justification for your reasoning.

A. $\dfrac{dy}{dx} = \cos(x)$, **B.** $\dfrac{dy}{dx} = \sin(y)$, **C.** $\dfrac{dy}{dx} = x + y$.

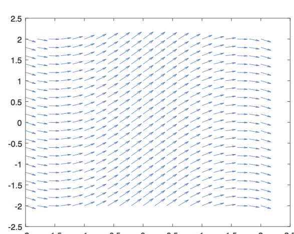


FIGURE 21. **I**

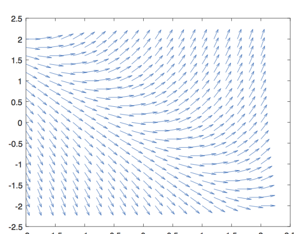


FIGURE 22. **II**

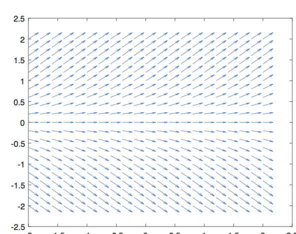


FIGURE 23. **III**

1.4. Approximate Solutions

We know a lot more about solutions of *linear* differential equations than what we know about solutions of *nonlinear* differential equations. While we have an explicit formula for the solutions to all linear equations—Theorem 1.2.3—there is no such formula for solutions to all nonlinear equations. In § 1.1 we solved only two particular cases of nonlinear equations—separable equations and Euler homogeneous equations. But these nonlinear equations are only a tiny part of all nonlinear equations.

After many years of trying and failing, people had to give up on the goal of finding a formula for solutions to *all* nonlinear equations. Instead, they focused on finding approximate solutions to nonlinear differential equations. In this section we focus on approximate solutions obtained using either the Taylor series expansions or the Picard iteration. Each of these techniques produces an infinite sequence of functions. Sometimes these sequences converge to a solution of the differential equation. In these cases, the further we move in the sequence of functions the closer we get to a solution of the differential equation. Because of this, the functions in the sequence are called approximate solutions.

If each function in the sequence is a better approximation than the previous function, we can get as close as we want to the exact solution using these approximations. If the solution of a differential equation describes a certain physical phenomena, then we can use these approximations to predict the behavior of the system as accurately as we wish.

Another use for these sequence of approximate solutions is to show that certain nonlinear differential equations actually have solutions. Such statements are called existence theorems for solutions of differential equations. In this section we use the Picard iteration to show that a certain class of nonlinear equations have solutions, and the solution is unique provided appropriate initial conditions. This result is called the Picard-Lindelöf theorem. We end this section comparing what we know about solutions of linear differential equations with solutions of nonlinear differential equations.

Before we start we show a few examples of linear and nonlinear equations.

Example 1.4.1 (Linear and Non-Linear Equations).

(a) The differential equation

$$\frac{y'(t)}{y(t)+3t} = 2t^2$$

is actually linear, because when we write it in the normal form

$$y'(t) = 2t^2\, y(t) + 6t^3$$

the right-hand side is linear in the second argument. So, we know a formula for the solutions of this equation.

(b) The differential equation

$$y'(t) = \frac{y^2(t) + t\, y(t) + t^2}{t^2 + y^2(t)}$$

is nonlinear. This equation is Euler homogeneous, since it can be written as

$$y' = \frac{(\frac{y}{t})^2 + (\frac{y}{t}) + 1}{1 + (\frac{y}{t})^2}.$$

We know that Euler homogeneous equations can be transformed into a separable equation and solved exactly.

(c) The differential equation

$$y'(t) = 2ty(t) + \ln\big(y(t)\big)$$

is nonlinear. However, the equation can not be transformed into a separable equation, and we do not know how to find a formula for its solutions.

◁

1.4.1. Taylor Series. We use the Taylor series of a function to find a sequence of approximate solutions of a first order differential equation. Recall the Taylor series expansion centered at $t = t_0$ of a function y,

$$\begin{aligned} y(t) &= \sum_{n=0}^{\infty} \frac{1}{n!}\, y^{(n)}(t_0)\,(t-t_0)^n \\ &= y(t_0) + y'(t_0)\,(t-t_0) + \frac{1}{2!}\, y''(t_0)\,(t-t_0)^2 + \cdots, \end{aligned}$$

where $n!$ is the n-th factorial, $y^{(n)}(t_0)$ is the n-th derivative of y evaluated at $t = t_0$, but we also denoted $y^{(0)} = y$ (the zero derivative is the original function), and $y^{(1)} = y'$, $y^{(2)} = y''$ (first and second derivatives are denoted as usual). We also used that $0! = 1$ and $1! = 1$. In this section we focus on the Taylor formula centered at $t_0 = 0$, which is

$$\begin{aligned} y(t) &= \sum_{n=0}^{\infty} \frac{1}{n!}\, y^{(n)}(0)\, t^n \\ &= y(0) + y'(0)\, t + \frac{1}{2!}\, y''(0)\, t^2 + \cdots. \end{aligned}$$

The first $n+1$ terms of the expansion above are called the n-th order Taylor approximation.

Definition 1.4.1. *The n-th order* ***Taylor approximation*** *centered at t_0 of a function y is given by*

$$\tau_n(t) = \sum_{k=0}^{n} \frac{1}{k!}\, y^{(k)}(t_0)\,(t-t_0)^k$$

Remark: The definition above implies a simple relation between τ_n and τ_{n-1},

$$\tau_n(t) = \tau_{n-1}(t) + \frac{1}{n!}\, y^{(n)}(t_0)\,(t-t_0)^n.$$

Taylor expansions have two main applications. One application is to simplify calculations by replacing a possible complicated function $y(t)$ by a simple polynomial approximation. A second application is to extend the definition of a function $y(t)$ from a real variable t to a more general type of variable, for example to a complex variable or a matrix variable. We will discuss both types of extensions, to a complex variable or to a matrix variable, of the exponential function later in this textbook.

In both applications above we know $y(t)$ and then we compute its derivatives and evaluate these derivatives at $t = t_0$. In this section we are interested in a different situation. In our case, the function $y(t)$ is not known. Instead, $y(t)$ is solution of a differential equation with an initial condition,

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0.$$

It turns out that the initial condition and the differential equation is enough to compute all the derivatives of the function $y(t)$ at the time of the initial condition, t_0.

Theorem 1.4.2 (Taylor Approximation)**.** *The initial value problem*

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0, \tag{1.4.1}$$

with $f(t, y)$ infinitely continuously differentiable in both variables, determines $\tau_n(t)$, the n-th order Taylor approximation of the solution $y(t)$ of (1.4.1), for any integer $n \geqslant 0$.

Proof of Theorem 1.4.2: The Taylor approximation formula is

$$\tau_n(t) = y(t_0) + y'(0)\,(t - t_0) + \frac{1}{2!}\,y''(t_0)\,(t - t_0)^2 + \cdots + \frac{1}{n!}\,y^{(n)}(t_0)\,(t - t_0)^n.$$

This formula says that to determine τ_n we need to know the function value at $t = t_0$ and all the derivatives values at $t = t_0$, that is, we need to know

$$y^{(k)}(t_0), \quad \text{for all} \quad k = 0, 1, 2, \cdots n.$$

But $y(t_0)$ is given by the initial condition, $y(t_0) = y_0$, therefore the initial condition fixes the zero order Taylor approximation,

$$\tau_0(t) = y_0.$$

The next Taylor approximation is

$$\tau_1(t) = \tau_0(t) + y'(t_0)\,(t - t_0).$$

The differential equation relates $y'(t)$ with $y(t)$ for all $t > t_0$, so in the limit $t \to t_0^+$ we get

$$y'(t_0) = f(t_0, y_0),$$

where we used again the initial condition $y(t_0) = y_0$. Therefore,

$$\tau_1(t) = y_0 + f(t_0, y_0)\,(t - t_0).$$

The next Taylor approximation is

$$\tau_2(t) = \tau_1(t) + \frac{1}{2!}\,y''(t_0)\,(t - t_0)^2.$$

Again, the differential equation relates $y'(t)$ with $y(t)$ for all $t > t_0$, so we can take one more t-derivative on both sides of the equation,

$$y''(t) = \frac{d}{dt} f(t, y(t)) \quad \Rightarrow \quad y''(t) = \partial_t f(t, y(t)) + \partial_y f(t, y(t))\,y'(t).$$

If we take the limit $t \to t_0^+$ in the last equation above and we recall that $y(t_0) = y_0$ and we already know the value of $y'(t_0)$, then we also know the value of $y''(t_0)$, since

$$y''(t_0) = \partial_t f(t_0, y_0) + \partial_y f(t_0, y_0)\,y'(t_0).$$

This expression gives us

$$\tau_2(t) = y_0 + f(t_0, y_0)\,(t - t_0) + \frac{1}{2}\big[\partial_t f(t_0, y_0) + \partial_y f(t_0, y_0)\,f(t_0, y_0)\big]\,(t - t_0)^2.$$

Let's compute one more Taylor approximation,

$$\tau_3(t) = \tau_2(t) + \frac{1}{3!}\,y^{(3)}(t_0)\,(t - t_0)^3,$$

where $y^{(3)} = y'''$. We first find $y^{(3)}(t)$ computing one more derivative in the equation for y'',

$$y^{(3)} = \frac{d}{dt}\big(\partial_t f + (\partial_y f)\,y'\big),$$

which gives us,

$$y^{(3)} = \partial_t\big(\partial_t f + (\partial_y f)\,y'\big) + \partial_y\big(\partial_t f + (\partial_y f)\,y'\big)\,y'$$

that is,

$$y^{(3)} = \partial_t^2 f + (\partial_t \partial_y f)\,y' + (\partial_y f)\,y'' + (\partial_y \partial_t f)\,y' + (\partial_y^2 f)\,(y')^2.$$

The last expression above can be simplified a bit as

$$y^{(3)} = \partial_t^2 f + 2\,(\partial_t \partial_y f)\,y' + (\partial_y f)\,y'' + (\partial_y^2 f)\,(y')^2.$$

If we take the limit $t \to t_0^+$ in the last equation above, and recalling we already know $y(t_0) = y_0$ and $y'(t_0)$, and $y''(t_0)$, then

$$y^{(3)}(t_0) = \partial_t^2 f(t_0, y_0) + 2\,\big(\partial_t \partial_y f(t_0, y_0)\big)\, y'(t_0) + \big(\partial_y f(t_0, y_0)\big)\, y''(t_0) + \big(\partial_y^2 f(t_0, y_0)\big)\,\big(y'(t_0)\big)^2,$$

is also know, which gives us $\tau_3(t)$. This process can be continued to compute $y^{(n)}(t_0)$ for all integers $n \geqslant 0$, which establishes the Theorem. □

Example 1.4.2. Use the Taylor series to find the first four approximate solutions of the linear initial value problem

$$y'(t) = 2\,y(t) + 3, \qquad y(0) = 1.$$

Solution: Recall the n-th order Taylor approximation centered at $t = 0$,

$$\tau_n(t) = y(0) + y'(0)\,t + \frac{1}{2!}\,y''(0)\,t^2 + \cdots + \frac{1}{n!}\,y^{(n)}(0)\,t^n.$$

Also recall that

$$\tau_n(t) = \tau_{n-1}(t) + \frac{1}{n!}\,y^{(n)}(0)\,t^n.$$

The initial condition provides $\tau_0(t)$, which is

$$\tau_0(t) = y(0) \quad \Rightarrow \quad \tau_0(t) = 1.$$

The next approximation is

$$\tau_1(t) = \tau_0(t) + y'(0)\,t.$$

We get $y'(0)$ from the differential equation,

$$y'(0) = 2\,y(0) + 3 \quad \Rightarrow \quad y'(0) = 5,$$

which gives us

$$\tau_1(t) = 1 + 5\,t.$$

The next approximation is

$$\tau_2(t) = \tau_1(t) + \frac{1}{2!}\,y''(0)\,t^2.$$

We get $y''(0)$ by differentiating the differential equation,

$$y''(t) = 2\,y'(t), \quad \Rightarrow \quad y''(0) = 2\,y'(0) \quad \Rightarrow \quad y''(0) = 10,$$

and recalling that $2! = 2$ we arrive at

$$\tau_2(t) = 1 + 5\,t + 5\,t^2.$$

The last approximation we compute here is

$$\tau_3(t) = \tau_2(t) + \frac{1}{3!}\,y'''(0)\,t^3.$$

We get $y'''(0)$ by differentiating the differential equation for y'' we computed above,

$$y'''(t) = 2\,y''(t), \quad \Rightarrow \quad y'''(0) = 2\,y''(0) \quad \Rightarrow \quad y''(0) = 20,$$

and recalling that $3! = 6$ we arrive at

$$\tau_3(t) = 1 + 5\,t + 5\,t + \frac{10}{3}\,t^3.$$

◁

In the case that the $\lim_{n\to\infty} \tau_n(t)$ converges and defines a continuously differentiable function on the t variable, then this function is a solution of the initial value problem (1.4.1).

Theorem 1.4.3 (Solution by Taylor Approximation)**.** *Let $\tau_n(t)$ be the Taylor approximation given in Theorem 1.4.2. If the limit $n \to \infty$ of $\tau_n(t)$ converges and*

$$y_T(t) = \lim_{n\to\infty} \tau_n(t)$$

is a continuously differentiable function, then $y_T(t)$ is a solution of the initial value problem in (1.4.1).

Proof of Theorem 1.4.3: It is not difficult to see that the functions

$$y_T(t) = \lim_{n\to\infty} \tau_n(t) \qquad \text{and} \qquad g(t) = f(t, y_T(t))$$

satisfy

$$g^{(n)}(t_0) = y_T^{(n+1)}(t_0), \qquad n = 0, 1, 2, \cdots . \tag{1.4.2}$$

The case $n = 0$ is obtained as follows: first recall that $y_T(t_0) = y_0$ and $y_T'(t_0) = f(t_0, y_0)$; second evaluate $g(t)$ at $t = t_0$, the result is

$$g(t_0) = f(t_0, y_T(t_0)) = f(t_0, y_0) = y_T'(t_0).$$

The case $n = 1$ is given by

$$g'(t_0) = \frac{d}{dt} f(t, y(t))\Big|_{t=t_0} = y_T''(t_0).$$

From here it is not difficult to see that the definitions of $y_T(t)$ and $g(t)$ imply Eq. (1.4.2). Using equation (1.4.2) in the Taylor expansion

$$g(t) = \sum_{k=0}^{\infty} \frac{1}{k!} g^{(k)}(t_0)\,(t - t_0)^k$$

we get

$$g(t) = \sum_{k=0}^{\infty} \frac{1}{k!} y_T^{(k+1)}(t_0)\,(t-t_0)^k = \Big(y_T(0) + \sum_{k=0}^{\infty} \frac{1}{(k+1)!} y_T^{(k+1)}(t_0)\,(t-t_0)^{(k+1)}\Big)' = y_T'(t),$$

where we used that $y_T(0)$ is a constant. This last equation above shows that

$$y_T'(t) = f(t, y_T(t)),$$

which establishes the Theorem. □

Example 1.4.3. We have seen in a previous section that the solutions of the initial value problem

$$y'(t) = a\,y(t) + b, \qquad y(0) = y_0,$$

with a, b constants, is given by

$$y(t) = \Big(y_0 + \frac{a}{b}\Big)\, e^{at} - \frac{b}{a}.$$

Use the Taylor approximation method to find the solution formula above.

Solution: Since a and b are constants,

$$y'(t) = a\,y(t) + b \quad \Rightarrow \quad y''(t) = a\,y'(t) \quad \Rightarrow \quad y^{(n+1)}(t) = a\,y^{(n)}(t).$$

The initial condition $y(0) = y_0$ and the equations above imply

$$\begin{aligned}
y'(0) &= ay_0 + b,\\
y''(0) &= a\,y'(0) = a\,(ay_0 + b),\\
y'''(0) &= a\,y''(0) = a^2(ay_0 + b),\\
&\ \vdots\\
y^{(n)}(0) &= a\,y^{(n-1)}(0) = a^{n-1}(ay_0 + b).
\end{aligned}$$

Therefore, the Taylor formula for the solution, $y_T(t)$ is

$$y_T(t) = y_0 + (ay_0 + b)\,t + a\,(ay_0 + b)\,\frac{t^2}{2!} + \cdots + + a^{n-1}\,(ay_0 + b)\,\frac{t^n}{n!} + \cdots .$$

If we do some simple algebraic manipulations we get

$$\begin{aligned}
y_T(t) &= y_0 + (ay_0 + b)\Big(t + \frac{a\,t^2}{2!} + \cdots + \frac{a^{n-1}\,t^n}{n!} + \cdots\Big)\\
&= y_0 + (ay_0 + b)\frac{1}{a}\Big(at + \frac{(at)^2}{2!} + \cdots + \frac{(at)^n}{n!} + \cdots\Big)\\
&= y_0 - (ay_0 + b)\frac{1}{a} + (ay_0 + b)\frac{1}{a}\Big(1 + at + \frac{(at)^2}{2!} + \cdots + \frac{(at)^n}{n!} + \cdots\Big)\\
&= y_0 - y_0 - \frac{b}{a} + \Big(y_0 + \frac{b}{a}\Big)\,e^{at}\\
&= \Big(y_0 + \frac{b}{a}\Big)\,e^{at} - \frac{b}{a}.
\end{aligned}$$

So, we have shown that the Taylor approximation method gives the solution formula

$$y_T(t) = \Big(y_0 + \frac{b}{a}\Big)\,e^{at} - \frac{b}{a}.$$

$\lhd$

1.4.2. Picard Iteration. Unlike the Taylor approximation, which is defined for a function, the Picard approximation is defined for an initial value problem.

Definition 1.4.4. *The* ***Picard iteration*** *of an initial value problem*

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0,$$

is the sequence of functions $y_n(t)$, for $n = 0, 1, 2, \cdots$, given as follows,

$$y_0(t) = y_0, \qquad y_n(t) = y_0 + \int_{t_0}^{t} f(s, y_{n-1}(s))\,ds, \quad n = 1, 2, \cdots .$$

Remark: The equation defining the Picard iteration is derived from the differential equation itself. Indeed, given the differential equation

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

integrate on both sides of that equation with respect to t,

$$\int_{t_0}^{t} y'(s)\,ds = \int_{t_0}^{t} f(s, y(s))\,ds \quad \Rightarrow \quad y(t) - y(t_0) = \int_{t_0}^{t} f(s, y(s))\,ds, \tag{1.4.3}$$

where on the last equation we used the Fundamental Theorem of Calculus. If we now use the initial condition, we get

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

It is this integral form of the original differential equation what we used to construct the Picard iteration. We will see later that the functions in the Picard iteration have the following property: the larger n the closer y_n is to the solution of the initial value problem.

In the examples below, we use simple the differential equations to show how to construct the picard iteration. The differential equations in these examples are linear equations, which we already know how to solve—either as separable equations or with the integrating factor method. We use simple equations because we want to show how to construct the Picard iteration, not how to solve a new type of equations. Furthermore, because the equations are so simple, we can actually compute the limit of the sequence, $\lim_{n\to\infty} y_n(t)$. Furthermore, we show that this limit is the actual solution of the differential equation computed with other methods. In real life applications usually this is not possible and the only thing we can do is to stop the Picard iteration for a value of n large enough.

Example 1.4.4. Use the Picard iteration to find the solution to

$$y' = 2\,y + 3 \qquad y(0) = 1.$$

Solution: We first transform the differential equation into an integral equation.

$$\int_0^t y'(s)\, ds = \int_0^t (2\,y(s) + 3)\, ds \quad \Rightarrow \quad y(t) - y(0) = \int_0^t (2\,y(s) + 3)\, ds.$$

Using the initial condition, $y(0) = 1$,

$$y(t) = 1 + \int_0^t (2\,y(s) + 3)\, ds.$$

We now define the sequence of approximate solutions:

$$y_0 = y(0) = 1, \quad y_{n+1}(t) = 1 + \int_0^t (2\,y_n(s) + 3)\, ds, \quad n \geqslant 0.$$

We now compute the first elements in the sequence. We said $y_0 = 1$, now y_1 is given by

$$n = 0, \quad y_1(t) = 1 + \int_0^t (2\,y_0(s) + 3)\, ds = 1 + \int_0^t 5\, ds = 1 + 5t.$$

So $y_1 = 1 + 5t$. Now we compute y_2,

$$y_2 = 1 + \int_0^t (2\,y_1(s) + 3)\, ds = 1 + \int_0^t \big(2(1+5s) + 3\big)\, ds \quad \Rightarrow \quad y_2 = 1 + \int_0^t \big(5 + 10s\big)\, ds = 1 + 5t + 5t^2.$$

So we've got $y_2(t) = 1 + 5t + 5t^2$. Now y_3,

$$y_3 = 1 + \int_0^t (2\,y_2(s) + 3)\, ds = 1 + \int_0^t \big(2(1 + 5s + 5s^2) + 3\big)\, ds$$

so we have,

$$y_3 = 1 + \int_0^t \big(5 + 10s + 10s^2\big)\, ds = 1 + 5t + 5t^2 + \frac{10}{3}\, t^3.$$

So we obtained $y_3(t) = 1 + 5t + 5t^2 + \frac{10}{3}\,t^3$. We now rewrite this expression so we can get a power series expansion that can be written in terms of simple functions. The first step is done already, to write the powers of t as t^n, for $n = 1, 2, 3$,

$$y_3(t) = 1 + 5t^1 + 5t^2 + \frac{5(2)}{3}\,t^3$$

We now multiply by one each term so we get the factorials $n!$ on each term

$$y_3(t) = 1 + 5\,\frac{t^1}{1!} + 5(2)\,\frac{t^2}{2!} + 5(2^2)\,\frac{t^3}{3!}$$

We then realize that we can rewrite the expression above in terms of power of $(2t)$, that is,

$$y_3(t) = 1 + \frac{5}{2}\,\frac{(2t)^1}{1!} + \frac{5}{2}\,\frac{(2t)^2}{2!} + \frac{5}{2}\,\frac{(2t)^3}{3!} = 1 + \frac{5}{2}\Big((2t) + \frac{(2t)^2}{2!} + \frac{(2t)^3}{3!}\Big).$$

From this last expression is simple to guess the n-th approximation

$$y_N(t) = 1 + \frac{5}{2}\Big((2t) + \frac{(2t)^2}{2!} + \frac{(2t)^3}{3!} + \cdots + \frac{(2t)^N}{N!}\Big) = 1 + \frac{5}{2}\sum_{k=1}^{N}\frac{(2t)^k}{k!}.$$

Recall now that the power series expansion for the exponential

$$e^{at} = \sum_{k=0}^{\infty}\frac{(at)^k}{k!} = 1 + \sum_{k=1}^{\infty}\frac{(at)^k}{k!} \quad\Rightarrow\quad \sum_{k=1}^{\infty}\frac{(at)^k}{k!} = (e^{at} - 1).$$

Then, the limit $N \to \infty$ is given by

$$y(t) = \lim_{N\to\infty} y_N(t) = 1 + \frac{5}{2}\sum_{k=1}^{\infty}\frac{(2t)^k}{k!} = 1 + \frac{5}{2}\big(e^{2t} - 1\big),$$

One last rewriting of the solution and we obtain

$$y(t) = \frac{5}{2}e^{2t} - \frac{3}{2}.$$

$\lhd$

Remark: The differential equation $y' = 2\,y + 3$ is of course linear, so the solution to the initial value problem in Example 1.4.4 can be obtained using the methods in Section 1.2,

$$e^{-2t}\,(y' - 2\,y) = e^{-2t}\,3 \quad\Rightarrow\quad e^{-2t}\,y = -\frac{3}{2}\,e^{-2t} + c \quad\Rightarrow\quad y(t) = c\,e^{2t} - \frac{3}{2};$$

and the initial condition implies

$$1 = y(0) = c - \frac{3}{2} \quad\Rightarrow\quad c = \frac{5}{2} \quad\Rightarrow\quad y(t) = \frac{5}{2}\,e^{2t} - \frac{3}{2}.$$

Example 1.4.5. Use the proof of Picard iteration to find the solution to

$$y' = a\,y + b \qquad y(0) = \hat{y}_0, \qquad a, b \in \mathbb{R}.$$

Solution: We first transform the differential equation into an integral equation.

$$\int_0^t y'(s)\,ds = \int_0^t (a\,y(s) + b)\,ds \quad\Rightarrow\quad y(t) - y(0) = \int_0^t (a\,y(s) + b)\,ds.$$

Using the initial condition, $y(0) = \hat{y}_0$,

$$y(t) = \hat{y}_0 + \int_0^t (a\, y(s) + b)\, ds.$$

We now define the sequence of approximate solutions:

$$y_0 = y(0) = \hat{y}_0, \quad y_{n+1}(t) = \hat{y}_0 + \int_0^t (a\, y_n(s) + b)\, ds, \quad n \geqslant 0.$$

We now compute the first elements in the sequence. We said $y_0 = \hat{y}_0$, now y_1 is given by

$$\begin{aligned} n = 0, \quad y_1(t) &= y_0 + \int_0^t (a\, y_0(s) + b)\, ds \\ &= \hat{y}_0 + \int_0^t (a\, \hat{y}_0 + b)\, ds \\ &= \hat{y}_0 + (a\, \hat{y}_0 + b)t. \end{aligned}$$

So $y_1 = \hat{y}_0 + (a\, \hat{y}_0 + b)t$. Now we compute y_2,

$$\begin{aligned} y_2 &= \hat{y}_0 + \int_0^t [a\, y_1(s) + b]\, ds \\ &= \hat{y}_0 + \int_0^t \big[a(\hat{y}_0 + (a\, \hat{y}_0 + b)s) + b\big]\, ds \\ &= \hat{y}_0 + (a\hat{y}_0 + b)t + (a\, \hat{y}_0 + b)\frac{at^2}{2} \end{aligned}$$

So we obtained $y_2(t) = \hat{y}_0 + (a\hat{y}_0 + b)t + (a\, \hat{y}_0 + b)\dfrac{at^2}{2}$. A similar calculation gives us y_3,

$$y_3(t) = \hat{y}_0 + (a\hat{y}_0 + b)t + (a\, \hat{y}_0 + b)\frac{at^2}{2} + (a\, \hat{y}_0 + b)\frac{a^2t^3}{3!}.$$

We now rewrite this expression so we can get a power series expansion that can be written in terms of simple functions. The first step is done already, to write the powers of t as t^n, for $n = 1, 2, 3$,

$$y_3(t) = \hat{y}_0 + (a\hat{y}_0 + b)\, \frac{(t)^1}{1!} + (a\, \hat{y}_0 + b)\, a\, \frac{t^2}{2!} + (a\, \hat{y}_0 + b)\, a^2\, \frac{t^3}{3!}.$$

We already have the factorials $n!$ on each term t^n. We now realize we can write the power functions as $(at)^n$ is we multiply eat term by one, as follows

$$y_3(t) = \hat{y}_0 + \frac{(a\hat{y}_0 + b)}{a}\, \frac{(at)^1}{1!} + \frac{(a\, \hat{y}_0 + b)}{a}\, \frac{(at)^2}{2!} + \frac{(a\, \hat{y}_0 + b)}{a}\, \frac{(at)^3}{3!}.$$

Now we can pull a common factor

$$y_3(t) = \hat{y}_0 + \Big(\hat{y}_0 + \frac{b}{a}\Big)\, \Big(\frac{(at)^1}{1!} + \frac{(at)^2}{2!} + \frac{(at)^3}{3!}\Big)$$

From this last expression is simple to guess the n-th approximation

$$\begin{aligned} y_N(t) &= \hat{y}_0 + \Big(\hat{y}_0 + \frac{b}{a}\Big)\, \Big(\frac{(at)^1}{1!} + \frac{(at)^2}{2!} + \frac{(at)^3}{3!} + \cdots + \frac{(at)^N}{N!}\Big) \\ \lim_{N\to\infty} y_N(t) &= \hat{y}_0 + \Big(\hat{y}_0 + \frac{b}{a}\Big)\, \sum_{k=1}^{\infty} \frac{(at)^k}{k!}. \end{aligned}$$

Recall now that the power series expansion for the exponential

$$e^{at} = \sum_{k=0}^{\infty} \frac{(at)^k}{k!} = 1 + \sum_{k=1}^{\infty} \frac{(at)^k}{k!} \quad \Rightarrow \quad \sum_{k=1}^{\infty} \frac{(at)^k}{k!} = (e^{at} - 1).$$

Notice that the sum in the exponential starts at $k = 0$, while the sum in y_n starts at $k = 1$. Then, the limit $n \to \infty$ is given by

$$\begin{aligned} y(t) &= \lim_{n\to\infty} y_n(t) \\ &= \hat{y}_0 + \Big(\hat{y}_0 + \frac{b}{a}\Big) \sum_{k=1}^{\infty} \frac{(at)^k}{k!} \\ &= \hat{y}_0 + \Big(\hat{y}_0 + \frac{b}{a}\Big) \big(e^{at} - 1\big), \end{aligned}$$

We have been able to add the power series and we have the solution written in terms of simple functions. One last rewriting of the solution and we obtain

$$y(t) = \Big(\hat{y}_0 + \frac{b}{a}\Big)\, e^{at} - \frac{b}{a}.$$

◁

Remark: We reobtained Eq. (1.2.5) in Theorem 1.2.2.

Example 1.4.6. Use the Picard iteration to find the solution of

$$y' = 5t\, y, \qquad y(0) = 1.$$

Solution: We first transform the differential equation into an integral equation.

$$\int_0^t y'(s)\, ds = \int_0^t 5s\, y(s)\, ds \quad \Rightarrow \quad y(t) - y(0) = \int_0^t 5s\, y(s)\, ds.$$

Using the initial condition, $y(0) = 1$,

$$y(t) = 1 + \int_0^t 5s\, y(s)\, ds.$$

We now define the sequence of approximate solutions:

$$y_0 = y(0) = 1, \quad y_{n+1}(t) = 1 + \int_0^t 5s\, y_n(s)\, ds, \quad n \geqslant 0.$$

We now compute the first four elements in the sequence. The first one is $y_0 = y(0) = 1$, the second one y_1 is given by

$$n = 0, \quad y_1(t) = 1 + \int_0^t 5s\, ds = 1 + \frac{5}{2}\, t^2.$$

So $y_1 = 1 + (5/2)t^2$. Now we compute y_2,

$$\begin{aligned} y_2 &= 1 + \int_0^t 5s\, y_1(s)\, ds \\ &= 1 + \int_0^t 5s\,\big(1 + \frac{5}{2}\, s^2\big)\, ds \\ &= 1 + \int_0^t \big(5s + \frac{5^2}{2}\, s^3\big)\, ds \\ &= 1 + \frac{5}{2}\, t^2 + \frac{5^2}{8}\, t^4. \end{aligned}$$

So we obtained $y_2(t) = 1 + \frac{5}{2}t^2 + \frac{5^2}{2^3}\, t^4$. A similar calculation gives us y_3,

$$\begin{aligned} y_3 &= 1 + \int_0^t 5s\, y_2(s)\, ds \\ &= 1 + \int_0^t 5s\,\big(1 + \frac{5}{2}\, s^2 + \frac{5^2}{2^3}\, s^4\big)\, ds \\ &= 1 + \int_0^t \big(5s + \frac{5^2}{2}\, s^3 + \frac{5^3}{2^3}\, s^5\big)\, ds \\ &= 1 + \frac{5}{2}\, t^2 + \frac{5^2}{8}\, t^4 + \frac{5^3}{2^3 6}\, t^6. \end{aligned}$$

So we obtained $y_3(t) = 1 + \frac{5}{2}t^2 + \frac{5^2}{2^3}\, t^4 + \frac{5^3}{2^4 3}\, t^6$. We now rewrite this expression so we can get a power series expansion that can be written in terms of simple functions. The first step is to write the powers of t as t^n, for $n = 1, 2, 3$,

$$y_3(t) = 1 + \frac{5}{2}(t^2)^1 + \frac{5^2}{2^3}\, (t^2)^2 + \frac{5^3}{2^4 3}\, (t^2)^3.$$

Now we multiply by one each term to get the right facctorials, $n!$ on each term,

$$y_3(t) = 1 + \frac{5}{2}\, \frac{(t^2)^1}{1!} + \frac{5^2}{2^2}\, \frac{(t^2)^2}{2!} + \frac{5^3}{2^3}\, \frac{(t^2)^3}{3!}.$$

No we realize that the factor $5/2$ can be written together with the powers of t^2,

$$y_3(t) = 1 + \frac{(\frac{5}{2}t^2)}{1!} + \frac{(\frac{5}{2}\, t^2)^2}{2!} + \frac{(\frac{5}{2}\, t^2)^3}{3!}.$$

From this last expression is simple to guess the n-th approximation

$$y_N(t) = 1 + \sum_{k=1}^{N} \frac{(\frac{5}{2}t^2)^k}{k!},$$

which can be proven by induction. Therefore,

$$y(t) = \lim_{N\to\infty} y_N(t) = 1 + \sum_{k=1}^{\infty} \frac{(\frac{5}{2}t^2)^k}{k!}.$$

Recall now that the power series expansion for the exponential

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

so we get

$$y(t) = 1 + (e^{\frac{5}{2}t^2} - 1) \quad \Rightarrow \quad y(t) = e^{\frac{5}{2}t^2}.$$

◁

Remark: The differential equation $y' = 5t\,y$ is of course separable, so the solution to the initial value problem in Example 1.4.6 can be obtained using the methods in Section 1.1,

$$\frac{y'}{y} = 5t \quad \Rightarrow \quad \ln(y) = \frac{5t^2}{2} + c. \quad \Rightarrow \quad y(t) = \tilde{c}\, e^{\frac{5}{2}t^2}.$$

We now use the initial condition,

$$1 = y(0) = \tilde{c} \quad \Rightarrow \quad c = 1,$$

so we obtain the solution

$$y(t) = e^{\frac{5}{2}t^2}.$$

Example 1.4.7. Use the Picard iteration to find the solution of

$$y' = 2t^4\, y, \qquad y(0) = 1.$$

Solution: We first transform the differential equation into an integral equation.

$$\int_0^t y'(s)\, ds = \int_0^t 2s^4\, y(s)\, ds \quad \Rightarrow \quad y(t) - y(0) = \int_0^t 2s^4\, y(s)\, ds.$$

Using the initial condition, $y(0) = 1$,

$$y(t) = 1 + \int_0^t 2s^4\, y(s)\, ds.$$

We now define the sequence of approximate solutions:

$$y_0 = y(0) = 1, \quad y_{n+1}(t) = 1 + \int_0^t 2s^4\, y_n(s)\, ds, \quad n \geqslant 0.$$

We now compute the first four elements in the sequence. The first one is $y_0 = y(0) = 1$, the second one y_1 is given by

$$n = 0, \quad y_1(t) = 1 + \int_0^t 2s^4\, ds = 1 + \frac{2}{5}\, t^5.$$

So $y_1 = 1 + (2/5)t^5$. Now we compute y_2,

$$\begin{aligned} y_2 &= 1 + \int_0^t 2s^4\, y_1(s)\, ds \\ &= 1 + \int_0^t 2s^4 \left(1 + \frac{2}{5}\, s^5\right) ds \\ &= 1 + \int_0^t \left(2s^4 + \frac{2^2}{5}\, s^9\right) ds \\ &= 1 + \frac{2}{5}\, t^5 + \frac{2^2}{5}\, \frac{1}{10} t^{10}. \end{aligned}$$

So we obtained $y_2(t) = 1 + \frac{2}{5}t^5 + \frac{2^2}{5^2}\,\frac{1}{2}\,t^{10}$. A similar calculation gives us y_3,

$$\begin{aligned} y_3 &= 1 + \int_0^t 2s^4\, y_2(s)\, ds \\ &= 1 + \int_0^t 2s^4 \Big(1 + \frac{2}{5}\,s^5 + \frac{2^2}{5^2}\,\frac{1}{2} s^{10}\Big)\, ds \\ &= 1 + \int_0^t \Big(2s^4 + \frac{2^2}{5}\,s^9 + \frac{2^3}{5^2}\,\frac{1}{2}\,s^{14}\Big)\, ds \\ &= 1 + \frac{2}{5}\,t^5 + \frac{2^2}{5}\,\frac{1}{10}\,t^{10} + \frac{2^3}{5^2}\,\frac{1}{2}\,\frac{1}{15}\,t^{15}. \end{aligned}$$

So we obtained $y_3(t) = 1 + \frac{2}{5}t^5 + \frac{2^2}{5^2}\,\frac{1}{2}\,t^{10} + \frac{2^3}{5^3}\,\frac{1}{2}\frac{1}{3}\,t^{15}$. We now try reorder terms in this last expression so we can get a power series expansion we can write in terms of simple functions. This is what we do:

$$\begin{aligned} y_3(t) &= 1 + \frac{2}{5}(t^5) + \frac{2^2}{5^3}\,\frac{(t^5)^2}{2} + \frac{2^3}{5^4}\,\frac{(t^5)^3}{6} \\ &= 1 + \frac{2}{5}\,\frac{(t^5)}{1!} + \frac{2^2}{5^2}\,\frac{(t^5)^2}{2!} + \frac{2^3}{5^3}\,\frac{(t^5)^3}{3!} \\ &= 1 + \frac{(\frac{2}{5}t^5)}{1!} + \frac{(\frac{2}{5}\,t^5)^2}{2!} + \frac{(\frac{2}{5}\,t^5)^3}{3!}. \end{aligned}$$

From this last expression is simple to guess the n-th approximation

$$y_N(t) = 1 + \sum_{n=1}^{N} \frac{(\frac{2}{5}t^5)^n}{n!},$$

which can be proven by induction. Therefore,

$$y(t) = \lim_{N\to\infty} y_N(t) = 1 + \sum_{n=1}^{\infty} \frac{(\frac{2}{5}t^5)^n}{n!}.$$

Recall now that the power series expansion for the exponential

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

so we get

$$y(t) = 1 + (e^{\frac{2}{5}t^5} - 1) \quad \Rightarrow \quad y(t) = e^{\frac{2}{5}t^5}.$$

◁

1.4.3. Picard vs Taylor. From the examples 1.4.2 and 1.4.4 we see that the first four Taylor and Picard approximations of solutions to the initial value problem

$$y' = 2\,y + 3, \qquad y(0) = 1$$

are exactly the same, that is,

$$\tau_n(t) = y_n(t), \quad \text{for all} \quad n = 0, 1, 2, 3.$$

In fact, our next result shows that both approximations are identical at all orders for all solutions of linear non-homogeneous equations with constant coefficient equations.

Theorem 1.4.5. *If the function y is solution of the initial value problem*

$$y'(t) = a\,y(t) + b, \qquad y(t_0) = y_0, \tag{1.4.4}$$

for any constants a, b, t_0, y_0, then

$$y_n(t) = \tau_n(t),$$

for all $t \in \mathbb{R}$ and $n = 0, 1, 2, \cdots$, where y_n is the n-order Picard approximation and τ_n is the n-order Taylor approximation centered at $t = t_0$ of the solution y.

Proof of Theorem 1.4.5: If y is the solution of the initial value in (1.4.4), its n-order Taylor expansion centered at $t = t_0$ is

$$\tau_n(t) = y_0 + y^{(1)}(t_0)\,(t - t_0) + \frac{y^{(2)}(t_0)}{2!}\,(t - t_0)^2 + \cdots + \frac{y^{(n)}(t_0)}{n!}\,(t - t_0)^n,$$

where $y^{(n)}$ is the n-th derivative of y. In particular, $\tau_0 = y_0$ and $\tau_1(t) = y_0 + y^{(1)}(t_0)\,(t - t_0)$. The Picard iteration of y is defined as follows, $y_0(t) = y_0$, and

$$y_n(t) = y_0 + \int_{t_0}^{t} \big(a\,y_{n-1}(s) + b\big)\,ds, \qquad n = 1, 2, \cdots.$$

From here we see that the zero-order of the Picard and Taylor approximations agree. Now, for $n = 1$ we get

$$y_1(t) = y_0 + \int_{t_0}^{t} \big(a\,y_0 + b\big)\,ds \quad \Rightarrow \quad y_1(t) = y_0 + (a\,y_0 + b)\,(t - t_0)$$

We now need to recall that y is solution of the differential equation in (1.4.4). This equation evalueated at $t = t_0$ says that

$$y'(t_0) = a\,y(t_0) + b. \quad \Rightarrow \quad y^{(1)}(t_0) = a\,y_0 + b.$$

Therefore, the first Picard approximation y_1 of y has the form

$$y_1(t) = y_0 + y^{(1)}(t_0)\,(t - t_0).$$

We conclude that $y_1(t) = \tau_1(t)$ for all $t \in \mathbb{R}$, that is, the order one Picard and Taylor approximations agree. Before we finish the proof we need a formula. Differentiate n-times the differential equation in (1.4.4) and evaluate the result at $t = t_0$; recall that a and b are constants, we get

$$y^{(n+1)}(t_0) = a\,y^{(n)}(t_0). \tag{1.4.5}$$

Now we are ready to finish the proof of the Theorem, and we do it by induction. We have shown that for $n = 0$ and $n = 1$ the Picard and Taylor approximations are the same. Now we prove the following:

$$y_n(t) = \tau_n(t) \quad \Rightarrow \quad y_{n+1}(t) = \tau_{n+1}(t).$$

Indeed, if $y_n(t) = \tau_n(t)$, then the Picard formula says

$$y_{n+1}(t) = y_0 + \int_{t_0}^{t} \big(a\,\tau_n(s) + b\big)\,ds$$

Using the formula for the Taylor expansion written at the beginning of the proof,

$$y_{n+1}(t) = y_0 + \int_{t_0}^{t} \Big(a\,\big(y_0 + y^{(1)}(t_0)\,(s - t_0) + \cdots + \frac{y^{(n)}(t_0)}{n!}\,(s - t_0)^n\big) + b\Big)\,ds.$$

If we reorder terms inside the integral, and we recall eq. (1.4.5), we get

$$\begin{aligned}
y_{(n+1)}(t) &= y_0 + \int_{t_0}^{t} \Big((a\,y_0 + b) + a\,y^{(1)}(t_0)\,(s-t_0) + \cdots + a\,\frac{y^{(n)}(t_0)}{n!}\,(s-t_0)^n \Big)\,ds \\
&= y_0 + \int_{t_0}^{t} \Big(y^{(1)}(t_0) + y^{(2)}(t_0)\,(s-t_0) + \cdots + \frac{y^{(n+1)}(t_0)}{n!}\,(s-t_0)^n \Big)\,ds \\
&= y_0 + y^{(1)}\,(t-t_0) + y^{(2)}(t_0)\frac{(t-t_0)^2}{2} + \cdots + \frac{y^{(n+1)}(t_0)}{n!}\,\frac{(t-t_0)^{(n+1)}}{(n+1)} \\
&= \tau_{n+1}(t).
\end{aligned}$$

We conclude that $y_n = \tau_n$ implies that $y_{n+1} = \tau_{n+1}$. This establishes the Theorem. □

In the interactive graph below we plot the Taylor approximation and the Taylor approximations for the initial value problem in examples 1.4.2 and 1.4.4. Here are a few instructions to use the interactive graph.

- The slider **Function** turns on-off the graph of the solution $y(t)$, displayed in purple.
- We graph in blue approximate solutions y_n of the differential equation constructed with the Picard iteration up to the order $n = 10$. The slider **Picard-App-Blue** turns on-off the Picard approximate solution.
- We graph in green the n-order Taylor expansion centered $t = 0$ of the solution of the differential equation, up to order $n = 10$. The slider **Taylor-App-Green** turns on-off the Taylor approximation of the solution.

Picard vs Taylor Approximations: Linear Case

Theorem 1.4.5 above says that for solutions of linear equations having constant coefficients the Picard and Taylor approximations of the solution are identical. This is not true for solutions of either linear equations with variable coefficients or nonlinear equations. In the next example we compute the first three approximations of both the Picard and Taylor approximations of solutions to a nonlinear differential equation, and we show that they are different.

Example 1.4.8. Show that the Picard and Taylor approximations of the solution $y(t)$ of the initial value problem below are different, where

$$y'(t) = y^2(t), \qquad y(0) = -1, \qquad t \geqslant 0.$$

Remark: This differential equation is separable, so we could solve it and find out that the solution of the initial value problem is

$$y(t) = -\frac{1}{(t+1)}. \tag{1.4.6}$$

Then, we could use this solution to construct the Taylor approximations centered at $t = 0$,

$$\tau_n(t) = y_0^{(0)} + y_0^{(1)}\,t + \cdots + y_0^{(n)}\,\frac{t^n}{n!}.$$

However, when we solve the example we are going to construct the Taylor approximation of the solution without using the expression of the actual solution given in (1.4.6).

Solution: We start computing the Taylor approximation of the solution $y(t)$ of the initial value problem in the example. The formula for the Taylor approximation is

$$\tau_n(t) = y_0^{(0)} + y_0^{(1)}\,t + \cdots + y_0^{(n)}\,\frac{t^n}{n!},$$

where $y_0^{(n)}$ is the n-derivative of y evaluated at $t = 0$, and the case $n = 0$ is just $y(0)$. From the initial condition we know that

$$y(0) = -1,$$

which gives us the first term in the Taylor approximation. To compute the second term we need $y'(0)$. But the differential equation says

$$y'(t) = y^2(t) \quad \Rightarrow \quad y'(0) = (y(0))^2 = (-1)^2 = 1 \quad \Rightarrow \quad y'(0) = 1,$$

which gives us the second term in the Taylor approximation. To compute the third term we need $y''(0)$. Notice that

$$y'' = (y')' = (y^2)' = 2\,y(t)\,y'(t) = 2\,y(t)\,y^2(t) \quad \Rightarrow \quad y''(t) = 2\,y^3(t).$$

If we evaluate this last expression at $t = 0$ we get

$$y''(0) = 2\,(-1)^3 = -2 \quad \Rightarrow \quad y''(0) = -2,$$

which gives us the third term in the Taylor expansion of the solution. Summarizing, we have the first three Taylor approximations of the solution,

$$\tau_0(t) = -1, \quad \tau_1(t) = -1 + t, \quad \tau_2(t) = -1 + t - t^2.$$

The Picard approximation of y is $y_0(t) = y(0)$, and then

$$y_{n+1}(t) = y(0) + \int_0^t y_n^2(s)\,ds.$$

Again, a straightforward calculation gives,

$$y_0(t) = -1 = \tau_0(t), \qquad y_1(t) = -1 + t = \tau_1(t).$$

But the approximation y_1 is different form τ_2. Indeed,

$$y_2(t) = y(0) + \int_0^t \big(y_1(s))^2\big)\,ds = -1 + \int_0^t (-1+s)^2\,ds = -1 + \int_0^t (1 - 2s + s^2)\,ds,$$

so we conclude that

$$y_2(t) = -1 + t - t^2 + \frac{t^3}{3} \quad \Rightarrow \quad y_2(t) = \tau_2(t) + \frac{t^3}{3}.$$

Therefore, $y_2 \neq \tau_2$. We decide which approximation is more precise in the following interactive graph. Here are a few instructions to use the interactive graph.

- The slider **Function** turns on-off the graph of the solution $y(t)$, displayed in purple.
- We graph in blue approximate solutions y_n of the differential equation constructed with the Picard iteration up to the order $n = 5$. The slider **Picard-App-Blue** turns on-off the Picard approximate solution.
- We graph in green the n-order Taylor expansion centered $t = 0$ of the solution of the differential equation, up to order $n = 5$. The slider **Taylor-App-Green** turns on-off the Taylor approximation of the solution.

Picard vs Taylor Approximations: Non-Linear Case - Explicit Solution

We can see in the interactive graph that the **Picard iteration approximates better the solution values than the Taylor series** expansion of that solution in a neighborhood of the initial condition. ◁

In the next example study a nonlinear differential equation, which we **do not know** how to find an explicit formula for the solution. Yet, we compute the Picard and Taylor approximations of the solutions and we can compare them.

Example 1.4.9. Show that the Picard and Taylor approximations of the solution $y(t)$ of the initial value problem below are different, where

$$y'(t) = y^2(t) + t, \qquad y(0) = -1, \qquad t \geqslant 0.$$

Solution: This differential equation is not linear, not separable, and not Euler Homogeneous. So we do not know how to find a solution y of that equation. But we can compute the Taylor and Picard approximations of the solution. The Taylor approximation is

$$\tau_n(t) = y_0^{(0)} + y_0^{(1)}\, t + \cdots + y_0^{(n)}\, \frac{t^n}{n!},$$

where $y_0^{(n)}$ is the n-derivative of y evaluated at $t = 0$, and the case $n = 0$ is just $y(0)$. We now use the initial condition $y(0) = -1$ and the equation itself to find all the $y_0^{(n)}$. Indeed,

$$y_0^{(1)} = y'(0) = \big(y(0)\big)^2 + 0 = (-1)^2 = 1 \quad \Rightarrow \quad y_0^{(1)} = 1.$$

This coefficient, and the previous one, gives us the Taylor approximation

$$\tau_1(t) = -1 + t.$$

To compute the coefficient $y_0^{(2)}$ we need to take one derivative to the differential equation,

$$y''(t) = 2y(t)\, y'(t) + 1.$$

Therefore,

$$y_0^{(2)} = y''(0) = 2y(0)\, y'(0) + 1 = 2\,(-1)\,(1) + 1 = -1 \quad \Rightarrow \quad y_0^{(2)} = -1.$$

This coefficient, and the previous ones, gives us the Taylor approximation

$$\tau_2(t) = -1 + t - t^2.$$

On the other hand, the Picard iteration is computed in the usual way, $y_0(t) = y(0)$, and

$$y_{n+1}(t) = \int_0^t \big(y_n^2(s) + s\big)\, ds.$$

So, $y_0(t) = -1 = \tau_0(t)$, and then

$$y_1(t) = -1 + \int_0^t \big((-1)^2 + s\big)\, ds \quad \Rightarrow \quad y_1(t) = -1 + t + \frac{t^2}{2}.$$

Therefore, $y_1 \neq \tau_1$. We can compute one more term in the Picard iteration,

$$\begin{aligned}
y_2(t) &= -1 + \int_0^t \Big(\Big(-1 + s + \frac{s^2}{2}\Big)^2 + s\Big)\, ds \\
&= -1 + \int_0^t \Big(1 + s^2 + \frac{s^4}{4} - 2s - s^2 + s^3 + s\Big)\, ds \\
&= -1 + \int_0^t \Big(1 - s + s^3 + \frac{s^4}{4}\Big)\, ds \\
&= -1 + t - \frac{t^2}{2} + \frac{t^4}{4} + \frac{t^5}{20}.
\end{aligned}$$

Again, $y_2 \neq \tau_2$. We decide which approximation is more precise in the following interactive graph. Here are a few instructions to use the interactive graph.

- Unlike the previous interactive graph, we do not have a slider **Function**, since we do not have an explicit expression for the solution of the differential equation.
- We graph in blue approximate solutions y_n of the differential equation constructed with the Picard iteration up to the order $n = 5$. The slider **Picard-App-Blue** turns on-off the Picard approximate solution.

- We graph in green the n-order Taylor expansion centered $t = 0$ of the solution of the differential equation, up to order $n = 5$. The slider **Taylor-App-Green** turns on-off the Taylor approximation of the solution.

Picard vs Taylor Approximations: Non-Linear Case - No Explicit Solution

We can see in the interactive graph that the **Picard iteration is different from the Taylor series** expansion of that solution in a neighborhood of the initial condition. ◁

1.4.4. Existence and Uniqueness of Solutions. The Picard iteration can be used to show that a large class of nonlinear differential equations, have solutions and that the solution is uniquely determined by appropriate initial conditions. This result is known as the Picard-Lindelöf Theorem.

Theorem 1.4.6 (Picard-Lindelöf)**.** *Consider the initial value problem*

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0. \tag{1.4.7}$$

If the function f is continuous in t and differentiable in y on some rectangle on the ty-plane containing the point (t_0, y_0) in its interior, then there is a unique solution y of the initial value problem in (1.4.7) on an open interval containing t_0.

Remark: We prove this theorem rewriting the differential equation as an integral equation for the unknown function y. Then we use this integral equation to construct a sequence of approximate solutions $\{y_n\}$ to the original initial value problem. Next we show that this sequence of approximate solutions has a unique limit as $n \to \infty$. We end the proof showing that this limit is the only solution of the original initial value problem. This proof follows [11] § 1.6 and Zeidler's [12] § 1.8. It is important to read the review on complete normed vector spaces, called Banach spaces, given in these references.

Proof of Theorem 1.4.6: We start writing the differential equation in 1.4.7 as an integral equation, hence we integrate on both sides of that equation with respect to t,

$$\int_{t_0}^{t} y'(s)\, ds = \int_{t_0}^{t} f(s, y(s))\, ds \quad \Rightarrow \quad y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds. \tag{1.4.8}$$

We have used the Fundamental Theorem of Calculus on the left-hand side of the first equation to get the second equation. And we have introduced the initial condition $y(t_0) = y_0$. We use this integral form of the original differential equation to construct a sequence of functions $\{y_n\}_{n=0}^{\infty}$. The domain of every function in this sequence is $D_a = [t_0 - a, t_0 + a]$ for some $a > 0$. The sequence is defined as follows,

$$y_{n+1}(t) = y_0 + \int_{t_0}^{t} f(s, y_n(s))\, ds, \qquad n \geqslant 0, \qquad y_0(t) = y_0. \tag{1.4.9}$$

We see that the first element in the sequence is the constant function determined by the initial conditions in (1.4.7). The iteration in (1.4.9) is called the Picard iteration. The central idea of the proof is to show that the sequence $\{y_n\}$ is a Cauchy sequence in the space $C(D_b)$ of uniformly continuous functions in the domain $D_b = [t_0 - b, t_0 + b]$ for a small enough $b > 0$. This function space is a Banach space under the norm

$$\|u\| = \max_{t \in D_b} |u(t)|.$$

See [11] and references therein for the definition of Cauchy sequences, Banach spaces, and the proof that $C(D_b)$ with that norm is a Banach space. We now show that the sequence

$\{y_n\}$ is a Cauchy sequence in that space. Any two consecutive elements in the sequence satisfy

$$\begin{aligned}
\|y_{n+1} - y_n\| &= \max_{t \in D_b} \Big| \int_{t_0}^{t} f(s, y_n(s))\, ds - \int_{t_0}^{t} f(s, y_{n-1}(s))\, ds \Big| \\
&\leqslant \max_{t \in D_b} \int_{t_0}^{t} \big| f(s, y_n(s)) - f(s, y_{n-1}(s)) \big|\, ds \\
&\leqslant k \max_{t \in D_b} \int_{t_0}^{t} |y_n(s) - y_{n-1}(s)|\, ds \\
&\leqslant kb\, \|y_n - y_{n-1}\|.
\end{aligned}$$

Denoting $r = kb$, we have obtained the inequality

$$\|y_{n+1} - y_n\| \leqslant r\, \|y_n - y_{n-1}\| \quad \Rightarrow \quad \|y_{n+1} - y_n\| \leqslant r^n\, \|y_1 - y_0\|.$$

Using the triangle inequality for norms and and the sum of a geometric series one compute the following,

$$\begin{aligned}
\|y_n - y_{n+m}\| &= \|y_n - y_{n+1} + y_{n+1} - y_{n+2} + \cdots + y_{n+(m-1)} - y_{n+m}\| \\
&\leqslant \|y_n - y_{n+1}\| + \|y_{n+1} - y_{n+2}\| + \cdots + \|y_{n+(m-1)} - y_{n+m}\| \\
&\leqslant (r^n + r^{n+1} + \cdots + r^{n+m})\, \|y_1 - y_0\| \\
&\leqslant r^n (1 + r + r^2 + \cdots + r^m)\, \|y_1 - y_0\| \\
&\leqslant r^n \Big(\frac{1 - r^m}{1 - r} \Big)\, \|y_1 - y_0\|.
\end{aligned}$$

Now choose the positive constant b such that $b < \min\{a, 1/k\}$, hence $0 < r < 1$. In this case the sequence $\{y_n\}$ is a Cauchy sequence in the Banach space $C(D_b)$, with norm $\|\ \|$, hence converges. Denote the limit by $y = \lim_{n\to\infty} y_n$. This function satisfies the equation

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

which says that y is not only continuous but also differentiable in the interior of D_b, hence y is solution of the initial value problem in (1.4.7). The proof of uniqueness of the solution follows the same argument used to show that the sequence above is a Cauchy sequence. Consider two solutions y and $\tilde{y}$ of the initial value problem above. That means,

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s)\, ds, \quad \tilde{y}(t) = y_0 + \int_{t_0}^{t} f(s, \tilde{y}(s)\, ds.$$

Therefore, their difference satisfies

$$\begin{aligned}
\|y - \tilde{y}\| &= \max_{t \in D_b} \Big| \int_{t_0}^{t} f(s, y(s))\, ds - \int_{t_0}^{t} f(s, \tilde{y}(s))\, ds \Big| \\
&\leqslant \max_{t \in D_b} \int_{t_0}^{t} \big| f(s, y(s)) - f(s, \tilde{y}(s)) \big|\, ds \\
&\leqslant k \max_{t \in D_b} \int_{t_0}^{t} |y(s) - \tilde{y}(s)|\, ds \\
&\leqslant kb\, \|y - \tilde{y}\|.
\end{aligned}$$

Since b is chosen so that $r = kb < 1$, we got that

$$\|y - \tilde{y}\| \leqslant r\, \|y - \tilde{y}\|, \quad r < 1 \quad \Rightarrow \quad \|y - \tilde{y}\| = 0 \quad \Rightarrow \quad y = \tilde{y}.$$

This establishes the Theorem. □

1.4.5. Linear vs Nonlinear Equations. The main result in § 1.2 was Theorem 1.2.3, which says that an initial value problem for a linear differential equation

$$y' = a(t)\, y + b(t), \qquad y(t_0) = y_0,$$

with a, b continuous functions on (t_1, t_2), and constants $t_0 \in (t_1, t_2)$ and $y_0 \in \mathbb{R}$, has the unique solution y on (t_1, t_2) given by

$$y(t) = e^{A(t)} \Big(y_0 + \int_{t_0}^{t} e^{-A(s)}\, b(s)\, ds \Big),$$

where we introduced the function $A(t) = \displaystyle\int_{t_0}^{t} a(s)\, ds$.

Example 1.4.10. Find the domain of the solution $y(t)$ of the initial value problem

$$(t-1)\, y' - \frac{\ln(t)}{(t-3)}\, y = \cos(2t), \qquad y(2) = 1.$$

Solution: We first write the equation above in the normal form,

$$y' = \frac{\ln(t)}{(t-1)(t-3)}\, y + \frac{\cos(2t)}{(t-1)},$$

This is a linear non-homogeneous equation,

$$y' = a(t)\, y + b(t)$$

where

$$a(t) = \frac{\ln(t)}{(t-1)(t-3)}, \qquad b(t) = \frac{\cos(2t)}{(t-1)}.$$

The coefficient $a(t)$ contains the function $\ln(t)$, which is defined only for $t \in (0, \infty)$. This same coefficient $a(t)$ is not defined for $t = 1$ and $t = 3$. The function $b(t)$ is not defined for $t = 1$. All this implies that the largest domain where both functions $a(t)$ and $b(t)$ are defined and are continuous is

$$D_1 = (0, 1) \cup (1, 3) \cup (3, \infty).$$

Which means that the solution, $y(t)$, may not be defined for $t \leqslant 0$, or at $t = 1$ or at $t = 3$. That is, we know for sure that the solution $y(t)$ of the linear differential equation above is defined either on

$$(0, 1) \qquad \text{or} \qquad (1, 3) \qquad \text{or} \qquad (3, \infty).$$

The initial condition in this problem is $y(2) = 1$, which means that the initial value of t is $t_0 = 2$ and the initial value of y is $y_0 = 1$. The important thing here is the value of t_0. Since $t_0 = 2 \in (1, 3)$, then Theorem 1.2.3 says that the domain where we know for sure the solution $y(t)$ is defined is

$$D = (1, 3).$$

$\lhd$

Remark: It is not clear whether the solution in the example above can be extended to a larger domain than $(1, 3)$. What the Theorem 1.2.3 says is that we are sure that the solution exists on the domain $D = (1, 3)$.

From the result above, Theorem 1.2.3, we can see that solutions to linear differential equations satisfy the following properties:

(a) There is an explicit expression for the solutions of a differential equations.

(b) For every initial condition $y_0 \in \mathbb{R}$ there exists a unique solution.
(c) For every initial condition $y_0 \in \mathbb{R}$ the solution $y(t)$ is defined for all (t_1, t_2).

Remark: None of these properties hold for solutions of nonlinear differential equations.

From the Picard-Lindelöf Theorem one can see that solutions to nonlinear differential equations satisfy the following properties:

(i) There is no explicit formula for the solution to every nonlinear differential equation.
(ii) Solutions to initial value problems for nonlinear equations may be non-unique when the function f does not satisfy the Lipschitz condition.
(iii) The domain of a solution y to a nonlinear initial value problem may change when we change the initial data y_0.

The next three examples (1.4.11)-(1.4.13) are particular cases of the statements in (i)-(iii). We start with an equation whose *solutions cannot be written in explicit form*.

Example 1.4.11. For every constant a_1, a_2, a_3, a_4, find all solutions y to the equation

$$y'(t) = \frac{t^2}{\big(y^4(t) + a_4\, y^3(t) + a_3\, y^2(t) + a_2\, y(t) + a_1\big)}. \tag{1.4.10}$$

Solution: The nonlinear differential equation above is separable, so we follow § 1.1 to find its solutions. First we rewrite the equation as

$$\big(y^4(t) + a_4\, y^3(t) + a_3\, y^2(t) + a_2\, y(t) + a_1\big)\, y'(t) = t^2.$$

Then we integrate on both sides of the equation,

$$\int \big(y^4(t) + a_4\, y^3(t) + a_3\, y^2(t) + a_2\, y(t) + a_1\big)\, y'(t)\, dt = \int t^2\, dt + c.$$

Introduce the substitution $u = y(t)$, so $du = y'(t)\, dt$,

$$\int \big(u^4 + a_4\, u^3 + a_3\, u^2 + a_2\, u + a_1\big)\, du = \int t^2\, dt + c.$$

Integrate the left-hand side with respect to u and the right-hand side with respect to t. Substitute u back by the function y, hence we obtain

$$\frac{1}{5}\, y^5(t) + \frac{a_4}{4}\, y^4(t) + \frac{a_3}{3}\, y^3(t) + \frac{a_2}{2}\, y(t) + a_1\, y(t) = \frac{t^3}{3} + c.$$

This is an implicit form for the solution y of the problem. The solution is the root of a polynomial degree five for all possible values of the polynomial coefficients. But it has been proven that there is no formula for the roots of a general polynomial degree bigger or equal five. We conclude that that there is no explicit expression for solutions y of Eq. (1.4.10). $\triangleleft$

We now give an example of the statement in (ii), that is, a differential equation which does not satisfy one of the hypotheses in Theorem 1.4.6. The function f has a discontinuity at a line in the ty-plane where the initial condition for the initial value problem is given. We then show that *such initial value problem has two solutions* instead of a unique solution.

Example 1.4.12. Find every solution y of the initial value problem

$$y'(t) = y^{1/3}(t), \qquad y(0) = 0. \tag{1.4.11}$$

Remark: The equation above is nonlinear, separable, and $f(t,y) = y^{1/3}$ has derivative

$$\partial_y f = \frac{1}{3}\,\frac{1}{y^{2/3}}.$$

Since the function $\partial_y f$ is not continuous at $y = 0$ and the initial condition in the problem above is at $y = 0$, this problem does not satisfies the hypotheses in Theorem 1.4.6.

Solution: The solution to the initial value problem in Eq. (1.4.11) exists but it is not unique, since we now show that it has two solutions. The first solution is

$$y_1(t) = 0.$$

The second solution can be computed as using the ideas from separable equations, that is,

$$\int \big[y(t)\big]^{-1/3}\, y'(t)\, dt = \int dt + c_0.$$

Then, the substitution $u = y(t)$, with $du = y'(t)\,dt$, implies that

$$\int u^{-1/3}\, du = \int dt + c_0.$$

Integrate and substitute back the function y. The result is

$$\frac{3}{2}\big[y(t)\big]^{2/3} = t + c_0 \quad \Rightarrow \quad y(t) = \Big[\frac{2}{3}(t + c_0)\Big]^{3/2}.$$

The initial condition above implies

$$0 = y(0) = \Big(\frac{2}{3}c_0\Big)^{3/2} \quad \Rightarrow \quad c_0 = 0,$$

so the second solution is:

$$y_2(t) = \Big(\frac{2}{3}t\Big)^{3/2}.$$

◁

Finally, an example of the statement in (iii). In this example we have an equation with *solutions defined in a domain that depends on the initial data.*

Example 1.4.13. Find the solution y to the initial value problem

$$y'(t) = y^2(t), \qquad y(0) = y_0.$$

Solution: This is a nonlinear separable equation, so we can again apply the ideas in Sect. 1.1. We first find all solutions of the differential equation,

$$\int \frac{y'(t)\,dt}{y^2(t)} = \int dt + c_0 \quad \Rightarrow \quad -\frac{1}{y(t)} = t + c_0 \quad \Rightarrow \quad y(t) = -\frac{1}{c_0 + t}.$$

We now use the initial condition in the last expression above,

$$y_0 = y(0) = -\frac{1}{c_0} \quad \Rightarrow \quad c_0 = -\frac{1}{y_0}.$$

So, the solution of the initial value problem above is:

$$y(t) = \frac{1}{\Big(\dfrac{1}{y_0} - t\Big)}.$$

This solution diverges at $t = 1/y_0$, so the domain of the solution y is not the whole real line $\mathbb{R}$. Instead, the domain is $\mathbb{R} - \{y_0\}$, so it depends on the values of the initial data y_0. ◁

In the next example we consider an equation of the form $y'(t) = f(t, y(t))$, where f does not satisfy the hypotheses in Theorem 1.4.6.

Example 1.4.14. Consider the nonlinear initial value problem

$$y'(t) = \frac{1}{(t-1)(t+1)(y(t)-2)(y(t)+3)},$$
$$y(t_0) = y_0. \tag{1.4.12}$$

Find the regions on the plane where the hypotheses in Theorem 1.4.6 are not satisfied.

Solution: In this case the function f is given by:

$$f(t, y) = \frac{1}{(t-1)(t+1)(y-2)(y+3)}, \tag{1.4.13}$$

so f is not defined on the lines

$$t = 1, \qquad t = -1, \qquad y = 2, \qquad y = -3.$$

See Fig. 24. Along these lines the hypotheses of Theorem 1.4.6 are not satisfied. Below we show two possible situations.

(a) If the initial data is $t_0 = 0$, $y_0 = 1$, then Theorem 1.4.6 implies that there exists a unique solution on any region $\hat{R}$ contained in the rectangle $R = (-1, 1) \times (-3, 2)$.

(b) If the initial data is $t = 0$, $y_0 = 2$, then the hypotheses of Theorem 1.4.6 are not satisfied and we do not know whether there is a solution to this initial value problem.

◁

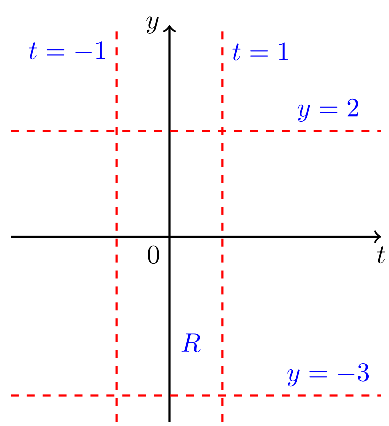


FIGURE 24. Red regions where the function f in Eq. (1.4.13) is not defined.

1.4.6. The Linearization Method. The linearization method is a slight modification of the analysis given in § 1.3.2, which was used to obtain a qualitative graph of a solution to an autonomous system without actually solving the differential equation. In subsection § 1.3.2 we had an autonomous differential equation

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

and we used the sign of the function $f(y)$ to determine the intervals where a solution of the differential equation, $y(t)$, is increasing or decreasing in time. In the linearization method we use the sign of $f'(y_c)$—the derivative of $f(y)$ evaluated at the equilibrium solutions y_c (also called critical points)—to determine the behavior of solutions $y(t)$ of the differential equation near the equilibrium solutions.

The linearization method is more restrictive than the previous method in § 1.3.2, because the information we obtain about the solution of the nonlinear system is only valid near the equilibrium solution. In the previous method, the increasing-decreasing information about the solution of the nonlinear system applies even far from equilibrium. However, there is one advantage of the linearization method—it can be easily generalized from a single equation to *a system* of nonlinear differential equations. We will study this generalization in § **??**.

Consider an autonomous differential equation

$$y'(t) = f(y(t)), \tag{1.4.14}$$

with $f(y)$ twice continuously differentiable. Let y_c an equilibrium solution of this equation (1.4.14), that is, y_c is a solution of

$$f(y_c) = 0.$$

Let us compute the Taylor expansion of $f(y)$ centered at y_c,

$$f(y) = f(y_c) + f'(y_c)\,(y - y_c) + o\big((y - y_c)^2\big).$$

We evaluate the variable y in this expansion with the function $y(t)$ and we put this expansion in the equation (1.4.14), that is,

$$y'(t) = f(y_c) + f'(y_c)\,(y(t) - y_c) + o\big((y(t) - y_c)^2\big).$$

Since y_c is an equilibrium solution we have $f(y_c) = 0$, and since y_c is a constant we have $y_c' = 0$, then

$$(y(t) - y_c)' = f'(y_c)\,(y(t) - y_c) + o\big((y(t) - y_c)^2\big).$$

If we introduce the function $\Delta y(t) = y(t) - y_c$, then the differential equation in (1.4.14) has the form

$$\Delta y'(t) = f'(y_c)\,\Delta y(t) + o\big(\Delta y^2\big).$$

This last equation says that the equation coefficients near an equilibrium solution y_c are close to the equation coefficients of the linear differential equation

$$u'(t) = f'(y_c)\,u(t).$$

This linear equation is called the linearization of equation (1.4.14).

Definition 1.4.7. *The **linearization** of the scalar autonomous equation*

$$y' = f(y)$$

at the equilibrium solution y_c is the linear system for a function $u(t)$ given by

$$u' = f'(y_c)\,u$$

Our first result is to summarize the calculation above, which is about the equation coefficients of the nonlinear system and its linearization.

Theorem 1.4.8 (Linearization Equation)**.** *If a nonlinear autonomous equation*

$$y' = f(y)$$

has a critical point y_c, then in a neighborhood of y_c the equation coefficients of this nonlinear system are close to the equation coefficients of its linearization at y_c, given by

$$u' = f'(y_c)\,u.$$

The proof of this theorem is the calculation we did above. This result is about the equations not their solutions. Here we say that the equation coefficients of the nonlinear systems and its linearization are close near the equilibrium solution. Our next result relates the solutions of the nonlinear system and its linearization near the equilibrium solution.

Theorem 1.4.9 (Linearization Solutions)**.** *Assume that the function $y(t)$ is the solution of the initial value problem*

$$y' = f(y), \qquad y(t_0) = y_0, \tag{1.4.15}$$

with $f(y)$ being twice continuously differentiable in an interval $[y_a, y_b]$, and this interval satisfying $y_0, y_c \in (y_a, y_b)$, where y_c is an equilibrium solution of the differential equation in (1.4.15). If the function $u(t)$ is the solution of the initial value problem

$$u' = f'(y_c)\, u, \qquad u(t_0) = y_0 - y_c, \tag{1.4.16}$$

then given an $\epsilon > 0$ there exists a $\delta > 0$ such that for $|t - t_0| < \delta$ and $|y_0 - y_c| < \delta$ then

$$|\Delta y(t) - u(t)| < \epsilon,$$

where $\Delta y(t) = y(t) - y_c$.

Remark: In other words, if the initial condition y_0 is close enough to the equilibrium solution y_c, then for a time t close enough to the initial time t_0 the solution $y(t)$ of the nonlinear initial value problem in (1.4.15) is close to the function $y_c + u(t)$, where $u(t)$ is the solution of the linearization initial value problem (1.4.16).

Proof of Theorem 1.4.9: Recall the Taylor expansion formula centered at a point x_0 for a function $g(x)$,

$$g(x) = g(x_0) + g'(x_0)\,(x - x_0) + \frac{g''(x_1)}{2}\,(x - x_0)^2,$$

where the point x_1 satisfies that $|x_1 - x_0| \leqslant |x - x_0|$. We now use this Taylor expansion formula in the solution of both the linearization equation (1.4.16) and the nonlinear equation (1.4.15). Since the solution of the linearization equation is

$$u(t) = \Delta y_0\, e^{k(t-t_0)},$$

where $k = f'(y_c)$ and $\Delta y_0 = y_0 - y_c$, then the Taylor expansion of $u(t)$ centered at t_0 is

$$u(t) = \Delta y_0 \Big(1 + k\,\Delta t + \frac{k^2}{2}\, e^{k(t_1 - t_0)}\,(\Delta t)^2 \Big)$$

where we introduced the point t_1 satisfying $|t_1 - t_0| \leqslant |t - t_0|$ and we also introduced the notation $\Delta t = t - t_0$. Notice that we do not have a formula for the solution $y(t)$ of the nonlinear equation. Nevertheless, we can always write the Taylor expansion centered at t_0 of this solution $y(t)$ as follows,

$$y(t) = y_0 + y'(t_0)\,\Delta t + \frac{y''(t_2)}{2}\,(\Delta t)^2$$

where we introduced the point t_2 satisfying $|t_2 - t_0| \leqslant |t - t_0|$. Since $y(t)$ is solution of the nonlinear initial value problem (1.4.15) we have

$$y'(t_0) = f(y(t_0)) = f(y_0),$$

and also we have

$$y''(t) = \frac{d}{dt} f(y(t)) = f'(y(t))\, y'(t) = f'(y(t))\, f(y(t)),$$

so we get that

$$y''(t_2) = f'(y(t_2))\, f(y(t_2)). \tag{1.4.17}$$

We will use this last equation a little later. Then, the function $\Delta y(t) = y(t) - y_c$, where y_c is the equilibrium solution used to compute $u(t)$, can be written as

$$\Delta y(t) = \Delta y_0 + f(y_0)\,\Delta t + \frac{y''(t_2)}{2}\,(\Delta t)^2.$$

We are interested in computing the difference

$$\begin{aligned}\Delta y(t) - u(t) &= \Delta y_0 + f(y_0)\,\Delta t + \frac{y''(t_2)}{2}\,(\Delta t)^2\\ &\quad - \Delta y_0\Big(1 + k\,\Delta t + \frac{k^2}{2}\,e^{k(t_1-t_0)}\,(\Delta t)^2\Big)\\ &= (f(y_0) - f'(y_c)\,\Delta y_0)\Delta t + \Big(\frac{y''(t_2)}{2} - \frac{k^2\Delta y_0}{2}\,e^{k(t_1-t_0)}\Big)(\Delta t)^2\end{aligned}$$

However, the Taylor expansion of $f(y)$ centered at the equilibrium solution y_c,

$$f(y_0) = f(y_c) + f'(y_c)\,\Delta y_0 + \frac{f''(y_1)}{2}\,(\Delta y_0)^2,$$

where we introduced y_1 such that $|y_1 - y_c| \leqslant |y_0 - y_c|$. Since $f(y_c) = 0$ we have that

$$f(y_0) - f'(y_c)\,\Delta y_0 = \frac{f''(y_1)}{2}\,(\Delta y_0)^2.$$

Using this equation we get

$$\Delta y(t) - u(t) = \frac{f''(y_1)}{2}\,(\Delta y_0)^2\Delta t + \Big(\frac{y''(t_2)}{2} - \frac{k^2}{2}\,e^{k(t_1-t_0)}\Big)(\Delta t)^2.$$

Then we can see that

$$|\Delta y(t) - u(t)| \leqslant \frac{1}{2}\Big(|f''(y_1)|\,(\Delta y_0)^2\,|\Delta t| + \big(|y''(t_2)| + |f'(y_c)|^2\,e^{|k||\Delta t|}\big)(\Delta t)^2\Big).$$

If we finally use Eq. (1.4.17), then we get

$$|\Delta y(t) - u(t)| \leqslant \frac{1}{2}\Big(|f''(y_1)|\,(\Delta y_0)^2\,|\Delta t| + \big(|f'(y(t_2))|\,|f(y(t_2))| + |f'(y_c)|^2\,e^{|f'(y_c)|\,|\Delta t|}\big)(\Delta t)^2\Big).$$

Since the function $f(y)$ is twice continuously differentiable in an interval $[y_a, y_b]$ such that $y_0, y_c \in (y_a, y_b)$, then f, f' and f'' are bounded in that interval, which means there exist positive constants M_0, M_1, M_2 such that

$$|f(y)| \leqslant M_0, \qquad |f'(y)| \leqslant M_1, \qquad |f''(y)| \leqslant M_2, \qquad \forall\, y \in [y_a, y_b].$$

This information in our inequality above gives

$$|\Delta y(t) - u(t)| \leqslant \frac{1}{2}\Big(M_2\,(\Delta y_0)^2\,|\Delta t| + \big(M_1\,M_0 + M_1^2\,e^{M_1\,|\Delta t|}\big)(\Delta t)^2\Big).$$

Therefore, given $0 < \delta \leqslant 1$ we restrict our values of t and y_0 such that

$$|\Delta t| \leqslant \delta, \qquad |\Delta y_0| \leqslant \delta.$$

Then, the inequality above has the form

$$|\Delta y(t) - u(t)| \leqslant \frac{1}{2}\big(M_2 + M_1\,M_0 + M_1^2\,e^{M_1}\big)\,\delta^2.$$

Therefore, given any $\epsilon > 0$ we choose δ to be

$$\delta \leq \min\Big\{1, \sqrt{\frac{2\epsilon}{M_2 + M_1 M_0 + M_1^2\,e^{M_1}}}\Big\}.$$

For such δ w conclude that

$$|\Delta y(t) - u(t)| \leq \epsilon.$$

This establishes the Theorem. □

1.4.7. Exercises.

1.4.1.- Use the Picard iteration to find the first four elements, y_0, y_1, y_2, and y_3, of the sequence $\{y_n\}_{n=0}^{\infty}$ of approximate solutions to the initial value problem

$$y' = 6y + 1, \quad y(0) = 0.$$

1.4.2.- Use the Picard iteration to find the information required below about the sequence $\{y_n\}_{n=0}^{\infty}$ of approximate solutions to the initial value problem

$$y' = 3y + 5, \quad y(0) = 1.$$

(a) The first 4 elements in the sequence, y_0, y_1, y_2, and y_3.
(b) The general term $c_k(t)$ of the approximation

$$y_n(t) = 1 + \sum_{k=1}^{n} \frac{c_k(t)}{k!}.$$

(c) Find the limit $y(t) = \lim_{n\to\infty} y_n(t)$.

1.4.3.- Find the domain where the solution of the initial value problems below is well-defined.

(a) $y' = \dfrac{-4t}{y}, \quad y(0) = y_0 > 0.$
(b) $y' = 2ty^2, \quad y(0) = y_0 > 0.$

1.4.4.- By looking at the equation coefficients, find a domain where the solution of the initial value problem below exists,

(a) $(t^2 - 4)\, y' + 2\ln(t)\, y = 3t$, and initial condition $y(1) = -2$.
(b) $y' = \dfrac{y}{t(t-3)}$, and initial condition $y(-1) = 2$.

1.4.5.- State where in the plane with points (t, y) the hypothesis of Theorem 1.4.6 are not satisfied.

(a) $y' = \dfrac{y^2}{2t - 3y}$.
(b) $y' = \sqrt{1 - t^2 - y^2}$.

CHAPTER 2

Second Order Linear Equations

Newton's second law of motion, $ma = f$, is maybe one of the first differential equations written. This is a second order equation, since the acceleration is the second time derivative of the particle position function. Second order differential equations are more difficult to solve than first order equations. In § 2.1 we compare results on linear first and second order equations. While there is an explicit formula for all solutions to first order linear equations, not such formula exists for all solutions to second order linear equations. The most one can get is the result in Theorem 2.1.9. In § 2.2 we find explicit formulas for all solutions to linear second order equations that are both homogeneous and with constant coefficients. These formulas are generalized to nonhomogeneous equations in § 2.3. In § **??** we solve special second order equations, which include Newton's equations in the case that the force depends only on the position function. In this case we see that the mechanical energy of the system is conserved. We also present in more detail the Reduction Order Method to find a new solution of a second order equation if we already know one solution of that equation.

2.1. General Properties

The differential equation that started the whole field of differential equations is Newton's second law of motion for a point particle—the force acting on the particle is equal to its mass times the acceleration of the particle. Newton's equation is a second order differential equation for the position of the particle as function of time. The equation is linear when the force is a linear function of the position and the velocity of the particle.

In this section we study general second order linear differential equations but we focus our examples on Newton's equation for systems moving in one space dimension under forces linear in the position and velocity. Our main example is a mass-spring system, where an object is attached to a spring and both oscillate along a straight line. An integral of Newton's equation defines the mechanical energy of the mass-spring system. We show that this energy is constant during the motion of springs oscillating without friction.

We then state Theorem 2.1.3, which says that second order linear equations with continuous coefficients always have solutions, and these solutions are defined on the same domain where the equation coefficients are continuous. Furthermore, the solution is uniquely determined by two appropriate initial conditions.

The equations for mass-spring systems are of a particular type, called homogeneous equations. We show that homogeneous equations satisfy the superposition property—the linear combination of two solutions is also a solution. This property is important to prove our second main result, Theorem 2.1.9, which is the closest we can get to a formula for solutions to second order linear homogeneous equations. This theorem says that to know all solutions to second order linear homogeneous equations we only need to know two solutions that are not proportional to each other, called fundamental solutions.

We end this section introducing the Wronskian of two functions, which happens to be nonzero when the functions are not proportional to each other. When the functions are solutions to a second order linear differential equation, then the Wronskian itself satisfies a first order linear equation. This result is called Abel's theorem and it shows that solutions with different initial conditions will be not proportional to each other.

2.1.1. Definitions and Examples.

We introduce *second order* differential equations and then the particular case of second order *linear* differential equations.

Definition 2.1.1. *A* ***second order differential equation*** *for $y(t)$ is*

$$y'' = f(t, y, y'). \tag{2.1.1}$$

The equation (2.1.1) is ***linear, non-homogeneous*** *iff*

$$y'' + a_1(t)\, y' + a_0(t)\, y = b(t), \tag{2.1.2}$$

where a_1, a_0, b are given functions on the interval $I \subset \mathbb{R}$. The equation (2.1.2):

(a) is ***homogeneous*** *iff the source $b(t) = 0$ for all $t \in \mathbb{R}$;*

(b) has ***constant coefficients*** *iff a_1 and a_0 are constants;*

(c) has ***variable coefficients*** *iff either a_1 or a_0 is not constant.*

Remarks:

(a) The homogeneous equations presented here are essentially different from the Euler homogeneous equations we studied in § 1.1.

(b) We define second order linear equations with constant coefficients when only a_1 and a_0 are constants, but b can be non-constant. This is a different definition from the first order linear equations with constant coefficients, where we required that the coefficient b be also constant.

Example 2.1.1.

(a) A second order, linear, homogeneous, constant coefficients equation is

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

(b) A second order, linear, nonhomogeneous, constant coefficients, equation is

$$y'' - 3y' + y = \cos(3t).$$

(c) A second order, linear, nonhomogeneous, variable coefficients equation is

$$y'' + 2t\, y' - \ln(t)\, y = e^{3t}.$$

(d) A second order, non-linear equation is

$$y'' + 2t\, y' - \ln(t)\, y^2 = e^{3t}.$$

◁

2.1.2. Newtonian Dynamics. Newton's second law of motion for a point particle having mass m moving in one space dimension, y, under a force f is an example of a second order differential equation. This equation is usually written as

$$ma = f,$$

where $a = y''$ is the particle's acceleration—the second time derivative of the position function y. The force acting on the particle can depend on time, on the position, and on the velocity of the particle. Then, Newton's equation can be written as

$$m\, y''(t) = f(t, y(t), y'(t)). \tag{2.1.3}$$

If the force acting on the particle is linear in the particle's position and velocity, then the differential equation in (2.1.3) is linear. Now we show a few examples of Newton's equation where the force is a linear function in the position and or velocity of the particle.

Example 2.1.2 (Mass-Spring, No Friction)**.** Consider a spring attached to a ceiling from its top end and having an object of mass m hanging from its bottom end, as pictured in Fig. 1. In this picture we have two springs, the one on the left is at rest at the equilibrium position, the one on the right is not at rest, since it is stretched out of the equilibrium position. We set y to be a vertical coordinate, with $y = 0$ at the equilibrium position of the mass-spring system and positive downwards. Newton's equation for this system is

$$m\, y'' = f_T,$$

where m is the mass of the object and f_T represents all the forces acting on the system,

$$f_T = f_g + f.$$

The first term is the weight of the object, $f_g = mg$, which is a positive term since it is directed downwards. We denoted by g the acceleration of gravity near the Earth surface, $g = 9.81$ meters/(seconds squared). The force done by the spring on the mass can be decomposed in two terms,

$$f = f_0 + f_s.$$

The force f_0 is directed upwards and compensates the weight of the object,

$$f_0 = -mg.$$

The force f_0 is responsible for keeping the mass-spring at the equilibrium position. The force f_s is the extra force done by the spring when it is stretched out of equilibrium. It is observed experimentally that the force f_s is proportional to the stretching of the spring away from equilibrium, y, and in the opposite direction of the stretching,

$$f_s = -k\, y, \qquad k > 0.$$

This stretching force f_s is called *Hooke's Law*, the positive constant k is the spring constant, with units of mass/(time squared). This constant characterizes the stiffness of the spring, the larger the constant the more stiff is the spring. Then, Newton's equation for this system, $m\, y'' = f_T$, has the form

$$m\, y'' = f_g + f_0 + f_s \quad \Rightarrow \quad m\, y'' = mg + f_0 - k\, y.$$

But $f_0 + mg = 0$, since these forces cancel each other, then

$$m\, y'' + k\, y = 0.$$

We see that this is a second order, linear, differential equation for the position $y(t)$ as function of time. ◁

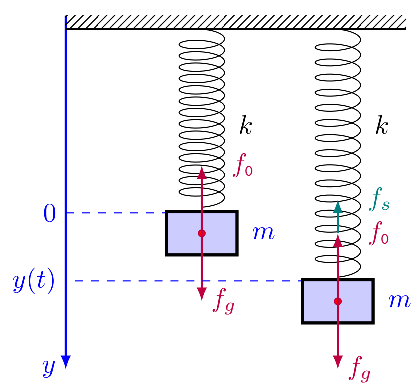


FIGURE 1. Mass-Spring System with coordinate system.

Example 2.1.3 (Mass-Spring with Friction)**.** Consider a mass-spring system as described in the example above. Suppose that the whole system is oscillating inside a liquid bath. In this case appears a damping force, from the friction between the oscillating mass and the liquid. The damping force is given by

$$f_d = -d\, y', \qquad d > 0.$$

The friction force damps the oscillations because it opposes the movement. Then, Newton's equation, $m\, y'' = f_T$, has a right-hand side $f_T = f_g + f_0 + f_s + f_d$. As in the previous example, the first two terms in the force cancel out, $f_g + f_0 = 0$, and we get

$$m\, y'' = -k\, y - d\, y' \quad \Rightarrow \quad m\, y'' + d\, y' + k\, y = 0.$$

We see from the last equation above that any second order linear differential equation with positive constant coefficients m, d, and k can always be identified with Newton's equation for a mass-spring system having spring constant k, mass m, and moving in one space dimension through a medium with damping constant d. ◁

Example 2.1.4 (Falling Particle)**.** Consider a particle moving vertically near the surface of the Earth. Discard any friction with the air and call $y(t)$ the particle position as function of time. The only force, f, acting on the particle is the gravitational force of the Earth, which near the Earth's surface is constant in time, independent of the particle's position. This force is given by

$$f = mg,$$

where the constant $g = 9.81 \ \ \mathrm{m/s}^2$ is called the acceleration of gravity near the surface of the Earth. Then, Newton's equation for such particle is

$$m\,y'' = -mg,$$

where we assumed that the particle position, y, is positive in the upward direction. ◁

2.1.3. Conservation of the Energy. If the force acting on a particle depends *only on the position* of the particle, then the velocity of the particle is an integrating factor for Newton's equation of motion. This means that Newton's equation multiplied by the velocity becomes the total time derivative of a function, called the mechanical energy of the particle. Newton's equation implies this energy remains constant during the motion.

Theorem 2.1.2 (Conservation of the Mechanical Energy)**.** *Let the position function $y(t)$ of a particle with mass m be a solution of Newton's equation*

$$m\,y'' = f.$$

If the force depends only on the particle's position,

$$f = f(y),$$

then there is a quantity, $E(t)$, constructed with the position $y(t)$ and velocity $v(t) = y'(t)$ of the particle that remains constant along the motion of the particle, that is,

$$E(t) = E(0).$$

This quantity is called the mechanical energy of the particle, and it is given by

$$E(t) = \frac{1}{2}\,m\,\big(v(t)\big)^2 + V(y(t)),$$

where we introduced the potential energy of the particle,

$$V(y) = -\int f(y)\,dy.$$

Remarks:

(a) The conservation of the mechanical energy holds for forces of the form

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

that is, the force is function only of the position.

(b) From the definition of the potential energy $V(y)$ we see that its y-derivative is related to the force,

$$f = -\frac{dV}{dy}.$$

(c) The function

$$K = \frac{1}{2}\,m\,v^2$$

is called the kinetic energy of the system. So the mechanical energy is given by the sum of kinetic energy (measuring actual movement) and potential energy (capacity to produce movement),

$$E = K + V.$$

Proof of Theorem 2.1.2: Since the force on the particle depends only on the particle position, $f = f(y)$, we can always compute its (negative) antiderivative,

$$V(y) = -\int f(y)\,dy \quad \Rightarrow \quad f = -\frac{dV}{dy}.$$

The function V is called the *potential energy* of the particle. If we write Newton's law of motion $m\,y'' = f$ in terms of the potential energy we get

$$m\,y'' = -\frac{dV}{dy},$$

where, as usual, prime means derivative with respect to time,

$$y'(t) = \frac{dy}{dt}.$$

Now multiply Newton's equation by the particle velocity, y',

$$m\,y'\,y'' = -\frac{dV}{dy}\,y'.$$

The chain rule for derivatives of a composition of functions says that

$$m\,y'\,y'' = \frac{1}{2}\,m\,\frac{d}{dt}\big((y')^2\big) \quad \text{and} \quad -\frac{dV}{dy}\,y' = -\frac{d}{dt}(V(y(t))$$

Therefore, Newton's equation can be written as a total derivative,

$$\frac{d}{dt}\Big(\frac{1}{2}\,m\,v^2 + V(y)\Big) = 0,$$

where $v = y'$. If we introduce the mechanical energy of the particle,

$$E = \frac{1}{2}\,m\,v^2 + V(y),$$

then Newton's equation implies

$$E'(t) = 0 \quad \Rightarrow \quad E(t) = E(0).$$

We see that the mechanical energy is conserved during the motion of the particle, $y(t)$, and it is equal to its initial value. This establishes the Theorem. □

Example 2.1.5 (Mass-Spring System Undamped)**.** Show that the mechanical energy of a mass-spring system, as pictured in Fig. 1, is conserved.

Solution: We showd in Example 2.1.2 that Newton's equation of a mass-spring oscillating without friction is

$$m\,y'' + k\,y = 0,$$

where m is the object mass and k is the spring constant. We could use the formula for the mechanical energy given in Theorem 2.1.2, but to understand better where this formula comes from we derive it again for this system. Multiply Newton's equation by the velocity y' and recall the chain rule for derivatives.

$$m\,y'\,y'' + k\,y\,y' = 0 \quad \Rightarrow \quad m\,\frac{d}{dt}\Big(\frac{(y')^2}{2}\Big) + k\frac{d}{dt}\Big(\frac{y^2}{2}\Big) = 0, \quad \Rightarrow \quad \frac{d}{dt}\Big(\frac{m}{2}\,(y')^2 + \frac{k}{2}\,y^2\Big) = 0.$$

Denote the velocity by $v = y'$, then the mechanical energy for a mass-spring system is

$$E(t) = \frac{1}{2}\, m\, v^2 + \frac{1}{2}\, k\, y^2,$$

and the third equation on the right above says this energy is conserved along the motion,

$$E(t) = E(0).$$

Let us introduce the *kinetic energy* and the *potential energy*, respectively,

$$K(v) = \frac{1}{2}\, m\, v^2, \qquad V(y) = \frac{1}{2}\, k\, y^2.$$

Notice that these functions are non-negative. Then, the conservation of the mechanical energy has the form

$$K(v) + V(y) = E(0).$$

If the kinetic energy increases, then the potential energy must decrease; and viceversa. Because each term is non-negative, the maximum value of the kinetic energy happens when the potential energy vanishes; and viceversa. ◁

Example 2.1.6 (Mass-Spring System Undamped)**.** An object of mass 10 grams is hanging from a spring with constant 20 grams per seconds square. Assume that the object is initially at rest and the spring is stretched 10 centimeters. Then, find both the maximum speed of the object, $v_{\max}$, and the maximum displacement, $y_{\max}$, achieved by the object while oscillating.

Solution: We know that the differential equation describing the object movement is

$$m\, y'' + k\, y = 0, \quad m = 10, \quad k = 20.$$

Unfortunately, we do not know how to solve this differential equation, yet. Fortunately, we do not need to solve this equation to answer the question above, because this system has a conserved energy,

$$E(t) = E(0),$$

where

$$E(t) = \frac{1}{2}\, m\, v^2 + \frac{1}{2}\, k\, y^2.$$

and $v = y'$. Using the data of the problem we get

$$E(t) = 5(v(t))^2 + 10(y(t))^2.$$

Since we know that at the initial time $t = 0$ we have

$$y(0) = 10, \quad v(0) = 0 \quad \Rightarrow \quad E(0) = 5(v(0))^2 + 10(y(0))^2 = 0 + 1000 \quad \Rightarrow \quad E(0) = 1000.$$

Since the energy is conserved we get that

$$\frac{1}{2}\, m\, v^2 + \frac{1}{2}\, k\, y^2 = 1000 \qquad \text{for all } t.$$

The left hand side has a maximum speed $v_{\max}$ when y^2 has the lowest possible value, and that is when $y = 0$. This happens when the object passes through the equilibrium position. At that position the speed is the maximum possible, given by

$$5\,(v_{\max})^2 + 10(0)^2 = 1000 \quad \Rightarrow \quad |v_{\max}| = 10\sqrt{2}.$$

The left hand side has a maximum displacement $y_{\max}$ when v^2 has the lowest possible value, and that is when $v = 0$. This happens when the object's velocity changes direction. At that time the object is at the maximum elongation, given by

$$5(0)^2 + 10\,(y_{\max})^2 = 1000 \quad \Rightarrow \quad |y_{\max}| = 10.$$

◁

In our next example we shoot a bullet in the vertical direction and we want to find the maximum altitude achieved by the bullet.

Example 2.1.7 (Mass Falling on Earth)**.** An object of mass m kilograms moves vertically to the ground under the action of the Earth gravitational acceleration near the surface, denoted as g, which has the value of $g \simeq 9.81$ meters per second square (although we do not need the exact value here). Denote by y vertical coordinate, *positive upwards*, and let $y = 0$ be at the earth surface. If the initial position of the object is $y(0) = y_0$ meters and its initial velocity is $y'(0) = v_0$ meters per second, find the maximum altitude $y_{\max}$ achieved by the object.

Solution: Once the object is in motion the only force acting on it is its own weight,

$$f_g = -mg,$$

where the negative sign indicates the force is directed downwards, which in our coordinate system is negative. Then, the differential equation describing the projectile movement is

$$m\, y'' = -mg,$$

with m the object mass and g the Earth gravitational acceleration. Although we know how to solve this differential equation for the function $y(t)$, we also can solve this problem using only the mechanical energy of this system. Multiply Newton's equation by y',

$$m\, y'\, y'' + mg\, y' = 0 \quad \Rightarrow \quad \frac{d}{dt}\Big(\frac{1}{2} m (y')^2 + mg\, y\Big) = 0.$$

As usual, denote the velocity by $v = y'$, then the energy is

$$E(t) = \frac{m}{2}\, v^2 + mg\, y.$$

The previous equation says that this energy is conserved along the motion, that is

$$E'(t) = 0 \quad \Rightarrow \quad E(t) = E(0).$$

Using the initial condition of the problem, $y(0) = y_0$ and $v(0) = v_0$ we get the initial energy

$$E(0) = \frac{m}{2}\, v_0^2 + mg\, y_0.$$

By looking at the energy we see that the maximum altitude achieved at a time $t_{\max}$ when the object velocity vanishes, therefore

$$E(t_{\max}) = E(0) \quad \Rightarrow \quad 0 + mg\, y_{\max} = \frac{m}{2}\, v_0^2 + mg\, y_0 \quad \Rightarrow \quad y_{\max} = \frac{v_0^2}{2g} + y_0.$$

Notice that the maximum altitude does not depend on the mass m of the object, it depends only on the initial velocity and the initial position. ◁

In the next example we compute the equation satisfied by the mechanical energy of a mass-spring system oscillating with friction.

Example 2.1.8 (Energy of Spring with Friction)**.** Consider a mass-spring system with friction, as described in Example 2.1.3. Find the equation satisfied by the mechanical energy of this mass-spring system.

Solution: Following Example 2.1.3 we denote by m the mass of the object hanging from the spring, k the spring constant, and d the liquid damping constant. We saw in that example that Newton's equation for this mass-spring system with friction is

$$m\, y'' + d\, y' + k\, y = 0.$$

To obtain the equation for the mechanical energy we proceed as in the proof of Theorem 2.1.2. We multiply Newton's equation by the integrating factor, the velocity y',

$$m\, y'\, y'' + d\,(y')^2 + k\, y\, y' = 0.$$

We use the chain rule to construct the mechanical energy function on the left-side and we move the term proportional to $(y')^2$ to the right-hand side,

$$\frac{1}{2}\, m\, \big((y')^2\big)' + \frac{1}{2}\, k\, \big(y^2\big)' = -d\,(y')^2.$$

We introduce the notation $v = y'$ and we get

$$\Big(\frac{1}{2}\, m\, v^2 + \frac{1}{2}\, k\, y^2\Big)' = -d\, v^2.$$

If we denote the mechanical energy for the mass-spring system as usual,

$$E(t) = \frac{1}{2}\, m\, v^2 + \frac{1}{2}\, k\, y^2,$$

then we found that

$$E'(t) = -d\, v^2 \leqslant 0.$$

This is the equation satisfied by the mechanical energy of a mass-spring system with friction. The right-hand side above is negative for nonzero velocity, meaning that the mechanical energy is a decreasing function of time, hence not conserved. ◁

Remark: A cornerstone principle in physics is that energy cannot be created nor destroyed, it is called the conservation of the energy. Although it seems that our result in Example 2.1.8 contradicts this principle, further study will reveal that it does not. It has happened many times in the history of physics that the conservation of the energy seems to fail; only to be found out later that the conservation of the energy is indeed true and the real problem was that we were not looking at the whole picture. This is exactly what is happening in Example 2.1.8. When an object oscillates in a viscous liquid the mechanical energy decreases because it is transformed into a different type of energy, heat. The temperature of the liquid and the object increase as the oscillations slow down. Since we are not taking into account the thermal energy in our previous example, our result only shows the decrease in the mechanical energy of the spring.

Example 2.1.9 (Motion in Viscous Liquid)**.** A bullet with mass m is shot horizontally with initial velocity v_0 into a tank containing a viscous liquid with damping constant d. Discarding any vertical movement, how long does it take until the kinetic energy of the bullet is 1% of the initial kinetic energy? (That is, the bullet practically stops.)

Solution: The gravitational force on the bullet is in the vertical direction, but we are discarding the movement in that direction. We focus only on the movement in the horizontal direction, so let's denote our position function as $x(t)$, positive in the direction the bullet is moving. The only force acting on the bullet in the horizontal direction is the friction with the liquid, $f_d = -d\, x'$. Newton's second law of motion says that

$$m\, x'' = -d\, x'.$$

We can obtain a formula for the kinetic energy if we multiply Newton's equation by x',

$$m\, x'\, x'' = -d\,(x')^2 \quad \Rightarrow \quad \frac{d}{dt}\Big(\frac{1}{2}\, m\,(x')^2\Big) = -d\,(x')^2.$$

Denote the velocity by $v = x'$, then

$$\frac{d}{dt}\Big(\frac{1}{2}\,m\,v^2\Big) = -\frac{2d}{m}\,\Big(\frac{1}{2}\,m\,v^2\Big) \quad \Rightarrow \quad \frac{d}{dt}K = -\frac{2d}{m}\,K,$$

where we introduced the bullet's kinetic energy $K = (m/2)v^2$. So, this kinetic energy satisfies the differential equation

$$K' + \frac{2d}{m}\,K = 0.$$

The solution of this equation is

$$K(t) = K(0)\,e^{-(2d/m)t}.$$

We need to find a time t_1 such that $K(t_1) = K(0)/100$, that is

$$\frac{K(0)}{100} = K(0)\,e^{-(2d/m)t_1} \quad \Rightarrow \quad \ln\Big(\frac{1}{100}\Big) = -\frac{2d}{m}\,t_1 \quad \Rightarrow \quad t_1 = \frac{m}{2d}\,\ln(100).$$

◁

2.1.4. Existence and Uniqueness of Solutions.

Second order linear differential equations have solutions in the case that the equation coefficients are continuous functions. And the solution of the equation is unique when we specify two appropriate initial conditions. The latter means that the two arbitrary integrations constants of the general solution can be uniquely determined by appropriately chosen initial conditions. In this short subsection we only mention this result without a proof.

Theorem 2.1.3 (Existence and Uniqueness). *Consider the initial value problem*

$$y'' + a_1(t)\,y' + a_0(t)\,y = b(t), \qquad y(t_0) = y_0, \quad y'(t_0) = y_1. \tag{2.1.4}$$

If the functions a_1, a_0, b are continuous on an open interval (t_1, t_2), then there exists a unique solution $y(t)$ of Eq. (2.1.4) defined on that interval (t_1, t_2) for every choice of the initial data $t_0 \in (t_1, t_2)$, and y_0, $y_1 \in \mathbb{R}$.

Remark: The fixed point argument used in the proof of Picard-Lindelöf's Theorem 1.4.6 can be extended to prove Theorem 2.1.3.

Example 2.1.10. Find the domain of the solution of the initial value problem

$$(t-1)\,y'' - 3t\,y' + \frac{4(t-1)}{(t-3)}\,y = t(t-1), \qquad y(2) = 1, \qquad y'(2) = 0.$$

Solution: We first write the equation above in the form given in the Theorem above,

$$y'' - \frac{3t}{(t-1)}\,y' + \frac{4}{(t-3)}\,y = t.$$

The equation coefficients are defined on the domain

$$(-\infty, 1) \cup (1, 3) \cup (3, \infty).$$

Which means that the solution may not be defined at $t = 1$ or $t = 3$. That is, we know for sure that the solution is defined on

$$(-\infty, 1) \qquad \text{or} \qquad (1, 3) \qquad \text{or} \qquad (3, \infty).$$

Since the initial condition is at $t_0 = 2 \in (1, 3)$, then the domain where we know for sure the solution is defined is

$$D = (1, 3).$$

◁

Remark: It is not clear whether the solution in the example above can be extended to a larger domain than $(1, 3)$. What the Theorem 2.1.3 says is that we are sure that the solution exists on the domain $(1, 3)$.

2.1.5. Properties of Homogeneous Equations. All the second order linear differential equations studied in the examples above have been homogeneous, as defined in Def. 2.1.1. In the rest of this section we study general concepts about homogeneous equations that will help us get as close as possible to a formula for their solutions. These concepts include the notion of an operator, linear operators, and the superposition property of solutions to homogeneous equations. We start introducing the notion of an operator.

Definition 2.1.4 (Operator)**.** *A second order linear differential* ***operator****, denoted as L, acting on twice continuously differentiable functions, y, is given by*

$$L(y) = y'' + a_1(t)\, y' + a_0(t)\, y, \tag{2.1.5}$$

where a_1, a_0, are given continuous functions.

Operators provide a convenient notation to write second order linear differential equations. The differential equation

$$y'' + a_1(t)\, y' + a_0(t)\, y = f(t)$$

can be written as

$$L(y) = f.$$

Example 2.1.11. Compute the operator $L(y) = t\, y'' + 2y' - \dfrac{8}{t}\, y$ acting on $y(t) = t^3$.

Solution: Since $y(t) = t^3$, then $y'(t) = 3t^2$ and $y''(t) = 6t$, hence

$$L(t^3) = t\,(6t) + 2(3t^2) - \frac{8}{t}\, t^3 \quad \Rightarrow \quad L(t^3) = 4t^2.$$

The function L acts on the function $y(t) = t^3$ and the result is the function $L(t^3) = 4t^2$. ◁

In the definition above we see that L operates on a function y and the result is a new function given by Eq. (2.1.5). For that reason L is called an *operator*, also a *transformation*. The name emphasizes that L is a special type of function, which operates on other functions, instead of usual functions that operate on numbers. The operator L above is also called a *differential operator*, since $L(y)$ contains derivatives of y. Furthermore, L is called a *second order* differential operator, since the highest derivative in L is a second order derivative. Lastly, the operator L above is called a linear operator, because it satisfies the following property.

Definition 2.1.5 (Linear Operator)**.** *An operator L is a* ***linear operator*** *iff for every pair of functions y_1, y_2 and constants c_1, c_2 holds*

$$L(c_1 y_1 + c_2 y_2) = c_1 L(y_1) + c_2 L(y_2). \tag{2.1.6}$$

Now we show that the operator L defined in Def. 2.1.4 is indeed a linear operator.

Theorem 2.1.6 (Linear Operator)**.** *The operator*

$$L(y) = y'' + a_1\, y' + a_0\, y,$$

as defined in Def. 2.1.4 is a linear operator.

Proof of Theorem 2.1.6: This is a straightforward calculation:

$$L(c_1y_1 + c_2y_2) = (c_1y_1 + c_2y_2)'' + a_1\,(c_1y_1 + c_2y_2)' + a_0\,(c_1y_1 + c_2y_2).$$

Recall that derivations is a linear operation and then reoorder terms in the following way,

$$L(c_1y_1 + c_2y_2) = \big(c_1y_1'' + a_1\,c_1y_1' + a_0\,c_1y_1\big) + \big(c_2y_2'' + a_1\,c_2y_2' + a_0\,c_2y_2\big).$$

Introduce the definition of L back on the right-hand side. We then conclude that

$$L(c_1y_1 + c_2y_2) = c_1L(y_1) + c_2L(y_2).$$

This establishes the Theorem. □

The linearity of an operator L translates into the superposition property of the solutions to the homogeneous equation $L(y) = 0$.

Theorem 2.1.7 (Superposition)**.** *If L is a linear operator and y_1, y_2 are solutions of the homogeneous equations $L(y_1) = 0$, $L(y_2) = 0$, then for every constants c_1, c_2 holds*

$$L(c_1\,y_1 + c_2\,y_2) = 0.$$

Remarks:

(a) This result is *not true* for nonhomogeneous equations. Indeed, given functions y_1 and y_2 solutions of the same non-homogeneous equation

$$L(y_1) = f, \qquad L(y_2) = f,$$

the function $(y_1 + y_2)$ satisfies a different differential equation,

$$L(y_1 + y_2) = L(y_1) + L(y_2) = f + f = 2f.$$

(b) The linearity of an operator L and the superposition property of solutions of the equation $L(y) = 0$ are deeply connected—like two sides of the same coin.

Proof of Theorem 2.1.7: Verify that the function $y = c_1y_1 + c_2y_2$ satisfies $L(y) = 0$ for every constants c_1, c_2, that is,

$$L(y) = L(c_1y_1 + c_2y_2) = c_1\,L(y_1) + c_2\,L(y_2) = c_1\,0 + c_2\,0 = 0.$$

This establishes the Theorem. □

We now introduce the notion of linearly dependent or independent functions.

Definition 2.1.8. *Consider two functions y_1, y_2 defined on an interval I. The functions are **linearly dependent** iff there is a constant, c, so that for all $t \in I$ holds*

$$y_1(t) = c\,y_2(t).$$

*Otherwise, the functions are **linearly independent**.*

Remarks:

(a) Two functions y_1, y_2 are linearly dependent when they are proportional to each other.
(b) The function $y_1 = 0$ is proportional to every other function y_2, since $y_1 = 0 = 0\,y_2$.
(c) If the functions y_1, y_2 satisfy $y_1 = t\,y_2$, then they are linearly independent, since they are not proportional to each other.

The definitions of linearly dependent or independent functions found in the literature are equivalent to the definition given here, but they are worded in a slight different way.

Often in the literature, two functions are called linearly dependent on the interval I iff there exist constants c_1, c_2, not both zero, such that for all $t \in I$ holds

$$c_1 y_1(t) + c_2 y_2(t) = 0.$$

Two functions are called linearly independent on the interval I iff they are not linearly dependent, that is, the only constants c_1 and c_2 that for all $t \in I$ satisfy the equation

$$c_1 y_1(t) + c_2 y_2(t) = 0$$

are the constants $c_1 = c_2 = 0$. This wording makes it simple to generalize these definitions to an arbitrary number of functions.

Example 2.1.12.

(a) Show that $y_1(t) = \sin(t)$, $y_2(t) = 2\sin(t)$ are linearly dependent.
(b) Show that $y_1(t) = \sin(t)$, $y_2(t) = t\sin(t)$ are linearly independent.

Solution:
Part (a): This is trivial, since $2y_1(t) - y_2(t) = 0$.

Part (b): Find constants c_1, c_2 such that for all $t \in \mathbb{R}$ holds

$$c_1 \sin(t) + c_2 t \sin(t) = 0.$$

Evaluating at $t = \pi/2$ and $t = 3\pi/2$ we obtain

$$c_1 + \frac{\pi}{2}\, c_2 = 0, \quad c_1 + \frac{3\pi}{2}\, c_2 = 0 \quad \Rightarrow \quad c_1 = 0, \quad c_2 = 0.$$

We conclude: The functions y_1 and y_2 are linearly independent. $\lhd$

The concepts of operator, linearity, superposition, linearly dependence, are needed to introduce our next result. If we know two linearly independent solutions of a second order linear *homogeneous* differential equation, then we know all possible solutions to that equation. *Any* other solution has to be a linear combination of the previous two solutions. It is crucial for this result that the equation be homogeneous. This is the closer we can get to a general formula for solutions to second order linear homogeneous differential equations.

Theorem 2.1.9 (General Solution)**.** *If y_1 and y_2 are linearly independent solutions of*

$$L(y) = 0 \tag{2.1.7}$$

on an interval $I \subset \mathbb{R}$, where $L(y) = y'' + a_1\, y' + a_0\, y$, and a_1, a_2 are continuous functions on I, then every solution y of Eq. (2.1.7) on the interval I can be written as a linear combination

$$y(t) = c_1\, y_1(t) + c_2\, y_2(t), \tag{2.1.8}$$

for appropriate values of the constants c_1, c_2.

Before we prove Theorem 2.1.9, it is convenient to state the following the definitions, which come naturally from this Theorem.

Definition 2.1.10.

(a) The functions y_1 and y_2 are ***fundamental solutions*** *of the equation $L(y) = 0$ iff these functions y_1, y_2 are linearly independent and satisfy the equations*

$$L(y_1) = 0, \qquad L(y_2) = 0.$$

(b) The ***general solution*** *of the equation $L(y) = 0$ is a family of functions given by*

$$y(t) = c_1\, y_1(t) + c_2\, y_2(t),$$

with c_1, c_2 arbitrary constants, and y_1, y_2 fundamental solutions of $L(y) = 0$.

Example 2.1.13. Show that $y_1 = e^t$ and $y_2 = e^{-2t}$ are fundamental solutions of

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

Solution: We first show that y_1 and y_2 are solutions to the differential equation, since

$$L(y_1) = y_1'' + y_1' - 2y_1 = e^t + e^t - 2e^t = (1 + 1 - 2)e^t = 0,$$

$$L(y_2) = y_2'' + y_2' - 2y_2 = 4\, e^{-2t} - 2\, e^{-2t} - 2e^{-2t} = (4 - 2 - 2)e^{-2t} = 0.$$

It is clear that y_1 and y_2 are linearly independent, since they are not proportional to each other. Anyway, we give a formal proof of this statement.

To show that y_1 and y_2 above are linearly independent we need show that the only constants c_1 and c_2 satisfying the equation $c_1\, y_1 + c_2\, y_2 = 0$ for all $t \in \mathbb{R}$ are the constants $c_1 = c_2 = 0$. To see that this is the case we write

$$c_1\, e^t + c_2\, e^{-2t} = 0$$

Since the equation above must hold for all $t \in \mathbb{R}$, its t-derivative must also hold,

$$c_1\, e^t - 2c_2\, e^{-2t} = 0.$$

Take $t = 0$ in both equations above,

$$0 = c_1 + c_2, \quad 0 = c_1 - 2c_2 \quad \Rightarrow \quad c_1 = c_2 = 0.$$

Since the only solution is $c_1 = c_2 = 0$, we conclude that y_1 and y_2 are fundamental solutions of the differential equation above. ◁

Remark: The fundamental solutions of an homogeneous equation are not unique. For example, it is not hard to show that another set of fundamental solutions for the equation in the example above are

$$y_1(t) = e^t + e^{-2t}, \qquad y_2(t) = e^t - e^{-2t}.$$

To prove Theorem 2.1.9 we need to introduce the Wronskian function and to verify some of its properties. In the following subsection we study the Wronskian function and we prove Abel's Theorem. We use these results in the proof of Theorem 2.1.9, and in the next subsection we prove them.

Proof of Theorem 2.1.9: We need to show that, given any fundamental solution pair, y_1, y_2, any other solution y to the homogeneous equation $L(y) = 0$ must be a unique linear combination of the fundamental solutions,

$$y(t) = c_1\, y_1(t) + c_2\, y_2(t), \tag{2.1.9}$$

for appropriately chosen constants c_1, c_2.

First, the superposition property implies that the function y above is solution of the homogeneous equation $L(y) = 0$ for every pair of constants c_1, c_2.

Second, given a function y, if there exist constants c_1, c_2 such that Eq. (2.1.9) holds, then these constants are unique. The reason is that functions y_1, y_2 are linearly independent. This can be seen from the following argument. If there are another constants $\tilde{c}_1$, $\tilde{c}_2$ so that

$$y(t) = \tilde{c}_1\, y_1(t) + \tilde{c}_2\, y_2(t),$$

then subtract the expression above from Eq. (2.1.9),

$$0 = (c_1 - \tilde{c}_1)\, y_1 + (c_2 - \tilde{c}_2)\, y_2 \quad \Rightarrow \quad c_1 - \tilde{c}_1 = 0, \quad c_2 - \tilde{c}_2 = 0,$$

where we used that y_1, y_2 are linearly independent. This second part of the proof can be obtained from the part three below, but it is a good idea to highlight it here.

So we only need to show that the expression in Eq. (2.1.9) contains all solutions. We need to show that we are not missing any other solution. In this third part of the argument enters Theorem 2.1.3. This Theorem says that, in the case of homogeneous equations, the initial value problem

$$L(y) = 0, \qquad y(t_0) = d_1, \quad y'(t_0) = d_2,$$

always has a unique solution. That means, a good parametrization of all solutions to the differential equation $L(y) = 0$ is given by the two constants, d_1, d_2 in the initial condition. To finish the proof of Theorem 2.1.9 we need to show that the constants c_1 and c_2 are also good to parametrize all solutions to the equation $L(y) = 0$. One way to show this, is to find an invertible map from the constants d_1, d_2, which we know parametrize all solutions, to the constants c_1, c_2. The map itself is simple to find, we just use the initial condition,

$$\begin{aligned} d_1 &= c_1\, y_1(t_0) + c_2\, y_2(t_0) \\ d_2 &= c_1\, y_1'(t_0) + c_2\, y_2'(t_0). \end{aligned}$$

We now need to show that this map is invertible. From linear algebra we know that this map acting on c_1, c_2 is invertible iff the determinant of the coefficient matrix is nonzero,

$$\begin{vmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{vmatrix} = y_1(t_0)\, y_2'(t_0) - y_1'(t_0) y_2(t_0) \neq 0.$$

This leads us to investigate the function

$$W_{12}(t) = y_1(t)\, y_2'(t) - y_1'(t) y_2(t)$$

This function is called the Wronskian of the two functions y_1, y_2. At the end of this section we prove Theorem 2.1.12, which says the following: If y_1, y_2 are fundamental solutions of $L(y) = 0$ on $I \subset \mathbb{R}$, then $W_{12}(t) \neq 0$ on I. Therefore, $W_{12}(t_0) \neq 0$, and then the map linking d_1, d_2 with c_1, c_2 is invertible, meaning the constants c_1, c_2 parametrize all solutions of the differential equation. This statement establishes the Theorem. □

2.1.6. The Wronskian Function.

We now introduce a function that provides information about the linear dependency of two functions y_1, y_2. This function is called the Wronskian to honor the polish scientist Josef Wronski, who first introduced it in 1821 while studying a different problem. In this subsection we prove the property of the Wronskian we used in the proof of Theorem 2.1.9. We start with the definition of the Wronskian and a couple of examples.

Definition 2.1.11. *The* ***Wronskian*** *of the differentiable functions y_1, y_2 is the function*

$$W_{12}(t) = y_1(t) y_2'(t) - y_1'(t) y_2(t).$$

Remark: If we introduce the matrix valued function

$$A(t) = \begin{bmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{bmatrix},$$

then the Wronskian can be written using the determinant of that 2×2 matrix,

$$W_{12}(t) = \det\big(A(t)\big) = y_1(t) y_2'(t) - y_1'(t) y_2(t).$$

An alternative notation is: $W_{12} = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}$.

Example 2.1.14. Find the Wronskian of the functions:
(a) $y_1(t) = \sin(t)$ and $y_2(t) = 2\sin(t)$. (ld)
(b) $y_1(t) = \sin(t)$ and $y_2(t) = t\sin(t)$. (li)

Solution:
Part (a): By the definition of the Wronskian:

$$W_{12}(t) = \begin{vmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{vmatrix} = \begin{vmatrix} \sin(t) & 2\sin(t) \\ \cos(t) & 2\cos(t) \end{vmatrix} = \sin(t) 2\cos(t) - \cos(t) 2\sin(t)$$

We conclude that $W_{12}(t) = 0$. Notice that y_1 and y_2 are linearly dependent.

Part (b): Again, by the definition of the Wronskian:

$$W_{12}(t) = \begin{vmatrix} \sin(t) & t\sin(t) \\ \cos(t) & \sin(t) + t\cos(t) \end{vmatrix} = \sin(t)\big[\sin(t) + t\cos(t)\big] - \cos(t) t\sin(t).$$

We conclude that $W_{12}(t) = \sin^2(t)$. Notice that y_1 and y_2 are linearly independent. ◁

In the proof of Theorem 2.1.9 we used the following property of the Wronskian.

Theorem 2.1.12 (Wronskian)**.** *If y_1, y_2 are fundamental solutions of $L(y) = 0$ on an open interval $I \subset \mathbb{R}$, then $W_{12}(t) \neq 0$ for every $t \in I$.*

The proof of this statement is at the end of this section, when prove Theorem 2.1.16. But before doing that we comment on the importance of the hypotheses in Theorem 2.1.12 and then we prove an auxiliary result, Abel's Theorem, before focusing on the proof of Theorem 2.1.12.

Remark: One of the hypotheses in the theorem above is that the functions y_1, y_2 must be solutions of an homogeneous second order linear differential equation, $L(y) = 0$. This hypothesis is important, without it the statement is not true. In other words, it is *not true* that "If y_1, y_2 are linearly independent on an open interval $I \subset \mathbb{R}$, then $W_{12}(t) \neq 0$ for all $t \in I$". In the following example we show two functions which are linearly independent and yet their Wronskian is zero.

Example 2.1.15. Show that the functions

$$y_1(t) = t^2, \quad \text{and} \quad y_2(t) = |t|\, t, \quad \text{for} \quad t \in \mathbb{R}$$

have Wronskian $W_{12} = 0$ and yet they are linearly independent.

Solution: First, we can see in Fig. 2 that these functions are linearly independent, since

$$y_1(t) = -y_2(t), \quad \text{for } t < 0, \quad \text{but} \quad y_1(t) = y_2(t), \quad \text{for } t > 0.$$

We see there is not c such that $y_1(t) = c\, y_2(t)$ for all $t \in \mathbb{R}$. Therefore, the functions y_1 and y_2 are linearly independent.

Second, these functions are differentiable in $\mathbb{R}$, so we can compute their Wronskian. For $t < 0$ we have

$$y_1(t) = -y_2(t) \quad \Rightarrow \quad W_{12} = y_1\, y_2' - y_1'\, y_2 = -y_2\, y_2' + y_2'\, y_2 = 0 \quad \text{for } t < 0.$$

For $t > 0$ we have

$$y_1(t) = y_2(t) \quad \Rightarrow \quad W_{12} = y_1\, y_2' - y_1'\, y_2 = y_2\, y_2' - y_2'\, y_2 = 0 \quad \text{for } t > 0.$$

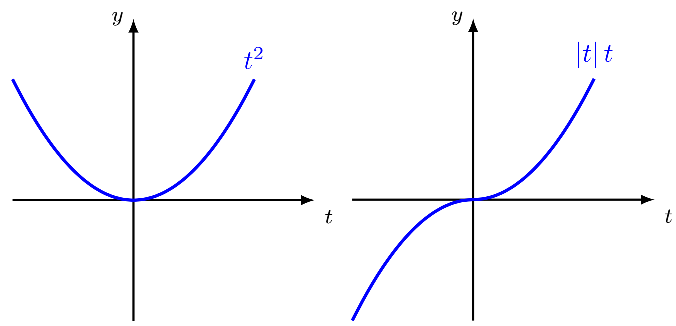


FIGURE 2. We graph the functions $y_1 = t^2$ and $y_2 = |t|\,t$.

Finally, we compute the Wronskian at $t = 0$, that is,

$$W_{12}(0) = y_1(0)\,y_2'(0) - y_1'(0)\,y_2(0).$$

It is clear that $y_1(0) = 0$, $y_2(0) = 0$, and $y_1'(0) = 0$. We only need to check that $y_2(t) = |t|\,t$ is differentiable at $t = 0$. We know that $y_2(t)$ is given by

$$y_2(t) = -t^2, \quad \text{for } \; t < 0, \quad \text{and} \quad y_2(t) = t^2, \quad \text{for } \; t > 0.$$

Then, the derivative of y_2 is well defined for $t < 0$ and for $t > 0$,

$$y_2'(t) = -2t, \quad \text{for } \; t < 0, \quad \text{and} \quad y_2'(t) = 2t, \quad \text{for } \; t > 0,$$

Therefore,

$$\lim_{t\to 0^-} y_2'(t) = 0 = \lim_{t\to 0^+} y_2'(t).$$

We conclude that $y_2'(0)$ exists and $y_2'(0) = 0$. Therefore, $W_{12}(t)$ vanishes for all $t \in \mathbb{R}$. $\triangleleft$

In the example above we showed that when

$$y_1 = y_2 \quad \text{or} \quad y_1 = -y_2 \quad \Rightarrow \quad W_{12} = 0.$$

This result is the particular case of a more general result. If two functions satisfy that $y_1 = c\,y_2$, for any constant c, then their Wronskian is zero.

Theorem 2.1.13 (Wronskian LD). *If y_1, y_2 are linearly dependent on $I \subset \mathbb{R}$, then*

$$W_{12} = 0 \quad \text{on} \quad I.$$

Proof of Theorem 2.1.13: Since the functions y_1, y_2 are linearly dependent, there exists a nonzero constant c such that $y_1 = c\,y_2$; hence holds,

$$W_{12} = y_1\,y_2' - y_1'\,y_2 = (c\,y_2)\,y_2' - (c\,y_2)'\,y_2 = 0.$$

This establishes the Theorem. □

Remark: It is often cited in the literature the contrapositive of Theorem 2.1.13. Recall that given an implication $A \Rightarrow B$, the contrapositive is No $B \Rightarrow$ No A. The contrapositive of a statement is equivalent to the original statement. We state the contrapositive Theorem 2.1.13 in the following Corollary.

Corollary 2.1.14 (Wronskian LD). *If functions y_1, y_2 defined on an interval $I \subset \mathbb{R}$ have Wronskian $W_{12}(t_0) \neq 0$ at a point $t_0 \in I$, then the functions y_1, y_2 are linearly independent on I.*

Let's go back to our main subject in this subsection, Theorem 2.1.12. We have seen in Example 2.1.15 that the linear independence of functions y_1, y_2 is not enough to show that their Wronskian is nonzero. We need to assume something else on functions y_1, y_2. In Theorem 2.1.12 we assume that these functions are solutions a differential equation.

We need one last result before proving Theorem 2.1.12. We now show that the Wronskian of two solutions of an homogeneous second order linear differential equation satisfies a differential equation of its own. The equation of the Wronskian is a first order linear equation, which can be solved. This result is known as Abel's Theorem.

Theorem 2.1.15 (Abel)**.** *If y_1, y_2 are twice continuously differentiable solutions of*

$$y'' + a_1(t)\, y' + a_0(t)\, y = 0, \tag{2.1.10}$$

where a_1, a_0 are continuous on $I \subset \mathbb{R}$, then the Wronskian W_{12} satisfies

$$W_{12}' + a_1(t)\, W_{12} = 0. \tag{2.1.11}$$

Therefore, for any $t_0 \in I$, the Wronskian W_{12} is given by the expression

$$W_{12}(t) = W_{12}(t_0)\, e^{-A_1(t)}, \tag{2.1.12}$$

where $A_1(t) = \displaystyle\int_{t_0}^{t} a_1(s)\, ds$.

Proof of Theorem 2.1.15: Compute the derivative of the Wronskian function,

$$W_{12}' = \big(y_1\, y_2' - y_1'\, y_2\big)' = y_1\, y_2'' - y_1''\, y_2.$$

Recall that both y_1 and y_2 are solutions to Eq. (2.1.10), meaning,

$$y_1'' = -a_1\, y_1' - a_0\, y_1, \qquad y_2'' = -a_1\, y_2' - a_0\, y_2.$$

Replace these expressions in the formula for W_{12}' above,

$$W_{12}' = y_1\,\big(-a_1\, y_2' - a_0\, y_2\big) - \big(-a_1\, y_1' - a_0\, y_1\big)\, y_2 \quad \Rightarrow \quad W_{12}' = -a_1\big(y_1\, y_2' - y_1'\, y_2\big)$$

So we obtain the equation

$$W_{12}' + a_1(t)\ W_{12} = 0.$$

This equation for W_{12} is a first order linear equation. The solution can be found using the method of integrating factors, given in Section 1.2, which gives Eq. 2.1.12. This establishes the Theorem. □

Before proving Theorem 2.1.12 we show one simple application of Abel's Theorem.

Example 2.1.16. Find the Wronskian of two solutions of the equation

$$t^2\, y'' - t(t+2)\, y' + (t+2)\, y = 0, \qquad t > 0.$$

Solution: Notice that we do not known the explicit expression for the solutions. Nevertheless, Theorem 2.1.15 says that we can compute their Wronskian. First, we have to rewrite the differential equation in the form given in that Theorem, namely,

$$y'' - \Big(\frac{2}{t} + 1\Big)\, y' + \Big(\frac{2}{t^2} + \frac{1}{t}\Big)\, y = 0.$$

Then, Theorem 2.1.15 says that the Wronskian satisfies the differential equation

$$W_{12}'(t) - \Big(\frac{2}{t} + 1\Big)\, W_{12}(t) = 0.$$

This is a first order, linear equation for W_{12}, so its solution can be computed using the method of integrating factors. That is, first compute the integral

$$-\int_{t_0}^{t}\Big(\frac{2}{s}+1\Big)\,ds = -2\ln\Big(\frac{t}{t_0}\Big) - (t-t_0)$$
$$= \ln\Big(\frac{t_0^2}{t^2}\Big) - (t-t_0).$$

Then, the integrating factor μ is given by

$$\mu(t) = \frac{t_0^2}{t^2}\,e^{-(t-t_0)},$$

which satisfies the condition $\mu(t_0)=1$. So the solution, W_{12} is given by

$$\Big(\mu(t)W_{12}(t)\Big)' = 0 \quad\Rightarrow\quad \mu(t)W_{12}(t) - \mu(t_0)W_{12}(t_0) = 0$$

so, the solution is

$$W_{12}(t) = W_{12}(t_0)\,\frac{t^2}{t_0^2}\,e^{(t-t_0)}.$$

If we call the constant $c = W_{12}(t_0)/[t_0^2 e^{t_0}]$, then the Wronskian has the simpler form

$$W_{12}(t) = c\,t^2 e^t.$$

$\lhd$

Finally, we are ready to prove Theorem 2.1.12. However, instead of proving it, we prove an equivalent statement—the contrapositive of Theorem 2.1.12.

Theorem 2.1.16 (Wronskian CP)**.** *If y_1, y_2 are solutions of $L(y)=0$ on $I\subset\mathbb{R}$ and there is a point $t_1\in I$ such that $W_{12}(t_1)=0$, then y_1, y_2 are linearly dependent on I.*

Proof of Theorem 2.1.16: We know that y_1, y_2 are solutions of $L(y)=0$. Then, Abel's Theorem says that their Wronskian W_{12} is given by

$$W_{12}(t) = W_{12}(t_0)\,e^{-A_1(t)},$$

for any $t_0\in I$. Chossing the point t_0 to be t_1, the point where by hypothesis $W_{12}(t_1)=0$, we get that

$$W_{12}(t) = 0 \quad \text{for all} \quad t\in I.$$

Knowing that the Wronskian vanishes identically on I, we can write

$$y_1\,y_2' - y_1'\,y_2 = 0,$$

on I. If either y_1 or y_2 is the function zero, then the set is linearly dependent. So we can assume that both are not identically zero. Let's assume there exists $\tau_1\in I$ such that $y_1(\tau_1)\neq 0$. By continuity, y_1 is nonzero in an open neighborhood $I_1\subset I$ of τ_1. So in that neighborhood we can divide the equation above by y_1^2,

$$\frac{y_1\,y_2' - y_1'\,y_2}{y_1^2} = 0 \quad\Rightarrow\quad \Big(\frac{y_2}{y_1}\Big)' = 0 \quad\Rightarrow\quad \frac{y_2}{y_1} = c, \qquad \text{on } I_1,$$

where $c\in\mathbb{R}$ is an arbitrary constant. So we conclude that y_1 is proportional to y_2 on the open set I_1. That means that the function $y(t) = y_2(t) - c\,y_1(t)$, satisfies

$$L(y)=0, \qquad y(\tau_1)=0, \quad y'(\tau_1)=0.$$

Therefore, the existence and uniqueness Theorem 2.1.3 says that $y(t)=0$ for all $t\in I$. This finally shows that y_1 and y_2 are linearly dependent. This establishes the Theorem. $\square$

By proving the contrapositive of Theorem 2.1.12 we have proven Theorem 2.1.12. Then, we have finished the proof of the General Solution Theorem 2.1.9.

2.1.7. Exercises.

2.1.1.- Find the longest interval where the solution y of the initial value problems below is defined. (Do not try to solve the differential equations.)

(a) $t^2 y'' + 6y = 2t$, $y(1) = 2$, $y'(1) = 3$.
(b) $(t - 6)y'' + 3ty' - y = 1$, $y(3) = -1$, $y'(3) = 2$.

2.1.2.- (a) Verify that $y_1(t) = t^2$ and $y_2(t) = 1/t$ are solutions to the differential equation

$$t^2 y'' - 2y = 0, \quad t > 0.$$

(b) Show that $y(t) = a\,t^2 + \dfrac{b}{t}$ is solution of the same equation for all constants $a, b \in \mathbb{R}$.

2.1.3.- If the graph of y, solution to a second order linear differential equation $L(y(t)) = 0$ on the interval $[a, b]$, is tangent to the t-axis at any point $t_0 \in [a, b]$, then find the solution y explicitly.

2.1.4.- Can the function $y(t) = \sin(t^2)$ be solution on an open interval containing $t = 0$ of a differential equation

$$y'' + a(t)\,y' + b(t)y = 0,$$

with continuous coefficients a and b? Explain your answer.

2.1.5.- Verify whether the functions y_1, y_2 below are a fundamental set for the differential equations given below:

(a) $y_1(t) = \cos(2t)$, $y_2(t) = \sin(2t)$,

$$y'' + 4y = 0.$$

(b) $y_1(t) = e^t$, $y_2(t) = t\,e^t$,

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

(c) $y_1(x) = x$, $y_2(t) = x\,e^x$,

$$x^2\,y'' - 2x(x + 2)\,y' + (x + 2)\,y = 0.$$

2.1.6.- Compute the Wronskian of the following functions:

(a) $f(t) = \sin(t)$, $g(t) = \cos(t)$.
(b) $f(x) = x$, $g(x) = x\,e^x$.
(c) $f(\theta) = \cos^2(\theta)$, $g(\theta) = 1 + \cos(2\theta)$.

2.1.7.- If the Wronskian of any two solutions of the differential equation

$$y'' + p(t)\,y' + q(t)\,y = 0$$

is constant, what does this imply about the coefficients p and q?

2.1.8.- Let $y(t) = c_1\,t + c_2\,t^2$ be the general solution of a second order linear differential equation $L(y) = 0$. By eliminating the constants c_1 and c_2 in terms of y and y', find a second order differential equation satisfied by y.

2.2. Homogenous Constant Coefficients Equations

The main result in § 2.1 is Theorem 2.1.9, which says that the closest we can get to a formula for the solutions of an homogeneous second order linear differential equation is Eq. (2.1.8). This general solution formula says that all solutions of the differential equation are linear combinations of two solutions not proportional to each other—fundamental solutions.

In this section we obtain the fundamental solutions in the particular case that the homogeneous second order linear equation has *constant coefficients*. Such problem reduces to solve for the roots of a degree-two polynomial, called the characteristic polynomial.

2.2.1. The Roots of the Characteristic Polynomial. The main result in Theorem 2.1.9 is that all the solutions of an homogeneous second order linear differential equation are linear combinations of two fundamental solutions. In this section we find fundamental solutions in the case that the equation has constant coefficients. Since the equation is so simple, we find such solutions by trial and error. Here is an example of how this works.

Example 2.2.1. Find fundamental solutions to the equation

$$y'' + 5y' + 6y = 0. \tag{2.2.1}$$

Solution: We guess solutions of the equation from a set of simple candidates, such as

$$y(t) = c, \qquad y(t) = t^n, \qquad y(t) = e^{rt}, \qquad \text{etc},$$

where c, n, and r are constants. It is simple to see that the only constant solution of the equation is $c = 0$, since

$$c'' + 5\,c' + 6\,c = 0 \quad \Rightarrow \quad c = 0.$$

Next we try with power functions $y(t) = t^n$. If $y(t) = t^n$ is a solution, then

$$n(n-1)\,t^{(n-2)} + 5n\,t^{(n-1)} + 6\,t^n = 0 \quad \Rightarrow \quad t^{(n-2)}\left(n(n-1) + 5n\,t + 6\,t^2\right) = 0$$

so we arrive at the equation

$$n(n-1) + 5n\,t + 6\,t^2 = 0, \quad \text{for all } t \in \mathbb{R}.$$

But the equation above is not true for any choice of n, therefore the functions $y(t) = t^n$ cannot be solutions of the differential equation. From this failed attempt we see that it would be promising to try with a test function having a derivative proportional to the original function,

$$y'(t) = r\,y(t).$$

Such function would be simplified from the equation. For that reason we now try with $y(t) = e^{rt}$. If we introduce this function in the differential equation we get

$$(e^{rt})'' + 5\,(e^{rt})' + 6\,e^{rt} = 0 \quad \Rightarrow \quad (r^2 + 5r + 6)\,e^{rt} = 0 \quad \Rightarrow \quad r^2 + 5r + 6 = 0. \tag{2.2.2}$$

We have eliminated the exponential and any t-dependence from the differential equation, and now the equation is a condition on the constant r. So we look for the appropriate values of r, which are the roots of a polynomial degree two,

$$r_\pm = \frac{1}{2}\left(-5 \pm \sqrt{25 - 24}\right) = \frac{1}{2}\,(-5 \pm 1) \quad \Rightarrow \quad \begin{cases} r_+ = -2, \\ r_- = -3. \end{cases}$$

We have obtained two different roots, which implies we have two different solutions,

$$y_1(t) = e^{-2t}, \qquad y_2(t) = e^{-3t}.$$

These solutions are not proportional to each other, so the are fundamental solutions to the differential equation in (2.2.1). Then, Theorem 2.1.9 in § 2.1 implies that we have found all possible solutions to the differential equation, which are given by

$$y(t) = c_1 \, e^{-2t} + c_2 \, e^{-3t}, \qquad c_1, c_2 \in \mathbb{R}. \tag{2.2.3}$$

◁

The exponential functions $y(t) = e^{rt}$ we tried in the example above will provide solutions to constant coefficient equations of the form

$$y'' + a_1 \, y' + a_0 \, y = 0,$$

for almost any choice of the constants a_1, a_0. Indeed,

$$(e^{rt})'' + a_1 \, (e^{rt})' + a_0 \, e^{rt} = 0 \quad \Rightarrow \quad \big(r^2 + a_1 \, r + a_0\big) \, e^{rt} = 0 \quad \Rightarrow \quad r^2 + a_1 \, r + a_0 = 0.$$

The polynomial on the equation on the far right is important and we will give it a name.

Definition 2.2.1. *The **characteristic polynomial** and **characteristic equation** of the second order linear homogeneous equation with constant coefficients*

$$y'' + a_1 y' + a_0 = 0,$$

are given by

$$p(r) = r^2 + a_1 r + a_0, \qquad p(r) = 0.$$

As we saw in Example 2.2.1, the roots of the characteristic polynomial are crucial to express the solutions of the differential equation above. The characteristic polynomial is a second degree polynomial with real coefficients, and the general expression for its roots is

$$r_\pm = \frac{1}{2}\Big(-a_1 \pm \sqrt{a_1^2 - 4a_0}\Big).$$

If the discriminant $(a_1^2 - 4a_0)$ is positive, zero, or negative, then the roots of p are different real numbers, only one real number, or a complex-conjugate pair of complex numbers. We summarize our results in the following statement.

Theorem 2.2.2 (Constant Coefficients)**.** *If $r_\pm$ are the roots of the characteristic polynomial to the second order linear homogeneous equation with constant coefficients*

$$y'' + a_1 y' + a_0 y = 0, \tag{2.2.4}$$

and if c_+, c_- are arbitrary constants, then we have the following results:

(a) If $r_+ \neq r_-$, real or complex, then the general solution of Eq. (2.2.4) is given by

$$y(t) = c_+ \, e^{r_+ t} + c_- \, e^{r_- t}.$$

(b) If $r_+ = r_- = r_0 \in \mathbb{R}$, then the general solution of Eq. (2.2.4) is given by

$$y(t) = c_+ \, e^{r_0 t} + c_- \, t \, e^{r_0 t}.$$

Furthermore, given real constants t_0, y_0 and y_1, there is a unique solution to the initial value problem given by Eq. (2.2.4) with the initial conditions $y(t_0) = y_0$ and $y'(t_0) = y_1$.

Remarks:

(a) The proof is to guess that functions $y(t) = e^{rt}$ must be solutions for appropriate values of the exponent constant r, the latter being roots of the characteristic polynomial. When the characteristic polynomial has two different roots, Theorem 2.1.9 implies we have all solutions. When the root is repeated we use the reduction of order method to find a second solution not proportional to the first one.

(b) At the end of the section we show a proof where we construct the fundamental solutions y_1, y_2 without guessing them. We do not need to use Theorem 2.1.9 in this second proof, which is based completely in a generalization of the reduction of order method.

Proof of Theorem 2.2.2: We guess that particular solutions to Eq. 2.2.4 must be exponential functions of the form $y(t) = e^{rt}$, because the exponential will cancel out from the equation and only a condition for r will remain. This is what happens,

$$(e^{rt})'' + a_1\,(e^{rt})' + a_0\,e^{rt} = 0 \quad \Rightarrow \quad \big(r^2 + a_1\,r + a_0\big)\,e^{rt} = 0 \quad \Rightarrow \quad r^2 + a_1\,r + a_0 = 0.$$

The last equation says that the appropriate values of the exponent are the root of the characteristic polynomial. We now have two cases. If $r_+ \neq r_-$ then the solutions

$$y_+(t) = e^{r_+t}, \quad y_-(t) = e^{r_-t},$$

are linearly independent, so the general solution to the differential equation is

$$y(t) = c_+\,e^{r_+t} + c_-\,e^{r_-t}.$$

If $r_+ = r_- = r_0$, then we have found only one solution $y_+(t) = e^{r_0t}$, and we need to find a second solution not proportional to y_+. This is what the reduction of order method is designed to do. We write the second solution as

$$y_-(t) = v(t)\,y_+(t) \quad \Rightarrow \quad y_-(t) = v(t)\,e^{r_0t},$$

and we put this expression in the differential equation (2.2.4),

$$\big(v'' + 2r_0v' + vr_0^2\big)\,e^{r_0t} + \big(v' + r_0v\big)\,a_1\,e^{r_0t} + a_0v\,e^{r_0t} = 0.$$

We cancel the exponential out of the equation and we reorder terms,

$$v'' + (2r_0 + a_1)\,v' + (r_0^2 + a_1r_0 + a_0)\,v = 0.$$

We now need to use that r_0 is a root of the characteristic polynomial, $r_0^2 + a_1r_0 + a_0 = 0$, so the last term in the equation above vanishes. But we also need to use that the root r_0 is repeated,

$$r_0 = -\frac{a_1}{2} \pm \frac{1}{2}\sqrt{a_1^2 - 4a_0} = -\frac{a_1}{2} \quad \Rightarrow \quad 2r_0 + a_1 = 0.$$

The equation on the right side above implies that the second term in the differential equation for v vanishes. So we get that

$$v'' = 0 \quad \Rightarrow \quad v(t) = c_1 + c_2t$$

and the second solution is

$$y_-(t) = (c_1 + c_2t)\,y_+(t).$$

If we choose the constant $c_2 = 0$, the function y_- is proportional to y_+. So we definitely want $c_2 \neq 0$. The other constant, c_1, only adds a term proportional to y_+, therefore we can choose it zero. So the simplest choice is $c_1 = 0$, $c_2 = 1$, and we get the fundamental solutions

$$y_+(t) = e^{r_0t}, \qquad y_-(t) = t\,e^{r_0t}.$$

So the general solution for the repeated root case is

$$y(t) = c_+\,e^{r_0t} + c_-t\,e^{r_0t}.$$

The furthermore part follows from solving a 2×2 linear system for the unknowns c_+ and c_-. The initial conditions for the case $r_+ \neq r_-$ are the following,

$$y_0 = y(t_0) = c_+\,e^{r_+t_0} + c_-\,e^{r_-t_0}, \qquad y_1 = y'(t_0) = r_+c_+\,e^{r_+t_0} + r_-c_-\,e^{r_-t_0}.$$

It is not difficult to verify that this system is always solvable and the solutions are

$$c_+ = -\frac{(r_-y_0 - y_1)}{(r_+ - r_-)}\,e^{-r_+t_0}, \qquad c_- = \frac{(r_+y_0 - y_1)}{(r_+ - r_-)}\,e^{-r_-t_0}.$$

The initial conditions for the case $r_- = r_- = r_0$ are the following,

$$y_0 = y(t_0) = (c_+ + c_- t_0)\, e^{r_0 t_0}, \qquad y_1 = y'(t_0) = c_-\, e^{r_0 t_0} + r_0 (c_+ + c_- t_0)\, e^{r_0 t_0}.$$

It is also not difficult to verify that this system is always solvable and the solutions are

$$c_+ = \big(y_0 - (y_1 - r_0 y_0)\, t_0\big)\, e^{-r_0 t_0}, \qquad c_- = (y_1 - r_0 y_0)\, e^{-r_0 t_0}.$$

This establishes the Theorem. □

Example 2.2.2. Consider an object of mass $m = 1$ grams hanging from a spring with spring constant $k = 6$ grams per second square moving in a fluid with damping constant $d = 5$ grams per second. Introduce a coordinate system, y, which is positive downwards and $y = 0$ is at the spring equilibrium position. Find the movement of this object if the initial position is $y(0) = 1$ centimeter and the initial velocity is $y'(0) = -1$ centimeter per second.

Solution: The movement of the object attached to the spring in that liquid is the solution of the initial value problem

$$y'' + 5y' + 6y = 0, \qquad y(0) = 1, \qquad y'(0) = -1.$$

Notice the initial velocity is negative, which means the initial velocity is in the upward direction. We know from Example 2.2.1 that the general solution of the differential equation above is

$$y(t) = c_+\, e^{-2t} + c_-\, e^{-3t}.$$

We now find the constants c_+ and c_- that satisfy the initial conditions above,

$$\left.\begin{aligned} 1 &= y(0) = c_+ + c_- \\ -1 &= y'(0) = -2c_+ - 3c_- \end{aligned}\right\} \quad \Rightarrow \quad \begin{cases} c_+ = 2, \\ c_- = -1. \end{cases}$$

Therefore, the unique solution to the initial value problem is

$$y(t) = 2e^{-2t} - e^{-3t}.$$

The solution is a combination of two decaying exponentials in such a way that the solution approaches the resting position in the limit $t \to \infty$ from the initial position without making any oscillation. This means that the fluid viscosity is really high and it dampens any oscillation in the spring. ◁

Example 2.2.3. Find the general solution, $y(t)$, of the differential equation

$$2y'' - 3y' + y = 0.$$

Solution: We look for every solutions of the form $y(t) = e^{rt}$, where r is solution of the characteristic equation

$$2r^2 - 3r + 1 = 0 \quad \Rightarrow \quad r = \frac{1}{4}\big(3 \pm \sqrt{9-8}\big) \quad \Rightarrow \quad \begin{cases} r_+ = 1, \\ r_- = \dfrac{1}{2}. \end{cases}$$

Therefore, the general solution of the equation above is

$$y(t) = c_+ e^{t} + c_- e^{t/2}.$$

◁

Example 2.2.4. Consider an object of mass $m = 9$ grams hanging from a spring with spring constant $k = 1$ grams per second square moving in a fluid with damping constant $d = 6$ grams per second. Introduce a coordinate system, y, which is positive downwards and $y = 0$ is at the spring equilibrium position. Find the movement of this object if the initial position is $y(0) = 1$ centimeter and the initial velocity is $y'(0) = 5/3$ centimeters per second.

Solution: The movement of the object is described by the solution to the initial value problem

$$9y'' + 6y' + y = 0, \qquad y(0) = 1, \qquad y'(0) = \frac{5}{3}.$$

The characteristic polynomial is $p(r) = 9r^2 + 6r + 1$, with roots given by

$$r_\pm = \frac{1}{18}\big(-6 \pm \sqrt{36 - 36}\big) \quad \Rightarrow \quad r_+ = r_- = -\frac{1}{3}.$$

Theorem 2.2.2 says that the general solution has the form

$$y(t) = c_+\, e^{-t/3} + c_- t\, e^{-t/3}.$$

We need to compute the derivative of the expression above to impose the initial conditions,

$$y'(t) = -\frac{c_+}{3}\, e^{-t/3} + c_-\Big(1 - \frac{t}{3}\Big)\, e^{-t/3},$$

then, the initial conditions imply that

$$\left.\begin{aligned} 1 &= y(0) = c_+, \\ \frac{5}{3} &= y'(0) = -\frac{c_+}{3} + c_- \end{aligned}\right\} \quad \Rightarrow \quad c_+ = 1, \qquad c_- = 2.$$

So, the solution to the initial value problem above is

$$y(t) = (1 + 2t)\, e^{-t/3}.$$

◁

Example 2.2.5. Consider an object of mass $m = 1$ grams hanging from a spring with spring constant $k = 13$ grams per second square moving in a fluid with damping constant $d = 4$ grams per second. Introduce a coordinate system, y, which is positive downwards and $y = 0$ is at the spring equilibrium position. Find the position function of this object for arbitrary initial position and velocity.

Solution: The position function y of the mass-spring system must be solution of Newton's equation of motion

$$y'' + 4y' + 13y = 0.$$

To find the solutions we first need to find the roots of the characteristic polynomial,

$$r^2 + 4r + 13 = 0 \quad \Rightarrow \quad r_\pm = \frac{1}{2}\big(-4 \pm \sqrt{16 - 52}\big) \quad \Rightarrow \quad r_\pm = \frac{1}{2}\big(-4 \pm \sqrt{36}\big),$$

so we obain the roots

$$r_\pm = -2 \pm 3i.$$

Since the roots of the characteristic polynomial are different, Theorem 2.2.2 says that the general solution of the differential equation above, which includes complex-valued solutions, can be written as follows,

$$y(t) = \tilde{c}_+\, e^{(-2+3i)t} + \tilde{c}_-\, e^{(-2-3i)t}, \qquad \tilde{c}_+, \tilde{c}_- \in \mathbb{C}.$$

This general solution describes all possible motions of the mass-spring system above. An equivalent description is possible in terms of only real-valued functions. In the next subsection we see how this latter description can be done. ◁

2.2.2. Real Solutions for Complex Roots. We study in more detail the solutions of the differential equation (2.2.4),

$$y'' + a_1 y' + a_0 y = 0,$$

in the case the characteristic polynomial has complex roots. Since these roots are given by

$$r_\pm = -\frac{a_1}{2} \pm \frac{1}{2}\sqrt{a_1^2 - 4a_0},$$

the roots are complex-valued in the case $a_1^2 - 4a_0 < 0$. We use the notation

$$r_\pm = \alpha \pm i\beta, \quad \text{with} \quad \alpha = -\frac{a_1}{2}, \quad \beta = \sqrt{a_0 - \frac{a_1^2}{4}}.$$

The fundamental solutions in Theorem 2.2.2 are the complex-valued functions

$$\tilde{y}_+ = e^{(\alpha+i\beta)t}, \qquad \tilde{y}_- = e^{(\alpha-i\beta)t}.$$

The general solution constructed from these solutions is

$$y(t) = \tilde{c}_+ \, e^{(\alpha+i\beta)t} + \tilde{c}_- \, e^{(\alpha-i\beta)t},$$

where the constants $\tilde{c}_1$, $\tilde{c}_2$ are complex-valued.

Usually, we are interested in real-valued solutions, for example when the differential equation describes a mass-spring system. The problem of having complex-valued fundamental solutions is that even real-valued solutions are expressed in terms of complex-valued quantities. Although it is not hard to find conditions on the complex constants $\tilde{c}_+$ and $\tilde{c}_-$ so that the function $y(t)$ above is real valued, the expression for the general solution is still complex-valued.

It is more convenient to write the general solution $y(t)$ in terms of real-valued fundamental solutions, say $y_1(t)$ and $y_2(t)$. In this case the general solution is

$$y(t) = c_1 \, y_1(t) + c_2 \, y_2(t),$$

and then real-valued solutions are given for c_1 and c_2 real, while complex solutions are given for c_1 and c_2 complex. For this reason we now provide a new set of fundamental solutions which is real-valued.

Theorem 2.2.3 (Real Valued Fundamental Solutions)**.** *If the differential equation*

$$y'' + a_1 \, y' + a_0 \, y = 0, \tag{2.2.5}$$

where a_1, a_0 are real constants, has characteristic polynomial with complex roots $r_\pm = \alpha \pm i\beta$ and complex valued fundamental solutions

$$\tilde{y}_+(t) = e^{(\alpha+i\beta)t}, \qquad \tilde{y}_-(t) = e^{(\alpha-i\beta)t},$$

then the equation also has real valued fundamental solutions given by

$$y_+(t) = e^{\alpha t}\cos(\beta t), \qquad y_-(t) = e^{\alpha t}\sin(\beta t).$$

Furthermore, the general solution of the Eq. (2.2.10) can be written either as

$$y(t) = \big(c_1 \cos(\beta t) + c_2 \sin(\beta t)\big)\, e^{\alpha t},$$

where c_1, c_2 are arbitrary constants, or as

$$y(t) = A\, e^{\alpha t} \cos(\beta t - \phi)$$

where $A > 0$ is the amplitude and $\phi \in [-\pi, \pi)$ is the phase shift of the solution.

Proof of Theorem 2.2.3: We start with the complex valued fundamental solutions

$$\tilde{y}_{+}(t) = e^{(\alpha+i\beta)t}, \qquad \tilde{y}_{-}(t) = e^{(\alpha-i\beta)t}.$$

We take the function $\tilde{y}_{+}$ and we use a property of complex exponentials,

$$\tilde{y}_{+}(t) = e^{(\alpha+i\beta)t} = e^{\alpha t}\, e^{i\beta t} = e^{\alpha t}\,\big(\cos(\beta t) + i\sin(\beta t)\big),$$

where on the last step we used Euler's formula $e^{i\theta} = \cos(\theta) + i\sin(\theta)$. Repeat this calculation for y_{-} we get,

$$\tilde{y}_{+}(t) = e^{\alpha t}\,\big(\cos(\beta t) + i\sin(\beta t)\big), \quad \tilde{y}_{-}(t) = e^{\alpha t}\,\big(\cos(\beta t) - i\sin(\beta t)\big).$$

If we recall the superposition property of linear homogeneous equations, Theorem 2.1.7, we know that any linear combination of the two solutions above is also a solution of the differential equation (2.2.10), in particular the combinations

$$y_{+}(t) = \frac{1}{2}\big(\tilde{y}_{+}(t) + \tilde{y}_{-}(t)\big), \quad y_{-}(t) = \frac{1}{2i}\big(\tilde{y}_{+}(t) - \tilde{y}_{-}(t)\big).$$

A straightforward computation gives

$$y_{+}(t) = e^{\alpha t}\cos(\beta t), \qquad y_{-}(t) = e^{\alpha t}\sin(\beta t).$$

Therefore, the general solution is

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

There is an equivalent way to express the general solution above given by

$$y(t) = A\, e^{\alpha t}\cos(\beta t - \phi).$$

These two expressions for the general solution $y(t)$ are equivalent because of the trigonometric identity

$$A\cos(\beta t - \phi) = A\cos(\beta t)\cos(\phi) + A\sin(\beta t)\sin(\phi),$$

which holds for all A and ϕ, and βt. Then, it is not difficult to see that

$$\left.\begin{aligned} c_1 &= A\cos(\phi)\\ c_2 &= A\sin(\phi)\end{aligned}\right\} \quad\Leftrightarrow\quad \left\{\begin{aligned} A &= \sqrt{c_1^2 + c_2^2}\\ \tan(\phi) &= \frac{c_2}{c_1}.\end{aligned}\right.$$

This establishes the Theorem. □

Example 2.2.6. Describe the movement of the object in Example 2.2.5 above, which satisfies Newton's equation

$$y'' + 4y' + 13y = 0,$$

with initial position of 2 centimeters and initial velocity of 2 centimeters per second.

Solution: We already found the roots of the characteristic polynomial,

$$r^2 + 4r + 13 = 0 \quad\Rightarrow\quad r_{\pm} = \frac{1}{2}\big(-4 \pm \sqrt{16 - 52}\big) \quad\Rightarrow\quad r_{\pm} = -2 \pm 3i.$$

So the complex-valued fundamental solutions are

$$\tilde{y}_{+}(t) = e^{(-2+3i)\,t}, \quad \tilde{y}_{-}(t) = e^{(-2-3i)\,t}.$$

Theorem 2.2.3 says that real-valued fundamental solutions are given by

$$y_{+}(t) = e^{-2t}\cos(3t), \quad y_{-}(t) = e^{-2t}\sin(3t).$$

So the real-valued general solution can be written as

$$y(t) = \big(c_{+}\cos(3t) + c_{-}\sin(3t)\big)\, e^{-2t}, \quad c_{+},\ c_{-} \in \mathbb{R}.$$

Soon we will need its derivative, which is

$$y'(t) = -2\big(c_+ \cos(3t) + c_- \sin(3t)\big)\, e^{-2t} + \big(-3\, c_+ \sin(3t) + 3\, c_- \cos(3t)\big)\, e^{-2t}.$$

We now use the initial conditions, $y(0) = 2$, and $y'(0) = 2$,

$$\left.\begin{aligned} 2 &= y(0) = c_+ \\ 2 &= y'(0) = -2c_+ + 3c_- \end{aligned}\right\} \quad \Rightarrow \quad c_+ = 2, \quad c_- = 2,$$

therefore the solution is

$$y(t) = \big(2\cos(3t) + 2\sin(3t)\big)\, e^{-2t}. \tag{2.2.6}$$

◁

Example 2.2.7. Write the solution of the Example 2.2.6 above in terms of the amplitude A and phase shift ϕ.

Solution: We rewrite the solution in Eq. (2.2.6) in terms of amplitude and phase shift

$$y(t) = A\, e^{-2t} \cos(3t - \phi).$$

We will need the derivative of the expression above,

$$y'(t) = -2A\, e^{-2t} \cos(3t - \phi) - 3A\, e^{-2t} \sin(3t - \phi).$$

Let us use again the initial conditions $y(0) = 2$, and $y'(0) = 2$,

$$\left.\begin{aligned} 2 &= y(0) = A\cos(-\phi) \\ 2 &= y'(0) = -2A\cos(-\phi) - 3A\sin(-\phi) \end{aligned}\right\} \quad \Rightarrow \quad \left\{\begin{aligned} A\cos(\phi) &= 2 \\ -2A\cos(\phi) + 3A\sin(\phi) &= 2. \end{aligned}\right.$$

Using the first equation in the second one we get

$$\left.\begin{aligned} A\cos(\phi) &= 2 \\ -4 + 3A\sin(\phi) &= 2 \end{aligned}\right\} \quad \Rightarrow \quad \left\{\begin{aligned} A\cos(\phi) &= 2 \\ A\sin(\phi) &= 2. \end{aligned}\right.$$

From here it is not too difficult to see that

$$A = \sqrt{2^2 + 2^2} = 2\sqrt{(2)}, \qquad \tan(\phi) = 1.$$

Since $\phi \in [-\pi, \pi)$, the equation $\tan(\phi) = 1$ has two solutions in that interval,

$$\phi_1 = \frac{\pi}{4}, \qquad \phi_2 = \frac{\pi}{4} - \pi = -\frac{3\pi}{4}.$$

But the ϕ we need satisfies that $\cos(\phi) > 0$ and $\sin(\phi) > 0$, which means $\phi = \dfrac{\pi}{4}$. Then,

$$y(t) = 2\sqrt{2}\, e^{-2t} \cos\Big(3t - \frac{\pi}{4}\Big).$$

◁

Remark: Sometimes it is difficult to remember the formula for real valued fundamental solutions. One way to obtain those solutions without remembering the formula is to repeat the proof of Theorem 2.2.3. Start with the complex-valued solution $\tilde{y}_+$ and use the properties of the complex exponential,

$$\tilde{y}_+(t) = e^{(-2+3i)t} = e^{-2t}\, e^{3it} = e^{-2t}\, \big(\cos(3t) + i\sin(3t)\big).$$

The real-valued fundamental solutions are the real and imaginary parts of this expression.

Example 2.2.8. Find real-valued fundamental solutions to the equation

$$y'' + 2\,y' + 6\,y = 0.$$

Solution: The roots of the characteristic polynomial $p(r) = r^2 + 2r + 6$ are

$$r_\pm = \frac{1}{2}\big[-2 \pm \sqrt{4-24}\big] = \frac{1}{2}\big[-2 \pm \sqrt{-20}\big] \quad \Rightarrow \quad r_\pm = -1 \pm i\sqrt{5}.$$

These are complex-valued roots, with

$$\alpha = -1, \qquad \beta = \sqrt{5}.$$

Real-valued fundamental solutions are

$$y_1(t) = e^{-t}\,\cos(\sqrt{5}\,t), \qquad y_2(t) = e^{-t}\,\sin(\sqrt{5}\,t).$$

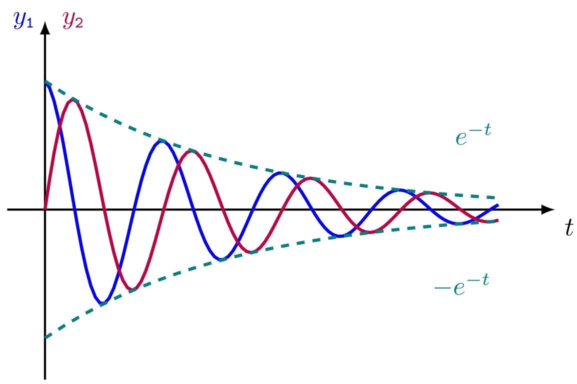


FIGURE 3. Solutions from Example 2.2.8.

The differential equation in this example, is a particular case of

$$m\,y'' + d\,y' + k\,y = 0$$

which describes the movement of a mass-spring system with mass m, spring constant k, oscillating in a liquid with damping constant d. In the case of this example we have $d/m = 2$ and $k/m = 6$. Second order differential equations with positive coefficients and with characteristic polynomials having complex roots, like the one in this example, describe physical processes related to damped oscillations. ◁

Example 2.2.9. Find the real valued general solution of

$$y'' + 5\,y = 0.$$

Solution: The characteristic polynomial is $p(r) = r^2 + 5$, with roots $r_\pm = \pm\sqrt{5}\,i$. In this case $\alpha = 0$, and $\beta = \sqrt{5}$. Real valued fundamental solutions are

$$y_+(t) = \cos(\sqrt{5}\,t), \qquad y_-(t) = \sin(\sqrt{5}\,t).$$

The real valued general solution is

$$y(t) = c_+\,\cos(\sqrt{5}\,t) + c_-\,\sin(\sqrt{5}\,t), \qquad c_+, c_- \in \mathbb{R}.$$

Physical processes that oscillate in time without dissipation could be described by differential equations like the one in this example. ◁

In the following example we solve an initial value problem for a mass-spring system oscillating in a medium *without friction*. We write the solution in terms of the amplitude and the phase shift of the oscillations.

Example 2.2.10 (No Friction)**.** Find the motion of a 2 kg mass attached to a spring with constant $k = 8$ kg/sec^2 moving in a medium without any friction, and having initial conditions $y(0) = -\sqrt{3}$ m and $y'(0) = 2$ m/sec.

Solution: Newton's law of motion for this mass is

$$m\,y'' + k\,y = 0$$

with $m = 2$, $k = 8$, that is,

$$y'' + 4y = 0.$$

The characteristic polynomial is $p(r) = r^2 + 4$ and its roos are

$$r_{\pm} = \pm 2i.$$

We can write the solution in terms of an amplitude and a phase shift,

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

We now use the initial conditions to find out the amplitude A and phase-shift ϕ. But first we need to compute the derivative,

$$y'(t) = -2A\sin(2t - \phi).$$

The initial conditions imply

$$-\sqrt{3} = y(0) = A\cos(-\phi) = A\cos(\phi) \quad \Rightarrow \quad A\cos(\phi) = -\sqrt{3}, \tag{2.2.7}$$

$$2 = y'(0) = -2A\sin(-\phi) = 2A\sin(\phi) \quad \Rightarrow \quad A\sin(\phi) = 1, \tag{2.2.8}$$

where we used the identities

$$\cos(-\phi) = \cos(\phi), \qquad \sin(-\phi) = -\sin(\phi).$$

The amplitude A can be obtained by first squaring both equations (2.2.7), (2.2.8), and then adding them,

$$A^2\big(\cos^2(\phi) + \sin^2(\phi)\big) = \big(-\sqrt{3}\big)^2 + 1^2 = 3 + 1 \quad \Rightarrow \quad A = 2,$$

where we used that $A > 0$ and

$$\cos^2(\phi) + \sin^2(\phi) = 1.$$

The phase-shift ϕ can be computed from Eq. (2.2.8) divided by (2.2.7),

$$\frac{A\sin(\phi)}{A\cos(\phi)} = -\frac{1}{\sqrt{3}} \quad \Rightarrow \quad \tan(\phi) = -\frac{1}{\sqrt{3}}.$$

Recall $\phi \in [-\pi, \pi)$ and the equation for the tangent has two solutions in that interval,

$$\tan(\phi) = -\frac{1}{\sqrt{3}} \quad \Rightarrow \quad \phi_1 = -\frac{\pi}{6}, \quad \text{or} \quad \phi_2 = \pi - \frac{\pi}{6} = \frac{5\pi}{6}.$$

In order to decide which solution is the phase-shift in our problem we notice that, since the amplitude is non-negative, the equations in (2.2.7), (2.2.8) imply that the phase-shift ϕ must satisfy

$$\cos(\phi) < 0, \qquad \sin(\phi) > 0.$$

Our candidates for the phase-shift, $\phi_1 = -\dfrac{\pi}{6}$ and $\phi_2 = \dfrac{5\pi}{6}$ satisfy

$$\begin{matrix} \cos(\phi_1) > 0 \\ \sin(\phi_1) < 0 \end{matrix} \quad \text{and} \quad \begin{matrix} \cos(\phi_2) < 0 \\ \sin(\phi_2) > 0. \end{matrix}$$

Therefore, the phase shift in our problem is

$$\phi = \phi_2 = \frac{5\pi}{6}.$$

Therefore we obtain the solution

$$y(t) = 2\cos\Big(2t - \frac{5\pi}{6}\Big).$$

◁

In the following example we solve an initial value problem for a mass-spring system oscillating in a medium *with friction*. We write the solution in terms of the amplitude and the phase shift of the oscillations.

Example 2.2.11 (With Friction)**.** Find the motion of a 5 kg mass attached to a spring with constant $k = 5$ kg/sec^2 moving in a medium with damping constant $d = 5$ kg/sec, with initial conditions $y(0) = \sqrt{3}$ m and $y'(0) = 0$ m/sec.

Solution: Newton's law of motion for this mass is

$$m\, y'' + d\, y' + k\, y = 0$$

with $m = 5$, $k = 5$, $d = 5$, that is,

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

The characteristic polynomial is $p(r) = r^2 + r + 1$ and its roos are

$$r_\pm = \frac{1}{2}\big(-1 \pm \sqrt{1-4}\big) \quad \Rightarrow \quad r_\pm = -\frac{1}{2} \pm i\,\frac{\sqrt{3}}{2}.$$

We can write the solution in terms of an amplitude and a phase shift,

$$y(t) = A\, e^{-t/2} \cos\Big(\frac{\sqrt{3}}{2}\, t - \phi\Big).$$

We now use the initial conditions to find out the amplitude A and phase-shift ϕ. But first we need to compute the derivative,

$$y'(t) = -\frac{1}{2}\, A\, e^{-t/2} \cos\Big(\frac{\sqrt{3}}{2}\, t - \phi\Big) - \frac{\sqrt{3}}{2}\, A\, e^{-t/2} \sin\Big(\frac{\sqrt{3}}{2}\, t - \phi\Big).$$

The initial conditions and the identities $\cos(-\phi) = \cos(\phi)$ and $\sin(-\phi) = -\sin(\phi)$ imply

$$\sqrt{3} = y(0) = A\cos(\phi), \qquad 0 = y'(0) = -\frac{1}{2}\, A\cos(\phi) + \frac{\sqrt{3}}{2}\, A\sin(\phi).$$

If we use the equation on the left in the equation on the right we get

$$0 = -\frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2}\, A\sin(\phi).$$

Therefore, the initial condition can be written as

$$A\cos(\phi) = \sqrt{3}, \quad A\sin(\phi) = 1. \tag{2.2.9}$$

The amplitude A can be obtained by squaring both equation and adding them,

$$A^2\big(\cos^2(\phi) + \sin^2(\phi)\big) = 3 + 1 \quad \Rightarrow \quad A = 2,$$

since $\cos^2(\phi) + \sin^2(\phi) = 1$. The phase-shift ϕ can be computed from the quotient of the equations above,

$$\frac{A\sin(\phi)}{A\cos(\phi)} = \frac{1}{\sqrt{3}} \quad \Rightarrow \quad \tan(\phi) = \frac{1}{\sqrt{3}}.$$

Recall $\phi \in [-\pi, \pi)$ and the equation for the tangent has two solutions in that interval,

$$\tan(\phi) = \frac{1}{\sqrt{3}} \quad \Rightarrow \quad \phi_1 = \frac{\pi}{6}, \quad \text{or} \quad \phi_2 = \frac{\pi}{6} - \pi = -\frac{5\pi}{6}.$$

In order to decide which solution is the phase-shift in our problem we notice that, since the amplitude is non-negative, the equations in (2.2.9) imply that the phase-shift ϕ must satisfy

$$\cos(\phi) > 0, \qquad \sin(\phi) > 0.$$

Our candidates for the phase-shift, $\phi_1 = \frac{\pi}{6}$ and $\phi_2 = -\frac{5\pi}{6}$ satisfy

$$\begin{matrix} \cos(\phi_1) > 0 \\ \sin(\phi_1) > 0 \end{matrix} \quad \text{and} \quad \begin{matrix} \cos(\phi_2) < 0 \\ \sin(\phi_2) < 0. \end{matrix}$$

Therefore, the phase shift in our problem is

$$\phi = \phi_1 = \frac{\pi}{6}.$$

Therefore we obtain the solution

$$y(t) = 2\,e^{-t/2}\cos\Big(\frac{\sqrt{3}}{2}\,t - \frac{\pi}{6}\Big).$$

◁

Example 2.2.12 (Variable Dissipation)**.** Consider a mass-spring system with mass $m = 1$ grams and spring constant $k = 49$ grams/sec^2 hanging vertically in a liquid with damping constant $d \geqslant 0$. If y is the vertical coordinate, positive downwards with $y = 0$ at the resting position, then this system is described by the differential equation

$$y'' + d\,y' + 49\,y = 0.$$

(a) Find the values of the damping constant d so that all solutions oscillate without slowing down.
(b) Find the values of the damping constant d so that all solutions oscillate and slowing down.
(c) Find the values of the damping constant d so that all solutions slowing down without any oscillation.

Solution: The general solution of the differential equation above is given by

$$y(t) = c_1\,e^{r_1 t} + c_2\,e^{r_2 t},$$

where r_1, r_2 are solutions of the characteristic equation

$$r^2 + d\,r + 49 = 0 \quad \Rightarrow \quad r_\pm = -\frac{d}{2} \pm \sqrt{\Big(\frac{d}{2}\Big)^2 - 7^2}.$$

(a) If all the solutions of the differential equation oscillate without slowing down, then these solutions must be combinations of sine and cosine functions. This happens when

$$d = 0,$$

because in this case the solutions of the characteristic equation are

$$r_\pm = \pm 7i,$$

and then the general solution is

$$y(t) = c_1 \cos(7t) + c_2 \sin(7t).$$

These solutions oscillate without slowing down as time grows.

(b) If all the solutions of the differential equation oscillate and slow down, then these solutions must be combinations of decaying exponentials times sine or cosine functions. This happens when

$$0 < \frac{d}{2} < 7,$$

because in this case

$$7^2 - \Big(\frac{d}{2}\Big)^2 > 0$$

and then the solutions of the characteristic equation are

$$r_\pm = -\frac{d}{2} \pm \sqrt{\Big(\frac{d}{2}\Big)^2 - 7^2} = -\frac{d}{2} \pm \sqrt{-\Big(7^2 - \Big(\frac{d}{2}\Big)^2\Big)},$$

which leads us to

$$r_\pm = -\frac{d}{2} \pm i\,\sqrt{7^2 - \Big(\frac{d}{2}\Big)^2}.$$

If we introduce the real numbers

$$\alpha = d/2, \qquad \beta = \sqrt{7^2 - \Big(\frac{d}{2}\Big)^2},$$

then the solutions of the characteristic equation are $r_\pm = -\alpha \pm i\beta$. In this case the general solution of the differential equation is

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

These solutions oscillate and slow down as time grows.

(c) If all the solutions of the differential equation slow down without oscillations, then these solutions must be combinations of decaying exponentials without sine nor cosine functions. This happens when

$$\frac{d}{2} > 7,$$

because in this case

$$\Big(\frac{d}{2}\Big)^2 - 7^2 > 0$$

and then the solutions of the characteristic equation are

$$r_\pm = -\frac{d}{2} \pm \sqrt{\Big(\frac{d}{2}\Big)^2 - 7^2} < 0,$$

which are both real and both negative. In this case the general solution of the differential equation is

$$y(t) = c_1 e^{r_+ t} + c_2\, e^{r_- t}.$$

These solutions slow down without any oscillation.

◁

2.2.3. Constructive Proof of Theorem 2.2.2. We now show an alternative proof for Theorem 2.2.2, which does not involve guessing fundamental solutions of the equation.

Proof of Theorem 2.2.2: The idea of the proof is to transform the original equation into an equation simpler to solve for a new unknown, then solve this simpler problem for the new unknown, and transform back the solution to the original function. We transform the problem by by writing the function y as a product of two functions, that is, $y(t) = u(t)\,v(t)$. If we choose the function v in an appropriate way, then the equation for the function u will be simpler to solve than the equation for y. In order to introduce $y = uv$ into the differential equation we need to compute its first and second derivatives,

$$y = uv \quad \Rightarrow \quad y' = u'v + v'u \quad \Rightarrow \quad y'' = u''v + 2u'v' + v''u.$$

Therefore, Eq. (2.2.4) implies that

$$(u''v + 2u'v' + v''u) + a_1\,(u'v + v'u) + a_0\,uv = 0,$$

that is,

$$\Big[u'' + \Big(a_1 + 2\frac{v'}{v}\Big)\,u' + a_0\,u\Big]v + (v'' + a_1\,v')\,u = 0. \tag{2.2.10}$$

We now choose the function v such that

$$a_1 + 2\frac{v'}{v} = 0 \quad \Leftrightarrow \quad \frac{v'}{v} = -\frac{a_1}{2}. \tag{2.2.11}$$

We choose a simple solution of this equation, given by

$$v(t) = e^{-a_1 t/2}.$$

Having this expression for v one can compute v' and v'', and it is simple to check that

$$v'' + a_1 v' = -\frac{a_1^2}{4}\,v. \tag{2.2.12}$$

Introducing the first equation in (2.2.11) and Eq. (2.2.12) into Eq. (2.2.10), and recalling that v is non-zero, we obtain the simplified equation for the function u, given by

$$u'' - k\,u = 0, \qquad k = \frac{a_1^2}{4} - a_0. \tag{2.2.13}$$

Eq. (2.2.13) for u is simpler than the original equation (2.2.4) for y since in the former there is no term with the first derivative of the unknown function. To solve Eq. (2.2.13) we repeat the idea followed to obtain this equation, that is, express function u as a product of two functions, and solve a simple problem of one of the functions. We first consider the harder case, which is when $k \neq 0$. In this case, let us express $u(t) = e^{\sqrt{k}t}\,w(t)$. Hence,

$$u' = \sqrt{k}e^{\sqrt{k}t}\,w + e^{\sqrt{k}t}\,w' \quad \Rightarrow \quad u'' = ke^{\sqrt{k}t}\,w + 2\sqrt{k}e^{\sqrt{k}t}\,w' + e^{\sqrt{k}t}\,w''.$$

Therefore, Eq. (2.2.13) for function u implies the following equation for function w

$$0 = u'' - ku = e^{\sqrt{k}t}\,(2\sqrt{k}\,w' + w'') \quad \Rightarrow \quad w'' + 2\sqrt{k}w' = 0.$$

Only derivatives of w appear in the latter equation, so denoting $x(t) = w'(t)$ we have to solve a simple equation

$$x' = -2\sqrt{k}\,x \quad \Rightarrow \quad x(t) = x_0 e^{-2\sqrt{k}t}, \qquad x_0 \in \mathbb{R}.$$

Integrating we obtain w as follows,

$$w' = x_0 e^{-2\sqrt{k}t} \quad \Rightarrow \quad w(t) = -\frac{x_0}{2\sqrt{k}}e^{-2\sqrt{k}t} + c_0.$$

Renaming $c_1 = -x_0/(2\sqrt{k})$, we obtain

$$w(t) = c_1 e^{-2\sqrt{k}t} + c_0 \quad \Rightarrow \quad u(t) = c_0 e^{\sqrt{k}t} + c_1 e^{-\sqrt{k}t}.$$

We then obtain the expression for the solution $y = uv$, given by

$$y(t) = c_0 e^{(-\frac{a_1}{2}+\sqrt{k})t} + c_1 e^{(-\frac{a_1}{2}-\sqrt{k})t}.$$

Since $k = (a_1^2/4 - a_0)$, the numbers

$$r_\pm = -\frac{a_1}{2} \pm \sqrt{k} \quad \Leftrightarrow \quad r_\pm = \frac{1}{2}\Big(-a_1 \pm \sqrt{a_1^2 - 4a_0}\Big)$$

are the roots of the characteristic polynomial

$$r^2 + a_1\, r + a_0 = 0,$$

we can express all solutions of the Eq. (2.2.4) as follows

$$y(t) = c_0 e^{r_+ t} + c_1 e^{r_- t}, \qquad k \neq 0.$$

Finally, consider the case $k = 0$. Then, Eq. (2.2.13) is simply given by

$$u'' = 0 \quad \Rightarrow \quad u(t) = (c_0 + c_1 t) \qquad c_0, c_1 \in \mathbb{R}.$$

Then, the solution y to Eq. (2.2.4) in this case is given by

$$y(t) = (c_0 + c_1 t)\, e^{-a_1 t/2}.$$

Since $k = 0$, the characteristic equation $r^2 + a_1\, r + a_0 = 0$ has only one root,

$$r_+ = r_- = -a_1/2,$$

the solution y above can be expressed as

$$y(t) = (c_0 + c_1 t)\, e^{r_+ t}, \qquad k = 0.$$

The Furthermore part is the same as in Theorem 2.2.2. This establishes the Theorem. □

2.2.4. Note On the Repeated Root Case. In the case that the characteristic polynomial of a differential equation has repeated roots there is an interesting argument to guess the solution y_-. The idea is to take a particular type of limit in solutions of differential equations with complex valued roots.

Consider the equation in (2.2.4) with a characteristic polynomial having complex valued roots given by $r_\pm = \alpha \pm i\beta$, with

$$\alpha = -\frac{a_1}{2}, \qquad \beta = \sqrt{a_0 - \frac{a_1^2}{4}}.$$

Real valued fundamental solutions in this case are given by

$$\hat{y}_+ = e^{\alpha\, t} \cos(\beta t), \quad \hat{y}_- = e^{\alpha\, t} \sin(\beta t).$$

We now study what happen to these solutions $\hat{y}_+$ and $\hat{y}_-$ in the following limit: The variable t is held constant, α is held constant, and $\beta \to 0$. The last two conditions are conditions on the equation coefficients, a_1, a_0. For example, we fix a_1 and we vary $a_0 \to a_1^2/4$ from above. Since $\cos(\beta t) \to 1$ as $\beta \to 0$ with t fixed, then keeping α fixed too, we obtain

$$\hat{y}_+(t) = e^{\alpha\, t} \cos(\beta t) \;\longrightarrow\; e^{\alpha\, t} = y_+(t).$$

Since $\dfrac{\sin(\beta t)}{\beta t} \to 1$ as $\beta \to 0$ with t constant, that is, $\sin(\beta t) \to \beta t$, we conclude that

$$\frac{\hat{y}_-(t)}{\beta} = \frac{\sin(\beta t)}{\beta}\, e^{\alpha\, t} = \frac{\sin(\beta t)}{\beta t}\, t\, e^{\alpha\, t} \;\longrightarrow\; t\, e^{\alpha\, t} = y_-(t).$$

The calculation above says that the function $\hat{y}_{\text{-}}/\beta$ is close to the function $y_{\text{-}}(t) = t\, e^{\alpha\, t}$ in the limit $\beta \to 0$, t held constant. This calculation provides a candidate, $y_{\text{-}}(t) = t\, y_{\text{+}}(t)$, of a solution to Eq. (2.2.4). It is simple to verify that this candidate is in fact solution of Eq. (2.2.4). Since $y_{\text{-}}$ is not proportional to $y_{\text{+}}$, we conclude the functions $y_{\text{+}}$, $y_{\text{-}}$ are a fundamental set for the differential equation in (2.2.4) in the case the characteristic polynomial has repeated roots.

2.2.5. Exercises.

2.2.1.- .

2.3. Nonhomogeneous Equations

All solutions of a linear *homogeneous* equation can be obtained from only two solutions that are linearly independent— fundamental solutions. Every other solution is a linear combination of these two. This is the general solution formula for homogeneous equations, and it is the main result in § 2.1, Theorem 2.1.9. This result is not longer true for *nonhomogeneous* equations. The superposition property, Theorem 2.1.7, which played an important part to get the general solution formula for homogeneous equations, is not true for nonhomogeneous equations.

We start this section proving a general solution formula for nonhomogeneous equations. We show that all the solutions of the nonhomogeneous equation are a translation by a fixed function of the solutions of the homogeneous equation. The fixed function is one solution— it doesn't matter which one—of the nonhomogenous equation, and it is called a particular solution of the nonhomogeneous equation.

Later in this section we show two different ways to compute the particular solution of a nonhomogeneous equation—the undetermined coefficients method and the variation of parameters method. In the former method we guess a particular solution from the expression of the source in the equation. The guess contains a few unknown constants, the undetermined coefficients, that must be determined by the equation. The undetermined method works for constant coefficients linear operators and simple source functions. The source functions and the associated guessed solutions are collected in a small table. This table is constructed by trial and error. In the latter method we have a formula to compute a particular solution in terms of the equation source, and fundamental solutions of the homogeneous equation. The variation of parameters method works with variable coefficients linear operators and general source functions. But the calculations to find the solution are usually not so simple as in the undetermined coefficients method.

2.3.1. The General Solution Formula.

The general solution formula for homogeneous equations, Theorem 2.1.9, is no longer true for nonhomogeneous equations. But there is a general solution formula for nonhomogeneous equations. Such formula involves three functions, two of them are fundamental solutions of the homogeneous equation, and the third function is any solution of the nonhomogeneous equation. Every other solution of the nonhomogeneous equation can be obtained from these three functions.

Theorem 2.3.1 (General Solution)**.** *Every solution y of the nonhomogeneous equation*

$$L(y) = f, \tag{2.3.1}$$

with $L(y) = y'' + a_1\, y' + a_0\, y$, where a_1, a_0, and f are continuous functions, is given by

$$y = c_1\, y_1 + c_2\, y_2 + y_p,$$

where the functions y_1 and y_2 are fundamental solutions of the homogeneous equation, $L(y_1) = 0$, $L(y_2) = 0$, and y_p is any solution of the nonhomogeneous equation $L(y_p) = f$.

Before we proof Theorem 2.3.1 we state the following definition, which comes naturally from this Theorem.

Definition 2.3.2. *The* ***general solution*** *of the nonhomogeneous equation $L(y) = f$ is a two-parameter family of functions*

$$y_{\text{gen}}(t) = c_1\, y_1(t) + c_2\, y_2(t) + y_p(t), \tag{2.3.2}$$

where the functions y_1 and y_2 are fundamental solutions of the homogeneous equation, $L(y_1) = 0$, $L(y_2) = 0$, and y_p is any solution of the nonhomogeneous equation $L(y_p) = f$.

Remark: The difference of any two solutions of the nonhomogeneous equation is actually a solution of the homogeneous equation. This is the key idea to prove Theorem 2.3.1. In other words, the solutions of the nonhomogeneous equation are a *translation by a fixed function*, y_p, of the solutions of the homogeneous equation.

Proof of Theorem 2.3.1: Let y be any solution of the nonhomogeneous equation $L(y) = f$. Recall that we already have one solution, y_p, of the nonhomogeneous equation, $L(y_p) = f$. We can now subtract the second equation from the first,

$$L(y) - L(y_p) = f - f = 0 \quad \Rightarrow \quad L(y - y_p) = 0.$$

The equation on the right is obtained from the linearity of the operator L. This last equation says that the difference of any two solutions of the nonhomogeneous equation is solution of the homogeneous equation. The general solution formula for homogeneous equations says that all solutions of the homogeneous equation can be written as linear combinations of a pair of fundamental solutions, y_1, y_2. So the exist constants c_1, c_2 such that

$$y - y_p = c_1\, y_1 + c_2\, y_2.$$

Since for every y solution of $L(y) = f$ we can find constants c_1, c_2 such that the equation above holds true, we have found a formula for all solutions of the nonhomogeneous equation. This establishes the Theorem. □

2.3.2. The Undetermined Coefficients Method.

The general solution formula in (2.3.2) is the most useful if there is a way to find a particular solution y_p of the nonhomogeneous equation $L(y_p) = f$. We now present a method to find such particular solution, the undetermined coefficients method. This method works for *linear operators L with constant coefficients* and for *simple source functions f*. Here is a summary of the undetermined coefficients method:

(1) Find fundamental solutions y_1, y_2 of the homogeneous equation $L(y) = 0$.
(2) Given the source functions f, guess the solutions y_p following the Table 1 below.
(3) If the function y_p given by the table satisfies $L(y_p) = 0$, then change the guess to $ty_{p.}$. If ty_p satisfies $L(ty_p) = 0$ as well, then change the guess to $t^2 y_p$.
(4) Find the undetermined constants k in the function y_p using the equation $L(y) = f$, where y is y_p, or ty_p or $t^2 y_p$.

$f(t)$ (Source) (K, m, a, b, given.)	$y_p(t)$ (Guess) (k not given.)
Ke^{at}	ke^{at}
$K_m t^m + \cdots + K_0$	$k_m t^m + \cdots + k_0$
$K_1 \cos(bt) + K_2 \sin(bt)$	$k_1 \cos(bt) + k_2 \sin(bt)$
$(K_m t^m + \cdots + K_0)\, e^{at}$	$(k_m t^m + \cdots + k_0)\, e^{at}$
$\big(K_1 \cos(bt) + K_2 \sin(bt)\big)\, e^{at}$	$\big(k_1 \cos(bt) + k_2 \sin(bt)\big)\, e^{at}$
$(K_m t^m + \cdots + K_0)\big(\tilde{K}_1 \cos(bt) + \tilde{K}_2 \sin(bt)\big)$	$(k_m t^m + \cdots + k_0)\big(\tilde{k}_1 \cos(bt) + \tilde{k}_2 \sin(bt)\big)$

TABLE 1. List of sources f and solutions y_p to the equation $L(y_p) = f$.

This is the undetermined coefficients method. It is a set of simple rules to find a particular solution y_p of an nonhomogeneous equation $L(y_p) = f$ in the case that the source function f is one of the entries in the Table 1. There are a few formulas in particular cases and a few generalizations of the whole method. We discuss them after a few examples.

Example 2.3.1 (First Guess Right)**.** Find all solutions to the nonhomogeneous equation

$$y'' - 3y' - 4y = 3\,e^{2t}.$$

Solution: From the problem we get $L(y) = y'' - 3y' - 4y$ and $f(t) = 3\,e^{2t}$.

(1): Find fundamental solutions y_+, y_- to the homogeneous equation $L(y) = 0$. Since the homogeneous equation has constant coefficients we find the characteristic equation

$$r^2 - 3r - 4 = 0 \quad \Rightarrow \quad r_+ = 4, \quad r_- = -1, \quad \Rightarrow \quad y_+(t) = e^{4t}, \quad y_- = (t) = e^{-t}.$$

(2): The table says: For $f(t) = 3e^{2t}$ guess $y_p(t) = k\,e^{2t}$. The constant k is the undetermined coefficient we must find.

(3): Since $y_p(t) = k\,e^{2t}$ is not solution of the homogeneous equation, we do not need to modify our guess. (Recall: $L(y) = 0$ iff exist constants c_+, c_- such that $y(t) = c_+\,e^{4t} + c_-\,e^{-t}$.)

(4): Introduce y_p into $L(y_p) = f$ and find k. So we do that,

$$(2^2 - 6 - 4)ke^{2t} = 3\,e^{2t} \quad \Rightarrow \quad -6k = 3 \quad \Rightarrow \quad k = -\frac{1}{2}.$$

We guessed that y_p must be proportional to the exponential e^{2t} in order to cancel out the exponentials in the equation above. We have obtained that

$$y_p(t) = -\frac{1}{2}\,e^{2t}.$$

The undetermined coefficients method gives us a way to compute a particular solution y_p of the nonhomogeneous equation. We now use the general solution theorem, Theorem 2.3.1, to write the general solution of the nonhomogeneous equation,

$$y_{\text{gen}}(t) = c_+\,e^{4t} + c_-\,e^{-t} - \frac{1}{2}\,e^{2t}.$$

$\lhd$

Remark: The step (4) in Example 2.3.1 is a particular case of the following statement.

Theorem 2.3.3. *Consider the equation $L(y) = f$, where $L(y) = y'' + a_1\,y' + a_0\,y$ has constant coefficients and p is its characteristic polynomial. If the source function is $f(t) = K\,e^{at}$, with $p(a) \neq 0$, then a particular solution of the nonhomogeneous equation is*

$$y_p(t) = \frac{K}{p(a)}\,e^{at}.$$

Proof of Theorem 2.3.3: Since the linear operator L has constant coefficients, let us write L and its associated characteristic polynomial p as follows,

$$L(y) = y'' + a_1 y' + a_0 y, \qquad p(r) = r^2 + a_1 r + a_0.$$

Since the source function is $f(t) = K\,e^{at}$, the Table 1 says that a good guess for a particular soution of the nonhomogneous equation is $y_p(t) = k\,e^{at}$. Our hypothesis is that this guess is not solution of the homogenoeus equation, since

$$L(y_p) = (a^2 + a_1 a + a_0)\,k\,e^{at} = p(a)\,k\,e^{at}, \quad \text{and} \quad p(a) \neq 0.$$

We then compute the constant k using the equation $L(y_p) = f$,

$$(a^2 + a_1 a + a_0)\, k\, e^{at} = K\, e^{at} \quad \Rightarrow \quad p(a)\, k\, e^{at} = K\, e^{at} \quad \Rightarrow \quad k = \frac{K}{p(a)}.$$

We get the particular solution $y_p(t) = \dfrac{K}{p(a)}\, e^{at}$. This establishes the Theorem. □

Remark: As we said, the step (4) in Example 2.3.1 is a particular case of Theorem 2.3.3,

$$y_p(t) = \frac{3}{p(2)}\, e^{2t} = \frac{3}{(2^2 - 6 - 4)}\, e^{2t} = \frac{3}{-6}\, e^{2t} \quad \Rightarrow \quad y_p(t) = -\frac{1}{2}\, e^{2t}.$$

In the following example our first guess for a particular solution y_p happens to be a solution of the homogenous equation.

Example 2.3.2 (First Guess Wrong)**.** Find all solutions to the nonhomogeneous equation

$$y'' - 3y' - 4y = 3\, e^{4t}.$$

Solution: If we write the equation as $L(y) = f$, with $f(t) = 3\, e^{4t}$, then the operator L is the same as in Example 2.3.1. So the solutions of the homogeneous equation $L(y) = 0$, are the same as in that example,

$$y_+(t) = e^{4t}, \qquad y_-(t) = e^{-t}.$$

The source function is $f(t) = 3\, e^{4t}$, so the Table 1 says that we need to guess $y_p(t) = k\, e^{4t}$. However, this function y_p is solution of the homogeneous equation, because

$$y_p = k\, y_+ \quad \Rightarrow \quad L(y_p) = 0.$$

We have to change our guess, as indicated in the undetermined coefficients method, step (3)

$$y_p(t) = kt\, e^{4t}.$$

This new guess is not solution of the homogeneous equation. So we proceed to compute the constant k. We introduce the guess into $L(y_p) = f$,

$$y_p' = (1 + 4t)\, k\, e^{4t}, \qquad y_p'' = (8 + 16t)\, k\, e^{4t} \quad \Rightarrow \quad \big[8 - 3 + (16 - 12 - 4)t\big]\, k\, e^{4t} = 3\, e^{4t},$$

therefore, we get that

$$5k = 3 \quad \Rightarrow \quad k = \frac{3}{5} \quad \Rightarrow \quad y_p(t) = \frac{3}{5}\, t\, e^{4t}.$$

The general solution theorem for nonhomogneneous equations says that

$$y_{\rm gen}(t) = c_+\, e^{4t} + c_-\, e^{-t} + \frac{3}{5}\, t\, e^{4t}.$$

◁

In the following example the equation source is a trigonometric function.

Example 2.3.3 (First Guess Right)**.** Find all solutions to the nonhomogeneous equation

$$y'' - 3y' - 4y = 2\sin(t).$$

Solution: If we write the equation as $L(y) = f$, with $f(t) = 2\sin(t)$, then the operator L is the same as in Example 2.3.1. So the solutions of the homogeneous equation $L(y) = 0$, are the same as in that example,

$$y_+(t) = e^{4t}, \qquad y_-(t) = e^{-t}.$$

Since the source function is $f(t) = 2\sin(t)$, the Table 1 says that we need to choose the function $y_p(t) = k_1\cos(t) + k_2\sin(t)$. This function y_p is not solution to the homogeneous equation. So we look for the constants k_1, k_2 using the differential equation,

$$y_p' = -k_1\sin(t) + k_2\cos(t), \qquad y_p'' = -k_1\cos(t) - k_2\sin(t),$$

and then we obtain

$$[-k_1\cos(t) - k_2\sin(t)] - 3[-k_1\sin(t) + k_2\cos(t)] - 4[k_1\cos(t) + k_2\sin(t)] = 2\sin(t).$$

Reordering terms in the expression above we get

$$(-5k_1 - 3k_2)\cos(t) + (3k_1 - 5k_2)\sin(t) = 2\sin(t).$$

The last equation must hold for all $t \in \mathbb{R}$. In particular, it must hold for $t = \pi/2$ and for $t = 0$. At these two points we obtain, respectively,

$$\left.\begin{aligned} 3k_1 - 5k_2 &= 2,\\ -5k_1 - 3k_2 &= 0, \end{aligned}\right\} \quad \Rightarrow \quad \begin{cases} k_1 = \dfrac{3}{17},\\[2mm] k_2 = -\dfrac{5}{17}. \end{cases}$$

So the particular solution to the nonhomogeneous equation is given by

$$y_p(t) = \frac{1}{17}\big[3\cos(t) - 5\sin(t)\big].$$

The general solution theorem for nonhomogeneous equations implies

$$y_{\rm gen}(t) = c_+ e^{4t} + c_- e^{-t} + \frac{1}{17}\big[3\cos(t) - 5\sin(t)\big].$$

$\lhd$

Example 2.3.4 (First Guess Right)**.** Find all solutions to the nonhomogeneous equation

$$y'' - 3y' - 4y = 3\,t^2.$$

Solution: If we write the equation as $L(y) = f$, with $f(t) = 3\,t^2$, then the operator L is the same as in Example 2.3.1. So the solutions of the homogeneous equation $L(y) = 0$, are the same as in that example,

$$y_+(t) = e^{4t}, \qquad y_-(t) = e^{-t}.$$

Since the source is $f(t) = 3\,t^2$, Table 1 says we need to choose the function

$$y_p(t) = k_2t^2 + k_1t + k_0.$$

This function y_p is not solution to the homogeneous equation. So we look for the constants k_2, k_1, k_0 using the differential equation. We start computing the first two derivative of y_p,

$$y_p' = 2k_2t + k_1, \qquad y_p'' = 2k_2,$$

and then put all that in the differential equation,

$$\big(2k_2\big) - 3\big(2k_2t + k_1\big) - 4\big(k_2t^2 + k_1t + k_0\big) = 3\,t^2.$$

Reordering terms in the expression above we get

$$\big(-4k_2 - 3\big)t^2 + \big(-6k_2 - 4k_1\big)t + \big(2k_2 - 3k_1 - 4k_0\big) = 0$$

The last equation must hold for all $t \in \mathbb{R}$. This implies that each coefficient must vanish,

$$\begin{aligned} 4k_2 + 3 &= 0\\ 6k_2 + 4k_1 &= 0\\ 2k_2 - 3k_1 - 4k_0 &= 0. \end{aligned}$$

(Proof: If we have the equation $at^2 + bt + c = 0$ for all t, then evaluating at $t = 0$ we get that $c = 0$; derivate the equation with respect to t and we get $2at + b = 0$ for all t, evaluate that at $t = 0$ and we get $b = 0$; derivate one more time and the get $2a = 0$, that is $a = 0$. End of Proof.) We solve this system from the top equation to the bottom, and we get

$$k_2 = -\frac{3}{4}, \qquad k_1 = \frac{9}{8}, \qquad k_0 = \frac{39}{32}.$$

Then, the particular solution to the nonhomogeneous equation is given by

$$y_p(t) = -\frac{3}{4}\,t^2 + \frac{9}{8}\,t + \frac{39}{32}.$$

The general solution theorem for nonhomogeneous equations implies

$$y_{\rm gen}(t) = c_+\, e^{4t} + c_-\, e^{-t} - \frac{3}{4}\,t^2 + \frac{9}{8}\,t + \frac{39}{32}.$$

◁

In the next example we show a few nonhomogeneous equations and the corresponding guesses for the particular solution y_p.

Example 2.3.5. We provide few more examples of nonhomogeneous equations and the appropriate guesses for the particular solutions.

(a) For $y'' - 3y' - 4y = 3e^{2t}\sin(t)$, guess, $y_p(t) = \big[k_1\cos(t) + k_2\sin(t)\big]\, e^{2t}$.

(b) For $y'' - 3y' - 4y = 2t^2\, e^{3t}$, guess, $y_p(t) = \big(k_2 t^2 + k_1 t + k_0\big)\, e^{3t}$.

(c) For $y'' - 3y' - 4y = 2t^2\, e^{4t}$, guess, $y_p(t) = \big(k_2 t^2 + k_1 t + k_0\big)\, t\, e^{4t}$.

(d) For $y'' - 3y' - 4y = 3t\,\sin(t)$, guess, $y_p(t) = (k_1 t + k_0)\,\big[\tilde k_1\cos(t) + \tilde k_2\sin(t)\big]$.

◁

Remark: Suppose that the source function f does not appear in Table 1, but f can be written as $f = f_1 + f_2$, with f_1 and f_2 in the table. In such case look for a particular solution $y_p = y_{p_1} + y_{p_2}$, where $L(y_{p_1}) = f_1$ and $L(y_{p_2}) = f_2$. Since the operator L is linear,

$$L(y_p) = L(y_{p_1} + y_{p_2}) = L(y_{p_1}) + L(y_{p_2}) = f_1 + f_2 = f \quad \Rightarrow \quad L(y_p) = f.$$

In our next example we describe the electric current flowing through an RLC-series electric circuit, which consists of a resistor R, an inductor L, a capacitor C, and a voltage source $V(t)$ connected in series as shown in Fig. 4.

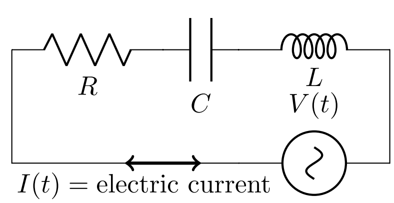


FIGURE 4. An RLC circuit.

This system is described by an integro-differential equation found by Kirchhoff, now called Kirchhoff's voltage law,

$$L\, I'(t) + R\, I(t) + \frac{1}{C}\int_{t_0}^{t} I(s)\, ds = V(t). \tag{2.3.3}$$

If we take one time derivative in the equation above we obtain a second order differential equation for the electric current,

$$L\,I''(t) + R\,I'(t) + \frac{1}{C}\,I(t) = V'(t).$$

This equation is usually rewritten as

$$I''(t) + 2\Big(\frac{R}{2L}\Big)\,I'(t) + \frac{1}{LC}\,I(t) = \frac{V'(t)}{L}.$$

If we introduce *damping frequency* $\omega_d = \dfrac{R}{2L}$ and the *natural frequency* $\omega_0 = \dfrac{1}{\sqrt{LC}}$, then Kirchhoff's law can be expressed as

$$I'' + 2\omega_d\,I' + \omega_0^2\,I = \frac{V'(t)}{L}.$$

We are now ready to solve the following example.

Example 2.3.6. Consider a RLC-series circuit with no resistor, capacitor C, inductor L and voltage source $V(t) = V_0\,\sin(\nu t)$, where $\nu \neq \omega_0 = \dfrac{1}{\sqrt{LC}}$. Find the electric current in the case $I(0) = 0$, $I'(0) = 0$.

Solution: Kirchhoff equation for this problem is

$$I'' + \omega_0^2 I = v_0 \nu\,\cos(\nu t)$$

where we denoted $v_0 = \dfrac{V_0}{L}$. We start finding the solutions of the homogeneous equation

$$I'' + \omega_0^2 I = 0.$$

The characteristic equation is $r^2 + \omega_0^2 = 0$, and the roots are $r_\pm = \pm\omega_0 i$, and real valued fundamental solutions are

$$I_+ = \cos(\omega_0 t), \qquad I_- = \sin(\omega_0 t).$$

For $\nu \neq \omega_0$ the source function is not solution of the homogeneous equation, so the correct guess for a particular solution of the nonhomogeneous equation is

$$I_p = c_1 \cos(\nu t) + c_2 \sin(\nu t).$$

If we put this function I_p into the nonhomogeneous equation we get

$$-\nu^2\big(c_1\cos(\nu t) + c_2\sin(\nu t)\big) + \omega_0^2\big(c_1\cos(\nu t) + c_2\sin(\nu t)\big) = v_0\nu\cos(\nu t).$$

If we reorder terms we get

$$\big(c_1(\omega_0^2 - \nu^2) - v_0\nu\big)\cos(\nu t) + c_2(\omega_0^2 - \nu^2)\sin(\nu t) = 0.$$

From here we get that

$$c_1(\omega_0^2 - \nu^2) - v_0\nu = 0, \qquad c_2(\omega_0^2 - \nu^2).$$

Since we are studying the case $\nu \neq \omega_0$, we conclude that

$$c_1 = \frac{v_0\nu}{(\omega_0^2 - \nu^2)}, \qquad c_2 = 0.$$

So, the particular solution is

$$I_p(t) = \frac{v_0\nu}{(\omega_0^2 - \nu^2)}\cos(\nu t).$$

The general solution of the nonhomogeneous equation is

$$I(t) = c_+ I_+ + c_- I_- + I_p \quad \Rightarrow \quad I(t) = c_+ \cos(\omega_0 t) + c_- \sin(\omega_0 t) + \frac{v_0 \nu}{(\omega_0^2 - \nu^2)} \cos(\nu t).$$

We now look for the solution that satisfies the initial conditions $I(0) = 0$, and $I'(0) = 0$. For the first condition we get

$$0 = I(0) = c_+ + \frac{v_0 \nu}{(\omega_0^2 - \nu^2)} \quad \Rightarrow \quad c_+ = -\frac{v_0 \nu}{(\omega_0^2 - \nu^2)}.$$

The other boundary condition implies

$$0 = I'(0) = c_- \omega_0 \quad \Rightarrow \quad c_- = 0.$$

So, the solution of the initial value problem for the electric current is

$$I(t) = \frac{v_0 \nu}{(\omega_0^2 - \nu^2)} \big(\cos(\nu t) - \cos(\omega_0 t) \big).$$

◁

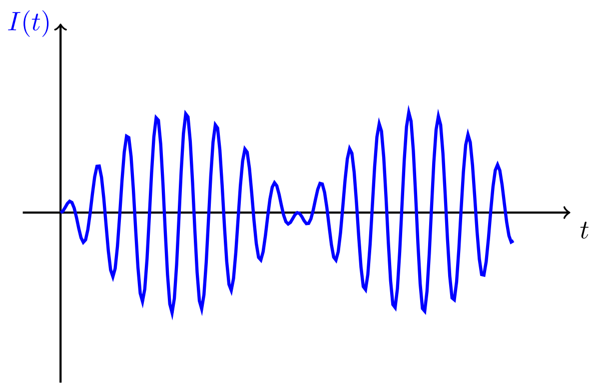


FIGURE 5. The I for ν close to ω_0, showing beating, when ν is close to ω_0.

Interactive Graph Link: Beating Phenomenon. Click on the interactive graph link here to see how the solution $I(t)$ changes when $\nu \to \omega_0$, exhibiting the beating phenomenon shown in Fig. 5.

2.3.3. The Variation of Parameters Method. This method provides a second way to find a particular solution y_p to a nonhomogeneous equation $L(y) = f$. We summarize this method in formula to compute y_p in terms of any pair of fundamental solutions to the homogeneous equation $L(y) = 0$. The variation of parameters method works with second order linear equations having *variable coefficients* and contiuous but otherwise *arbitrary sources*. When the source function of a nonhomogeneous equation is simple enough to appear in Table 1 the undetermined coefficients method is a quick way to find a particular solution to the equation. When the source is more complicated, one usually turns to the variation of parameters method, with its more involved formula for a particular solution.

Theorem 2.3.4 (Variation of Parameters)**.** *A particular solution to the equation*

$$L(y) = f,$$

with $L(y) = y'' + a_1(t)\, y' + a_0(t)\, y$ *and* a_1, a_0, f *continuous functions, is given by*

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

where y_1, y_2 are fundamental solutions of the homogeneous equation $L(y) = 0$ and the functions u_1, u_2 are defined by

$$u_1(t) = \int -\frac{y_2(t) f(t)}{W_{y_1 y_2}(t)}\, dt, \qquad u_2(t) = \int \frac{y_1(t) f(t)}{W_{y_1 y_2}(t)}\, dt, \tag{2.3.4}$$

where $W_{y_1 y_2}$ is the Wronskian of y_1 and y_2.

The proof is a generalization of the reduction order method. Recall that the reduction order method is a way to find a second solution y_2 of an homogeneous equation if we already know one solution y_1. One writes $y_2 = u\, y_1$ and the original equation $L(y_2) = 0$ provides an equation for u. This equation for u is simpler than the original equation for y_2 because the function y_1 satisfies $L(y_1) = 0$.

The formula for y_p can be seen as a generalization of the reduction order method. We write y_p in terms of both fundamental solutions y_1, y_2 of the homogeneous equation,

$$y_p(t) = u_1(t)\, y_1(t) + u_2(t)\, y_2(t).$$

We put this y_p in the equation $L(y_p) = f$ and we find an equation relating u_1 and u_2. It is important to realize that we have added one new function to the original problem. The original problem is to find y_p. Now we need to find u_1 and u_2, but we still have only one equation to solve, $L(y_p) = f$. The problem for u_1, u_2 cannot have a unique solution. So we are completely free to add a second equation to the original equation $L(y_p) = f$. We choose the second equation so that we can solve for u_1 and u_2.

Proof of Theorem 2.3.4: We look for a particular solution y_p of the form

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

We hope that the equations for u_1, u_2 will be simpler to solve than the equation for y_p. But we started with one unknown function and now we have two unknown functions. So we are free to add one more equation to fix u_1, u_2. We choose

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

In other words, we choose $u_2 = \displaystyle\int -\frac{y_1}{y_2}\, u_1'\, dt$. Let's put this y_p into $L(y_p) = f$. We need y_p' (and recall, $u_1'\, y_1 + u_2'\, y_2 = 0$)

$$y_p' = u_1'\, y_1 + u_1\, y_1' + u_2'\, y_2 + u_2\, y_2' \quad \Rightarrow \quad y_p' = u_1\, y_1' + u_2\, y_2'.$$

and we also need y_p'',

$$y_p'' = u_1'\, y_1' + u_1\, y_1'' + u_2'\, y_2' + u_2\, y_2''.$$

So the equation $L(y_p) = f$ is

$$(u_1'\, y_1' + u_1\, y_1'' + u_2'\, y_2' + u_2\, y_2'') + a_1(u_1\, y_1' + u_2\, y_2') + a_0(u_1\, y_1 + u_2\, y_2) = f$$

We reorder a few terms and we see that

$$u_1'\, y_1' + u_2'\, y_2' + u_1\,(y_1'' + a_1\, y_1' + a_0\, y_1) + u_2\,(y_2'' + a_1\, y_2' + a_0\, y_2) = f.$$

The functions y_1 and y_2 are solutions to the homogeneous equation,

$$y_1'' + a_1\, y_1' + a_0\, y_1 = 0, \qquad y_2'' + a_1\, y_2' + a_0\, y_2 = 0,$$

so u_1 and u_2 must be solution of a simpler equation that the one above, given by

$$u_1'\, y_1' + u_2'\, y_2' = f. \tag{2.3.5}$$

So we end with the equations

$$u_1'\,y_1' + u_2'\,y_2' = f$$
$$u_1'\,y_1 + u_2'\,y_2 = 0.$$

And this is a 2×2 algebraic linear system for the unknowns u_1', u_2'. It is hard to overstate the importance of the word "algebraic" in the previous sentence. From the second equation above we compute u_2' and we introduce it in the first equation,

$$u_2' = -\frac{y_1}{y_2}\,u_1' \quad \Rightarrow \quad u_1'y_1' - \frac{y_1y_2'}{y_2}\,u_1' = f \quad \Rightarrow \quad u_1'\Big(\frac{y_1'y_2 - y_1y_2'}{y_2}\Big) = f.$$

Recall that the Wronskian of two functions is $W_{12} = y_1y_2' - y_1'y_2$, we get

$$u_1' = -\frac{y_2 f}{W_{12}} \quad \Rightarrow \quad u_2' = \frac{y_1 f}{W_{12}}.$$

These equations are the derivative of Eq. (2.3.4). Integrate them in the variable t and choose the integration constants to be zero. We get Eq. (2.3.4). This establishes the Theorem. $\square$

Remark: The integration constants in the expressions for u_1, u_2 can always be chosen to be zero. To understand the effect of the integration constants in the function y_p, let us do the following. Denote by u_1 and u_2 the functions in Eq. (2.3.4), and given any real numbers c_1 and c_2 define

$$\tilde{u}_1 = u_1 + c_1, \qquad \tilde{u}_2 = u_2 + c_2.$$

Then the corresponding solution $\tilde{y}_p$ is given by

$$\tilde{y}_p = \tilde{u}_1\,y_1 + \tilde{u}_2\,y_2 = u_1\,y_1 + u_2\,y_2 + c_1\,y_1 + c_2\,y_2 \quad \Rightarrow \quad \tilde{y}_p = y_p + c_1\,y_1 + c_2\,y_2.$$

The two solutions $\tilde{y}_p$ and y_p differ by a solution to the homogeneous differential equation. So both functions are also solution to the nonhomogeneous equation. One is then free to choose the constants c_1 and c_2 in any way. We chose them in the proof above to be zero.

Example 2.3.7. Find the general solution of the nonhomogeneous equation

$$y'' - 5y' + 6y = 2\,e^t.$$

Solution: The formula for y_p in Theorem 2.3.4 requires we know fundamental solutions to the homogeneous problem. So we start finding these solutions first. Since the equation has constant coefficients, we compute the characteristic equation,

$$r^2 - 5r + 6 = 0 \quad \Rightarrow \quad r_\pm = \frac{1}{2}\big(5 \pm \sqrt{25-24}\big) \quad \Rightarrow \quad \begin{cases} r_+ = 3, \\ r_- = 2. \end{cases}$$

So, the functions y_1 and y_2 in Theorem 2.3.4 are in our case given by

$$y_1(t) = e^{3t}, \qquad y_2(t) = e^{2t}.$$

The Wronskian of these two functions is given by

$$W_{y_1y_2}(t) = (e^{3t})(2\,e^{2t}) - (3\,e^{3t})(e^{2t}) \quad \Rightarrow \quad W_{y_1y_2}(t) = -e^{5t}.$$

We are now ready to compute the functions u_1 and u_2. Notice that Eq. (2.3.4) the following differential equations

$$u_1' = -\frac{y_2 f}{W_{y_1y_2}}, \qquad u_2' = \frac{y_1 f}{W_{y_1y_2}}.$$

So, the equation for u_1 is the following,

$$u_1' = -e^{2t}(2\,e^t)(-e^{-5t}) \quad \Rightarrow \quad u_1' = 2\,e^{-2t} \quad \Rightarrow \quad u_1 = -e^{-2t},$$

$$u_2' = e^{3t}(2\,e^{t})(-e^{-5t}) \quad \Rightarrow \quad u_2' = -2\,e^{-t} \quad \Rightarrow \quad u_2 = 2\,e^{-t},$$

where we have chosen the constant of integration to be zero. The particular solution we are looking for is given by

$$y_p = (-e^{-2t})(e^{3t}) + (2\,e^{-t})(e^{2t}) \quad \Rightarrow \quad y_p = e^{t}.$$

Then, the general solution theorem for nonhomogeneous equation implies

$$y_{\rm gen}(t) = c_+\,e^{3t} + c_-\,e^{2t} + e^{t} \qquad c_+, c_- \in \mathbb{R}.$$

◁

Example 2.3.8. Find a particular solution to the differential equation

$$t^2 y'' - 2y = 3t^2 - 1,$$

knowing that $y_1 = t^2$ and $y_2 = 1/t$ are solutions to the homogeneous equation $t^2 y'' - 2y = 0$.

Solution: We first rewrite the nonhomogeneous equation above in the form given in Theorem 2.3.4. In this case we must divide the whole equation by t^2,

$$y'' - \frac{2}{t^2}\,y = 3 - \frac{1}{t^2} \quad \Rightarrow \quad f(t) = 3 - \frac{1}{t^2}.$$

We now proceed to compute the Wronskian of the fundamental solutions y_1, y_2,

$$W_{y_1 y_2}(t) = (t^2)\Big(\frac{-1}{t^2}\Big) - (2t)\Big(\frac{1}{t}\Big) \quad \Rightarrow \quad W_{y_1 y_2}(t) = -3.$$

We now use the equation in (2.3.4) to obtain the functions u_1 and u_2,

$$\begin{aligned} u_1' &= -\frac{1}{t}\Big(3 - \frac{1}{t^2}\Big)\,\frac{1}{-3} & \qquad u_2' &= (t^2)\Big(3 - \frac{1}{t^2}\Big)\,\frac{1}{-3} \\ &= \frac{1}{t} - \frac{1}{3}\,t^{-3} \quad \Rightarrow \quad u_1 = \ln(t) + \frac{1}{6}t^{-2}, & &= -t^2 + \frac{1}{3} \quad \Rightarrow \quad u_2 = -\frac{1}{3}\,t^3 + \frac{1}{3}t. \end{aligned}$$

A particular solution to the nonhomogeneous equation above is $\tilde{y}_p = u_1 y_1 + u_2 y_2$, that is,

$$\begin{aligned} \tilde{y}_p &= \Big[\ln(t) + \frac{1}{6}t^{-2}\Big](t^2) + \frac{1}{3}(-t^3 + t)(t^{-1}) \\ &= t^2 \ln(t) + \frac{1}{6} - \frac{1}{3}\,t^2 + \frac{1}{3} \\ &= t^2 \ln(t) + \frac{1}{2} - \frac{1}{3}\,t^2 \\ &= t^2 \ln(t) + \frac{1}{2} - \frac{1}{3}\,y_1(t). \end{aligned}$$

However, a simpler expression for a solution of the nonhomogeneous equation above is

$$y_p = t^2 \ln(t) + \frac{1}{2}.$$

◁

Remark: Sometimes it could be difficult to remember the formulas for functions u_1 and u_2 in (2.3.4). In such case one can always go back to the place in the proof of Theorem 2.3.4 where these formulas come from, the system

$$\begin{aligned} u_1' y_1' + u_2' y_2' &= f \\ u_1' y_1 + u_2' y_2 &= 0. \end{aligned}$$

The system above could be simpler to remember than the equations in (2.3.4). We end this Section using the equations above to solve the problem in Example 2.3.8. Recall that the solutions to the homogeneous equation in Example 2.3.8 are $y_1(t) = t^2$, and $y_2(t) = 1/t$, while the source function is $f(t) = 3 - 1/t^2$. Then, we need to solve the system

$$t^2\, u_1' + u_2' \frac{1}{t} = 0,$$
$$2t\, u_1' + u_2' \frac{(-1)}{t^2} = 3 - \frac{1}{t^2}.$$

This is an algebraic linear system for u_1' and u_2'. Those are simple to solve. From the equation on top we get u_2' in terms of u_1', and we use that expression on the bottom equation,

$$u_2' = -t^3\, u_1' \quad \Rightarrow \quad 2t\, u_1' + t\, u_1' = 3 - \frac{1}{t^2} \quad \Rightarrow \quad u_1' = \frac{1}{t} - \frac{1}{3t^3}.$$

Substitue back the expression for u_1' in the first equation above and we get u_2'. We get,

$$u_1' = \frac{1}{t} - \frac{1}{3t^3}$$
$$u_2' = -t^2 + \frac{1}{3}.$$

We should now integrate these functions to get u_1 and u_2 and then get the particular solution $\tilde{y}_p = u_1 y_1 + u_2 y_2$. We do not repeat these calculations, since they are done Example 2.3.8.

2.3.4. Exercises.

2.3.1.- .

2.3.2.- .

2.4. Forced Oscillations

In this section we study the movement of a mass-spring system having a natural frequency ω_0 and oscillating in a medium with or without friction while moving under the effects of an external force, which itself oscillates in time with a driving frequency ω. We begin studying the case without friction and we show two main results. First, if the driving frequency ω approaches the natural frequency ω_0, the solution develops a pulsating modulation in the amplitude, called beats. Second, if the driving frequency is equal to the natural frequency, $\omega = \omega_0$, the solution amplitude diverges in time, and this behavior is called resonance. Then we add friction to this system and we see that the divergence in time of the resonant solution is tamed by the friction effects.

2.4.1. Description of the Problem. Consider a mass-spring system as in § 2.1, with spring constant $k > 0$ and an attached object of mass $m > 0$ oscillating vertically in a medium with friction coefficient $d \geqslant 0$. We introduce a vertical coordinate y, positive downwards, with $y = 0$ at the rest position of the mass-spring system, as in Fig. 6.

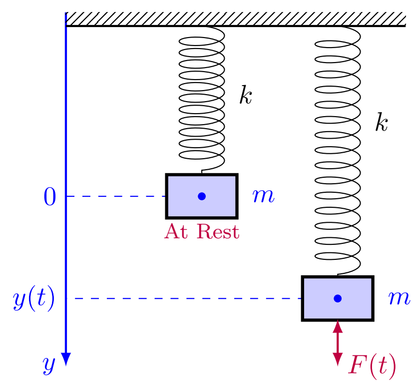


FIGURE 6. Mass-Spring System with an external force $F(t)$.

The intrinsic forces in this system are the weight of the mass-spring, the force of the spring, and the friction force. The weight force is given by

$$f_g = mg,$$

which is positive since it points downwards and g is the acceleration of gravity at the Earth surface, $g = 9.81\mathrm{m/s}^2$. The force of the spring can be split in two parts, $f_0 + f_s$, where

$$f_0 = -mg, \qquad f_s = -k\,y.$$

We see that f_0 balances the weight of the mass-spring while f_s is the extra force made by the spring when it is extended by an amount y away from equilibrium. This force is opposite to the spring displacement, called Hooke's law. The friction force is given by

$$f_d = -d\,y',$$

that is, it points in the opposite direction from the object's velocity. The equation of the mass-spring system is given by Newton's equation, which in absence of any other forces is given by

$$m\,y'' = f_g + f_0 + f_s + f_d,$$

and this equation can be simplified into the form used in § 2.1,

$$m\,y'' + d\,y' + k\,y = 0.$$

In this section we add an external force to this system, $F(t)$, then Newton's equation is

$$m\,y'' + d\,y' + k\,y = F(t).$$

We will study the case when the external force is an oscillatory force,

$$F(t) = F_0 \cos(\omega t),$$

where we call ω the *driving frequency* of the force. The final form of Newton's equation in this section is

$$m\,y'' + d\,y' + k\,y = F_0 \cos(\omega t). \tag{2.4.1}$$

There is another frequency that will play an important role in this section called the natural frequency of the mass-spring system. The *natural frequency* of the system is the frequency that the spring would oscillate if there are no external forces and no friction. Newton's equation in this case is

$$m\,y'' + k\,y = 0.$$

From our work on § 2.2 it is simple to see that the general solution of this system is

$$y_h(t) = c_1 \cos(\omega_0 t) + c_2 \sin(\omega_0 t),$$

where $\omega_0 = \sqrt{\dfrac{k}{m}}$ is the natural frequency of this mass-spring system.

In this section we study solutions of Eq. (2.4.1) in different cases.

(a) No friction, which means $d = 0$. We study two solutions in this case:
- The *nonresonant solution*, which is the case $\omega \neq \omega_0$, that is, the force driving frequency is different from the natural frequency of the system.
- The *resonant solution*, which is the case $\omega = \omega_0$, that is the force driving frequency is the same as the natural frequency of the system.

(b) Friction, which means $d > 0$. In this case we have solution formulas for both cases, $\omega \neq \omega_0$ and $\omega = \omega_0$. We classify these solutions according to the friction coefficient: Small friction, critical friction, and large friction. Each of these three types of solutions are decomposed into two parts:
- The *transient part of the solution:* This part of the solution is due to the initial conditions and it goes to zero in time because of the dissipation affects from the friction forces.
- The *steady part of the solution:* This part of the solution is due to the always acting force, which as time grows it approaches a nonzero equilibrium with the dissipation effects of the friction forces.

Notice that the solutions without friction do not have a transient part, they only have a steady part, since there is no friction to slow down any part of the solution.

2.4.2. No Friction. Consider the mass-spring system described by Newton's equation (2.4.1). Assume we discard any friction effects, so we set $d = 0$. Newton's equation is then given by

$$m\,y'' + k\,y = F_0 \cos(\omega t).$$

We simplify the equation dividing it by the object mass m,

$$y'' + \omega_0^2\, y = f_0 \cos(\omega t),$$

where we introduced the natural frequency $\omega_0 = \sqrt{k/m}$ and we the rescaled force coefficient $f_0 = F_0/m$. As we mentioned earlier, there are two main types of solutions to this last equation above. Solutions where $\omega \neq \omega_0$, called nonresonant solutions, and solutions where $\omega = \omega_0$, called resonant solutions. We start computing the nonresonant solution for some special initial conditions.

Theorem 2.4.1 (Nonresonant Solution)**.** *The solution $y(t)$ of the initial value problem*

$$y'' + \omega_0^2\, y = f_0 \cos(\omega t), \qquad \omega \neq \omega_0, \qquad y(0) = 0, \quad y'(0) = 0, \tag{2.4.2}$$

called a ***nonresonant*** *solution, is given by*

$$y(t) = \frac{f_0}{(\omega_0^2 - \omega^2)} \big(\cos(\omega t) - \cos(\omega_0 t)\big). \tag{2.4.3}$$

Proof of Theorem 2.4.1: We follow the ideas given in § 2.2 and we first find the general solution of the homogeneous equation

$$y'' + \omega_0^2\, y = 0.$$

We try solutions of the form $y(t) = r^{rt}$, the exponent is solution of the characteristic equation

$$r^2 + \omega_0^2 = 0 \quad \Rightarrow \quad r_\pm = \pm \omega_0\, i.$$

We know from § 2.2 that the general solution of this homogeneous equation is

$$y_h(t) = c_1 \cos(\omega_0 t) + c_2 \sin(\omega_0 t).$$

Now we follow the work we did in § 2.3 and we find a particular solution $y_p(t)$ of the non-homogeneous equation

$$y'' + \omega_0^2\, y = f_0 \cos(\omega t).$$

Since $\omega \neq \omega_0$, then we guess

$$y_p(t) = k \cos(\omega t).$$

This is a good guess, since this $y_p(t)$ is not solution of the homogeneous equation. Then we can put this $y_p(t)$ into the non-homogeneous equation and find the undetermined constant k. Then, we get

$$(-\omega^2 + \omega_0^2)\, k \cos(\omega t) = f_0 \cos(\omega t) \quad \Rightarrow \quad k = \frac{f_0}{\omega_0^2 - \omega^2}.$$

Therefore, a particular solution $y_p(t)$ of the non-homogenous equation is

$$y_p(t) = \frac{f_0}{(\omega_0^2 - \omega^2)} \cos(\omega t).$$

The general solution of the differential equation in Eq. 2.4.2 is

$$y(t) = y_h(t) + y_p(t) \quad \Rightarrow \quad y(t) = c_1 \cos(\omega_0 t) + c_2 \sin(\omega_0 t) + \frac{f_0}{(\omega_0^2 - \omega^2)} \cos(\omega t).$$

The initial condition $y(0) = 0$ implies

$$0 = y(0) = c_1 + 0 + \frac{f_0}{(\omega_0^2 - \omega^2)} \quad \Rightarrow \quad c_1 = -\frac{f_0}{(\omega_0^2 - \omega^2)}$$

The other condition is for $y'(t)$, which is given by

$$y'(t) = -\omega_0 c_1 \sin(\omega_0 t) + \omega_0 c_2 \cos(\omega_0 t) - \omega \frac{f_0}{(\omega_0^2 - \omega^2)} \sin(\omega t).$$

The initial condition $y'(0) = 0$ implies

$$0 = y'(0) = 0 + \omega_0 c_2 + 0 \quad \Rightarrow \quad c_2 = 0,$$

since $\omega_0 \neq 0$. Therefore, we obtained the solution of the initial value problem,

$$y(t) = \frac{f_0}{(\omega_0^2 - \omega^2)} \big(\cos(\omega t) - \cos(\omega_0 t)\big).$$

This establishes the Theorem. □

The nonresonant solution in Eq. (2.4.3) develops a particular behavior called beats in the case that the driving frequency ω is close enough to the natural frequency ω_0. Below is a precise definition of what we mean by close enough.

Definition 2.4.2. *The function $y(t)$ given by Eq. (2.4.3) develops **beats** when the driving frequency ω and the natural frequency ω_0 are close in the sense*

$$\frac{|\omega - \omega_0|}{\omega_0} \lessapprox 0.01. \tag{2.4.4}$$

In Fig. 7 we graph the solution in Eq. (2.4.3) for a value of ω satisfying the beats condition in Eq. (2.4.4). We see that the oscillations in the solution have a lower frequency modulation in the amplitude. This modulation in amplitude produces a pulsation in the solution.

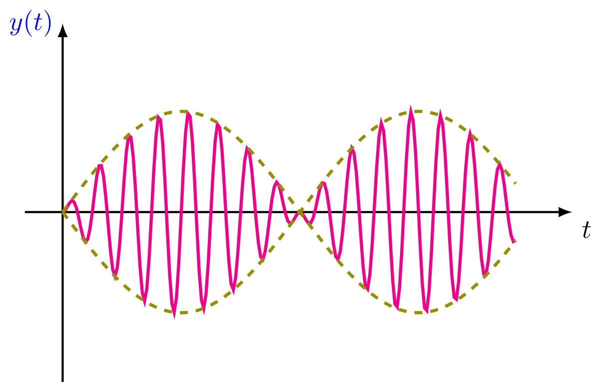


FIGURE 7. The spring oscillation $y(t)$ for driving frequency ω close to the natural frequency ω_0. This graph shows the modulation in amplitude called beats.

This modulation in the amplitude is hard to see in the expression for the solution given by Eq. (2.4.3). However, this amplitude modulation can be seen analytically if we rewrite the nonresonant solution in Eq. (2.4.3), which is a difference of cosines, in terms of a product of sines.

Corollary 2.4.3. *The nonresonant solution in Eq. (2.4.3) can be rewritten as*

$$y(t) = \frac{2f_0}{(\omega_0^2 - \omega^2)} \sin\Big(\frac{(\omega_0 - \omega)}{2}\, t\Big) \sin\Big(\frac{(\omega_0 + \omega)}{2}\, t\Big) \tag{2.4.5}$$

Proof of Corollary 2.4.3: We only need to use the trigonometric identity

$$\cos(\alpha) - \cos(\beta) = -2\sin\Big(\frac{\alpha - \beta}{2}\Big) \sin\Big(\frac{\alpha + \beta}{2}\Big)$$

with $\alpha = \omega t$ and $\beta = \omega_0 t$, which gives us

$$y(t) = \frac{-2f_0}{(\omega_0^2 - \omega^2)} \sin\Big(\frac{(\omega - \omega_0)}{2}\, t\Big) \sin\Big(\frac{(\omega_0 + \omega)}{2}\, t\Big).$$

This establishes the Corollary. □

The nonresonant solution given as in Eq. (2.4.5) can be written as a slow modulated amplitude, $A(t)$, times a fast oscillation,

$$y(t) = A(t)\,\sin\Big(\frac{(\omega_0+\omega)}{2}\,t\Big),$$

where

$$A(t) = \frac{2f_0}{(\omega_0^2-\omega^2)}\,\sin\Big(\frac{(\omega_0-\omega)}{2}\,t\Big).$$

The rapid oscillations given in the solid purple line have the large frequency $(\omega_0+\omega)/2$ while slow modulation in amplitude given by the dashed olive line have a low frequency $(\omega_0-\omega)/2$. The latter is a low frequency oscillation because we are assuming that ω is close to ω_0, which means $\omega_0-\omega$ is close to zero. The name beats may originate in sound waves. When a given sound has an amplitude modulation as in Fig. 7 one hears a pulsation over the original sound, hence the name beats.

2.4.3. Resonant Solution. In Theorem 2.4.1 we assumed that the driving frequency is different from the natural frequency of the mass-spring. Now we study the case when the external force oscillates exactly at the natural frequency of the spring. Notice that the nonresonant solution found in Theorem 2.4.1 is not defined for $\omega=\omega_0$, because this solution takes the form 0/0. We need to redo the calculation to compute the solution specifically for the case $\omega=\omega_0$.

Theorem 2.4.4 (Resonant Solution)**.** *The solution $y(t)$ of the initial value problem*

$$y''+\omega_0^2\,y = f_0\cos(\omega_0 t), \qquad y(0)=0, \quad y'(0)=0, \tag{2.4.6}$$

*called a **resonant** solution, is given by*

$$y(t) = \frac{f_0}{2\omega_0}\,t\,\sin(\omega_0 t). \tag{2.4.7}$$

We see that the resonant solution oscillates with the natural frequency of the mass-spring and with the amplitude of the oscillations increasing linearly in time, forever. In a real mass-spring system this means that the amplitude of the oscillations increases until the spring deforms or breaks.

Proof of Theorem 2.4.4: In the proof of Theorem 2.4.1 we found the general solution of the homogeneous equation

$$y''+\omega_0^2\,y=0,$$

which are given by

$$y_h(t) = c_1\cos(\omega_0 t) + c_2\sin(\omega_0 t).$$

Now we follow again the work we did in § 2.3 and we find a particular solution $y_p(t)$ of the non-homogeneous equation

$$y''+\omega_0^2\,y = f_0\cos(\omega_0 t). \tag{2.4.8}$$

Since the driving frequency is ω_0 we cannot guess

$$y_p(t) = k\cos(\omega_0 t)$$

because this $y_p(t)$ is solution of the homogeneous equation. Instead, we guess

$$y_p(t) = k_1\,t\cos(\omega_0 t) + k_2\,t\,\sin(\omega_0 t).$$

This is a good guess, since this last $y_p(t)$ is not solution of the homogeneous equation and we will be able to find the constants k_1, k_2. We find these constants in the usual way, called the undetermined coefficients method. We put this $y_p(t)$ into the non-homogeneous equation and find the undetermined constants k_1 and k_2. First we need to compute $y_p'(t)$ and then $y_p''(t)$,

$$y_p'(t) = k_1 \cos(\omega_0 t) - k_1 \omega_0 t \sin(\omega_0 t) + k_2 \sin(\omega_0 t) + k_2 \omega_0 t \cos(\omega_0 t),$$

$$y_p''(t) = -2k_1 \omega_0 \sin(\omega_0 t) - k_1 \omega_0^2 t \cos(\omega_0 t) + 2k_2 \omega_0 \cos(\omega_0 t) - k_2 \omega_0^2 t \sin(\omega_0 t).$$

We use the expressions for y_p and y_p'' in the non-homogeneous equation in (2.4.8),

$$-2k_1 \omega_0 \sin(\omega_0 t) - k_1 \omega_0^2 t \cos(\omega_0 t) + 2k_2 \omega_0 \cos(\omega_0 t) - k_2 \omega_0^2 t \sin(\omega_0 t)$$
$$+\omega_0^2 \big(k_1\, t \cos(\omega_0 t) + k_2\, t\, \sin(\omega_0 t)\big) = f_0 \cos(\omega_0 t).$$

After a few simplifications we get

$$-2k_1 \omega_0 \sin(\omega_0 t) + 2k_2 \omega_0 \cos(\omega_0 t) = f_0 \cos(\omega t) \quad \Rightarrow \quad k_1 = 0, \qquad k_2 = \frac{f_0}{2\omega_0}.$$

Therefore, a particular solution $y_p(t)$ of the non-homogenous equation is

$$y_p(t) = \frac{f_0}{2\omega_0}\, t\, \sin(\omega_0 t).$$

Notice that this particular solution already satisfies the initial conditions

$$y_p(0) = 0, \qquad y_p'(0) = 0.$$

Therefore, this $y_p(t)$ is the unique solution of the initial value problem. This establishes the Theorem. □

The resonant solution in Eq. (2.4.7) can be written as

$$y_p(t) = A(t)\, \sin(\omega_0 t), \qquad A(t) = \frac{f_0}{2\omega_0}\, t.$$

We see that the amplitude $A(t)$ of this solution increases with time, t, and diverges in the limit $t \to \infty$. This mathematical increase in the amplitude means that the real life mass-spring increases its amplitude until the spring no longer acts as a spring, because it deforms or it breaks. In the Fig. 8 we graph the resonant solution in blue and a beats solution in red. The beats solution has the same parameters as the resonant solution plus a driving frequency close to the natural frequency of the mass-spring.

The graphs in Fig. 8 suggest how the resonant solution can be constructed as a limit of the beats solution when the driving frequency ω approaches the natural frequency ω_0. A more precise statement is given in the following result.

Theorem 2.4.5. *Consider the nonresonant and resonant solutions*

$$y_{NR}(t) = \frac{2f_0}{(\omega_0^2 - \omega^2)}\, \sin\Big(\frac{(\omega_0 - \omega)}{2}\, t\Big)\, \sin\Big(\frac{(\omega_0 + \omega)}{2}\, t\Big), \qquad y_R(t) = \frac{f_0}{2\omega_0}\, t\, \sin(\omega_0 t).$$

Then, for every fixed value of time, t, holds

$$\lim_{\omega \to \omega_0} y_{NR}(t) = y_R(t).$$

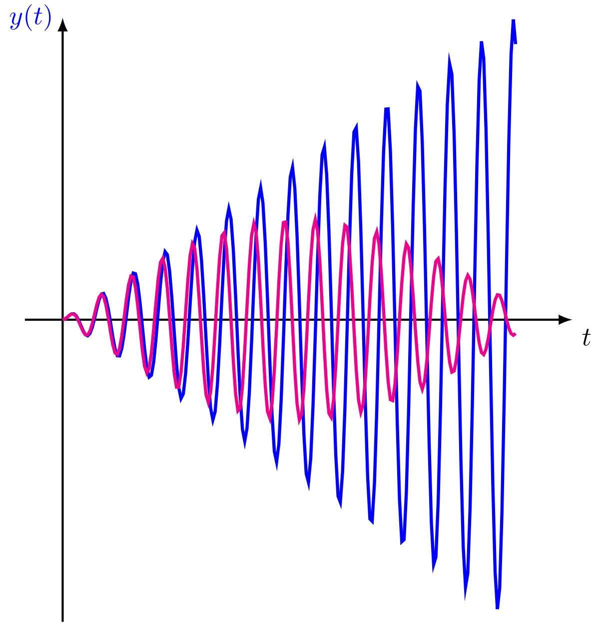


FIGURE 8. The resonant solution in blue and a beats solution in red. This graph shows how the amplitude of the oscillations in the resonant solution increases in time without bound.

Proof of Theorem 2.4.5: We only need to rewrite the nonresonant solution as follows

$$\begin{aligned} y_{NR}(t) &= \frac{2f_0}{(\omega_0-\omega)(\omega_0+\omega)}\,\sin\Big(\frac{(\omega_0-\omega)}{2}\,t\Big)\,\sin\Big(\frac{(\omega_0+\omega)}{2}\,t\Big) \\ &= \frac{2f_0}{(\omega_0+\omega)}\,\Big(\frac{t}{2}\Big)\,\Bigg(\frac{\sin\big(\frac{(\omega_0-\omega)}{2}\,t\big)}{\frac{(\omega_0-\omega)}{2}\,t}\Bigg)\,\sin\Big(\frac{(\omega_0+\omega)}{2}\,t\Big) \\ &\qquad\downarrow \qquad\qquad\qquad\quad \downarrow \qquad\qquad\qquad \downarrow \\ &\quad\ \frac{f_0}{\omega_0}\quad \Big(\frac{t}{2}\Big)\qquad\quad 1 \qquad\qquad\quad \sin(\omega_0\,t) \qquad = y_R(t), \end{aligned}$$

where we used that for a fixed t we have

$$\frac{\sin\big(\frac{(\omega_0-\omega)}{2}\,t\big)}{\frac{(\omega_0-\omega)}{2}\,t} \to 1 \quad \text{and} \quad (\omega_0+\omega)\to 2\omega_0 \quad \text{as} \quad \omega\to\omega_0.$$

Therefore, for a fixed value of time t we have shown that

$$y_{NR}(t) \to y_R(t) \quad \text{as} \quad \omega\to\omega_0.$$

This establishes the Theorem. □

2.4.4. Damped Forced Oscillations. Once again, consider the mass-spring system described by Newton's equation (2.4.1),

$$m\,y'' + d\,y' + k\,y = F_0\cos(\omega t).$$

We simplify the equation dividing by the object mass m,

$$y'' + 2\alpha\,y' + \omega_0^2\,y = f_0\cos(\omega t),$$

where we introduced the natural frequency ω_0, the rescaled friction coefficient α, and the rescaled force coefficient f_0, as follows,

$$\omega_0 = \sqrt{\frac{k}{m}}, \qquad \alpha = \frac{d}{2m}, \qquad f_0 = \frac{F_0}{m}.$$

Now we show the general solution of this equation, which depend on the values of the friction constant α.

Theorem 2.4.6 (Damped Forced Oscillations)**.** *The general solution of the differential equation*

$$y'' + 2\alpha\, y' + \omega_0^2\, y = f_0 \cos(\omega t),$$

is given by the formula

$$y(t) = y_T(t) + y_S(t),$$

*where the **transient** part of the solution is given by either of the three expression below, depending on the value of α,*

$$y_T(t) = e^{-\alpha t}\big(c_1\cos(\beta t) + c_2 \sin(\beta t)\big), \qquad \beta = \sqrt{\omega_0^2 - \alpha^2}, \quad 0 < \alpha < \omega_0, \quad \textit{(small friction)}$$
$$y_T(t) = e^{-\alpha t}\big(c_1 + c_2\, t\big), \qquad \alpha = \omega_0, \quad \textit{(critical friction)}$$
$$y_T(t) = e^{-\alpha t}\big(c_1\cosh(\gamma t) + c_2 \sinh(\gamma t)\big), \quad \gamma = \sqrt{\alpha^2 - \omega_0^2}, \qquad \alpha > \omega_0, \quad \textit{(large friction)}$$

*were c_1, c_2 are arbitrary constants, and the **steady** part of the solution for any value of the friction parameter $\alpha \geqslant 0$ is given by*

$$y_S(t) = \frac{f_0}{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2}\big((\omega_0^2 - \omega^2)\cos(\omega t) + 2\alpha\omega \sin(\omega t)\big).$$

We see that the motion of the mass-spring system can be split into two main parts: a transient part made by the initial conditions of the system that goes to zero as time grows, and a steady part made by the external force that remains finite and balances the effects of friction.

Proof of Theorem 2.4.6: In order to find the general solution of the equation

$$y'' + 2\alpha\, y' + \omega_0^2\, y = f_0 \cos(\omega t)$$

we start finding the general solution of the homogeneous equation

$$y'' + 2\alpha\, y' + \omega_0^2\, y = 0.$$

We try exponentials, $y(t) = e^{rt}$, which gives us the characteristic equation for the exponent,

$$r^2 + 2\alpha\, r + \omega_0^2 = 0 \quad \Rightarrow \quad r_\pm = -\alpha \pm \sqrt{\alpha^2 - \omega_0^2}.$$

The three cases mentioned in the theorem, small friction, critical friction, and large friction arise from these values of $r_\pm$, respectively,

$$r_\pm = -\alpha \pm \beta i \qquad \beta = \sqrt{\omega_0^2 - \alpha^2}, \qquad 0 < \alpha < \omega_0, \qquad \text{(small friction)}$$
$$r_\pm = -\alpha, \qquad \alpha = \omega_0, \qquad \text{(critical friction)}$$
$$r_\pm = -\alpha \pm \gamma, \qquad \gamma = \sqrt{\alpha^2 - \omega_0^2}, \qquad \alpha > \omega_0, \qquad \text{(large friction).}$$

Therefore the corresponding general solutions of the homogeneous equation, called here $y_T(t)$, are given by

$$y_T(t) = e^{-\alpha t}\big(c_1\cos(\beta t) + c_2\sin(\beta t)\big),$$
$$y_T(t) = e^{-\alpha t}\big(c_1 + c_2\, t\big),$$
$$y_T(t) = e^{-\alpha t}\big(c_1\cosh(\gamma t) + c_2\sinh(\gamma t)\big).$$

Now we only need to find a particular solution of the nonhomogeneous equation

$$y'' + 2\alpha\, y' + \omega_0^2\, y = f_0\cos(\omega t).$$

We use the undetermined coefficients method and we guess a particular solution, $y_S(t)$,

$$y_S(t) = k_1\cos(\omega t) + k_2\sin(\omega t),$$

which is a good guess for any value of $\alpha > 0$. The first and second derivatives are

$$y_S'(t) = \omega\big(-k_1\sin(\omega t) + k_2\cos(\omega t)\big)$$
$$y_S''(t) = \omega^2\big(-k_1\cos(\omega t) - k_2\sin(\omega t)\big).$$

If we put this function and its derivatives in the nonhomogeneous equation we get

$$\omega^2\big(-k_1\cos(\omega t) - k_2\sin(\omega t)\big) + 2\alpha\omega\big(-k_1\sin(\omega t) + k_2\cos(\omega t)\big)$$
$$+\omega_0^2\big(k_1\cos(\omega t) + k_2\sin(\omega t)\big) = f_0\cos(\omega t).$$

If we reorder a few terms we arrive to the equation

$$\big(k_1(\omega_0^2 - \omega^2) + 2\alpha\omega k_2 - f_0\big)\cos(\omega t) + \big(k_2(\omega_0^2 - \omega^2) - 2\alpha\omega k_1\big)\sin(\omega t) = 0.$$

Since this last equation must hold for every value of the variable t, this means

$$k_1(\omega_0^2 - \omega^2) + 2\alpha\omega k_2 - f_0 = 0,$$
$$k_2(\omega_0^2 - \omega^2) - 2\alpha\omega k_1 = 0.$$

We solve these equations for the constants k_1 and k_2,

$$k_1 = \frac{(\omega_0^2 - \omega^2)\, f_0}{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2},$$
$$k_2 = \frac{2\alpha\omega\, f_0}{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2}.$$

Therefore, we have found a particular solution

$$y_S(t) = \frac{f_0}{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2}\big((\omega_0^2 - \omega^2)\cos(\omega t) + 2\alpha\omega\sin(\omega t)\big).$$

This establishes the Theorem. □

We can see in Theorem 2.4.6 that the transient part of the solution, $y_T(t)$, for any value of the friction constant $\alpha > 0$, has a factor $e^{-\alpha t}$. This factor is the reason we call this part of the solution a transient solution. This factor makes $y_T(t) \to 0$ as $t \to \infty$. Even in the large friction case, where we have $\cosh(\gamma t)$ and $\sinh(\gamma t)$ in the transient solution, the exponential factor $e^{-\alpha t}$ drives this part of the solution to zero as t grows, since by definition we have $\gamma = \sqrt{\alpha^2 - \omega_0^2} < \alpha$.

We also see in Theorem 2.4.6 that the steady part of the solution $y_S(t)$, which has no exponential decay, contains oscillatory functions with the force driving frequency ω. This solution is well defined even in the resonance case, $\omega = \omega_0$, and it is given by

$$y_S(t) = \frac{f_0}{2\alpha\omega_0}\sin(\omega_0 t). \tag{2.4.9}$$

The friction coefficient plays an important role in this situation, because this resonant solution above diverges when the friction coefficient approaches zero.

Sometimes it is convenient to write the steady part of the solution in terms of amplitude and phase shift.

Corollary 2.4.7. *The function*

$$y_S(t) = \frac{f_0}{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2} \big((\omega_0^2 - \omega^2)\cos(\omega t) + 2\alpha\omega \sin(\omega t) \big) \tag{2.4.10}$$

can be written as

$$y_S(t) = A\cos(\omega t - \phi),$$

where the amplitude A *and phase shift* ϕ *are given by*

$$A = \frac{f_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2}},$$

$$\tan(\phi) = \frac{2\alpha\omega}{\omega_0^2 - \omega^2} \quad \Rightarrow \quad \begin{cases} \phi = \arctan\Big(\dfrac{2\alpha\omega}{\omega_0^2 - \omega^2}\Big), & \omega_0 > \omega, \\ \phi = \pi + \arctan\Big(\dfrac{2\alpha\omega}{\omega_0^2 - \omega^2}\Big), & \omega_0 < \omega. \end{cases}$$

Proof of Corollary 2.4.7: The trigonometric identity

$$\cos(a - b) = \cos(a)\cos(b) + \sin(a)\sin(b)$$

implies that

$$y_S(t) = A\cos(\phi)\cos(\omega t) + A\sin(\phi)\sin(\omega t).$$

Comparing the expression above with Eq. (2.4.10) we get

$$A\cos(\phi) = \frac{(\omega_0^2 - \omega^2)\, f_0}{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2}, \qquad A\sin(\phi) = \frac{2\alpha\omega\, f_0}{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2}. \tag{2.4.11}$$

If we add up the squares of these two equations we get a formula for A, because

$$A^2 = \frac{\big((\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2\big)\, f_0^2}{\big((\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2\big)^2} \quad \Rightarrow \quad A = \frac{f_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2}},$$

where we used that $\cos^2(\phi) + \sin^2(\phi) = 1$. The formula for ϕ comes from

$$\tan(\phi) = \frac{A\sin(\phi)}{A\cos(\phi)} = \frac{2\alpha\omega}{\omega_0^2 - \omega^2}.$$

Notice that for $\omega < \omega_0$ the formulas in Eq. (2.4.11) say that $\cos(\phi) > 0$ and $\sin(\phi) > 0$, which means that

$$\phi = \arctan\Big(\frac{2\alpha\omega}{\omega_0^2 - \omega^2}\Big).$$

But in the case $\omega > \omega_0$ the formulas in Eq. (2.4.11) say that $\cos(\phi) < 0$ and $\sin(\phi) > 0$, and since every phase shift is defined in the interval $\phi \in (-\pi, \pi]$, all this means that

$$\phi = \pi + \arctan\Big(\frac{2\alpha\omega}{\omega_0^2 - \omega^2}\Big).$$

This establishes the Corollary. □

In Fig. 9 we show the graph of the function $A(\omega)$, the amplitude of the steady part of the solution as function of the driving frequency. We can see that at resonance, when $\omega = \omega_0$, the amplitude $A(\omega)$ is well-defined and it achieves its maximum value. We can use the expression of the solution $y_S(t)$ at resonance given in Eq. (2.4.9) to see that this maximum value, $A(\omega_0)$, satisfies that $A(\omega_0) \to \infty$ as the friction coefficient $\alpha \to 0$.

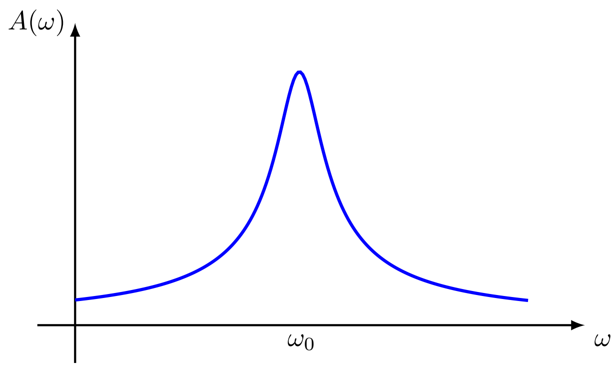


FIGURE 9. The amplitude given in the Corollary 2.4.7 as function of the driving frequency, $A(\omega)$.

In the last part of this section we study the behavior of the damped forced oscillation solutions found in Theorem 2.4.6 in the case that:

- we are at resonance, that is the driving frequency is equal to the natural frequency of the mass-spring, $\omega = \omega_0$;
- the friction coefficient α approaches zero.

Since we are interested in $\alpha \to 0$, we only need the solution in Theorem 2.4.6 for small friction, which is given by

$$y(t) = e^{-\alpha t}\big(c_1 \cos(\beta t) + c_2 \sin(\beta t)\big) + \frac{f_0}{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2}\big((\omega_0^2 - \omega^2)\cos(\omega t) + 2\alpha\omega \sin(\omega t)\big),$$

where $\beta = \sqrt{\omega_0^2 - \alpha^2}$. We choose the constants c_1, c_2 so that this solution satisfies the homogeneous initial conditions $y(0) = 0$ and $y'(0) = 0$, which are the same initial conditions we had in the case without friction. It is not difficult to verify that the solution satisfying these homogeneous initial conditions is

$$y(t) = \frac{f_0}{(\omega_0^2 - \omega^2)^2 + (2\alpha\omega)^2}\Big(e^{-\alpha t}\big(-(\omega_0^2 - \omega^2)\cos(\beta t) - \frac{\alpha}{\beta}(\omega_0^2 + \omega^2)\sin(\beta t)\big) + \big((\omega_0^2 - \omega^2)\cos(\omega t) + 2\alpha\omega \sin(\omega t)\big)\Big).$$

We are interested to see what happens at resonance, $\omega = \omega_0$, and the solution above reduces to the damped resonant solution

$$y_{DR}(t) = \frac{f_0}{2\alpha}\Big(-e^{-\alpha t}\frac{\sin(\beta t)}{\beta} + \frac{\sin(\omega_0 t)}{\omega_0}\Big).$$

We see the tin the case of no friction, $\alpha = 0$, the solution above is not defined, since for $\alpha = 0$ we have $\beta = \omega_0$, which means we get an expression of the form

$$y_{DR}(t)\Big|_{\alpha=0} = \frac{0}{0}.$$

It turns out that this damped resonant solution $y_{DR}(t)$ approaches the resonant solution without friction in Eq.(2.4.7) in the limit $\alpha \to 0$.

Theorem 2.4.8. *Consider the damped resonant and undamped resonant solutions*

$$y_{DR}(t) = \frac{f_0}{2\alpha}\Big(-e^{-\alpha t}\frac{\sin(\beta t)}{\beta} + \frac{\sin(\omega_0 t)}{\omega_0}\Big), \qquad y_R(t) = \frac{f_0}{2\omega_0}\, t\, \sin(\omega_0 t).$$

Then, for every fixed value of time, t, holds

$$\lim_{\alpha\to 0} y_{DR}(t) = y_R(t).$$

Proof of Theorem 2.4.8: In this calculation we need the expansions,

$$\sqrt{1-x^2} = 1 + \frac{x^2}{2} - \frac{x^4}{8} + \mathcal{O}(x^6), \qquad e^x = 1 + x + \frac{x^2}{2} + \mathcal{O}(x^3),$$

$$\frac{1}{\sqrt{1-x^2}} = 1 + \frac{x^2}{2} + \frac{x^4}{8} + \mathcal{O}(x^6), \qquad \sin(x) = x - \frac{x^3}{6} + \mathcal{O}(x^5).$$

Using these expansions we get

$$\beta = \omega_0\sqrt{1-\frac{\alpha^2}{\omega_0^2}} = \omega_0\Big(1 - \frac{\alpha^2}{2\omega_0^2} + \mathcal{O}(\alpha^4)\Big),$$

$$\frac{1}{\beta} = \frac{1}{\omega_0}\frac{1}{\sqrt{1-\frac{\alpha^2}{\omega_0^2}}} = \frac{1}{\omega_0}\Big(1 + \frac{\alpha^2}{2\omega_0^2} + \mathcal{O}(\alpha^4)\Big),$$

$$e^{-\alpha t} = 1 - \alpha t + \frac{\alpha^2 t^2}{2} + \mathcal{O}(\alpha^3).$$

If we use these expansions in $y_{DR}(t)$ we get

$$y_{DR}(t) = \frac{f_0}{2\alpha}\Big(-\big(1-\alpha t+\frac{\alpha^2t^2}{2}\big)\sin\big(\omega_0 t\,(1-\frac{\alpha^2}{2\omega_0^2})\big)\frac{1}{\omega_0}\big(1+\frac{\alpha^2}{2\omega_0^2}\big)+\frac{\sin(\omega_0 t)}{\omega_0}\Big)+\mathcal{O}(\alpha^3). \quad (2.4.12)$$

Now we need the formula

$$\sin(a-b) = \sin(a)\cos(b) - \sin(b)\cos(a),$$

which implies

$$\begin{aligned}\sin\big(\omega_0 t\,(1-\frac{\alpha^2}{2\omega_0^2})\big) &= \sin(\omega_0 t)\cos\big(\frac{\alpha^2 t}{2\omega_0}\big) - \sin\big(\frac{\alpha^2 t}{2\omega_0}\big)\cos(\omega_0 t)\\ &= \sin(\omega_0 t) - \frac{\alpha^2 t}{2\omega_0}\cos(\omega_0 t) + \mathcal{O}(\alpha^4), \qquad (2.4.13)\end{aligned}$$

where in the last step we used the expansions

$$\sin(x) = x - \frac{x^3}{6} + \mathcal{O}(x^5), \qquad \cos(x) = 1 - \frac{x^2}{4} + \mathcal{O}(x^4).$$

If we use Eq. (2.4.13) in Eq. (2.4.12) we get

$$y_{DR}(t) = \frac{f_0}{2\omega_0\alpha}\Big(\big(-1+\alpha t - \frac{\alpha^2t^2}{2}\big)\big(\sin(\omega_0 t) - \frac{\alpha^2 t}{2\omega_0}\cos(\omega_0 t)\big)\big(1+\frac{\alpha^2}{2\omega_0^2}\big) + \sin(\omega_0 t)\Big) + \mathcal{O}(\alpha^2).$$

Expanding the terms in the first product above we get

$$y_{DR}(t) = \frac{f_0}{2\omega_0\alpha}\Big(\big(-\sin(\omega_0 t) + \frac{\alpha^2 t}{2\omega_0}\cos(\omega_0 t) + \alpha t\sin(\omega_0 t) - \frac{\alpha^2t^2}{2}\sin(\omega_0 t)\big)\big(1+\frac{\alpha^2}{2\omega_0^2}\big) + \quad \sin(\omega_0 t)\Big) + \mathcal{O}(\alpha^2).$$

Expanding the terms in the last product above we get

$$\begin{aligned} y_{DR}(t) = \frac{f_0}{2\omega_0\alpha}\Big(-\cancel{\sin(\omega_0 t)} + \frac{\alpha^2 t}{2\omega_0}\cos(\omega_0 t) + \alpha t\sin(\omega_0 t) - \frac{\alpha^2 t^2}{2}\sin(\omega_0 t) - \frac{\alpha^2}{2\omega_0^2}\sin(\omega_0 t) \\ + \cancel{\sin(\omega_0 t)}\Big) + \quad \mathcal{O}(\alpha^2), \\ = \frac{f_0}{2\omega_0\alpha}\Big(\alpha t\sin(\omega_0 t) + \alpha^2\big(\frac{t}{2\omega_0}\cos(\omega_0 t) - \frac{1}{2\omega_0^2}\sin(\omega_0 t)\,(1+\omega_0^2 t^2)\big)\Big) + \mathcal{O}(\alpha^2). \end{aligned}$$

We now cancel the factors α and we get our final expression

$$y_{DR}(t) = \frac{f_0}{2\omega_0}\,t\sin(\omega_0 t) + \alpha\,\frac{f_0}{4\omega_0^3}\big(\omega_0 t\cos(\omega_0 t) - (1+\omega_0^2 t^2)\sin(\omega_0 t)\big) + \mathcal{O}(\alpha^2).$$

From this last expression is simple to see that

$$\lim_{\alpha\to 0} y_{DR}(t) = \frac{f_0}{2\omega_0}\,t\sin(\omega_0 t) = y_R(t).$$

This establishes the Theorem. □

2.4.5. Exercises.

2.4.1.- .

2.4.2.- .

CHAPTER 3

The Laplace Transform Method

The Laplace Transform is a transformation, meaning that it changes a function into a new function. Actually, it is a linear transformation, because it converts a linear combination of functions into a linear combination of the transformed functions. Even more interesting, the Laplace Transform converts derivatives into multiplications. These two properties make the Laplace Transform very useful to solve linear differential equations with constant coefficients. The Laplace Transform converts such differential equation for an unknown function into an algebraic equation for the transformed function. Usually it is easy to solve the algebraic equation for the transformed function. Then one converts the transformed function back into the original function. This function is the solution of the differential equation.

Solving a differential equation using a Laplace Transform is radically different from all the methods we have used so far. This method, as we will use it here, is relatively new. The Laplace Transform we define here was first used in 1910, but its use grew rapidly after 1920, specially to solve differential equations. Transformations like the Laplace Transform were known much earlier. Pierre Simon de Laplace used a similar transformation in his studies of probability theory, published in 1812, but analogous transformations were used even earlier by Euler around 1737.

3.1. Introduction to the Laplace Transform

The Laplace transform is an integral transformation that takes a function of a real variable, say t, multiplies that function by an exponential e^{-st}, where s is a real or complex variable, and integrates this product on $t \in [0, \infty)$. This transformation is an example of a broad type of transformations called integral transformations, which include the Fourier transform, the Laplace-Carson transform, the Mellin transform, and many others.

Transformations along these lines were first introduced by Leonhard Euler in 1737. Later on, in 1769, Euler used one of these transformations to solve second order linear differential equations with constant coefficients. Something along these lines is what we are going to do in this chapter. Joseph-Louis Lagrange used similar transformations in 1759 to solve the wave equation—a partial differential equation describing the propagation of waves in a medium, such as pressure waves on air which is the origin of sounds. Pierre-Simon Laplace (1749-1827) started working on integral transformations in 1779. Laplace returned to this subject during the period of 1782-1785 when he laid the main groundwork on integral transforms together with some applications to solve differential equations. That's why the integral transform we study in this section is named after Laplace. Many others worked in this subject, we only mention here Gustav Doetsch (1892-1977) who developed the modern version of the Laplace transform in 1937 and popularized its use on problems from physics and engineering.

3.1.1. Overview of the Method. Suppose we want to solve Newton's equation of motion for the position function $y(t)$ of a particle subject the second order linear differential equation

$$y'' + a_1\, y' + a_0\, y = f(t)$$

with constant coefficients a_1, a_0, and an impulsive force $f(t)$, and initial conditions

$$y(0) = y_0, \qquad y'(0) = y_1.$$

Impulsive forces are zero for all times except at a single point t_0 when they are infinite in a particular way—they transfer a finite amount of momentum (mass times velocity) to the particle. An example of an impulsive force is the force done by a hammer hitting a pendulum. The particle is the object hanging in the pendulum, and the force by the hammer happens at a single point in time, the intensity of the force at that time is very large, but the amount of momentum transferred to the system is finite. We will study impulsive forces in Section 3.4.

One way to describe impulsive forces mathematically is by integration on a finite time interval, where we can use their property that the amount of momentum transferred by the force is finite. Since integration in time will be part of the calculations used to solve the equation above, we may use integrations by parts in the terms that have derivatives of $y(t)$. So, to solve the initial value problem above we first multiply the equation by a function $\mu(t)$,

$$\mu(t)\,(y'' + a_1\, y' + a_0\, y) = \mu(t)\, f(t)$$

then we integrate in time from the initial time $t = 0$ to infinity,

$$\int_0^\infty \mu(t)\,(y'' + a_1\, y' + a_0\, y)\, dt = \int_0^\infty \mu(t)\, f(t)\, dt,$$

and now we move the derivatives from $y(t)$ to $\mu(t)$ using integration by parts. This idea could work if the function $\mu(t)$ has simple derivatives and helps the integral to converge in the interval $[0, \infty)$. Different choices of the function $\mu(t)$ define different transforms. The Laplace transform is the case when we choose the function $\mu(t)$ as

$$\mu(t) = e^{-st},$$

where s is any constant, real or complex. When s is real and positive or complex with positive real part, this exponential function will help make the integrals convergent. Also, this exponential function is simple to derivate,

$$\frac{d}{dt}(e^{-st}) = -s\, e^{-st}.$$

If we go back to our initial value problem above we have,

$$\int_0^\infty e^{-st}\,(y'' + a_1\, y' + a_0\, y)\, dt = \int_0^\infty e^{-st}\, f(t)\, dt.$$

For simplicity, let us consider the case here of homogeneous initial conditions,

$$y(0) = 0, \qquad y'(0) = 0,$$

and let's assume the solution $y(t)$ grows slower than an exponential as $t \to \infty$. In this case, we can integrate by parts the equation above and we get

$$\int_0^\infty e^{-st}\,(s^2\, y(t) + s\, a_1\, y(t) + a_0\, y)\, dt = \int_0^\infty e^{-st}\, f(t)\, dt,$$

where we used that the coefficients a_1, a_0 are constants. Equivalently,

$$(s^2 + a_1\, s + a_0) \int_0^\infty e^{-st}\, y(t)\, dt = \int_0^\infty e^{-st}\, f(t)\, dt.$$

We see that the integration by parts has changed *derivatives into multiplications* by s. So, we have transformed the *differential* equation for $y(t)$ into an *algebraic* equation for the transformed function $Y(s)$, where

$$Y(s) = \int_0^\infty e^{-st}\, y(t)\, dt.$$

We will call $Y(s)$ the Laplace transform of $y(t)$ and also use the notation

$$\mathcal{L}[y(t)] = Y(s).$$

The idea above will allow us to find the transformed function $Y(s)$ and then invert the transformation and find the function $y(t)$ solution of the original initial value problem. We can summarize these steps as follows,

$$\mathcal{L}\begin{bmatrix}\text{differential}\\ \text{eq. for } y.\end{bmatrix} \xrightarrow{(1)} \begin{bmatrix}\text{Algebraic}\\ \text{eq. for } \mathcal{L}[y].\end{bmatrix} \xrightarrow{(2)} \begin{bmatrix}\text{Solve the}\\ \text{algebraic}\\ \text{eq. for } \mathcal{L}[y].\end{bmatrix} \xrightarrow{(3)} \begin{bmatrix}\text{Transform back}\\ \text{to obtain } y.\\ \text{(Use the table.)}\end{bmatrix}$$

The last step above, transforming back from $Y(s)$ to $y(t)$, will be done using a Laplace transform table that we will construct by hand. For example, we will show later on that the Laplace transform changes

$$f(t) = \sin(at) \quad \text{into} \quad F(s) = \frac{a}{s^2 + a^2},$$

which will be one entry in our Laplace transform table.

As we see above, this idea will work on differential equations of any order as long as the equation is *linear with constant coefficients*, so we can carry out the integrations by parts. This concludes our brief overview of the ideas that led to the Laplace transform and how we use it to solve differential equations. In the next subsection we formally introduce the Laplace transform, we compute the Laplace transform of a few functions, then we show the main properties of this transformation, and finally we solve a first order initial value problem using the Laplace transform method.

3.1.2. The Laplace Transform. The Laplace transform is a transformation, meaning that it converts a function into a new function. We have seen transformations earlier in these notes. In Chapter 2 we used the transformation

$$L[y(t)] = y''(t) + a_1\, y'(t) + a_0\, y(t),$$

so that a second order linear differential equation with source f could be written as $L[y] = f$. There are simpler transformations, for example the differentiation operation itself,

$$D[f(t)] = f'(t).$$

Not all transformations involve differentiation. There are integral transformations, for example integration itself,

$$I[f(t)] = \int_0^x f(t)\, dt.$$

Of particular importance in many applications are integral transformations of the form

$$T[f(t)] = \int_a^b K(s,t)\, f(t)\, dt,$$

where K is a fixed function of two variables, called the *kernel* of the transformation, and a, b are real numbers or $\pm\infty$. The Laplace transform is a transfomation of this type, where the kernel is $K(s,t) = e^{-st}$, the constant $a = 0$, and $b = \infty$.

Definition 3.1.1. *The* ***Laplace transform*** *of a function f defined on $D_f = (0, \infty)$ is*

$$F(s) = \int_0^\infty e^{-st} f(t)\, dt, \tag{3.1.1}$$

defined for all $s \in D_F \subset \mathbb{R}$ where the integral converges.

In these note we use an alternative notation for the Laplace transform that emphasizes that the Laplace transform is a transformation: $\mathcal{L}[f] = F$, that is

$$\mathcal{L}[\quad] = \int_0^\infty e^{-st}\, (\quad)\, dt.$$

So, the Laplace transform will be denoted as either $\mathcal{L}[f]$ or F, depending whether we want to emphasize the transformation itself or the result of the transformation. We will also use the notation $\mathcal{L}[f(t)]$, or $\mathcal{L}[f](s)$, or $\mathcal{L}[f(t)](s)$, whenever the independent variables t and s are relevant in any particular context.

The Laplace transform is an improper integral—an integral on an unbounded domain. Improper integrals are defined as a limit of definite integrals,

$$\int_{t_0}^\infty g(t)\, dt = \lim_{N\to\infty} \int_{t_0}^N g(t)\, dt.$$

An improper integral *converges* iff the limit exists, otherwise the integral *diverges*.

Now we are ready to compute our first Laplace transform.

Example 3.1.1. Compute the Laplace transform of the function $f(t) = 1$, that is, $\mathcal{L}[1]$.

Solution: Following the definition,

$$\mathcal{L}[1] = \int_0^\infty e^{-st}\, dt = \lim_{N\to\infty} \int_0^N e^{-st}\, dt.$$

The definite integral above is simple to compute, but it depends on the values of s. For $s = 0$ we get

$$\lim_{N\to\infty} \int_0^N dt = \lim_{n\to\infty} N = \infty.$$

So, the improper integral diverges for $s = 0$. For $s \neq 0$ we get

$$\lim_{N\to\infty} \int_0^N e^{-st}\, dt = \lim_{N\to\infty} -\frac{1}{s}\, e^{-st}\Big|_0^N = \lim_{N\to\infty} -\frac{1}{s}\,(e^{-sN} - 1).$$

For $s < 0$ we have $s = -|s|$, hence

$$\lim_{N\to\infty} -\frac{1}{s}\,(e^{-sN} - 1) = \lim_{N\to\infty} -\frac{1}{s}\,(e^{|s|N} - 1) = -\infty.$$

So, the improper integral diverges for $s < 0$. In the case that $s > 0$ we get

$$\lim_{N\to\infty} -\frac{1}{s}\,(e^{-sN} - 1) = \frac{1}{s}.$$

If we put all these result together we get

$$\mathcal{L}[1] = \frac{1}{s}, \qquad s > 0.$$

◁

Example 3.1.2. Compute $\mathcal{L}[e^{at}]$, where $a \in \mathbb{R}$.

Solution: We start with the definition of the Laplace transform,

$$\mathcal{L}[e^{at}] = \int_0^\infty e^{-st}(e^{at})\, dt = \int_0^\infty e^{-(s-a)t}\, dt.$$

In the case $s = a$ we get

$$\mathcal{L}[e^{at}] = \int_0^\infty 1\, dt = \infty,$$

so the improper integral diverges. In the case $s \neq a$ we get

$$\begin{aligned} \mathcal{L}[e^{at}] &= \lim_{N\to\infty} \int_0^N e^{-(s-a)t}\, dt, \qquad s \neq a, \\ &= \lim_{N\to\infty} \Big[\frac{(-1)}{(s-a)}\, e^{-(s-a)t}\Big|_0^N\Big] \\ &= \lim_{N\to\infty} \Big[\frac{(-1)}{(s-a)}\,(e^{-(s-a)N} - 1)\Big]. \end{aligned}$$

Now we have to remaining cases. The first case is:

$$s - a < 0 \quad \Rightarrow \quad -(s-a) = |s-a| > 0 \quad \Rightarrow \quad \lim_{N\to\infty} e^{-(s-a)N} = \infty,$$

so the integral diverges for $s < a$. The other case is:

$$s - a > 0 \quad \Rightarrow \quad -(s-a) = -|s-a| < 0 \quad \Rightarrow \quad \lim_{N\to\infty} e^{-(s-a)N} = 0,$$

so the integral converges only for $s > a$ and the Laplace transform is given by

$$\mathcal{L}[e^{at}] = \frac{1}{(s-a)}, \qquad s > a.$$

◁

Example 3.1.3. Compute $\mathcal{L}[te^{at}]$, where $a \in \mathbb{R}$.

Solution: In this case the calculation is more complicated than above, since we need to integrate by parts. We start with the definition of the Laplace transform,

$$\mathcal{L}[te^{at}] = \int_0^\infty e^{-st} te^{at}\, dt = \lim_{N\to\infty} \int_0^N te^{-(s-a)t}\, dt.$$

This improper integral diverges for $s = a$, so $\mathcal{L}[te^{at}]$ is not defined for $s = a$. From now on we consider only the case $s \neq a$. In this case we can integrate by parts,

$$\mathcal{L}[te^{at}] = \lim_{N\to\infty} \Big[-\frac{1}{(s-a)} te^{-(s-a)t}\Big|_0^N + \frac{1}{s-a} \int_0^N e^{-(s-a)t}\, dt \Big],$$

that is,

$$\mathcal{L}[te^{at}] = \lim_{N\to\infty} \Big[-\frac{1}{(s-a)} te^{-(s-a)t}\Big|_0^N - \frac{1}{(s-a)^2} e^{-(s-a)t}\Big|_0^N \Big]. \tag{3.1.2}$$

In the case that $s < a$ the first term above diverges,

$$\lim_{N\to\infty} -\frac{1}{(s-a)} N\, e^{-(s-a)N} = \lim_{N\to\infty} -\frac{1}{(s-a)} N\, e^{|s-a|N} = \infty,$$

therefore $\mathcal{L}[te^{at}]$ is not defined for $s < a$. In the case $s > a$ the first term on the right hand side in (3.1.2) vanishes, since

$$\lim_{N\to\infty} -\frac{1}{(s-a)} N\, e^{-(s-a)N} = 0, \qquad \frac{1}{(s-a)} t\, e^{-(s-a)t}\big|_{t=0} = 0.$$

Regarding the other term, and recalling that $s > a$,

$$\lim_{N\to\infty} -\frac{1}{(s-a)^2} e^{-(s-a)N} = 0, \qquad \frac{1}{(s-a)^2} e^{-(s-a)t}\big|_{t=0} = \frac{1}{(s-a)^2}.$$

Therefore, we conclude that

$$\mathcal{L}[te^{at}] = \frac{1}{(s-a)^2}, \qquad s > a.$$

◁

Example 3.1.4. Compute $\mathcal{L}[\sin(at)]$, where $a \in \mathbb{R}$.

Solution: In this case we need to compute

$$\mathcal{L}[\sin(at)] = \int_0^\infty e^{-st} \sin(at)\, dt = \lim_{N\to\infty} \int_0^N e^{-st} \sin(at)\, dt$$

The definite integral above can be computed integrating by parts twice,

$$\int_0^N e^{-st} \sin(at)\, dt = -\frac{1}{s} \big[e^{-st} \sin(at)\big]\Big|_0^N - \frac{a}{s^2} \big[e^{-st} \cos(at)\big]\Big|_0^N - \frac{a^2}{s^2} \int_0^N e^{-st} \sin(at)\, dt,$$

which implies that

$$\Big(1 + \frac{a^2}{s^2}\Big) \int_0^N e^{-st} \sin(at)\, dt = -\frac{1}{s} \big[e^{-st} \sin(at)\big]\Big|_0^N - \frac{a}{s^2} \big[e^{-st} \cos(at)\big]\Big|_0^N.$$

then we get

$$\int_0^N e^{-st} \sin(at)\, dt = \frac{s^2}{(s^2+a^2)} \Big[-\frac{1}{s} \big[e^{-st} \sin(at)\big]\Big|_0^N - \frac{a}{s^2} \big[e^{-st} \cos(at)\big]\Big|_0^N \Big].$$

and finally we get

$$\int_0^N e^{-st} \sin(at)\, dt = \frac{s^2}{(s^2+a^2)} \Big[-\frac{1}{s} \big[e^{-sN} \sin(aN) - 0\big] - \frac{a}{s^2} \big[e^{-sN} \cos(aN) - 1\big] \Big].$$

One can check that the limit $N \to \infty$ on the right hand side above does not exist for $s \leqslant 0$, so $\mathcal{L}[\sin(at)]$ does not exist for $s \leqslant 0$. In the case $s > 0$ it is not difficult to see that

$$\int_0^{\infty} e^{-st} \sin(at)\, dt = \Big(\frac{s^2}{s^2 + a^2}\Big) \Big[\frac{1}{s}\,(0 - 0) - \frac{a}{s^2}\,(0 - 1)\Big]$$

so we obtain the final result

$$\mathcal{L}[\sin(at)] = \frac{a}{s^2 + a^2}, \qquad s > 0.$$

◁

In Table 1 below we present a short list of Laplace transforms. They can be computed in the same way we computed the the Laplace transforms in the examples above.

$f(t)$	$F(s) = \mathcal{L}[f(t)]$	D_F
$f(t) = 1$	$F(s) = \dfrac{1}{s}$	$s > 0$
$f(t) = e^{at}$	$F(s) = \dfrac{1}{(s - a)}$	$s > a$
$f(t) = t^n$	$F(s) = \dfrac{n!}{s^{(n+1)}}$	$s > 0$
$f(t) = \sin(at)$	$F(s) = \dfrac{a}{s^2 + a^2}$	$s > 0$
$f(t) = \cos(at)$	$F(s) = \dfrac{s}{s^2 + a^2}$	$s > 0$
$f(t) = \sinh(at)$	$F(s) = \dfrac{a}{s^2 - a^2}$	$s > \lvert a \rvert$
$f(t) = \cosh(at)$	$F(s) = \dfrac{s}{s^2 - a^2}$	$s > \lvert a \rvert$
$f(t) = t^n e^{at}$	$F(s) = \dfrac{n!}{(s - a)^{(n+1)}}$	$s > a$
$f(t) = e^{at} \sin(bt)$	$F(s) = \dfrac{b}{(s - a)^2 + b^2}$	$s > a$
$f(t) = e^{at} \cos(bt)$	$F(s) = \dfrac{(s - a)}{(s - a)^2 + b^2}$	$s > a$
$f(t) = e^{at} \sinh(bt)$	$F(s) = \dfrac{b}{(s - a)^2 - b^2}$	$s - a > \lvert b \rvert$
$f(t) = e^{at} \cosh(bt)$	$F(s) = \dfrac{(s - a)}{(s - a)^2 - b^2}$	$s - a > \lvert b \rvert$

TABLE 1. List of a few Laplace transforms.

3.1.3. Main Properties. Since we are more or less confident on how to compute a Laplace transform, we can start asking deeper questions. For example, what type of functions have a Laplace transform? It turns out that a large class of functions, those that are piecewise continuous on $[0, \infty)$ and bounded by an exponential. This last property is particularly important and we give it a name.

Definition 3.1.2. *A function f defined on $[0, \infty)$ is of* ***exponential order*** *s_0, where s_0 is any real number, iff there exist positive constants k, T such that*

$$|f(t)| \leqslant k\, e^{s_0 t} \quad \text{for all} \quad t > T. \tag{3.1.3}$$

Remarks:

(a) When the precise value of the constant s_0 is not important we will say that f is of exponential order.

(b) An example of a function that is not of exponential order is $f(t) = e^{t^2}$.

This definition helps to describe a set of functions having Laplace transform. Piecewise continuous functions on $[0, \infty)$ of exponential order have Laplace transforms.

Theorem 3.1.3 (Convergence of LT)**.** *If a function f defined on $[0, \infty)$ is piecewise continuous and of exponential order s_0, then the $\mathcal{L}[f]$ exists for all $s > s_0$ and there exists a positive constant k such that*

$$\big|\mathcal{L}[f]\big| \leqslant \frac{k}{s - s_0}, \qquad s > s_0.$$

Proof of Theorem 3.1.3: From the definition of the Laplace transform we know that

$$\mathcal{L}[f] = \lim_{N \to \infty} \int_0^N e^{-st} f(t)\, dt.$$

The definite integral on the interval $[0, N]$ exists for every $N > 0$ since f is piecewise continuous on that interval, no matter how large N is. We only need to check whether the integral converges as $N \to \infty$. This is the case for functions of exponential order, because

$$\Big|\int_0^N e^{-st} f(t)\, dt\Big| \leqslant \int_0^N e^{-st} |f(t)|\, dt \leqslant \int_0^N e^{-st} k e^{s_0 t}\, dt = k \int_0^N e^{-(s-s_0)t}\, dt.$$

Therefore, for $s > s_0$ we can take the limit as $N \to \infty$,

$$\big|\mathcal{L}[f]\big| \leqslant \lim_{N \to \infty} \Big|\int_0^N e^{-st} f(t)\, dt\Big| \leqslant k\, \mathcal{L}[e^{s_0 t}] = \frac{k}{(s - s_0)}.$$

Therefore, the comparison test for improper integrals implies that the Laplace transform $\mathcal{L}[f]$ exists at least for $s > s_0$, and it also holds that

$$\big|\mathcal{L}[f]\big| \leqslant \frac{k}{s - s_0}, \qquad s > s_0.$$

This establishes the Theorem. □

The next result says that the Laplace transform is a linear transformation. This means that the Laplace transform of a linear combination of functions is the linear combination of their Laplace transforms.

Theorem 3.1.4 (Linearity)**.** *If $\mathcal{L}[f]$ and $\mathcal{L}[g]$ exist, then for all a, $b \in \mathbb{R}$ holds*

$$\mathcal{L}[af + bg] = a\, \mathcal{L}[f] + b\, \mathcal{L}[g].$$

Proof of Theorem 3.1.4: Since integration is a linear operation, so is the Laplace transform, as this calculation shows,

$$\begin{aligned}\mathcal{L}[af+bg] &= \int_0^\infty e^{-st}\big[af(t)+bg(t)\big]\,dt\\ &= a\int_0^\infty e^{-st}f(t)\,dt + b\int_0^\infty e^{-st}g(t)\,dt\\ &= a\,\mathcal{L}[f] + b\,\mathcal{L}[g].\end{aligned}$$

This establishes the Theorem. □

Example 3.1.5. Compute $\mathcal{L}[3t^2 + 5\cos(4t)]$.

Solution: From the Theorem above and the Laplace transform in Table **??** we know that

$$\begin{aligned}\mathcal{L}[3t^2+5\cos(4t)] &= 3\,\mathcal{L}[t^2] + 5\,\mathcal{L}[\cos(4t)]\\ &= 3\Big(\frac{2}{s^3}\Big) + 5\Big(\frac{s}{s^2+4^2}\Big), \qquad s>0\\ &= \frac{6}{s^3} + \frac{5s}{s^2+4^2}.\end{aligned}$$

Therefore,

$$\mathcal{L}[3t^2+5\cos(4t)] = \frac{5s^4+6s^2+96}{s^3(s^2+16)}, \qquad s>0.$$

◁

The Laplace transform can be used to solve differential equations. The Laplace transform converts a differential equation into an algebraic equation. This is so because the Laplace transform converts derivatives into multiplications. Here is the precise result.

Theorem 3.1.5 (Derivative into Multiplication)**.** *If a function f is continuously differentiable on $[0,\infty)$ and of exponential order s_0, then $\mathcal{L}[f']$ exists for $s > s_0$ and*

$$\mathcal{L}[f'] = s\,\mathcal{L}[f] - f(0), \qquad s > s_0. \tag{3.1.4}$$

Proof of Theorem 3.1.5: The main calculation in this proof is to compute

$$\mathcal{L}[f'] = \lim_{N\to\infty}\int_0^N e^{-st}\,f'(t)\,dt.$$

We start computing the definite integral above. Since f' is continuous on $[0,\infty)$, that definite integral exists for all positive N, and we can integrate by parts,

$$\begin{aligned}\int_0^N e^{-st}f'(t)\,dt &= \Big[\Big(e^{-st}f(t)\Big)\Big|_0^N - \int_0^N (-s)e^{-st}f(t)\,dt\Big]\\ &= e^{-sN}f(N) - f(0) + s\int_0^N e^{-st}f(t)\,dt.\end{aligned}$$

We now compute the limit of this expression above as $N\to\infty$. Since f is continuous on $[0,\infty)$ of exponential order s_0, we know that

$$\lim_{N\to\infty}\int_0^N e^{-st}f(t)\,dt = \mathcal{L}[f], \qquad s > s_0.$$

Let us use one more time that f is of exponential order s_0. This means that there exist positive constants k and T such that $|f(t)| \leqslant k\,e^{s_0 t}$, for $t > T$. Therefore,

$$\lim_{N\to\infty} e^{-sN}f(N) \leqslant \lim_{N\to\infty} k\,e^{-sN}e^{s_0N} = \lim_{N\to\infty} k\,e^{-(s-s_0)N} = 0, \qquad s > s_0.$$

These two results together imply that $\mathcal{L}[f']$ exists and holds

$$\mathcal{L}[f'] = s\,\mathcal{L}[f] - f(0), \qquad s > s_0.$$

This establishes the Theorem. □

Example 3.1.6. Verify the result in Theorem 3.1.5 for the function $f(t) = \cos(bt)$.

Solution: We need to compute the left hand side and the right hand side of Eq. (3.1.4) and verify that we get the same result. We start with the left hand side,

$$\mathcal{L}[f'] = \mathcal{L}[-b\sin(bt)] = -b\,\mathcal{L}[\sin(bt)] = -b\,\frac{b}{s^2+b^2} \quad \Rightarrow \quad \mathcal{L}[f'] = -\frac{b^2}{s^2+b^2}.$$

We now compute the right hand side,

$$s\,\mathcal{L}[f] - f(0) = s\,\mathcal{L}[\cos(bt)] - 1 = s\,\frac{s}{s^2+b^2} - 1 = \frac{s^2 - s^2 - b^2}{s^2+b^2},$$

so we get

$$s\,\mathcal{L}[f] - f(0) = -\frac{b^2}{s^2+b^2}.$$

We conclude that $\mathcal{L}[f'] = s\,\mathcal{L}[f] - f(0)$. ◁

It is not difficult to generalize Theorem 3.1.5 to higher order derivatives.

Theorem 3.1.6 (Higher Derivatives into Multiplication)**.** *If a function f is n-times continuously differentiable on $[0,\infty)$ and of exponential order s_0, then $\mathcal{L}[f''], \cdots, \mathcal{L}[f^{(n)}]$ exist for $s > s_0$ and*

$$\mathcal{L}[f''] = s^2\,\mathcal{L}[f] - s\,f(0) - f'(0) \tag{3.1.5}$$

$$\vdots$$

$$\mathcal{L}[f^{(n)}] = s^n\,\mathcal{L}[f] - s^{(n-1)}\,f(0) - \cdots - f^{(n-1)}(0). \tag{3.1.6}$$

Proof of Theorem 3.1.6: We need to use Eq. (3.1.4) n times. We start with the Laplace transform of a second derivative,

$$\begin{aligned}\mathcal{L}[f''] &= \mathcal{L}[(f')'] \\ &= s\,\mathcal{L}[f'] - f'(0) \\ &= s\big(s\,\mathcal{L}[f] - f(0)\big) - f'(0) \\ &= s^2\,\mathcal{L}[f] - s\,f(0) - f'(0).\end{aligned}$$

The formula for the Laplace transform of an nth derivative is computed by induction on n. We assume that the formula is true for $n-1$,

$$\mathcal{L}[f^{(n-1)}] = s^{(n-1)}\,\mathcal{L}[f] - s^{(n-2)}\,f(0) - \cdots - f^{(n-2)}(0).$$

Since $\mathcal{L}[f^{(n)}] = \mathcal{L}[(f')^{(n-1)}]$, the formula above on f' gives

$$\begin{aligned}\mathcal{L}[(f')^{(n-1)}] &= s^{(n-1)}\,\mathcal{L}[f'] - s^{(n-2)}\,f'(0) - \cdots - (f')^{(n-2)}(0) \\ &= s^{(n-1)}\,\big(s\,\mathcal{L}[f] - f(0)\big) - s^{(n-2)}\,f'(0) - \cdots - f^{(n-1)}(0) \\ &= s^{(n)}\,\mathcal{L}[f] - s^{(n-1)}\,f(0) - s^{(n-2)}\,f'(0) - \cdots - f^{(n-1)}(0).\end{aligned}$$

This establishes the Theorem. □

Example 3.1.7. Verify Theorem 3.1.6 for f'', where $f(t) = \cos(bt)$.

Solution: We need to compute the left hand side and the right hand side in the first equation in Theorem (3.1.6), and verify that we get the same result. We start with the left hand side,

$$\mathcal{L}[f''] = \mathcal{L}[-b^2 \cos(bt)] = -b^2\, \mathcal{L}[\cos(bt)] = -b^2\, \frac{s}{s^2+b^2} \quad \Rightarrow \quad \mathcal{L}[f''] = -\frac{b^2 s}{s^2+b^2}.$$

We now compute the right hand side,

$$s^2\, \mathcal{L}[f] - s\, f(0) - f'(0) = s^2\, \mathcal{L}[\cos(bt)] - s - 0 = s^2\, \frac{s}{s^2+b^2} - s = \frac{s^3 - s^3 - b^2 s}{s^2+b^2},$$

so we get

$$s^2\, \mathcal{L}[f] - s\, f(0) - f'(0) = -\frac{b^2 s}{s^2+b^2}.$$

We conclude that $\mathcal{L}[f''] = s^2\, \mathcal{L}[f] - s\, f(0) - f'(0)$. ◁

The Laplace transform also satisfies a converse to Theorem 3.1.5, since multiplications can be transformed into derivatives.

Theorem 3.1.7 (Multiplication into Derivative)**.** *If a function f is of exponential order s_0 with a Laplace transform $F(s) = \mathcal{L}[f(t)]$, then $\mathcal{L}[t\, f(t)]$ exists for $s > s_0$ and*

$$\mathcal{L}[t\, f(t)] = -F'(s), \qquad s > s_0. \tag{3.1.7}$$

Proof of Theorem 3.1.7: From the definition of the Laplace Transform we see that

$$\begin{aligned}
\mathcal{L}[t\, f(t)] &= \int_0^\infty e^{-st}\, t\, f(t)\, dt \\
&= \int_0^\infty \frac{d}{ds}\Big(-e^{-st}\Big)\, f(t)\, dt \\
&= -\frac{d}{ds} \int_0^\infty e^{-st}\, f(t)\, dt \\
&= -\frac{d}{ds} \mathcal{L}[f(t)] \\
&= -F'(s).
\end{aligned}$$

This establishes the Theorem. □

The result in Theorem 3.1.7 can be generalized to higher powers.

Theorem 3.1.8 (Higher Powers into Derivative)**.** *If a function f is of exponential order s_0 with a Laplace transform $F(s) = \mathcal{L}[f(t)]$, then $\mathcal{L}[t^n\, f(t)]$ exists for $s > s_0$ and*

$$\mathcal{L}[t^n\, f(t)] = (-1)^n F^{(n)}(s), \qquad s > s_0, \tag{3.1.8}$$

where we denoted $F^{(n)} = \dfrac{d^n}{ds^n} F$.

Proof of Theorem 3.1.8: We use induction one more time. The case $n = 1$ is done in Theorem 3.1.7. We now assume that

$$\mathcal{L}[t^n\, f(t)] = (-1)^n \frac{d^n}{ds^n} \mathcal{L}[f(t)],$$

and we try to show that a similar formula holds for $n+1$. But this is the case, since

$$\begin{aligned}\mathcal{L}[t^{(n+1)}\,f(t)] &= \mathcal{L}\big[t^n\big(t\,f(t)\big)\big]\\ &= (-1)^n\,\frac{d^n}{ds^n}\mathcal{L}[t\,f(t)],\end{aligned}$$

since $t\,f(t)$ satisfies the hypotheses in Theorem 3.1.7, since $f(t)$ does. Then we use Theorem 3.1.7 one more time,

$$\begin{aligned}\mathcal{L}[t^{(n+1)}\,f(t)] &= (-1)^n\,\frac{d^n}{ds^n}\mathcal{L}[t\,f(t)],\\ &= (-1)^n\,\frac{d^n}{ds^n}(-1)\frac{d}{ds}\mathcal{L}[f(t)],\\ &= (-1)^{(n+1)}\,\frac{d^{(n+1)}}{ds^{(n+1)}}\mathcal{L}[f(t)],\\ &= (-1)^{(n+1)}\,F^{(n+1)}(s).\end{aligned}$$

This establishes the Theorem. □

3.1.4. Solving Differential Equations. The Laplace transform can be used to solve differential equations. We Laplace transform the whole equation, which converts the differential equation for y into an algebraic equation for $\mathcal{L}[y]$. We solve the Algebraic equation and we transform back.

$$\mathcal{L}\begin{bmatrix}\text{differential}\\ \text{eq. for } y.\end{bmatrix} \xrightarrow{(1)} \begin{matrix}\text{Algebraic}\\ \text{eq. for } \mathcal{L}[y].\end{matrix} \xrightarrow{(2)} \begin{matrix}\text{Solve the}\\ \text{algebraic}\\ \text{eq. for } \mathcal{L}[y].\end{matrix} \xrightarrow{(3)} \begin{matrix}\text{Transform back}\\ \text{to obtain } y.\\ \text{(Use the table.)}\end{matrix}$$

Example 3.1.8. Use the Laplace transform to find y solution of

$$y'' + 9\,y = 0, \qquad y(0) = y_0, \quad y'(0) = y_1.$$

Remark: Notice we already know what the solution of this problem is. Following § 2.2 we need to find the roots of

$$p(r) = r^2 + 9 \quad\Rightarrow\quad r_\pm = \pm 3\,i,$$

and then we get the general solution

$$y(t) = c_+\cos(3t) + c_-\sin(3t).$$

Then the initial condition will say that

$$y(t) = y_0\,\cos(3t) + \frac{y_1}{3}\,\sin(3t).$$

We now solve this problem using the Laplace transform method.

Solution: We now use the Laplace transform method:

$$\mathcal{L}[y'' + 9y] = \mathcal{L}[0] = 0.$$

The Laplace transform is a linear transformation,

$$\mathcal{L}[y''] + 9\,\mathcal{L}[y] = 0.$$

But the Laplace transform converts derivatives into multiplications,

$$s^2\,\mathcal{L}[y] - s\,y(0) - y'(0) + 9\,\mathcal{L}[y] = 0.$$

This is an algebraic equation for $\mathcal{L}[y]$. It can be solved by rearranging terms and using the initial condition,

$$(s^2+9)\,\mathcal{L}[y] = s\,y_0 + y_1 \quad \Rightarrow \quad \mathcal{L}[y] = y_0\,\frac{s}{(s^2+9)} + y_1\,\frac{1}{(s^2+9)}.$$

But from the Laplace transform table we see that

$$\mathcal{L}[\cos(3t)] = \frac{s}{s^2+3^2}, \qquad \mathcal{L}[\sin(3t)] = \frac{3}{s^2+3^2},$$

therefore,

$$\mathcal{L}[y] = y_0\,\mathcal{L}[\cos(3t)] + y_1\,\frac{1}{3}\,\mathcal{L}[\sin(3t)].$$

Once again, the Laplace transform is a linear transformation,

$$\mathcal{L}[y] = \mathcal{L}\Big[y_0\,\cos(3t) + \frac{y_1}{3}\,\sin(3t)\Big].$$

We obtain that

$$y(t) = y_0\,\cos(3t) + \frac{y_1}{3}\,\sin(3t).$$

$\lhd$

3.1.5. Exercises.

3.1.1.- .

3.1.2.- .

3.2. The Initial Value Problem

We will use the Laplace transform to solve differential equations. The main idea is,

$$\mathcal{L}\begin{bmatrix}\text{differential eq.}\\ \text{for } y(t).\end{bmatrix} \xrightarrow{(1)} \begin{matrix}\text{Algebraic eq.}\\ \text{for } \mathcal{L}[y(t)].\end{matrix} \xrightarrow{(2)} \begin{matrix}\text{Solve the}\\ \text{algebraic eq.}\\ \text{for } \mathcal{L}[y(t)].\end{matrix} \xrightarrow{(3)} \begin{matrix}\text{Transform back}\\ \text{to obtain } y(t).\\ \text{(Use the table.)}\end{matrix}$$

We will use the Laplace transform to solve differential equations with *constant coefficients*. Although the method can be used with variable coefficients equations, the calculations could be very complicated in such a case.

The Laplace transform method works with *very general source functions*, including step functions, which are discontinuous, and Dirac's deltas, which are generalized functions.

3.2.1. Solving Differential Equations.

As we see in the sketch above, we start with a differential equation for a function y. We first compute the Laplace transform of the whole differential equation. Then we use the linearity of the Laplace transform, Theorem 3.1.4, and the property that derivatives are converted into multiplications, Theorem 3.1.5, to transform the differential equation into an algebraic equation for $\mathcal{L}[y]$. Let us see how this works in a simple example, a first order linear equation with constant coefficients—we already solved it in § **??**.

Example 3.2.1. Use the Laplace transform to find the solution y to the initial value problem

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

Solution: In § 1.2 we saw one way to solve this problem, using the integrating factor method. One can check that the solution is $y(t) = 3e^{-2t}$. We now use the Laplace transform. First, compute the Laplace transform of the differential equation,

$$\mathcal{L}[y' + 2y] = \mathcal{L}[0] = 0.$$

Theorem 3.1.4 says the Laplace transform is a linear operation, that is,

$$\mathcal{L}[y'] + 2\,\mathcal{L}[y] = 0.$$

Theorem 3.1.5 relates derivatives and multiplications, as follows,

$$\Big(s\,\mathcal{L}[y] - y(0)\Big) + 2\,\mathcal{L}[y] = 0 \quad \Rightarrow \quad (s+2)\mathcal{L}[y] = y(0).$$

In the last equation we have been able to transform the original differential equation for y into an algebraic equation for $\mathcal{L}[y]$. We can solve for the unknown $\mathcal{L}[y]$ as follows,

$$\mathcal{L}[y] = \frac{y(0)}{s+2} \quad \Rightarrow \quad \mathcal{L}[y] = \frac{3}{s+2},$$

where in the last step we introduced the initial condition $y(0) = 3$. From the list of Laplace transforms given in §. 3.1 we know that

$$\mathcal{L}[e^{at}] = \frac{1}{s-a} \quad \Rightarrow \quad \frac{3}{s+2} = 3\mathcal{L}[e^{-2t}] \quad \Rightarrow \quad \frac{3}{s+2} = \mathcal{L}[3\,e^{-2t}].$$

So we arrive at $\mathcal{L}[y(t)] = \mathcal{L}[3\,e^{-2t}]$. Here is where we need one more property of the Laplace transform. We show right after this example that

$$\mathcal{L}[y(t)] = \mathcal{L}[3\,e^{-2t}] \quad \Rightarrow \quad y(t) = 3\,e^{-2t}.$$

This property is called one-to-one. Hence the only solution is $y(t) = 3\,e^{-2t}$. ◁

3.2.2. One-to-One Property. Let us repeat the method we used to solve the differential equation in Example 3.2.1. We first computed the Laplace transform of the whole differential equation. Then we use the linearity of the Laplace transform, Theorem 3.1.4, and the property that derivatives are converted into multiplications, Theorem 3.1.5, to transform the differential equation into an algebraic equation for $\mathcal{L}[y]$. We solved the algebraic equation and we got an expression of the form

$$\mathcal{L}[y(t)] = H(s),$$

where we have collected all the terms that come from the Laplace transformed differential equation into the function H. We then used a Laplace transform table to find a function h such that

$$\mathcal{L}[h(t)] = H(s).$$

We arrived to an equation of the form

$$\mathcal{L}[y(t)] = \mathcal{L}[h(t)].$$

Clearly, $y = h$ is one solution of the equation above, hence a solution to the differential equation. We now show that there are no solutions to the equation $\mathcal{L}[y] = \mathcal{L}[h]$ other than $y = h$. The reason is that the Laplace transform on continuous functions of exponential order is an one-to-one transformation, also called injective.

Theorem 3.2.1 (One-to-One). *If f, g are continuous on $[0, \infty)$ of exponential order, then*

$$\mathcal{L}[f] = \mathcal{L}[g] \quad \Rightarrow \quad f = g.$$

Remarks:

(a) The result above holds for continuous functions f and g. But it can be extended to piecewise continuous functions. In the case of piecewise continuous functions f and g satisfying $\mathcal{L}[f] = \mathcal{L}[g]$ one can prove that $f = g + h$, where h is a null function, meaning that $\int_0^T h(t)\,dt = 0$ for all $T > 0$. See Churchill's textbook [5], page 14.

(b) Once we know that the Laplace transform is a one-to-one transformation, we can define the inverse transformation in the usual way.

Definition 3.2.2. *The* ***inverse Laplace transform****, denoted $\mathcal{L}^{-1}$, of a function F is*

$$\mathcal{L}^{-1}[F(s)] = f(t) \quad \Leftrightarrow \quad F(s) = \mathcal{L}[f(t)].$$

Remarks: There is an explicit formula for the inverse Laplace transform, which involves an integral on the complex plane,

$$\mathcal{L}^{-1}[F(s)]\Big|_t = \frac{1}{2\pi i} \lim_{c\to\infty} \int_{a-ic}^{a+ic} e^{st}\, F(s)\, ds.$$

See for example Churchill's textbook [5], page 176. However, we do not use this formula in these notes, since it involves integration on the complex plane.

Proof of Theorem 3.2.1: The proof is based on a clever change of variables and on Weierstrass Approximation Theorem of continuous functions by polynomials. Before we get to the change of variable we need to do some rewriting. Introduce the function $u = f - g$, then the linearity of the Laplace transform implies

$$\mathcal{L}[u] = \mathcal{L}[f - g] = \mathcal{L}[f] - \mathcal{L}[g] = 0.$$

What we need to show is that the function u vanishes identically. Let us start with the definition of the Laplace transform,

$$\mathcal{L}[u] = \int_0^\infty e^{-st}\, u(t)\, dt.$$

We know that f and g are of exponential order, say s_0, therefore u is of exponential order s_0, meaning that there exist positive constants k and T such that

$$\big|u(t)\big| < k\, e^{s_0 t}, \qquad t > T.$$

Evaluate $\mathcal{L}[u]$ at $\tilde{s} = s_1 + n + 1$, where s_1 is any real number such that $s_1 > s_0$, and n is any positive integer. We get

$$\mathcal{L}[u]\Big|_{\tilde{s}} = \int_0^\infty e^{-(s_1+n+1)t}\, u(t)\, dt = \int_0^\infty e^{-s_1 t}\, e^{-(n+1)t}\, u(t)\, dt.$$

We now do the substitution $y = e^{-t}$, so $dy = -e^{-t}\, dt$,

$$\mathcal{L}[u]\Big|_{\tilde{s}} = \int_1^0 y^{s_1}\, y^n\, u\big(-\ln(y)\big)\,(-dy) = \int_0^1 y^{s_1}\, y^n\, u\big(-\ln(y)\big)\, dy.$$

Introduce the function $v(y) = y^{s_1}\, u(\big(-\ln(y)\big)$, so the integral is

$$\mathcal{L}[u]\Big|_{\tilde{s}} = \int_0^1 y^n\, v(y)\, dy. \tag{3.2.1}$$

We know that $\mathcal{L}[u]$ exists because u is continuous and of exponential order, so the function v does not diverge at $y = 0$. To double check this, recall that $t = -\ln(y) \to \infty$ as $y \to 0^+$, and u is of exponential order s_0, hence

$$\lim_{y\to 0^+} |v(y)| = \lim_{t\to\infty} e^{-s_1 t} |u(t)| < \lim_{t\to\infty} e^{-(s_1 - s_0)t} = 0.$$

Our main hypothesis is that $\mathcal{L}[u] = 0$ for all values of s such that $\mathcal{L}[u]$ is defined, in particular $\tilde{s}$. By looking at Eq. (3.2.1) this means that

$$\int_0^1 y^n\, v(y)\, dy = 0, \qquad n = 1, 2, 3, \cdots.$$

The equation above and the linearity of the integral imply that this function v is perpendicular to every polynomial p, that is

$$\int_0^1 p(y)\, v(y)\, dy = 0, \tag{3.2.2}$$

for every polynomial p. Knowing that, we can do the following calculation,

$$\int_0^1 v^2(y)\, dy = \int_0^1 \big(v(y) - p(y)\big)\, v(y)\, dy + \int_0^1 p(y)\, v(y)\, dy.$$

The last term in the second equation above vanishes because of Eq. (3.2.2), therefore

$$\begin{aligned} \int_0^1 v^2(y)\, dy &= \int_0^1 \big(v(y) - p(y)\big)\, v(y)\, dy \\ &\leqslant \int_0^1 \big|v(y) - p(y)\big|\, |v(y)|\, dy \\ &\leqslant \max_{y\in[0,1]} |v(y)| \int_0^1 \big|v(y) - p(y)\big|\, dy. \end{aligned} \tag{3.2.3}$$

We remark that the inequality above is true for every polynomial p. Here is where we use the Weierstrass Approximation Theorem, which essentially says that every continuous function on a closed interval can be approximated by a polynomial.

Theorem 3.2.3 (Weierstrass Approximation). *If f is a continuous function on a closed interval $[a,b]$, then for every $\epsilon > 0$ there exists a polynomial q_ϵ such that*

$$\max_{y\in[a,b]} |f(y) - q_\epsilon(y)| < \epsilon.$$

The proof of this theorem can be found on a real analysis textbook. Weierstrass result implies that, given v and $\epsilon > 0$, there exists a polynomial p_ϵ such that the inequality in (3.2.3) has the form

$$\int_0^1 v^2(y)\,dy \leqslant \max_{y\in[0,1]} |v(y)| \int_0^1 \big|v(y) - p_\epsilon(y)\big|\,dy \leqslant \max_{y\in[0,1]} |v(y)|\,\epsilon.$$

Since ϵ can be chosen as small as we please, we get

$$\int_0^1 v^2(y)\,dy = 0.$$

But v is continuous, hence $v = 0$, meaning that $f = g$. This establishes the Theorem. □

3.2.3. Partial Fractions. We are now ready to start using the Laplace transform to solve second order linear differential equations with constant coefficients. The differential equation for y will be transformed into an algebraic equation for $\mathcal{L}[y]$. We will then arrive to an equation of the form $\mathcal{L}[y(t)] = H(s)$. We will see, already in the first example below, that usually this function H does not appear in Table 1. We will need to rewrite H as a linear combination of simpler functions, each one appearing in Table 1. One of the more used techniques to do that is called Partial Fractions. Let us solve the next example.

Example 3.2.2. Use the Laplace transform to find the solution y to the initial value problem

$$y'' - y' - 2y = 0, \qquad y(0) = 1, \qquad y'(0) = 0.$$

Solution: First, compute the Laplace transform of the differential equation,

$$\mathcal{L}[y'' - y' - 2y] = \mathcal{L}[0] = 0.$$

Theorem 3.1.4 says that the Laplace transform is a linear operation,

$$\mathcal{L}[y''] - \mathcal{L}[y'] - 2\,\mathcal{L}[y] = 0.$$

Then, Theorem 3.1.5 relates derivatives and multiplications,

$$\Big[s^2\,\mathcal{L}[y] - s\,y(0) - y'(0)\Big] - \Big[s\,\mathcal{L}[y] - y(0)\Big] - 2\,\mathcal{L}[y] = 0,$$

which is equivalent to the equation

$$(s^2 - s - 2)\,\mathcal{L}[y] = (s-1)\,y(0) + y'(0).$$

Once again we have transformed the original differential equation for y into an algebraic equation for $\mathcal{L}[y]$. Introduce the initial condition into the last equation above, that is,

$$(s^2 - s - 2)\,\mathcal{L}[y] = (s-1).$$

Solve for the unknown $\mathcal{L}[y]$ as follows,

$$\mathcal{L}[y] = \frac{(s-1)}{(s^2 - s - 2)}.$$

The function on the right hand side above does not appear in Table 1. We now use *partial fractions* to find a function whose Laplace transform is the right hand side of the equation above. First find the roots of the polynomial in the denominator,

$$s^2 - s - 2 = 0 \quad \Rightarrow \quad s_\pm = \frac{1}{2}\big[1 \pm \sqrt{1+8}\big] \quad \Rightarrow \quad \begin{cases} s_+ = 2, \\ s_- = -1, \end{cases}$$

that is, the polynomial has two real roots. In this case we factorize the denominator,

$$\mathcal{L}[y] = \frac{(s-1)}{(s-2)(s+1)}.$$

The partial fraction decomposition of the right-hand side in the equation above is the following: Find constants a and b such that

$$\frac{(s-1)}{(s-2)(s+1)} = \frac{a}{s-2} + \frac{b}{s+1}.$$

A simple calculation shows

$$\frac{(s-1)}{(s-2)(s+1)} = \frac{a}{s-2} + \frac{b}{s+1} = \frac{a(s+1)+b(s-2)}{(s-2)(s+1)} = \frac{s(a+b)+(a-2b)}{(s-2)(s+1)}.$$

Hence constants a and b must be solutions of the equations

$$(s-1) = s(a+b) + (a-2b) \quad \Rightarrow \quad \begin{cases} a+b = 1, \\ a-2b = -1. \end{cases}$$

The solution is $a = \dfrac{1}{3}$ and $b = \dfrac{2}{3}$. Hence,

$$\mathcal{L}[y] = \frac{1}{3}\,\frac{1}{(s-2)} + \frac{2}{3}\,\frac{1}{(s+1)}.$$

From the list of Laplace transforms given in § **??**, Table 1, we know that

$$\mathcal{L}[e^{at}] = \frac{1}{s-a} \quad \Rightarrow \quad \frac{1}{s-2} = \mathcal{L}[e^{2t}], \qquad \frac{1}{s+1} = \mathcal{L}[e^{-t}].$$

So we arrive at the equation

$$\mathcal{L}[y] = \frac{1}{3}\,\mathcal{L}[e^{2t}] + \frac{2}{3}\,\mathcal{L}[e^{-t}] = \mathcal{L}\Big[\frac{1}{3}\big(e^{2t} + 2\,e^{-t}\big)\Big]$$

We conclude that

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

◁

The Partial Fraction Method is usually introduced in a second course of Calculus to integrate rational functions. We need it here to use Table 1 to find Inverse Laplace transforms. The method applies to rational functions

$$R(s) = \frac{Q(s)}{P(s)},$$

where P, Q are polynomials and the degree of the numerator is less than the degree of the denominator. In the example above

$$R(s) = \frac{(s-1)}{(s^2-s-2)}.$$

One starts rewriting the polynomial in the denominator as a product of polynomials degree two or one. In the example above,

$$R(s) = \frac{(s-1)}{(s-2)(s+1)}.$$

One then rewrites the rational function as an addition of simpler rational functions. In the example above,

$$R(s) = \frac{a}{(s-2)} + \frac{b}{(s+1)}.$$

We now solve a few examples to recall the different partial fraction cases that can appear when solving differential equations.

Example 3.2.3. Use the Laplace transform to find the solution y to the initial value problem

$$y'' - 4y' + 4y = 0, \qquad y(0) = 1, \qquad y'(0) = 1.$$

Solution: First, compute the Laplace transform of the differential equation,

$$\mathcal{L}[y'' - 4y' + 4y] = \mathcal{L}[0] = 0.$$

Theorem 3.1.4 says that the Laplace transform is a linear operation,

$$\mathcal{L}[y''] - 4\,\mathcal{L}[y'] + 4\,\mathcal{L}[y] = 0.$$

Theorem 3.1.5 relates derivatives with multiplication,

$$\Big[s^2\,\mathcal{L}[y] - s\,y(0) - y'(0)\Big] - 4\,\Big[s\,\mathcal{L}[y] - y(0)\Big] + 4\,\mathcal{L}[y] = 0,$$

which is equivalent to the equation

$$(s^2 - 4s + 4)\,\mathcal{L}[y] = (s-4)\,y(0) + y'(0).$$

Introduce the initial conditions $y(0) = 1$ and $y'(0) = 1$ into the equation above,

$$(s^2 - 4s + 4)\,\mathcal{L}[y] = s - 3.$$

Solve the algebraic equation for $\mathcal{L}[y]$,

$$\mathcal{L}[y] = \frac{(s-3)}{(s^2 - 4s + 4)}.$$

We now want to find a function y whose Laplace transform is the right hand side in the equation above. In order to see if partial fractions will be needed, we now find the roots of the polynomial in the denominator,

$$s^2 - 4s + 4 = 0 \quad \Rightarrow \quad s_\pm = \frac{1}{2}\big[4 \pm \sqrt{16-16}\big] \quad \Rightarrow \quad s_+ = s_- = 2.$$

that is, the polynomial has a single real root, so $\mathcal{L}[y]$ can be written as

$$\mathcal{L}[y] = \frac{(s-3)}{(s-2)^2}.$$

This expression is already in the partial fraction decomposition. We now rewrite the right hand side of the equation above in a way it is simple to use the Laplace transform table in § **??**,

$$\mathcal{L}[y] = \frac{(s-2) + 2 - 3}{(s-2)^2} = \frac{(s-2)}{(s-2)^2} - \frac{1}{(s-2)^2} \quad \Rightarrow \quad \mathcal{L}[y] = \frac{1}{s-2} - \frac{1}{(s-2)^2}.$$

From the list of Laplace transforms given in Table 1, § **??** we know that

$$\mathcal{L}[e^{at}] = \frac{1}{s-a} \quad \Rightarrow \quad \frac{1}{s-2} = \mathcal{L}[e^{2t}],$$

$$\mathcal{L}[te^{at}] = \frac{1}{(s-a)^2} \quad \Rightarrow \quad \frac{1}{(s-2)^2} = \mathcal{L}[t\, e^{2t}].$$

So we arrive at the equation

$$\mathcal{L}[y] = \mathcal{L}[e^{2t}] - \mathcal{L}[t\, e^{2t}] = \mathcal{L}\big[e^{2t} - t\, e^{2t}\big] \quad \Rightarrow \quad y(t) = e^{2t} - t\, e^{2t}.$$

◁

Example 3.2.4. Use the Laplace transform to find the solution y to the initial value problem

$$y'' - 4y' + 4y = 3\, e^{t}, \qquad y(0) = 0, \qquad y'(0) = 0.$$

Solution: First, compute the Laplace transform of the differential equation,

$$\mathcal{L}[y'' - 4y' + 4y] = \mathcal{L}[3\, e^{t}] = 3\left(\frac{1}{s-1}\right).$$

The Laplace transform is a linear operation,

$$\mathcal{L}[y''] - 4\, \mathcal{L}[y'] + 4\, \mathcal{L}[y] = \frac{3}{s-1}.$$

The Laplace transform relates derivatives with multiplication,

$$\Big[s^2\, \mathcal{L}[y] - s\, y(0) - y'(0)\Big] - 4\, \Big[s\, \mathcal{L}[y] - y(0)\Big] + 4\, \mathcal{L}[y] = \frac{3}{s-1},$$

But the initial conditions are $y(0) = 0$ and $y'(0) = 0$, so

$$(s^2 - 4s + 4)\, \mathcal{L}[y] = \frac{3}{s-1}.$$

Solve the algebraic equation for $\mathcal{L}[y]$,

$$\mathcal{L}[y] = \frac{3}{(s-1)(s^2 - 4s + 4)}.$$

We use *partial fractions* to simplify the right-hand side above. We start finding the roots of the polynomial in the denominator,

$$s^2 - 4s + 4 = 0 \quad \Rightarrow \quad s_\pm = \frac{1}{2}\big[4 \pm \sqrt{16 - 16}\big] \quad \Rightarrow \quad s_+ = s_- = 2.$$

that is, the polynomial has a single real root, so $\mathcal{L}[y]$ can be written as

$$\mathcal{L}[y] = \frac{3}{(s-1)(s-2)^2}.$$

The partial fraction decomposition of the right-hand side above is

$$\frac{3}{(s-1)(s-2)^2} = \frac{a}{(s-1)} + \frac{b\, s + c}{(s-2)^2} = \frac{a\, (s-2)^2 + (b\, s + c)(s-1)}{(s-1)(s-2)^2}$$

From the far right and left expressions above we get

$$3 = a\, (s-2)^2 + (b\, s + c)(s-1) = a\, (s^2 - 4s + 4) + b\, s^2 - b\, s + c\, s - c$$

Expanding all terms above, and reordering terms, we get

$$(a + b)\, s^2 + (-4a - b + c)\, s + (4a - c - 3) = 0.$$

Since this polynomial in s vanishes for all $s \in \mathbb{R}$, we get that

$$\left.\begin{aligned} a + b &= 0, \\ -4a - b + c &= 0, \\ 4a - c - 3 &= 0. \end{aligned}\right\} \quad \Rightarrow \quad \begin{cases} a = 3 \\ b = -3 \\ c = 9. \end{cases}$$

So we get
$$\mathcal{L}[y] = \frac{3}{(s-1)(s-2)^2} = \frac{3}{s-1} + \frac{-3s+9}{(s-2)^2}$$
One last trick is needed on the last term above,
$$\frac{-3s+9}{(s-2)^2} = \frac{-3(s-2+2)+9}{(s-2)^2} = \frac{-3(s-2)}{(s-2)^2} + \frac{-6+9}{(s-2)^2} = -\frac{3}{(s-2)} + \frac{3}{(s-2)^2}.$$
So we finally get
$$\mathcal{L}[y] = \frac{3}{s-1} - \frac{3}{(s-2)} + \frac{3}{(s-2)^2}.$$
From our Laplace transforms Table we know that
$$\mathcal{L}[e^{at}] = \frac{1}{s-a} \quad \Rightarrow \quad \frac{1}{s-2} = \mathcal{L}[e^{2t}],$$
$$\mathcal{L}[te^{at}] = \frac{1}{(s-a)^2} \quad \Rightarrow \quad \frac{1}{(s-2)^2} = \mathcal{L}[te^{2t}].$$
So we arrive at the formula
$$\mathcal{L}[y] = 3\,\mathcal{L}[e^t] - 3\,\mathcal{L}[e^{2t}] + 3\,\mathcal{L}[t\,e^{2t}] = \mathcal{L}\big[3\,(e^t - e^{2t} + t\,e^{2t})\big]$$
So we conclude that $y(t) = 3\,(e^t - e^{2t} + t\,e^{2t})$. ◁

Example 3.2.5. Use the Laplace transform to find the solution y to the initial value problem
$$y'' - 4y' + 4y = 3\sin(2t), \qquad y(0) = 1, \qquad y'(0) = 1.$$

Solution: First, compute the Laplace transform of the differential equation,
$$\mathcal{L}[y'' - 4y' + 4y] = \mathcal{L}[3\sin(2t)].$$
The right hand side above can be expressed as follows,
$$\mathcal{L}[3\sin(2t)] = 3\,\mathcal{L}[\sin(2t)] = 3\,\frac{2}{s^2+2^2} = \frac{6}{s^2+4}.$$
Theorem 3.1.4 says that the Laplace transform is a linear operation,
$$\mathcal{L}[y''] - 4\,\mathcal{L}[y'] + 4\,\mathcal{L}[y] = \frac{6}{s^2+4},$$
and Theorem 3.1.5 relates derivatives with multiplications,
$$\Big[s^2\,\mathcal{L}[y] - s\,y(0) - y'(0)\Big] - 4\,\Big[s\,\mathcal{L}[y] - y(0)\Big] + 4\,\mathcal{L}[y] = \frac{6}{s^2+4}.$$
Reorder terms,
$$(s^2 - 4s + 4)\,\mathcal{L}[y] = (s-4)\,y(0) + y'(0) + \frac{6}{s^2+4}.$$
Introduce the initial conditions $y(0) = 1$ and $y'(0) = 1$,
$$(s^2 - 4s + 4)\,\mathcal{L}[y] = s - 3 + \frac{6}{s^2+4}.$$
Solve this algebraic equation for $\mathcal{L}[y]$, that is,
$$\mathcal{L}[y] = \frac{(s-3)}{(s^2-4s+4)} + \frac{6}{(s^2-4s+4)(s^2+4)}.$$
From the Example above we know that $s^2 - 4s + 4 = (s-2)^2$, so we obtain
$$\mathcal{L}[y] = \frac{1}{s-2} - \frac{1}{(s-2)^2} + \frac{6}{(s-2)^2(s^2+4)}. \tag{3.2.4}$$

From the previous example we know that

$$\mathcal{L}[e^{2t} - te^{2t}] = \frac{1}{s-2} - \frac{1}{(s-2)^2}. \tag{3.2.5}$$

We know use *partial fractions* to simplify the third term on the right hand side of Eq. (3.2.4). The appropriate partial fraction decomposition for this term is the following: Find constants a, b, c, d, such that

$$\frac{6}{(s-2)^2\,(s^2+4)} = \frac{as+b}{s^2+4} + \frac{c}{(s-2)} + \frac{d}{(s-2)^2}$$

Take common denominator on the right hand side above, and one obtains the system

$$\begin{aligned} a + c &= 0,\\ -4a + b - 2c + d &= 0,\\ 4a - 4b + 4c &= 0,\\ 4b - 8c + 4d &= 6. \end{aligned}$$

The solution for this linear system of equations is the following:

$$a = \frac{3}{8}, \qquad b = 0, \qquad c = -\frac{3}{8}, \qquad d = \frac{3}{4}.$$

Therefore,

$$\frac{6}{(s-2)^2\,(s^2+4)} = \frac{3}{8}\,\frac{s}{s^2+4} - \frac{3}{8}\,\frac{1}{(s-2)} + \frac{3}{4}\,\frac{1}{(s-2)^2}$$

We can rewrite this expression above in terms of the Laplace transforms given in Table 1, in Sect. **??**, as follows,

$$\frac{6}{(s-2)^2\,(s^2+4)} = \frac{3}{8}\,\mathcal{L}[\cos(2t)] - \frac{3}{8}\,\mathcal{L}[e^{2t}] + \frac{3}{4}\,\mathcal{L}[te^{2t}],$$

and using the linearity of the Laplace transform,

$$\frac{6}{(s-2)^2\,(s^2+4)} = \mathcal{L}\Big[\frac{3}{8}\,\cos(2t) - \frac{3}{8}\,e^{2t} + \frac{3}{4}\,te^{2t}\Big]. \tag{3.2.6}$$

Finally, introducing Eqs. (3.2.5) and (3.2.6) into Eq. (3.2.4) we obtain

$$\mathcal{L}[y(t)] = \mathcal{L}\Big[(1-t)\,e^{2t} + \frac{3}{8}\,(-1+2t)\,e^{2t} + \frac{3}{8}\,\cos(2t)\Big].$$

Since the Laplace transform is an invertible transformation, we conclude that

$$y(t) = (1-t)\,e^{2t} + \frac{3}{8}\,(2t-1)\,e^{2t} + \frac{3}{8}\,\cos(2t).$$

◁

3.2.4. Higher Order IVP. The Laplace transform method can be used with linear differential equations of higher order than second order, as long as the equation coefficients are constant. Below we show how we can solve a fourth order equation.

Example 3.2.6. Use the Laplace transform to find the solution y to the initial value problem

$$y^{(4)} - 4y = 0, \qquad \begin{aligned} y(0) &= 1, & y'(0) &= 0,\\ y''(0) &= -2, & y'''(0) &= 0. \end{aligned}$$

Solution: Compute the Laplace transform of the differential equation,

$$\mathcal{L}[y^{(4)} - 4y] = \mathcal{L}[0] = 0.$$

The Laplace transform is a linear operation,

$$\mathcal{L}[y^{(4)}] - 4\,\mathcal{L}[y] = 0,$$

and the Laplace transform relates derivatives with multiplications,

$$\Big[s^4\,\mathcal{L}[y] - s^3\,y(0) - s^2\,y'(0) - s\,y''(0) - y'''(0)\Big] - 4\,\mathcal{L}[y] = 0.$$

From the initial conditions we get

$$\Big[s^4\,\mathcal{L}[y] - s^3 - 0 + 2s - 0\Big] - 4\,\mathcal{L}[y] = 0 \quad \Rightarrow \quad (s^4-4)\mathcal{L}[y] = s^3 - 2s \quad \Rightarrow \quad \mathcal{L}[y] = \frac{(s^3-2s)}{(s^4-4)}.$$

In this case we are lucky, because

$$\mathcal{L}[y] = \frac{s(s^2-2)}{(s^2-2)(s^2+2)} = \frac{s}{(s^2+2)}.$$

Since

$$\mathcal{L}[\cos(at)] = \frac{s}{s^2+a^2},$$

we get that

$$\mathcal{L}[y] = \mathcal{L}[\cos(\sqrt{2}t)] \quad \Rightarrow \quad y(t) = \cos(\sqrt{2}t).$$

$\lhd$

3.2.5. Exercises.

3.2.1.- .

3.2.2.- .

3.3. Discontinuous Sources

The Laplace transform method can be used to solve linear differential equations with discontinuous sources. In this section we review the simplest discontinuous function—the step function—and we use steps to construct more general piecewise continuous functions. Then, we compute the Laplace transform of a step function. But the main result in this section are the translation identities, Theorem 3.3.3. These identities, together with the Laplace transform table in § 3.1, can be very useful to solve differential equations with discontinuous sources.

3.3.1. Step Functions.

We start with a definition of a step function.

Definition 3.3.1. *The* ***step function*** *at $t = 0$ is denoted by u and given by*

$$u(t) = \begin{cases} 0 & t < 0, \\ 1 & t \geqslant 0. \end{cases} \tag{3.3.1}$$

Example 3.3.1. Graph the step u, $u_c(t) = u(t - c)$, and $u_{-c}(t) = u(t + c)$, for $c > 0$.

Solution: The step function u and its right and left translations are plotted in Fig. 1.

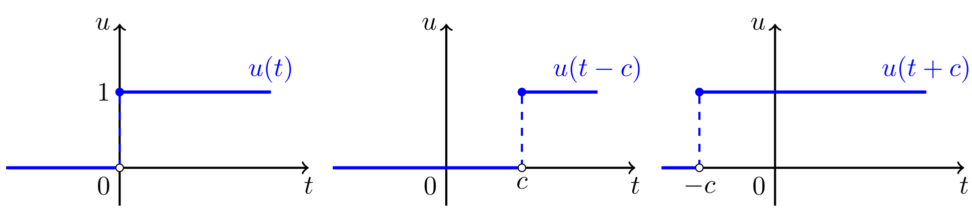


FIGURE 1. The graph of the step function given in Eq. (3.3.1), a right and a left translation by a constant $c > 0$, respectively, of this step function.

◁

Recall that given a function with values $f(t)$ and a positive constant c, then $f(t-c)$ and $f(t+c)$ are the function values of the right translation and the left translation, respectively, of the original function f. In Fig. 2 we plot the graph of functions $f(t) = e^{at}$, $g(t) = u(t)\, e^{at}$ and their respective right translations by $c > 0$.

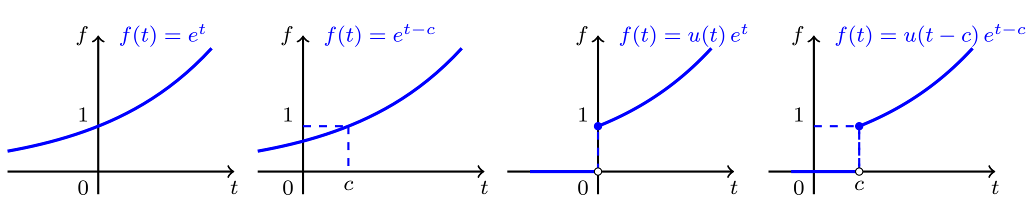


FIGURE 2. The function $f(t) = e^t$, its right translation by $c > 0$, the function $f(t) = u(t)\, e^{at}$ and its right translation by c.

Right and left translations of step functions are useful to construct bump functions.

Example 3.3.2. Graph the bump function $b(t) = u(t-a) - u(t-b)$, where $a < b$.

Solution: The bump function we need to graph is

$$b(t) = u(t-a) - u(t-b) \quad \Leftrightarrow \quad b(t) = \begin{cases} 0 & t < a, \\ 1 & a \leqslant t < b \\ 0 & t \geqslant b. \end{cases} \tag{3.3.2}$$

The graph of a bump function is given in Fig. 3, constructed from two step functions. Step and bump functions are useful to construct more general piecewise continuous functions.

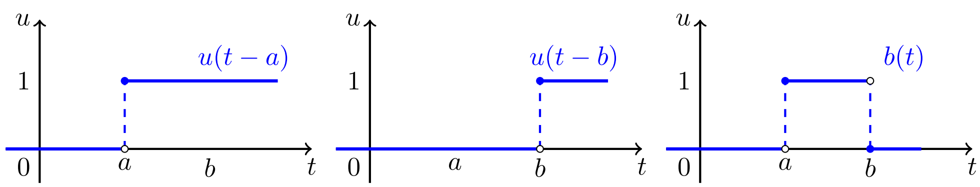


FIGURE 3. A bump function b constructed with translated step functions.

◁

Example 3.3.3. Graph the function

$$f(t) = \big[u(t-1) - u(t-2)\big]\, e^{at}.$$

Solution: Recall that the function

$$b(t) = u(t-1) - u(t-2),$$

is a bump function with sides at $t = 1$ andf $t = 2$. Then, the function

$$f(t) = b(t)\, e^{at},$$

is nonzero where b is nonzero, that is on $[1, 2)$, and on that domain it takes values e^{at}. The graph of f is given in Fig. 4. ◁

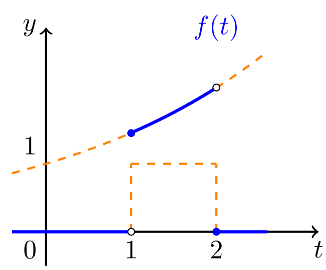


FIGURE 4. Function f.

3.3.2. The Laplace Transform of Steps.

We compute the Laplace transform of a step function using the definition of the Laplace transform.

Theorem 3.3.2. *For every number $c \in \mathbb{R}$ and and every $s > 0$ holds*

$$\mathcal{L}[u(t-c)] = \begin{cases} \dfrac{e^{-cs}}{s} & \text{for} \quad c \geqslant 0, \\ \dfrac{1}{s} & \text{for} \quad c < 0. \end{cases}$$

Proof of Theorem 3.3.2: Consider the case $c \geqslant 0$. The Laplace transform is

$$\mathcal{L}[u(t-c)] = \int_0^\infty e^{-st} u(t-c)\, dt = \int_c^\infty e^{-st}\, dt,$$

where we used that the step function vanishes for $t < c$. Now compute the improper integral,

$$\mathcal{L}[u(t-c)] = \lim_{N\to\infty} -\frac{1}{s}\big(e^{-Ns} - e^{-cs}\big) = \frac{e^{-cs}}{s} \quad \Rightarrow \quad \mathcal{L}[u(t-c)] = \frac{e^{-cs}}{s}.$$

Consider now the case of $c < 0$. The step function is identically equal to one in the domain of integration of the Laplace transform, which is $[0,\infty)$, hence

$$\mathcal{L}[u(t-c)] = \int_0^\infty e^{-st} u(t-c)\,dt = \int_0^\infty e^{-st}\,dt = \mathcal{L}[1] = \frac{1}{s}.$$

This establishes the Theorem. □

Example 3.3.4. Compute $\mathcal{L}[3\,u(t-2)]$.

Solution: The Laplace transform is a linear operation, so

$$\mathcal{L}[3\,u(t-2)] = 3\,\mathcal{L}[u(t-2)],$$

and the Theorem 3.3.2 above implies that $\mathcal{L}[3\,u(t-2)] = \dfrac{3\,e^{-2s}}{s}$. ◁

Remarks:

(a) The LT is an invertible transformation in the set of functions we work in our class.

(b) $\mathcal{L}[f] = F \quad \Leftrightarrow \quad \mathcal{L}^{-1}[F] = f$.

Example 3.3.5. Compute $\mathcal{L}^{-1}\Big[\dfrac{e^{-3s}}{s}\Big]$.

Solution: Theorem 3.3.2 says that $\dfrac{e^{-3s}}{s} = \mathcal{L}[u(t-3)]$, so $\mathcal{L}^{-1}\Big[\dfrac{e^{-3s}}{s}\Big] = u(t-3)$. ◁

3.3.3. Translation Identities.

We now introduce two properties relating the Laplace transform and translations. The first property relates the Laplace transform of a translation with a multiplication by an exponential. The second property can be thought as the inverse of the first one.

Theorem 3.3.3 (Translation Identities)**.** *If $\mathcal{L}[f(t)](s)$ exists for $s > a$, then*

$$\mathcal{L}[u(t-c)f(t-c)] = e^{-cs}\,\mathcal{L}[f(t)], \qquad s > a, \qquad c \geqslant 0 \tag{3.3.3}$$

$$\mathcal{L}[e^{ct} f(t)] = \mathcal{L}[f(t)](s-c), \qquad s > a + c, \qquad c \in \mathbb{R}. \tag{3.3.4}$$

Example 3.3.6. Take $f(t) = \cos(t)$ and write the equations given the Theorem above.

Solution:

$$\mathcal{L}[\cos(t)] = \frac{s}{s^2+1} \quad \Rightarrow \quad \begin{cases} \mathcal{L}[u(t-c)\,\cos(t-c)] = e^{-cs}\,\dfrac{s}{s^2+1} \\[2mm] \mathcal{L}[e^{ct}\cos(t)] = \dfrac{(s-c)}{(s-c)^2+1}. \end{cases}$$

◁

Remarks:

(a) We can highlight the main idea in the theorem above as follows:

$$\mathcal{L}\big[\text{right-translation } (uf)\big] = (\exp)\,\big(\mathcal{L}[f]\big),$$
$$\mathcal{L}\big[(\exp)\,(f)\big] = \text{translation}\big(\mathcal{L}[f]\big).$$

(b) Denoting $F(s) = \mathcal{L}[f(t)]$, then an equivalent expression for Eqs. (3.3.3)-(3.3.4) is

$$\mathcal{L}[u(t-c)f(t-c)] = e^{-cs}\,F(s),$$
$$\mathcal{L}[e^{ct}f(t)] = F(s-c).$$

(c) The inverse form of Eqs. (3.3.3)-(3.3.4) is given by,

$$\mathcal{L}^{-1}[e^{-cs}\,F(s)] = u(t-c)f(t-c), \tag{3.3.5}$$
$$\mathcal{L}^{-1}[F(s-c)] = e^{ct}f(t). \tag{3.3.6}$$

(d) Eq. (3.3.4) holds for all $c \in \mathbb{R}$, while Eq. (3.3.3) holds only for $c \geqslant 0$.

(e) Show that in the case that $c < 0$ the following equation holds,

$$\mathcal{L}[u(t+|c|)f(t+|c|)] = e^{|c|s}\Big(\mathcal{L}[f(t)] - \int_0^{|c|} e^{-st}\,f(t)\,dt\Big).$$

Proof of Theorem 3.3.3: The proof is again based in a change of the integration variable. We start with Eq. (3.3.3), as follows,

$$\begin{aligned}
\mathcal{L}[u(t-c)f(t-c)] &= \int_0^\infty e^{-st}u(t-c)f(t-c)\,dt\\
&= \int_c^\infty e^{-st}f(t-c)\,dt, \qquad \tau = t-c, \quad d\tau = dt, \quad c \geqslant 0,\\
&= \int_0^\infty e^{-s(\tau+c)}f(\tau)\,d\tau\\
&= e^{-cs}\int_0^\infty e^{-s\tau}f(\tau)\,d\tau\\
&= e^{-cs}\,\mathcal{L}[f(t)], \qquad s > a.
\end{aligned}$$

The proof of Eq. (3.3.4) is a bit simpler, since

$$\mathcal{L}\big[e^{ct}f(t)\big] = \int_0^\infty e^{-st}e^{ct}f(t)\,dt = \int_0^\infty e^{-(s-c)t}f(t)\,dt = \mathcal{L}[f(t)](s-c),$$

which holds for $s - c > a$. This establishes the Theorem. □

Example 3.3.7. Compute $\mathcal{L}\big[u(t-2)\,\sin(a(t-2))\big]$.

Solution: Both $\mathcal{L}[\sin(at)] = \dfrac{a}{s^2+a^2}$ and $\mathcal{L}[u(t-c)f(t-c)] = e^{-cs}\,\mathcal{L}[f(t)]$ imply

$$\mathcal{L}\big[u(t-2)\,\sin(a(t-2))\big] = e^{-2s}\,\mathcal{L}[\sin(at)] = e^{-2s}\,\frac{a}{s^2+a^2}.$$

We conclude: $\mathcal{L}\big[u(t-2)\,\sin(a(t-2))\big] = \dfrac{a\,e^{-2s}}{s^2+a^2}$. ◁

Example 3.3.8. Compute $\mathcal{L}\big[e^{3t}\,\sin(at)\big]$.

Solution: Since $\mathcal{L}[e^{ct}f(t)] = \mathcal{L}[f](s-c)$, then we get

$$\mathcal{L}\big[e^{3t}\,\sin(at)\big] = \frac{a}{(s-3)^2+a^2}, \qquad s > 3.$$

$\lhd$

Example 3.3.9. Compute both $\mathcal{L}\big[u(t-2)\,\cos\big(a(t-2)\big)\big]$ and $\mathcal{L}\big[e^{3t}\,\cos(at)\big]$.

Solution: Since $\mathcal{L}\big[\cos(at)\big] = \dfrac{s}{s^2+a^2}$, then

$$\mathcal{L}\big[u(t-2)\,\cos\big(a(t-2)\big)\big] = e^{-2s}\,\frac{s}{(s^2+a^2)}, \qquad \mathcal{L}\big[e^{3t}\,\cos(at)\big] = \frac{(s-3)}{(s-3)^2+a^2}.$$

$\lhd$

Example 3.3.10. Find the Laplace transform of the function

$$f(t) = \begin{cases} 0 & t < 1, \\ (t^2-2t+2) & t \geqslant 1. \end{cases} \tag{3.3.7}$$

Solution: The idea is to rewrite function f so we can use the Laplace transform Table 1, in § 3.1 to compute its Laplace transform. Since the function f vanishes for all $t<1$, we use step functions to write f as

$$f(t) = u(t-1)(t^2-2t+2).$$

Now, we would like to use the translation identity in Eq. (3.3.3), but we can not, since the step is a function of $(t-1)$ while the polynomial is a function of t. We need to rewrite the polynomial as a function of $(t-1)$. So we add and subtract 1 in the appropriate places

$$t^2-2t+2 = \big((t-1+1)^2 - 2(t-1+1)+2\big).$$

Recall the identity $(a+b)^2 = a^2+2ab+b^2$, and use it in the quadratic term above for $a=(t-1)$ and $b=1$. We get

$$(t-1+1)^2 = (t-1)^2 + 2(t-1) + 1^2.$$

This identity into the polynomial above implies

$$t^2-2t+2 = \big((t-1)^2 + 2(t-1) + 1 - 2(t-1) - 2 + 2\big) \quad\Rightarrow\quad t^2-2t+2 = (t-1)^2+1.$$

The polynomial is a parabola t^2 translated to the right and up by one. This is a discontinuous function, as it can be seen in Fig. 5.

So the function f can be written as follows,

$$f(t) = u(t-1)\,(t-1)^2 + u(t-1).$$

Since we know that $\mathcal{L}[t^2] = \dfrac{2}{s^3}$, then Eq. (3.3.3) implies

$$\mathcal{L}[f(t)] = \mathcal{L}[u(t-1)\,(t-1)^2] + \mathcal{L}[u(t-1)] = e^{-s}\,\frac{2}{s^3} + e^{-s}\,\frac{1}{s}$$

so we get

$$\mathcal{L}[f(t)] = \frac{e^{-s}}{s^3}\big(2+s^2\big).$$

$\lhd$

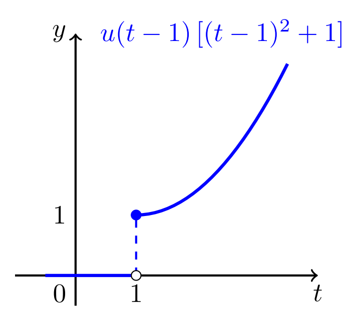


FIGURE 5. Function f given in Eq. (3.3.7).

Example 3.3.11. Find the function f such that $\mathcal{L}[f(t)] = \dfrac{e^{-4s}}{s^2+5}$.

Solution: Notice that

$$\mathcal{L}[f(t)] = e^{-4s}\Big(\frac{1}{s^2+5}\Big) \quad \Rightarrow \quad \mathcal{L}[f(t)] = \frac{1}{\sqrt{5}}\, e^{-4s}\Big(\frac{\sqrt{5}}{s^2+\big(\sqrt{5}\big)^2}\Big).$$

Recall that $\mathcal{L}[\sin(at)] = \dfrac{a}{(s^2+a^2)}$, then

$$\mathcal{L}[f(t)] = \frac{1}{\sqrt{5}}\, e^{-4s}\,\mathcal{L}[\sin(\sqrt{5}t)].$$

But the translation identity

$$e^{-cs}\,\mathcal{L}[f(t)] = \mathcal{L}[u(t-c)f(t-c)]$$

implies

$$\mathcal{L}[f(t)] = \frac{1}{\sqrt{5}}\,\mathcal{L}\big[u(t-4)\sin\big(\sqrt{5}\,(t-4)\big)\big],$$

hence we obtain

$$f(t) = \frac{1}{\sqrt{5}}\,u(t-4)\sin\big(\sqrt{5}\,(t-4)\big).$$

◁

Example 3.3.12. Find the function $f(t)$ such that $\mathcal{L}[f(t)] = \dfrac{(s-1)}{(s-2)^2+3}$.

Solution: We first rewrite the right-hand side above as follows,

$$\begin{aligned}
\mathcal{L}[f(t)] &= \frac{(s-1-1+1)}{(s-2)^2+3}\\
&= \frac{(s-2)}{(s-2)^2+3} + \frac{1}{(s-2)^2+3}\\
&= \frac{(s-2)}{(s-2)^2+\big(\sqrt{3}\big)^2} + \frac{1}{\sqrt{3}}\,\frac{\sqrt{3}}{(s-2)^2+\big(\sqrt{3}\big)^2}\\
&= \mathcal{L}[\cos(\sqrt{3}\,t)](s-2) + \frac{1}{\sqrt{3}}\,\mathcal{L}[\sin(\sqrt{3}\,t)](s-2).
\end{aligned}$$

But the translation identity $\mathcal{L}[f(t)](s-c) = \mathcal{L}[e^{ct}f(t)]$ implies

$$\mathcal{L}[f(t)] = \mathcal{L}\big[e^{2t}\cos\big(\sqrt{3}\,t\big)\big] + \frac{1}{\sqrt{3}}\mathcal{L}\big[e^{2t}\sin\big(\sqrt{3}\,t\big)\big].$$

So, we conclude that

$$f(t) = \frac{e^{2t}}{\sqrt{3}}\Big[\sqrt{3}\,\cos\big(\sqrt{3}\,t\big) + \sin\big(\sqrt{3}\,t\big)\Big].$$

◁

Example 3.3.13. Find $\mathcal{L}^{-1}\Big[\dfrac{2e^{-3s}}{s^2-4}\Big]$.

Solution: Since $\mathcal{L}^{-1}\Big[\dfrac{a}{s^2-a^2}\Big] = \sinh(at)$ and $\mathcal{L}^{-1}\big[e^{-cs}\,\hat{f}(s)\big] = u(t-c)\,f(t-c)$, then

$$\mathcal{L}^{-1}\Big[\frac{2e^{-3s}}{s^2-4}\Big] = \mathcal{L}^{-1}\Big[e^{-3s}\,\frac{2}{s^2-4}\Big] \quad \Rightarrow \quad \mathcal{L}^{-1}\Big[\frac{2e^{-3s}}{s^2-4}\Big] = u(t-3)\,\sinh\big(2(t-3)\big).$$

◁

Example 3.3.14. Find a function f such that $\mathcal{L}[f(t)] = \dfrac{e^{-2s}}{s^2+s-2}$.

Solution: Since the right hand side above does not appear in the Laplace transform Table in § 3.1, we need to simplify it in an appropriate way. The plan is to rewrite the denominator of the rational function $1/(s^2+s-2)$, so we can use partial fractions to simplify this rational function. We first find out whether this denominator has real or complex roots:

$$s_\pm = \frac{1}{2}\big[-1 \pm \sqrt{1+8}\big] \quad \Rightarrow \quad \begin{cases} s_+ = 1, \\ s_- = -2. \end{cases}$$

We are in the case of real roots, so we rewrite

$$s^2 + s - 2 = (s-1)\,(s+2).$$

The partial fraction decomposition in this case is given by

$$\frac{1}{(s-1)\,(s+2)} = \frac{a}{(s-1)} + \frac{b}{(s+2)} = \frac{(a+b)\,s + (2a-b)}{(s-1)\,(s+2)} \quad \Rightarrow \quad \begin{cases} a+b=0, \\ 2a-b=1. \end{cases}$$

The solution is $a = 1/3$ and $b = -1/3$, so we arrive to the expression

$$\mathcal{L}[f(t)] = \frac{1}{3}\,e^{-2s}\,\frac{1}{s-1} - \frac{1}{3}\,e^{-2s}\,\frac{1}{s+2}.$$

Recalling that

$$\mathcal{L}[e^{at}] = \frac{1}{s-a},$$

and Eq. (3.3.3) we obtain the equation

$$\mathcal{L}[f(t)] = \frac{1}{3}\,\mathcal{L}\big[u(t-2)\,e^{(t-2)}\big] - \frac{1}{3}\,\mathcal{L}\big[u(t-2)\,e^{-2(t-2)}\big]$$

which leads to the conclusion:

$$f(t) = \frac{1}{3}\,u(t-2)\Big[e^{(t-2)} - e^{-2(t-2)}\Big].$$

◁

3.3.4. Solving Differential Equations.

The last three examples in this section show how to use the methods presented above to solve differential equations with discontinuous source functions.

Example 3.3.15. Use the Laplace transform to find the solution of the initial value problem

$$y' + 2y = u(t-4), \qquad y(0) = 3.$$

Solution: We compute the Laplace transform of the whole equation,

$$\mathcal{L}[y'] + 2\,\mathcal{L}[y] = \mathcal{L}[u(t-4)] = \frac{e^{-4s}}{s}.$$

From the previous section we know that

$$\big[s\,\mathcal{L}[y] - y(0)\big] + 2\,\mathcal{L}[y] = \frac{e^{-4s}}{s} \quad \Rightarrow \quad (s+2)\,\mathcal{L}[y] = y(0) + \frac{e^{-4s}}{s}.$$

We introduce the initial condition $y(0) = 3$ into equation above,

$$\mathcal{L}[y] = \frac{3}{(s+2)} + e^{-4s}\,\frac{1}{s(s+2)} \quad \Rightarrow \quad \mathcal{L}[y] = 3\,\mathcal{L}\big[e^{-2t}\big] + e^{-4s}\,\frac{1}{s(s+2)}.$$

We need to invert the Laplace transform on the last term on the right hand side in equation above. We use the partial fraction decomposition on the rational function above, as follows

$$\frac{1}{s(s+2)} = \frac{a}{s} + \frac{b}{(s+2)} = \frac{a(s+2)+bs}{s(s+2)} = \frac{(a+b)\,s+(2a)}{s(s+2)} \quad \Rightarrow \quad \begin{cases} a+b=0, \\ \quad 2a=1. \end{cases}$$

We conclude that $a = 1/2$ and $b = -1/2$, so

$$\frac{1}{s(s+2)} = \frac{1}{2}\Big[\frac{1}{s} - \frac{1}{(s+2)}\Big].$$

We then obtain

$$\begin{aligned} \mathcal{L}[y] &= 3\,\mathcal{L}\big[e^{-2t}\big] + \frac{1}{2}\Big[e^{-4s}\,\frac{1}{s} - e^{-4s}\,\frac{1}{(s+2)}\Big] \\ &= 3\,\mathcal{L}\big[e^{-2t}\big] + \frac{1}{2}\Big(\mathcal{L}[u(t-4)] - \mathcal{L}\big[u(t-4)\,e^{-2(t-4)}\big]\Big). \end{aligned}$$

Hence, we conclude that

$$y(t) = 3e^{-2t} + \frac{1}{2}\,u(t-4)\Big[1 - e^{-2(t-4)}\Big].$$

◁

Example 3.3.16. Use the Laplace transform to find the solution to the initial value problem

$$y'' + y' + \frac{5}{4}\,y = b(t), \qquad y(0) = 0, \qquad y'(0) = 0, \qquad b(t) = \begin{cases} 1 & 0 \leqslant t < \pi \\ 0 & \quad t \geqslant \pi. \end{cases} \tag{3.3.8}$$

Solution: From Fig. 6, the source function b can be written as

$$b(t) = u(t) - u(t-\pi).$$

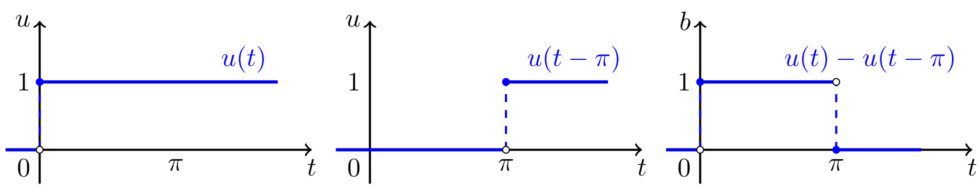


FIGURE 6. The graph of the u, its translation and b as given in Eq. (3.3.8).

The last expression for b is particularly useful to find its Laplace transform,

$$\mathcal{L}[b(t)] = \mathcal{L}[u(t)] - \mathcal{L}[u(t-\pi)] = \frac{1}{s} + e^{-\pi s}\,\frac{1}{s} \quad \Rightarrow \quad \mathcal{L}[b(t)] = (1 - e^{-\pi s})\,\frac{1}{s}.$$

Now Laplace transform the whole equation,

$$\mathcal{L}[y''] + \mathcal{L}[y'] + \frac{5}{4}\,\mathcal{L}[y] = \mathcal{L}[b].$$

Since the initial condition are $y(0) = 0$ and $y'(0) = 0$, we obtain

$$\Big(s^2 + s + \frac{5}{4}\Big)\,\mathcal{L}[y] = \big(1 - e^{-\pi s}\big)\,\frac{1}{s} \quad \Rightarrow \quad \mathcal{L}[y] = \big(1 - e^{-\pi s}\big)\,\frac{1}{s\,\Big(s^2 + s + \frac{5}{4}\Big)}.$$

Introduce the function

$$H(s) = \frac{1}{s\,\Big(s^2 + s + \frac{5}{4}\Big)} \quad \Rightarrow \quad y(t) = \mathcal{L}^{-1}[H(s)] - \mathcal{L}^{-1}[e^{-\pi s}\,H(s)].$$

That is, we only need to find the inverse Laplace transform of H. We use partial fractions to simplify the expression of H. We first find out whether the denominator has real or complex roots:

$$s^2 + s + \frac{5}{4} = 0 \quad \Rightarrow \quad s_\pm = \frac{1}{2}\big[-1 \pm \sqrt{1-5}\big],$$

so the roots are complex valued. An appropriate partial fraction decomposition is

$$H(s) = \frac{1}{s\left(s^2+s+\frac{5}{4}\right)} = \frac{a}{s} + \frac{(bs+c)}{\left(s^2+s+\frac{5}{4}\right)}$$

Therefore, we get

$$1 = a\left(s^2+s+\frac{5}{4}\right) + s\,(bs+c) = (a+b)\,s^2 + (a+c)\,s + \frac{5}{4}a.$$

This equation implies that a, b, and c, satisfy the equations

$$a+b=0, \quad a+c=0, \quad \frac{5}{4}\,a = 1.$$

The solution is, $a = \dfrac{4}{5}$, $b = -\dfrac{4}{5}$, $c = -\dfrac{4}{5}$. Hence, we have found that,

$$H(s) = \frac{1}{\left(s^2+s+\frac{5}{4}\right)s} = \frac{4}{5}\left[\frac{1}{s} - \frac{(s+1)}{\left(s^2+s+\frac{5}{4}\right)}\right]$$

Complete the square in the denominator,

$$s^2+s+\frac{5}{4} = \left[s^2 + 2\Big(\frac{1}{2}\Big)s + \frac{1}{4}\right] - \frac{1}{4} + \frac{5}{4} = \Big(s+\frac{1}{2}\Big)^2 + 1.$$

Replace this expression in the definition of H, that is,

$$H(s) = \frac{4}{5}\left[\frac{1}{s} - \frac{(s+1)}{\left[\left(s+\frac{1}{2}\right)^2+1\right]}\right]$$

Rewrite the polynomial in the numerator,

$$(s+1) = \Big(s+\frac{1}{2}+\frac{1}{2}\Big) = \Big(s+\frac{1}{2}\Big) + \frac{1}{2},$$

hence we get

$$H(s) = \frac{4}{5}\left[\frac{1}{s} - \frac{\left(s+\frac{1}{2}\right)}{\left[\left(s+\frac{1}{2}\right)^2+1\right]} - \frac{1}{2}\,\frac{1}{\left[\left(s+\frac{1}{2}\right)^2+1\right]}\right].$$

Use the Laplace transform table to get $H(s)$ equal to

$$H(s) = \frac{4}{5}\Big[\mathcal{L}[1] - \mathcal{L}\big[e^{-t/2}\cos(t)\big] - \frac{1}{2}\mathcal{L}[e^{-t/2}\sin(t)]\Big],$$

equivalently

$$H(s) = \mathcal{L}\Big[\frac{4}{5}\Big(1 - e^{-t/2}\cos(t) - \frac{1}{2}e^{-t/2}\sin(t)\Big)\Big].$$

Denote

$$h(t) = \frac{4}{5}\Big[1 - e^{-t/2}\cos(t) - \frac{1}{2}e^{-t/2}\sin(t)\Big]. \quad \Rightarrow \quad H(s) = \mathcal{L}[h(t)].$$

Recalling $\mathcal{L}[y(t)] = H(s) + e^{-\pi s}\,H(s)$, we obtain $\mathcal{L}[y(t)] = \mathcal{L}[h(t)] + e^{-\pi s}\,\mathcal{L}[h(t)]$, that is,

$$y(t) = h(t) + u(t-\pi)h(t-\pi).$$

◁

Example 3.3.17. Use the Laplace transform to find the solution to the initial value problem

$$y'' + y' + \frac{5}{4}\,y = g(t), \qquad y(0) = 0, \qquad y'(0) = 0, \qquad g(t) = \begin{cases} \sin(t) & 0 \leqslant t < \pi \\ 0 & t \geqslant \pi. \end{cases} \tag{3.3.9}$$

Solution: From Fig. 7, the source function g can be written as the following product,

$$g(t) = \big[u(t) - u(t-\pi)\big]\, \sin(t),$$

since $u(t) - u(t-\pi)$ is a box function, taking value one in the interval $[0, \pi]$ and zero on the complement. Finally, notice that the equation $\sin(t) = -\sin(t-\pi)$ implies that the function g can be expressed as follows,

$$g(t) = u(t)\, \sin(t) - u(t-\pi)\, \sin(t) \quad \Rightarrow \quad g(t) = u(t)\, \sin(t) + u(t-\pi)\, \sin(t-\pi).$$

The last expression for g is particularly useful to find its Laplace transform,

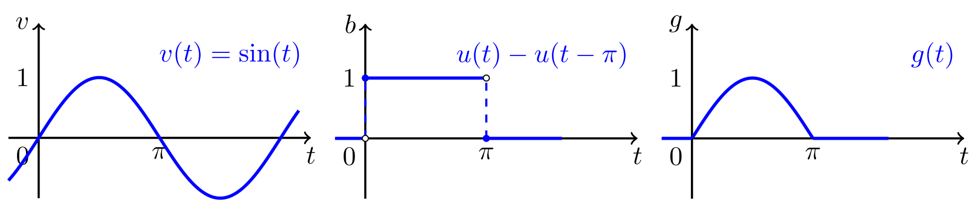


FIGURE 7. The graph of the sine function, a square function $u(t) - u(t-\pi)$ and the source function g given in Eq. (3.3.9).

$$\mathcal{L}[g(t)] = \frac{1}{(s^2+1)} + e^{-\pi s}\, \frac{1}{(s^2+1)}.$$

With this last transform is not difficult to solve the differential equation. As usual, Laplace transform the whole equation,

$$\mathcal{L}[y''] + \mathcal{L}[y'] + \frac{5}{4}\, \mathcal{L}[y] = \mathcal{L}[g].$$

Since the initial condition are $y(0) = 0$ and $y'(0) = 0$, we obtain

$$\Big(s^2 + s + \frac{5}{4}\Big)\, \mathcal{L}[y] = \big(1 + e^{-\pi s}\big)\, \frac{1}{(s^2+1)} \quad \Rightarrow \quad \mathcal{L}[y] = \big(1 + e^{-\pi s}\big)\, \frac{1}{\Big(s^2 + s + \frac{5}{4}\Big)\, (s^2+1)}.$$

Introduce the function

$$H(s) = \frac{1}{\Big(s^2 + s + \frac{5}{4}\Big)\, (s^2+1)} \quad \Rightarrow \quad y(t) = \mathcal{L}^{-1}[H(s)] + \mathcal{L}^{-1}[e^{-\pi s}\, H(s)].$$

That is, we only need to find the Inverse Laplace transform of H. We use partial fractions to simplify the expression of H. We first find out whether the denominator has real or complex roots:

$$s^2 + s + \frac{5}{4} = 0 \quad \Rightarrow \quad s_\pm = \frac{1}{2}\big[-1 \pm \sqrt{1-5}\big],$$

so the roots are complex valued. An appropriate partial fraction decomposition is

$$H(s) = \frac{1}{(s^2 + s + \frac{5}{4})\, (s^2+1)} = \frac{(as+b)}{(s^2 + s + \frac{5}{4})} + \frac{(cs+d)}{(s^2+1)}.$$

Therefore, we get

$$1 = (as+b)(s^2+1) + (cs+d)\Big(s^2 + s + \frac{5}{4}\Big),$$

equivalently,

$$1 = (a + c)\, s^3 + (b + c + d)\, s^2 + \Big(a + \frac{5}{4}\, c + d\Big)\, s + \Big(b + \frac{5}{4}\, d\Big).$$

This equation implies that a, b, c, and d, are solutions of

$$a + c = 0, \quad b + c + d = 0, \quad a + \frac{5}{4}\, c + d = 0, \quad b + \frac{5}{4}\, d = 1.$$

Here is the solution to this system:

$$a = \frac{16}{17}, \qquad b = \frac{12}{17}, \qquad c = -\frac{16}{17}, \qquad d = \frac{4}{17}.$$

We have found that,

$$H(s) = \frac{4}{17}\,\Big[\frac{(4s+3)}{\big(s^2 + s + \frac{5}{4}\big)} + \frac{(-4s+1)}{(s^2+1)}\Big].$$

Complete the square in the denominator,

$$s^2 + s + \frac{5}{4} = \Big[s^2 + 2\Big(\frac{1}{2}\Big)\, s + \frac{1}{4}\Big] - \frac{1}{4} + \frac{5}{4} = \Big(s + \frac{1}{2}\Big)^2 + 1.$$

$$H(s) = \frac{4}{17}\,\Big[\frac{(4s+3)}{\big[\big(s+\frac{1}{2}\big)^2 + 1\big]} + \frac{(-4s+1)}{(s^2+1)}\Big].$$

Rewrite the polynomial in the numerator,

$$(4s + 3) = 4\Big(s + \frac{1}{2} - \frac{1}{2}\Big) + 3 = 4\Big(s + \frac{1}{2}\Big) + 1,$$

hence we get

$$H(s) = \frac{4}{17}\,\Big[4\,\frac{\big(s+\frac{1}{2}\big)}{\big[\big(s+\frac{1}{2}\big)^2+1\big]} + \frac{1}{\big[\big(s+\frac{1}{2}\big)^2+1\big]} - 4\,\frac{s}{(s^2+1)} + \frac{1}{(s^2+1)}\Big].$$

Use the Laplace transform Table in 1 to get $H(s)$ equal to

$$H(s) = \frac{4}{17}\Big[4\,\mathcal{L}\big[e^{-t/2}\cos(t)\big] + \mathcal{L}\big[e^{-t/2}\sin(t)\big] - 4\,\mathcal{L}[\cos(t)] + \mathcal{L}[\sin(t)]\Big],$$

equivalently

$$H(s) = \mathcal{L}\Big[\frac{4}{17}\Big(4e^{-t/2}\cos(t) + e^{-t/2}\sin(t) - 4\cos(t) + \sin(t)\Big)\Big].$$

Denote

$$h(t) = \frac{4}{17}\Big[4e^{-t/2}\cos(t) + e^{-t/2}\sin(t) - 4\cos(t) + \sin(t)\Big] \quad \Rightarrow \quad H(s) = \mathcal{L}[h(t)].$$

Recalling $\mathcal{L}[y(t)] = H(s) + e^{-\pi s}\, H(s)$, we obtain $\mathcal{L}[y(t)] = \mathcal{L}[h(t)] + e^{-\pi s}\,\mathcal{L}[h(t)]$, that is,

$$y(t) = h(t) + u(t-\pi)h(t-\pi).$$

◁

3.3.5. Exercises.

3.3.1.- .

3.3.2.- .

3.4. Generalized Sources

We introduce a generalized function—the Dirac delta. We define the Dirac delta as a limit $n \to \infty$ of a particular sequence of functions, $\{\delta_n\}$. We will see that this limit is a function on the domain $\mathbb{R} - \{0\}$, but it is not a function on $\mathbb{R}$. For that reason we call this limit a generalized function—the Dirac delta generalized function.

We will show that each element in the sequence $\{\delta_n\}$ has a Laplace transform, and this sequence of Laplace transforms $\{\mathcal{L}[\delta_n]\}$ has a limit as $n \to \infty$. We use this limit of Laplace transforms to define the Laplace transform of the Dirac delta.

We will solve differential equations having the Dirac delta generalized function as source. Such differential equations appear often when one describes physical systems with impulsive forces—forces acting on a very short time but transferring a finite momentum to the system. Dirac's delta is tailored to model impulsive forces.

3.4.1. Sequence of Functions and the Dirac Delta. A sequence of functions is a sequence whose elements are functions. If each element in the sequence is a continuous function, we say that this is a sequence of continuous functions. Given a sequence of functions $\{y_n\}$, we compute the $\lim_{n\to\infty} y_n(t)$ for a fixed t. The limit depends on t, so it is a function of t, and we write it as

$$\lim_{n\to\infty} y_n(t) = y(t).$$

The domain of the limit function y is smaller or equal to the domain of the y_n. The limit of a sequence of continuous functions may or may not be a continuous function.

Example 3.4.1. The limit of the sequence below is a continuous function,

$$\Big\{ f_n(t) = \sin\Big(\Big(1 + \frac{1}{n}\Big)t\Big)\Big\} \to \sin(t) \quad \text{as} \quad n \to \infty.$$

As usual in this section, the limit is computed for each fixed value of t. ◁

However, not every sequence of continuous functions has a continuous function as a limit.

Example 3.4.2. Consider now the following sequence, $\{u_n\}$, for $n \geqslant 1$,

$$u_n(t) = \begin{cases} 0, & t < 0 \\ nt, & 0 \leqslant t \leqslant \dfrac{1}{n} \\ 1, & t > \dfrac{1}{n}. \end{cases} \tag{3.4.1}$$

This is a sequence of continuous functions whose limit is a discontinuous function. From the few graphs in Fig. 8 we can see that the limit $n \to \infty$ of the sequence above is a step function, indeed, $\lim_{n\to\infty} u_n(t) = \tilde{u}(t)$, where

$$\tilde{u}(t) = \begin{cases} 0 & \text{for} \quad t \leqslant 0, \\ 1 & \text{for} \quad t > 0. \end{cases}$$

We used a tilde in the name $\tilde{u}$ because this step function is not the same we defined in the previous section. The step u in § 3.3 satisfied $u(0) = 1$. ◁

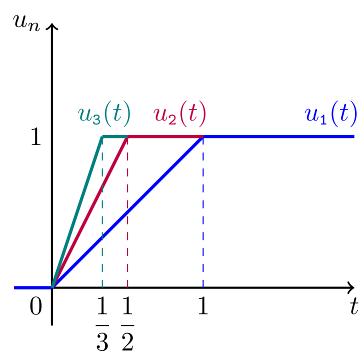


FIGURE 8. A few functions in the sequence $\{u_n\}$.

Exercise: Find a sequence $\{u_n\}$ so that its limit is the step function u defined in § 3.3.

Although every function in the sequence $\{u_n\}$ is continuous, the limit $\tilde{u}$ is a discontinuous function. It is not difficult to see that one can construct sequences of continuous functions having no limit at all. A similar situation happens when one considers sequences of piecewise discontinuous functions. In this case the limit could be a continuous function, a piecewise discontinuous function, or not a function at all.

We now introduce a particular sequence of piecewise discontinuous functions with domain $\mathbb{R}$ such that the limit as $n \to \infty$ does not exist for all values of the independent variable t. The limit of the sequence is not a function with domain $\mathbb{R}$. In this case, the limit is a new type of object that we will call Dirac's delta generalized function. Dirac's delta is the limit of a sequence of particular bump functions.

Definition 3.4.1. *The **Dirac delta** generalized function is the limit*

$$\delta(t) = \lim_{n\to\infty} \delta_n(t),$$

for every fixed $t \in \mathbb{R}$ of the sequence functions $\{\delta_n\}_{n=1}^{\infty}$,

$$\delta_n(t) = n\left[u(t) - u\big(t - \frac{1}{n}\big)\right]. \tag{3.4.2}$$

The sequence of bump functions introduced above can be rewritten as follows,

$$\delta_n(t) = \begin{cases} 0, & t < 0 \\ n, & 0 \leqslant t < \dfrac{1}{n} \\ 0, & t \geqslant \dfrac{1}{n}. \end{cases}$$

We then obtain the equivalent expression,

$$\delta(t) = \begin{cases} 0 & \text{for} \quad t \neq 0, \\ \infty & \text{for} \quad t = 0. \end{cases}$$

Remark: It can be shown that there exist infinitely many sequences $\{\tilde{\delta}_n\}$ such that their limit as $n \to \infty$ is Dirac's delta. For example, another sequence is

$$\tilde{\delta}_n(t) = n\left[u\big(t + \frac{1}{2n}\big) - u\big(t - \frac{1}{2n}\big)\right]$$
$$= \begin{cases} 0, & t < -\dfrac{1}{2n} \\ n, & -\dfrac{1}{2n} \leqslant t \leqslant \dfrac{1}{2n} \\ 0, & t > \dfrac{1}{2n}. \end{cases}$$

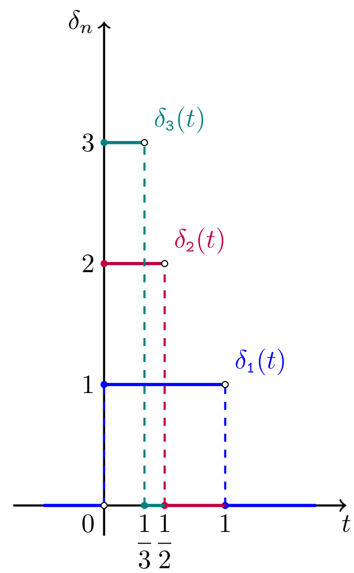


FIGURE 9. A few functions in the sequence $\{\delta_n\}$.

The Dirac delta generalized function is the function identically zero on the domain $\mathbb{R} - \{0\}$. Dirac's delta is not defined at $t = 0$, since the limit diverges at that point. If we shift each element in the sequence by a real number c, then we define

$$\delta(t - c) = \lim_{n\to\infty} \delta_n(t - c), \qquad c \in \mathbb{R}.$$

This shifted Dirac's delta is identically zero on $\mathbb{R} - \{c\}$ and diverges at $t = c$. If we shift the graphs given in Fig. 9 by any real number c, one can see that

$$\int_c^{c+1} \delta_n(t-c)\,dt = 1$$

for every $n \geqslant 1$. Therefore, the sequence of integrals is the constant sequence, $\{1, 1, \cdots\}$, which has a trivial limit, 1, as $n \to \infty$. This says that the divergence at $t = c$ of the sequence $\{\delta_n\}$ is of a very particular type. The area below the graph of the sequence elements is always the same. We can say that this property of the sequence provides the main defining property of the Dirac delta generalized function.

Using a limit procedure one can generalize several operations from a sequence to its limit. For example, translations, linear combinations, and multiplications of a function by a generalized function, integration and Laplace transforms.

Definition 3.4.2. *We introduce the following operations on the Dirac delta:*

$$f(t)\,\delta(t-c) + g(t)\,\delta(t-c) = \lim_{n\to\infty} \big[f(t)\,\delta_n(t-c) + g(t)\,\delta_n(t-c)\big],$$
$$\int_a^b \delta(t-c)\,dt = \lim_{n\to\infty} \int_a^b \delta_n(t-c)\,dt,$$
$$\mathcal{L}[\delta(t-c)] = \lim_{n\to\infty} \mathcal{L}[\delta_n(t-c)].$$

Remark: The notation in the definitions above could be misleading. In the left hand sides above we use the same notation as we use on functions, although Dirac's delta is not a function on $\mathbb{R}$. Take the integral, for example. When we integrate a function f, the integration symbol means "take a limit of Riemann sums", that is,

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} \sum_{i=0}^{n} f(x_i)\,\Delta x, \qquad x_i = a + i\,\Delta x, \qquad \Delta x = \frac{b-a}{n}.$$

However, when f is a generalized function in the sense of a limit of a sequence of functions $\{f_n\}$, then by the integration symbol we mean to compute a different limit,

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} \int_a^b f_n(t)\,dt.$$

We use the same symbol, the integration, to mean two different things, depending whether we integrate a function or a generalized function. This remark also holds for all the operations we introduce on generalized functions, specially the Laplace transform, that will be often used in the rest of this section.

3.4.2. Computations with the Dirac Delta.

Once we have the definitions of operations involving the Dirac delta, we can actually compute these limits. The following statement summarizes few interesting results. The first formula below says that the infinity we found in the definition of Dirac's delta is of a very particular type; that infinity is such that Dirac's delta is integrable, in the sense defined above, with integral equal one.

Theorem 3.4.3. *For every $c \in \mathbb{R}$ and $\epsilon > 0$ holds,* $\displaystyle\int_{c-\epsilon}^{c+\epsilon} \delta(t-c)\,dt = 1.$

Proof of Theorem 3.4.3: The integral of a Dirac's delta generalized function is computed as a limit of integrals,

$$\int_{c-\epsilon}^{c+\epsilon} \delta(t-c)\,dt = \lim_{n\to\infty} \int_{c-\epsilon}^{c+\epsilon} \delta_n(t-c)\,dt.$$

If we choose $n > 1/\epsilon$, equivalently $1/n < \epsilon$, then the domain of the functions in the sequence is inside the interval $(c-\epsilon, c+\epsilon)$, and we can write

$$\int_{c-\epsilon}^{c+\epsilon} \delta(t-c)\,dt = \lim_{n\to\infty} \int_{c}^{c+\frac{1}{n}} n\,dt, \qquad \text{for} \qquad \frac{1}{n} < \epsilon.$$

Then it is simple to compute

$$\int_{c-\epsilon}^{c+\epsilon} \delta(t-c)\,dt = \lim_{n\to\infty} n\Big(c + \frac{1}{n} - c\Big) = \lim_{n\to\infty} 1 = 1.$$

This establishes the Theorem. □

The next result is also deeply related with the defining property of the Dirac delta—the sequence functions have all graphs of unit area.

Theorem 3.4.4. *If f is continuous on (a,b) and $c \in (a,b)$, then* $\displaystyle\int_a^b f(t)\,\delta(t-c)\,dt = f(c)$.

Proof of Theorem 3.4.4: We again compute the integral of a Dirac's delta as a limit of a sequence of integrals,

$$\begin{aligned}\int_a^b \delta(t-c)\,f(t)\,dt &= \lim_{n\to\infty} \int_a^b \delta_n(t-c)\,f(t)\,dt\\ &= \lim_{n\to\infty} \int_a^b n\Big[u(t-c) - u\big(t-c-\frac{1}{n}\big)\Big]\,f(t)\,dt\\ &= \lim_{n\to\infty} \int_c^{c+\frac{1}{n}} n\,f(t)\,dt, \qquad \frac{1}{n} < (b-c),\end{aligned}$$

To get the last line we used that $c \in [a,b]$. Let F be any primitive of f, so $F(t) = \int f(t)\,dt$. Then we can write,

$$\begin{aligned}\int_a^b \delta(t-c)\,f(t)\,dt &= \lim_{n\to\infty} n\big[F\big(c+\frac{1}{n}\big) - F(c)\big]\\ &= \lim_{n\to\infty} \frac{1}{\left(\frac{1}{n}\right)}\big[F\big(c+\frac{1}{n}\big) - F(c)\big]\\ &= F'(c)\\ &= f(c).\end{aligned}$$

This establishes the Theorem. □

In our next result we compute the Laplace transform of the Dirac delta. This result is a simple consequence of the previous theorem.

Theorem 3.4.5. *For all $s \in \mathbb{R}$ holds* $\mathcal{L}[\delta(t-c)] = \begin{cases} e^{-cs} & \text{for} \quad c \geqslant 0,\\ 0 & \text{for} \quad c < 0.\end{cases}$

Proof of Theorem 3.4.5: We use the previous theorem on the integral that defines a Laplace transform. Although the previous theorem applies to definite integrals, not to improper integrals, it can be extended to cover improper integrals. So, we use Theorem 3.4.4 with $f(t) = e^{-st}$, and we get

$$\mathcal{L}[\delta(t-c)] = \int_0^\infty e^{-st}\,\delta(t-c)\,dt = \begin{cases} e^{-cs} & \text{for} \quad c \geqslant 0, \\ 0 & \text{for} \quad c < 0, \end{cases}$$

This establishes the Theorem. □

From this result is simple to get the following generalization.

Theorem 3.4.6. *For all $s \in \mathbb{R}$ holds* $\mathcal{L}[g(t)\,\delta(t-c)] = \begin{cases} g(c)\,e^{-cs} & \textit{for} \quad c \geqslant 0, \\ 0 & \textit{for} \quad c < 0. \end{cases}$

We now give a second proof of Theorem 3.4.4, based in the definition of the Laplace transform and the same type of limits used in the proof of Theorem 3.4.4.

Proof of Theorem 3.4.6: We use again the previous theorem on the integral that defines a Laplace transform. So, we use Theorem 3.4.4 with $f(t) = e^{-st}\,g(t)$, and we get

$$\mathcal{L}[g(t)\delta(t-c)] = \int_0^\infty e^{-st}\,g(t)\,\delta(t-c)\,dt = \begin{cases} g(c)\,e^{-cs} & \text{for} \quad c \geqslant 0, \\ 0 & \text{for} \quad c < 0, \end{cases}$$

This establishes the Theorem. □

Second Proof of Theorem 3.4.5: The Laplace transform of a Dirac's delta is computed as a limit of Laplace transforms,

$$\begin{aligned} \mathcal{L}[\delta(t-c)] &= \lim_{n\to\infty} \mathcal{L}[\delta_n(t-c)] \\ &= \lim_{n\to\infty} \mathcal{L}\Big[n\big[u(t-c) - u\big(t-c-\frac{1}{n}\big)\big]\Big] \\ &= \lim_{n\to\infty} \int_0^\infty n\big[u(t-c) - u\big(t-c-\frac{1}{n}\big)\big]\,e^{-st}\,dt. \end{aligned}$$

The case $c < 0$ is simple. For $\dfrac{1}{n} < |c|$ holds

$$\mathcal{L}[\delta(t-c)] = \lim_{n\to\infty} \int_0^\infty 0\,dt \quad \Rightarrow \quad \mathcal{L}[\delta(t-c)] = 0, \qquad \text{for } s \in \mathbb{R}, \quad c < 0.$$

Consider now the case $c \geqslant 0$. We then have,

$$\mathcal{L}[\delta(t-c)] = \lim_{n\to\infty} \int_c^{c+\frac{1}{n}} n\,e^{-st}\,dt.$$

For $s = 0$ we get

$$\mathcal{L}[\delta(t-c)] = \lim_{n\to\infty} \int_c^{c+\frac{1}{n}} n\,dt = 1 \quad \Rightarrow \quad \mathcal{L}[\delta(t-c)] = 1 \quad \text{for } s = 0, \quad c \geqslant 0.$$

In the case that $s \neq 0$ we get,

$$\begin{aligned} \mathcal{L}[\delta(t-c)] &= \lim_{n\to\infty} \int_c^{c+\frac{1}{n}} n\,e^{-st}\,dt \\ &= \lim_{n\to\infty} -\frac{n}{s}\big(e^{-cs} - e^{-(c+\frac{1}{n})s}\big) \\ &= e^{-cs} \lim_{n\to\infty} \frac{(1 - e^{-\frac{s}{n}})}{\Big(\dfrac{s}{n}\Big)}. \end{aligned}$$

The limit on the last line above is a singular limit of the form $\frac{0}{0}$, so we can use the l'Hôpital rule to compute it, that is,

$$\lim_{n\to\infty} \frac{(1 - e^{-\frac{s}{n}})}{\left(\dfrac{s}{n}\right)} = \lim_{n\to\infty} \frac{\left(-\dfrac{s}{n^2}\, e^{-\frac{s}{n}}\right)}{\left(-\dfrac{s}{n^2}\right)} = \lim_{n\to\infty} e^{-\frac{s}{n}} = 1.$$

We then obtain,

$$\mathcal{L}[\delta(t-c)] = e^{-cs} \quad \text{for } s \neq 0, \quad c \geqslant 0.$$

This establishes the Theorem. □

3.4.3. Applications of the Dirac Delta. Dirac's delta generalized functions describe *impulsive forces* in mechanical systems, such as the force done by a stick hitting a marble. An impulsive force acts on an infinitely short time and transmits a finite momentum to the system.

Example 3.4.3. Use Newton's equation of motion and Dirac's delta to describe the change of momentum when a particle is hit by a hammer.

Solution: A point particle with mass m, moving on one space direction, x, with a force F acting on it is described by

$$ma = F \quad \Leftrightarrow \quad mx''(t) = F(t, x(t)),$$

where $x(t)$ is the particle position as function of time, $a(t) = x''(t)$ is the particle acceleration, and we will denote $v(t) = x'(t)$ the particle velocity. We saw in § 2.1 that Newton's second law of motion is a second order differential equation for the position function x. Now it is more convenient to use the *particle momentum*, $p = mv$, to write the Newton's equation,

$$mx'' = mv' = (mv)' = F \quad \Rightarrow \quad p' = F.$$

So the force F changes the momentum, P. If we integrate on an interval $[t_1, t_2]$ we get

$$\Delta p = p(t_2) - p(t_1) = \int_{t_1}^{t_2} F(t, x(t))\, dt.$$

Suppose that an impulsive force is acting on a particle at t_0 transmitting a finite momentum, say p_0. This is where the Dirac delta is uselful for, because we can write the force as

$$F(t) = p_0\, \delta(t - t_0),$$

then $F = 0$ on $\mathbb{R} - \{t_0\}$ and the momentum transferred to the particle by the force is

$$\Delta p = \int_{t_0-\Delta t}^{t_0+\Delta t} p_0\, \delta(t - t_0)\, dt = p_0.$$

The momentum transferred is $\Delta p = p_0$, but the force is identically zero on $\mathbb{R} - \{t_0\}$. We have transferred a finite momentum to the particle by an interaction at a single time t_0. ◁

In the next example we use a hammer to stop the spring on its track, so it won't move any more. The force done by the hammer is an impulsive force that can be modeled with a Dirac's delta generalized function. The timing and the magnitude of this impulsive force have to be tuned so that the spring stops and stays still.

Example 3.4.4 (Stopping a Spring with a Hammer)**.** Consider a mass-spring system with natural frequency ω_0 and discard any friction. Let us denote by $y(t)$ the spring displacement from the equilibrium position, positive downwards, as function of time t. At a time t_0 the system is struck by a hammer. This system is described by the differential equation

$$y'' + \omega_0^2\, y = f_0\, \delta(t - t_0),$$

where $\delta(t-t_0)$ is the Dirac's delta singular at $t = t_0$. If the spring starts moving with an initial displacement $y(0) = y_0$ and zero initial velocity, $y'(0) = 0$, then what is the shortest time t_0 and the force intensity f_0, so that the spring stops moving at all after $t = t_0$?

Solution: It is usually a good idea to have an intuitive understanding of a solution to a physical problem before trying to find an analytical, precise, mathematical solution to that problem. Intuitively, to stop the spring with a hammer we should do the following.

- The only way to stop an oscillating spring is to hit it exactly at the instant in time when the object attached to the spring is passing through the rest position at $y(t_n) = 0$.
- There are multiple times, t_n, when you can hit the spring and stop it—whenever the spring is passing through the rest position.
- You hit the spring at any other position, and the spring force will move it again.
- The magnitude of the force has to be carefully adjusted so the spring stops forever.
- This force should depend on the initial position of the spring—the longer the initial displacement the larger the force.
- This force should depend on the spring stiffness—the harder the spring the larger the force.

We know what to do, so let's do it. First we use the Laplace transform method to find the solution $y(t)$ of the initial value problem

$$y'' + \omega_0^2\, y = f_0\, \delta(t - t_0), \qquad y(0) = y_0, \qquad y'(0) = 0, \tag{3.4.3}$$

where ω_0 and t_0 are non-negative constants while f_0 and y_0 are arbitrary constants.

$$\mathcal{L}[y''] + \omega_0^2 \mathcal{L}[y] = f_0 \mathcal{L}[\delta(t - t_0)],$$

which gives us

$$s^2 \mathcal{L}[y] - s\, y(0) - y'(0) + \omega_0^2 \mathcal{L}[y] = f_0\, e^{-t_0 s}.$$

The initial conditions of the problem imply that

$$(s^2 + \omega_0^2)\mathcal{L}[y] = y_0 s + f_0\, e^{-t_0 s} \quad \Rightarrow \quad \mathcal{L}[y] = \frac{y_0 s + f_0\, e^{-t_0 s}}{s^2 + \omega_0^2}.$$

It is convenient to rewrite the expression for $\mathcal{L}[y]$ above a follows,

$$\mathcal{L}[y] = y_0 \Big(\frac{s}{s^2 + \omega_0^2}\Big) + \frac{f_0}{\omega_0}\, e^{-t_0 s} \Big(\frac{\omega_0}{s^2 + \omega_0^2}\Big).$$

This last expression is very convenient, since now it is more clear that

$$\mathcal{L}[y] = y_0\, \mathcal{L}[\cos(\omega_0 t)] + \frac{f_0}{\omega_0}\, e^{-t_0 s}\, \mathcal{L}[\sin(\omega_0 t)],$$

which leads us to

$$\mathcal{L}[y] = y_0\, \mathcal{L}[\cos(\omega_0 t)] + \frac{f_0}{\omega_0}\, \mathcal{L}[u(t - t_0) \sin(\omega_0 (t - t_0))].$$

So we got an expression for the solution of the initial value problem (3.4.3) above,

$$y(t) = y_0 \cos(\omega_0 t) + \frac{f_0}{\omega_0}\, u(t - t_0) \sin(\omega_0 (t - t_0)), \tag{3.4.4}$$

where $u(t-t_0)$ is the step function with step at $t=t_0$. It is convenient to rewrite the sine in the second term above as follows,

$$\begin{aligned}\sin(\omega_0(t-t_0)) &= \sin(\omega_0 t - \omega_0 t_0)\\ &= \sin(\omega_0 t)\cos(\omega_0 t_0) - \sin(\omega_0 t_0)\cos(\omega_0 t),\end{aligned}$$

where in the last step we use the trigonometric identity

$$\sin(\alpha-\beta) = \sin(\alpha)\cos(\beta) - \sin(\beta)\cos(\alpha).$$

Let's choose the time t_0, when the hammer strikes, to be such that

$$\omega_0 t_0 = (2n-1)\,\frac{\pi}{2}.$$

In this case we get

$$\sin(\omega_0(t-t_0)) = (-1)^n \cos(\omega_0 t), \tag{3.4.5}$$

where we used that

$$\cos\Big((2n-1)\frac{\pi}{2}\Big) = 0, \qquad \sin\Big((2n-1)\frac{\pi}{2}\Big) = (-1)^{n+1}.$$

If we use Eq. (3.4.5) in the solution (3.4.4), then we get

$$y(t) = \Big(y_0 + (-1)^n\,\frac{f_0}{\omega_0}\,u(t-t_0)\Big)\cos(\omega_0 t), \qquad t_0 = (2n-1)\,\frac{\pi}{2\omega_0}.$$

Let's write the solution above in a more clear way, with the solution expression for $t<t_0$ and for $t \geqslant t_0$, which are given by

$$\begin{aligned} y(t) &= y_0\cos(\omega_0 t) & & t<t_0,\\ y(t) &= \Big(y_0 + (-1)^n\,\frac{f_0}{\omega_0}\Big)\cos(\omega_0 t), & & t \geqslant t_0.\end{aligned}$$

The spring is going to stop for all $t \geqslant t_0$ if the second function above vanishes identically. This will happen for a particular choice of the force magnitude, f_0, namely,

$$f_0 = (-1)^{n+1} y_0\,\omega_0,$$

which has to be applied at the precise time of

$$t_0 = (2n-1)\,\frac{\pi}{2\omega_0},$$

for some $n=1$ or $n=2$, etc. Choose the value of n, then the corresponding values above for t_0 and f_0. A spring-mass system with these parameters describes a spring that will stop after $n/2$ oscillations. Notice the force changes signs with n, because after each half oscillation the spring approaches the equilibrium position from different sides, so the force must be applied on the correct side to stop the spring.

◁

3.4.4. The Impulse Response Function. We now want to solve differential equations with the Dirac delta as a source. But there is a particular type of solutions that will be important later on—solutions to initial value problems with the Dirac delta source and zero initial conditions. We give these solutions a particular name.

Definition 3.4.7. *The* ***impulse response function*** *at the point $c \geqslant 0$ of the constant coefficients linear operator $L(y) = y'' + a_1\,y' + a_0\,y$, is the solution y_δ of*

$$L(y_\delta) = \delta(t-c), \qquad y_\delta(0) = 0, \qquad y_\delta'(0) = 0.$$

Remark: Impulse response functions are also called *fundamental solutions*.

Theorem 3.4.8. *The function y_δ is the impulse response function at $c \geqslant 0$ of the constant coefficients operator $L(y) = y'' + a_1\, y' + a_0\, y$ iff holds*

$$y_\delta = \mathcal{L}^{-1}\Big[\frac{e^{-cs}}{p(s)}\Big].$$

where p is the characteristic polynomial of L.

Remark: The impulse response function y_δ at $c = 0$ satifies

$$y_\delta = \mathcal{L}^{-1}\Big[\frac{1}{p(s)}\Big].$$

Proof of Theorem 3.4.8: Compute the Laplace transform of the differential equation for for the impulse response function y_δ,

$$\mathcal{L}[y''] + a_1\, \mathcal{L}[y'] + a_0\, \mathcal{L}[y] = \mathcal{L}[\delta(t-c)] = e^{-cs}.$$

Since the initial data for y_δ is trivial, we get

$$(s^2 + a_1 s + a_0)\, \mathcal{L}[y] = e^{-cs}.$$

Since $p(s) = s^2 + a_1 s + a_0$ is the characteristic polynomial of L, we get

$$\mathcal{L}[y] = \frac{e^{-cs}}{p(s)} \quad \Leftrightarrow \quad y(t) = \mathcal{L}^{-1}\Big[\frac{e^{-cs}}{p(s)}\Big].$$

All the steps in this calculation are if and only ifs. This establishes the Theorem. □

Example 3.4.5. Find the impulse response function at $t = 0$ of the linear operator

$$L(y) = y'' + 2y' + 2y.$$

Solution: We need to find the solution y_δ of the initial value problem

$$y_\delta'' + 2y_\delta' + 2y_\delta = \delta(t), \qquad y_\delta(0) = 0, \qquad y_\delta'(0) = 0.$$

Since the souce is a Dirac delta, we have to use the Laplace transform to solve this problem. So we compute the Laplace transform on both sides of the differential equation,

$$\mathcal{L}[y_\delta''] + 2\, \mathcal{L}[y_\delta'] + 2\, \mathcal{L}[y_\delta] = \mathcal{L}[\delta(t)] = 1 \quad \Rightarrow \quad (s^2 + 2s + 2)\, \mathcal{L}[y_\delta] = 1,$$

where we have introduced the initial conditions on the last equation above. So we obtain

$$\mathcal{L}[y_\delta] = \frac{1}{(s^2 + 2s + 2)}.$$

The denominator in the equation above has complex valued roots, since

$$s_\pm = \frac{1}{2}\big[-2 \pm \sqrt{4-8}\big],$$

therefore, we complete squares $s^2 + 2s + 2 = (s+1)^2 + 1$. We need to solve the equation

$$\mathcal{L}[y_\delta] = \frac{1}{\big[(s+1)^2 + 1\big]} = \mathcal{L}[e^{-t}\sin(t)] \quad \Rightarrow \quad y_\delta(t) = e^{-t}\sin(t).$$

◁

Example 3.4.6. Find the impulse response function at $t = c \geqslant 0$ of the linear operator

$$L(y) = y'' + 2\,y' + 2\,y.$$

Solution: We need to find the solution y_δ of the initial value problem

$$y_\delta'' + 2\,y_\delta' + 2\,y_\delta = \delta(t-c), \quad y_\delta(0) = 0, \quad y_\delta'(0) = 0.$$

We have to use the Laplace transform to solve this problem because the source is a Dirac's delta generalized function. So, compute the Laplace transform of the differential equation,

$$\mathcal{L}[y_\delta''] + 2\,\mathcal{L}[y_\delta'] + 2\,\mathcal{L}[y_\delta] = \mathcal{L}[\delta(t-c)].$$

Since the initial conditions are all zero and $c \geqslant 0$, we get

$$(s^2 + 2s + 2)\,\mathcal{L}[y_\delta] = e^{-cs} \quad \Rightarrow \quad \mathcal{L}[y_\delta] = \frac{e^{-cs}}{(s^2 + 2s + 2)}.$$

Find the roots of the denominator,

$$s^2 + 2s + 2 = 0 \quad \Rightarrow \quad s_\pm = \frac{1}{2}\left[-2 \pm \sqrt{4-8}\right]$$

The denominator has complex roots. Then, it is convenient to complete the square in the denominator,

$$s^2 + 2s + 2 = \left[s^2 + 2\Big(\frac{2}{2}\Big)\,s + 1\right] - 1 + 2 = (s+1)^2 + 1.$$

Therefore, we obtain the expression,

$$\mathcal{L}[y_\delta] = \frac{e^{-cs}}{(s+1)^2 + 1}.$$

Recall that $\mathcal{L}[\sin(t)] = \dfrac{1}{s^2+1}$, and $\mathcal{L}[f](s-c) = \mathcal{L}[e^{ct}\,f(t)]$. Then,

$$\frac{1}{(s+1)^2+1} = \mathcal{L}[e^{-t}\,\sin(t)] \quad \Rightarrow \quad \mathcal{L}[y_\delta] = e^{-cs}\,\mathcal{L}[e^{-t}\,\sin(t)].$$

Since for $c \geqslant 0$ holds $e^{-cs}\,\mathcal{L}[f](s) = \mathcal{L}[u(t-c)\,f(t-c)]$, we conclude that

$$y_\delta(t) = u(t-c)\,e^{-(t-c)}\,\sin(t-c).$$

◁

Example 3.4.7. Find the solution y to the initial value problem

$$y'' - y = -20\,\delta(t-3), \qquad y(0) = 1, \qquad y'(0) = 0.$$

Solution: The source is a generalized function, so we need to solve this problem using the Lapace transform. So we compute the Laplace transform of the differential equation,

$$\mathcal{L}[y''] - \mathcal{L}[y] = -20\,\mathcal{L}[\delta(t-3)] \quad \Rightarrow \quad (s^2-1)\,\mathcal{L}[y] - s = -20\,e^{-3s},$$

where in the second equation we have already introduced the initial conditions. We arrive to the equation

$$\mathcal{L}[y] = \frac{s}{(s^2-1)} - 20\,e^{-3s}\,\frac{1}{(s^2-1)} = \mathcal{L}[\cosh(t)] - 20\,\mathcal{L}[u(t-3)\,\sinh(t-3)],$$

which leads to the solution

$$y(t) = \cosh(t) - 20\,u(t-3)\,\sinh(t-3).$$

◁

Example 3.4.8. Find the solution to the initial value problem

$$y'' + 4y = \delta(t-\pi) - \delta(t-2\pi), \qquad y(0)=0, \qquad y'(0)=0.$$

Solution: We again Laplace transform both sides of the differential equation,

$$\mathcal{L}[y''] + 4\,\mathcal{L}[y] = \mathcal{L}[\delta(t-\pi)] - \mathcal{L}[\delta(t-2\pi)] \quad \Rightarrow \quad (s^2+4)\,\mathcal{L}[y] = e^{-\pi s} - e^{-2\pi s},$$

where in the second equation above we have introduced the initial conditions. Then,

$$\begin{aligned}
\mathcal{L}[y] &= \frac{e^{-\pi s}}{(s^2+4)} - \frac{e^{-2\pi s}}{(s^2+4)}\\
&= \frac{e^{-\pi s}}{2}\,\frac{2}{(s^2+4)} - \frac{e^{-2\pi s}}{2}\,\frac{2}{(s^2+4)}\\
&= \frac{1}{2}\mathcal{L}\Big[u(t-\pi)\,\sin\big[2(t-\pi)\big]\Big] - \frac{1}{2}\mathcal{L}\Big[u(t-2\pi)\,\sin\big[2(t-2\pi)\big]\Big].
\end{aligned}$$

The last equation can be rewritten as follows,

$$y(t) = \frac{1}{2}\,u(t-\pi)\,\sin\big[2(t-\pi)\big] - \frac{1}{2}\,u(t-2\pi)\,\sin\big[2(t-2\pi)\big],$$

which leads to the conclusion that

$$y(t) = \frac{1}{2}\,\big[u(t-\pi) - u(t-2\pi)\big]\,\sin(2t).$$

$\lhd$

3.4.5. Comments on Generalized Sources.

We have used the Laplace transform to solve differential equations with the Dirac delta as a source function. It may be convenient to understand a bit more clearly what we have done, since the Dirac delta is not an ordinary function but a generalized function defined by a limit. Consider the following example.

Example 3.4.9. Find the impulse response function at $t = c > 0$ of the linear operator

$$L(y) = y'.$$

Solution: We need to solve the initial value problem

$$y'(t) = \delta(t-c), \qquad y(0) = 0.$$

In other words, we need to find a primitive of the Dirac delta. However, Dirac's delta is not even a function. Anyway, let us compute the Laplace transform of the equation, as we did in the previous examples,

$$\mathcal{L}[y'(t)] = \mathcal{L}[\delta(t-c)] \quad \Rightarrow \quad s\,\mathcal{L}[y(t)] - y(0) = e^{-cs} \quad \Rightarrow \quad \mathcal{L}[y(t)] = \frac{e^{-cs}}{s}.$$

But we know that

$$\frac{e^{-cs}}{s} = \mathcal{L}[u(t-c)] \quad \Rightarrow \quad \mathcal{L}[y(t)] = \mathcal{L}[u(t-c)] \quad \Rightarrow \quad y(t) = u(t-c).$$

$\lhd$

Looking at the differential equation $y'(t) = \delta(t-c)$ and at the solution $y(t) = u(t-c)$ one could like to write them together as

$$u'(t-c) = \delta(t-c). \tag{3.4.6}$$

But this is not correct, because the step function is a discontinuous function at $t = c$, hence not differentiable. What we have done is something different. We have found a sequence of functions u_n with the properties,

$$\lim_{n\to\infty} u_n(t-c) = u(t-c), \qquad \lim_{n\to\infty} u_n'(t-c) = \delta(t-c),$$

and we have called $y(t) = u(t-c)$. This is what we actually do when we solve a differential equation with a source defined as a limit of a sequence of functions, such as the Dirac delta. The Laplace transform method used on differential equations with generalized sources allows us to solve these equations without the need to write any sequence, which are hidden in the definitions of the Laplace transform of generalized functions. Let us solve the problem in the Example 3.4.9 one more time, but this time let us show where all the sequences actually are.

Example 3.4.10. Find the solution to the initial value problem

$$y'(t) = \delta(t-c), \qquad y(0) = 0, \qquad c > 0, \tag{3.4.7}$$

Solution: Recall that the Dirac delta is defined as a limit of a sequence of bump functions,

$$\delta(t-c) = \lim_{n\to\infty} \delta_n(t-c), \qquad \delta_n(t-c) = n\Big[u(t-c) - u\Big(t-c-\frac{1}{n}\Big)\Big], \qquad n = 1, 2, \cdots .$$

The problem we are actually solving involves a sequence and a limit,

$$y'(t) = \lim_{n\to\infty} \delta_n(t-c), \qquad y(0) = 0.$$

We start computing the Laplace transform of the differential equation,

$$\mathcal{L}[y'(t)] = \mathcal{L}[\lim_{n\to\infty} \delta_n(t-c)].$$

We have defined the Laplace transform of the limit as the limit of the Laplace transforms,

$$\mathcal{L}[y'(t)] = \lim_{n\to\infty} \mathcal{L}[\delta_n(t-c)].$$

If the solution is at least piecewise differentiable, we can use the property

$$\mathcal{L}[y'(t)] = s\,\mathcal{L}[y(t)] - y(0).$$

Assuming that property, and the initial condition $y(0) = 0$, we get

$$\mathcal{L}[y(t)] = \frac{1}{s} \lim_{n\to\infty} \mathcal{L}[\delta_n(t-c)] \quad \Rightarrow \quad \mathcal{L}[y(t)] = \lim_{n\to\infty} \frac{\mathcal{L}[\delta_n(t-c)]}{s}.$$

Introduce now the function $y_n(t) = u_n(t-c)$, given in Eq. (3.4.1), which for each n is the only continuous, piecewise differentiable, solution of the initial value problem

$$y_n'(t) = \delta_n(t-c), \qquad y_n(0) = 0.$$

It is not hard to see that this function u_n satisfies

$$\mathcal{L}[u_n(t)] = \frac{\mathcal{L}[\delta_n(t-c)]}{s}.$$

Therefore, using this formula back in the equation for y we get,

$$\mathcal{L}[y(t)] = \lim_{n\to\infty} \mathcal{L}[u_n(t)].$$

For continuous functions we can interchange the Laplace transform and the limit,

$$\mathcal{L}[y(t)] = \mathcal{L}[\lim_{n\to\infty} u_n(t)].$$

So we get the result,

$$y(t) = \lim_{n\to\infty} u_n(t) \quad \Rightarrow \quad y(t) = u(t-c).$$

We see above that we have found something more than just $y(t) = u(t-c)$. We have found

$$y(t) = \lim_{n\to\infty} u_n(t-c),$$

where the sequence elements u_n are continuous functions with $u_n(0) = 0$ and

$$\lim_{n\to\infty} u_n(t-c) = u(t-c), \qquad \lim_{n\to\infty} u_n'(t-c) = \delta(t-c),$$

Finally, derivatives and limits cannot be interchanged for u_n,

$$\lim_{n\to\infty} \big[u_n'(t-c)\big] \neq \big[\lim_{n\to\infty} u_n(t-c)\big]'$$

so it makes no sense to talk about y'. ◁

When the Dirac delta is defined by a sequence of functions, as we did in this section, the calculation needed to find impulse response functions must involve sequence of functions and limits. The Laplace transform method used on generalized functions allows us to hide all the sequences and limits. This is true not only for the derivative operator $L(y) = y'$ but for any second order differential operator with constant coefficients.

Definition 3.4.9. *A* ***solution*** *of the initial value problem with a Dirac's delta source*

$$y'' + a_1\, y' + a_0\, y = \delta(t-c), \qquad y(0) = y_0, \qquad y'(0) = y_1, \tag{3.4.8}$$

where a_1, a_0, y_0, y_1, and $c \in \mathbb{R}$, are given constants, is a function

$$y(t) = \lim_{n\to\infty} y_n(t),$$

where the functions y_n, with $n \geqslant 1$, are the unique solutions to the initial value problems

$$y_n'' + a_1\, y_n' + a_0\, y_n = \delta_n(t-c), \qquad y_n(0) = y_0, \qquad y_n'(0) = y_1, \tag{3.4.9}$$

and the source δ_n satisfy $\lim_{n\to\infty} \delta_n(t-c) = \delta(t-c)$.

The definition above makes clear what do we mean by a solution to an initial value problem having a generalized function as source, when the generalized function is defined as the limit of a sequence of functions. The following result says that the Laplace transform method used with generalized functions hides all the sequence computations.

Theorem 3.4.10. *The function y is solution of the initial value problem*

$$y'' + a_1\, y' + a_0\, y = \delta(t-c), \qquad y(0) = y_0, \qquad y'(0) = y_1, \qquad c \geqslant 0,$$

iff its Laplace transform satisfies the equation

$$\big(s^2\,\mathcal{L}[y] - sy_0 - y_1\big) + a_1\,\big(s\,\mathcal{L}[y] - y_0\big) - a_0\,\mathcal{L}[y] = e^{-cs}.$$

This Theorem tells us that to find the solution y to an initial value problem when the source is a Dirac's delta we have to apply the Laplace transform to the equation and perform the same calculations as if the Dirac delta were a function. This is the calculation we did when we computed the impulse response functions.

Proof of Theorem 3.4.10: Compute the Laplace transform on Eq. (3.4.9),

$$\mathcal{L}[y_n''] + a_1\,\mathcal{L}[y_n'] + a_0\,\mathcal{L}[y_n] = \mathcal{L}[\delta_n(t-c)].$$

Recall the relations between the Laplace transform and derivatives and use the initial conditions,

$$\mathcal{L}[y_n''] = s^2\,\mathcal{L}[y_n] - sy_0 - y_1, \qquad \mathcal{L}[y'] = s\,\mathcal{L}[y_n] - y_0,$$

and use these relation in the differential equation,

$$(s^2 + a_1 s + a_0)\,\mathcal{L}[y_n] - sy_0 - y_1 - a_1 y_0 = \mathcal{L}[\delta_n(t-c)],$$

Since δ_n satisfies that $\lim_{n\to\infty} \delta_n(t-c) = \delta(t-c)$, an argument like the one in the proof of Theorem 3.4.5 says that for $c \geqslant 0$ holds

$$\mathcal{L}[\delta_n(t-c)] = \mathcal{L}[\delta(t-c)] \quad \Rightarrow \quad \lim_{n\to\infty} \mathcal{L}[\delta_n(t-c)] = e^{-cs}.$$

Then

$$(s^2 + a_1 s + a_0) \lim_{n\to\infty} \mathcal{L}[y_n] - s y_0 - y_1 - a_1 y_0 = e^{-cs}.$$

Interchanging limits and Laplace transforms we get

$$(s^2 + a_1 s + a_0)\, \mathcal{L}[y] - s y_0 - y_1 - a_1 y_0 = e^{-cs},$$

which is equivalent to

$$\big(s^2\, \mathcal{L}[y] - s y_0 - y_1\big) + a_1\, \big(s\, \mathcal{L}[y] - y_0\big) - a_0\, \mathcal{L}[y] = e^{-cs}.$$

This establishes the Theorem. □

3.4.6. Exercises.

3.4.1.- .

3.4.2.- .

CHAPTER 4

Overview of Linear Algebra

We overview a few concepts of linear algebra such as vectors, dot product, matrices, matrix operations, determinants, inverse matrix formulas, eigenvalues and eigenvectors of a matrix, and diagonalizable matrices.

4.1. Orthogonal Vectors

In this section we review familiar concepts such as vectors in two or three dimensional space. Wen then introduce the notion of the dot product of two vectors, which allows us to characterize orthogonal vectors and to compute expansions of vectors in terms of orthogonal vectors. In a later chapter we generalize these ideas to the vector space of functions. We will introduce the idea of orthogonal functions and we will compute expansions of a given function in terms of orthogonal functions. A famous examples of these orthogonal expansions is the Fourier series expansion of a function.

4.1.1. The Vector Space $\mathbb{F}^n$. Before we start we need a bit of notation. Let us introduce $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, meaning $\mathbb{F}$ is either the set of real numbers $\mathbb{R}$ or the set of complex numbers $\mathbb{C}$. Then the space of all n-vectors, real or complex, will be denoted as $\mathbb{F}^n$, with $n = 1, 2, 3, \cdots$. A vector space $\mathbb{F}^n$ is the set of all vectors with n components together with the operation linear combination of vectors.

Definition 4.1.1. *The* ***vector space*** $\mathbb{F}^n$*, with* $n = 1, 2, 3, \cdots$*, is a set with an operation. The set is formed by all* n*-vectors* $\boldsymbol{u} \in \mathbb{F}^n$*, denoted as either columns or rows,*

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \langle u_1, \cdots, u_n \rangle,$$

where the numbers $u_i \in \mathbb{F}$*, for* $i = 1, \cdots, n$ *are called the vector components. The operation is the linear combination, that is, for all* n*-vectors* $\boldsymbol{u}$, $\boldsymbol{v} \in \mathbb{F}^n$*, and all scalars (numbers that are not components of vectors)* a, $b \in \mathbb{F}$*, holds*

$$a\boldsymbol{u} + b\boldsymbol{v} = a \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + b \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} au_1 + bv_1 \\ \vdots \\ au_n + bv_n \end{bmatrix}.$$

The operation of linear combination includes two important particular cases, the case $b = 0$ and the case $a = b = 1$, which are

$$a\boldsymbol{u} = \begin{bmatrix} au_1 \\ \vdots \\ au_n \end{bmatrix}, \qquad \boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix}.$$

There is a clear geometrical meaning for these two particular cases above. The geometrical meaning of multiplying a vector by a scalar, a, is to compress (for $|a| < 1$) or stretch (for $|a| > 1$) or reverse the direction (for $a < 0$) of that vector. The geometrical meaning of the addition of two vectors is called the parallelogram law of addition, which is how forces applied to objects behave. In the following example we can see the geometrical interpretation of both operations included in a linear combination.

Example 4.1.1 (Linear Combination)**.** In the case $n = 2$ we get the vector space $\mathbb{R}^2$. examples of vectors in $\mathbb{R}^2$ are the vectors $\boldsymbol{u}$, $\boldsymbol{v}$ below,

$$\boldsymbol{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

We can represent these vectors in the plane, as it is shown in Fig. 1. We can also compute the following linear combination $\boldsymbol{u} + 2\boldsymbol{v}$,

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

The graphical interpretation of this linear combination is the parallelogram law shown in Fig. 2, where we called $\boldsymbol{w}$ the result of the linear combination.

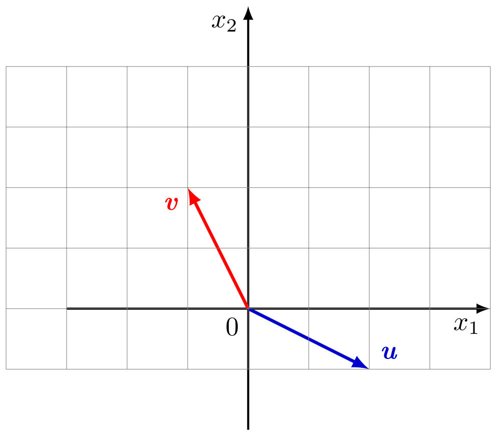


FIGURE 1. Graphical representation of the vectors above.

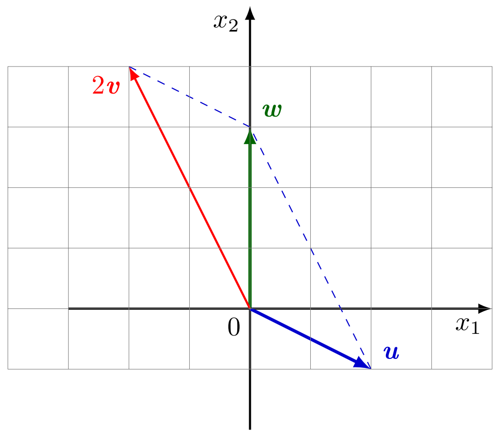


FIGURE 2. Graphical representation of the linear combination above.

◁

The operation of linear combination leads us to the concept of linear dependence or independence of a set of vectors. A set of vectors is linearly dependent when at least one of the vectors is a linear combination of some or all of the other vectors. On the contrary, a set of vectors is linearly independent if none of these vectors is a linear combination of any or all of the other vectors. It is not so straightforward to write these definitions in a precise way.

Definition 4.1.2. *A finite set of vectors $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$, with $k \geqslant 1$, in a vector space is* ***linearly dependent*** *iff there exists a set of scalars $\{c_1, \cdots, c_k\}$, not all of them zero, such that,*

$$c_1 \boldsymbol{v}_1 + \cdots + c_k \boldsymbol{v}_k = \boldsymbol{0}. \tag{4.1.1}$$

The set $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$ is called ***linearly independent*** *iff Eq. (4.1.1) implies that every scalar vanishes, $c_1 = \cdots = c_k = 0$.*

Example 4.1.2 (linear Dependence)**.** Determine whether the vectors below are linearly dependent or independent,

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}.$$

Solution: It is clear that

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \Rightarrow \quad 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Since the constants $c_1 = 2$, $c_2 = 3$, and $c_3 = -1$ are non-zero, the vectors above are linearly dependent. ◁

Example 4.1.3 (linear Dependence)**.** Determine whether the vectors below are linearly dependent or independent,

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Solution: We need to find constants c_1, c_2, c_3 so that

$$c_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix} + c_3 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which leads us to the following system of linear equations for the constants,

$$c_1 - c_2 + 3\,c_3 = 0 \tag{4.1.2}$$

$$2\,c_1 + 2\,c_2 + 2\,c_3 = 0 \tag{4.1.3}$$

$$3\,c_1 + 5\,c_2 + c_3 = 0. \tag{4.1.4}$$

The first equation above, Eq. (4.1.2), says

$$c_1 = c_2 - 3\,c_3,$$

while the second equation above, Eq. (4.1.3), says

$$c_1 = -c_2 - c_3.$$

These two equations together say

$$c_2 - 3\,c_3 = -c_2 - c_3 \quad \Rightarrow \quad c_2 = c_3.$$

Replacing this last equation in any of the previous equations for c_1 we get

$$c_1 = -2\,c_3$$

We now use these last two equations in Eq. (4.1.4), that is,

$$3\,(-2\,c_3) + 5\,(c_3) + c_3 = 0 \quad \Rightarrow \quad (-6+6)\,c_3 = 0 \quad \Rightarrow \quad 0 = 0.$$

This means that Eq. (4.1.4) is satisfied for all values of c_3. Therefore, we have found that the solutions c_1, c_2, c_3 are

$$c_1 = -2\,c_2, \qquad c_2 = c_3, \qquad c_3\text{: free.}$$

For example, choosing $c_3 = 1$ we have that

$$-2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Therefore, the vectors are linearly dependent. ◁

Example 4.1.4 (Linear Independence)**.** Show that the vectors are linearly independent,

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}.$$

Solution: First, it is clear that these two vectors point in different directions on the plane, so they should be linearly independent. But let's use the definition above. Let's set

$$c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and let's show that the only solution is $c_1 = c_2 = 0$. Indeed

$$\left.\begin{aligned} c_1 + 2\,c_2 = 0, \\ c_1 + 3\,c_3 = 0, \end{aligned}\right\} \quad \Rightarrow \quad c_1 = 0, \quad c_2 = 0.$$

Therefore, the vectors above are linearly independent. ◁

Example 4.1.5 (linear Dependence)**.** Determine whether the vectors below are linearly dependent or independent,

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}.$$

Solution: We need to find constants c_1, c_2, c_3 so that

$$c_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which leads us to the following system of linear equations for the constants

$$c_1 + 3\,c_2 + c_3 = 0 \tag{4.1.5}$$
$$2\,c_1 + 2\,c_2 + c_3 = 0 \tag{4.1.6}$$
$$3\,c_1 + c_2 + 2\,c_3 = 0. \tag{4.1.7}$$

Eq. (4.1.5) implies

$$c_1 = -3\,c_2 - c_3,$$

and this equation into Eq.(4.1.6) gives us

$$2\,(-3\,c_2 - c_3) + 2\,c_2 + c_3 = 0 \quad \Rightarrow \quad -4\,c_2 - c_3 = 0 \quad \Rightarrow \quad c_3 = -4\,c_2.$$

This last equation into the equation for c_1 implies

$$c_1 = c_2.$$

If we put these last two equations for c_1 and c_3 into Eq. (4.1.7) we get

$$3\,c_2 + c_2 + 2\,(-4\,c_2) = 0 \quad \Rightarrow \quad -4\,c_2 = 0 \quad \Rightarrow \quad c_2 = 0,$$

which in turns implies

$$c_1 = 0, \qquad c_3 = 0.$$

Therefore, the vectors are linearly independent. ◁

The concept of linear independence allows us to introduce the notion of how big is a vector space.

Definition 4.1.3. *The **dimension** of a vector space $\mathbb{F}^n$ is the maximum number of vectors that are linearly independent.*

It is not difficult to see that the dimension of the vector space $\mathbb{F}^n$ is indeed n. A linearly independent set of n vectors in $\mathbb{F}^n$ is called a ***base*** of $\mathbb{F}^n$. Again, it is not difficult to see that any vector in the space $\mathbb{F}^n$ is a linear combination of the vectors in a base.

4.1.2. Orthogonal Vectors. Two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ are orthogonal, also called perpendicular, if the angle between them is $\pi/2$. This geometrical condition can be translated into an analytical condition by introducing a new operation between vectors—the dot product. The dot product of two vectors is a scalar proportional to the projection of one of the vectors onto the other. Then, two vectors being orthogonal is equivalent to having zero dot product, that is, no projection onto each other.

The definition of the dot product of two real vectors is slightly different from the definition of the dot product of two complex vectors. To avoid this complication we restrict to real vector spaces $\mathbb{R}^n$ for the rest of this section.

Definition 4.1.4. *The **dot product** of vectors $\boldsymbol{u} = \langle u_1, \cdots, u_n\rangle$, $\boldsymbol{v} = \langle v_1, \cdots, v_n\rangle \in \mathbb{R}^n$ is*

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + \cdots + u_n v_n.$$

The dot product of two vectors is a scalar, which can be positive, zero of negative. The dot product satisfies the following properties.

Theorem 4.1.5. *For every vectors $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w} \in \mathbb{R}^n$ and every scalars a, $b \in \mathbb{R}$ holds,*

(a) Positivity: $\boldsymbol{u} \cdot \boldsymbol{u} = 0$ iff $\boldsymbol{u} = 0$; and $\boldsymbol{u} \cdot \boldsymbol{u} > 0$ for $\boldsymbol{u} \neq 0$.
(b) Symmetry: $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$.
(c) Linearity: $(a\boldsymbol{u} + b\boldsymbol{v}) \cdot \boldsymbol{w} = a\,(\boldsymbol{u} \cdot \boldsymbol{w}) + b\,(\boldsymbol{v} \cdot \boldsymbol{w})$.

Remark: The proof of these properties is simple and left as an exercise.

The dot product of two vectors is deeply related with the projections of one vector onto the other. To see this geometrical meaning of the dot product we need the following result.

Theorem 4.1.6. *The dot product of vectors $\boldsymbol{u}$, $\boldsymbol{v}$ $\in \mathbb{R}^n$ can be written as*

$$\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\|\,\|\boldsymbol{v}\| \cos(\theta),$$

where $\|\boldsymbol{u}\|$, $\|\boldsymbol{v}\|$ are the length of the vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $\theta \in [0, \pi]$ is the angle between the vectors.

Remark: Recall that the *length* (or magnitude) of a vector $\boldsymbol{u}$ is

$$\|\boldsymbol{u}\| = \sqrt{(u_1)^2 + \cdots + (u_n)^2}.$$

We now show the proof of the Theorem above in the case of $n = 3$. We leave the proof of the general case as an exercise.

Proof of Theorem 4.1.6 for $n = 3$: The law of cosines for a triangle with sides given by vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $(\boldsymbol{u} - \boldsymbol{v})$, as shown in Fig. 3, is

$$\|\boldsymbol{u} - \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - 2\|\boldsymbol{u}\|\boldsymbol{v}\| \cos(\theta)$$

If we write this formula in components we get

$$(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2 = (u_1^2 + u_2^2 + u_3^2) + (v_1^2 + v_2^2 + v_3^2) - 2\|\boldsymbol{u}\|\boldsymbol{v}\| \cos(\theta).$$

If we expand the squares on the left-hand side above and we cancel terms we get

$$-2u_1v_1 - 2u_2v_2 - 2u_3v_3 = -2\|\boldsymbol{u}\|\boldsymbol{v}\| \cos(\theta).$$

We now use on the left-hand side above the definition of the dot product of vectors $\boldsymbol{u}$, $\boldsymbol{v}$,

$$\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\|\boldsymbol{v}\| \cos(\theta).$$

This establishes the Theorem. □

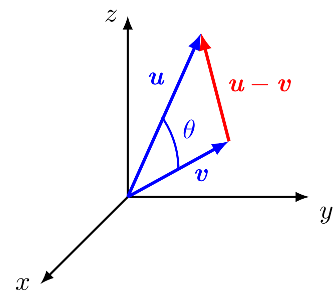


FIGURE 3. Vectors used in the proof of Theorem 4.1.6.

From the dot product expression in Theorem 4.1.6 we can see that the dot product of two vectors vanishes if and only if the angle in between the vectors is $\theta = \pi/2$. So the dot product is a good way to determine whether two vectors are perpendicular. We summarize this result in a theorem.

Theorem 4.1.7. *The nonzero vectors $\boldsymbol{u}$, $\boldsymbol{v} \in \mathbb{R}^n$ are orthogonal iff $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.*

Proof of Theorem 4.1.7: Suppose that $\boldsymbol{u}$, $\boldsymbol{v}$ satisfy that $\boldsymbol{u} \cdot \boldsymbol{v} = 0$. This is equivalent to

$$\|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos(\theta) = 0.$$

Since the vectors are nonzero, this is equivalent to

$$\cos(\theta) = 0.$$

But for $\theta \in [0, \pi]$ this is equivalent to $\theta = \pi/2$. This establishes the Theorem. □

We can go a little deeper in the geometrical meaning of the dot product. In fact, the dot product of two vectors is proportional to the scalar projection of one vector onto the other.

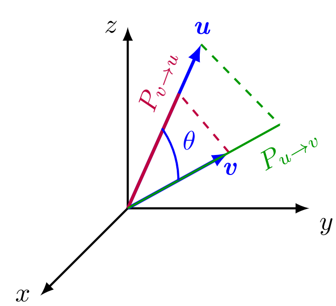


FIGURE 4. In this picture in $\mathbb{R}^3$ we show the scalar projection of $\boldsymbol{u}$ onto $\boldsymbol{v}$, which is the length of the solid segment in green. We also show the scalar projection of $\boldsymbol{v}$ onto $\boldsymbol{u}$, which is the length of the solid segment in purple.

For example, if we denote by $P_{u\to v}$ the scalar projection of $\boldsymbol{u}$ onto $\boldsymbol{v}$, then

$$P_{u\to v} = \|\boldsymbol{u}\| \cos(\theta) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{v}\|}.$$

Analogously, the scalar projection of $\boldsymbol{v}$ onto $\boldsymbol{u}$ is

$$P_{v\to u} = \|\boldsymbol{v}\|\cos(\theta) = \frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|}.$$

Notice that the length of a vector $\boldsymbol{u}$ can also be written in terms of the dot product,

$$\|\boldsymbol{u}\| = \sqrt{\boldsymbol{u}\cdot\boldsymbol{u}}.$$

A vector $\boldsymbol{u}$ is a *unit vector* if and only if

$$\boldsymbol{u}\cdot\boldsymbol{u} = 1.$$

And any vector $\boldsymbol{v}$ can be rescaled into a unit vector by dividing by its magnitude. So, the vector $\boldsymbol{u}$ below is a unit vector in the direction of the vector $\boldsymbol{v}$,

$$\boldsymbol{u} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}.$$

A set of vectors is an *orthogonal set* if all the vectors in the set are mutually perpendicular. An *orthonormal set* is an orthogonal set where all the vectors are unit vectors.

Example 4.1.6. The set of vectors $\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$ used in physics is an orthonormal set in $\mathbb{R}^3$.

Solution: These are the vectors

$$\boldsymbol{i} = \langle 1,0,0\rangle, \qquad \boldsymbol{j} = \langle 0,1,0\rangle, \qquad \boldsymbol{k} = \langle 0,0,1\rangle.$$

As we see in Fig. 5, they are mutually perpendicular and have unit magnitude, that is,

$$\begin{aligned} \boldsymbol{i}\cdot\boldsymbol{j} &= 0, & \boldsymbol{i}\cdot\boldsymbol{i} &= 1,\\ \boldsymbol{i}\cdot\boldsymbol{k} &= 0, & \boldsymbol{j}\cdot\boldsymbol{j} &= 1,\\ \boldsymbol{j}\cdot\boldsymbol{k} &= 0. & \boldsymbol{k}\cdot\boldsymbol{k} &= 1. \end{aligned}$$

◁

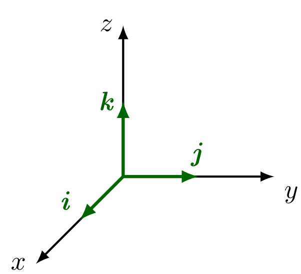


FIGURE 5. The vectors $\boldsymbol{i}$,$\boldsymbol{j}$,$\boldsymbol{k}$.

Any rotation of the vectors $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ (that is the three vectors are rotated in the same way) is also an orthonormal set of vectors.

Example 4.1.7 (Orthogonal and Unit Vectors)**.** Show that the vectors

$$\boldsymbol{v}_1 = \langle 1,1,1\rangle,\ \boldsymbol{v}_2 = \langle -2,1,1\rangle,\ \boldsymbol{v}_3 = \langle 0,-3,3\rangle$$

are mutually orthogonal. Find unit vectors in the direction of these vectors.

Solution: Let's see the orthogonality conditions.

$$\begin{aligned}\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 &= \langle 1,1,1\rangle \cdot \langle -2,1,1\rangle = -2+1+1 = 0,\\ \boldsymbol{v}_1 \cdot \boldsymbol{v}_3 &= \langle 1,1,1\rangle \cdot \langle 0,-3,3\rangle = -3+3 = 0,\\ \boldsymbol{v}_2 \cdot \boldsymbol{v}_3 &= \langle -2,1,1\rangle \cdot \langle 0,-3,3\rangle = -3+3 = 0,\end{aligned}$$

so, these vectors are mutually orthogonal. This set is not orthonormal, since

$$\|\boldsymbol{v}_1\| = \sqrt{3}, \quad \|\boldsymbol{v}_2\| = \sqrt{6}, \quad \|\boldsymbol{v}_3\| = \sqrt{18} = 3\sqrt{2}.$$

Therefore, unit vectors in the direction of the vectors above are

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{3}}\langle 1,1,1\rangle, \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{6}}\langle -2,1,1\rangle, \quad \boldsymbol{u}_3 = \frac{1}{\sqrt{2}}\langle 0,-1,1\rangle.$$

◁

Example 4.1.8 (Orthogonal Vectors)**.** Find constants c_1 and c_2 so that the vectors

$$\boldsymbol{v}_1 = \langle 1,2,5\rangle, \quad \boldsymbol{v}_2 = \langle -2,1,0\rangle, \quad \boldsymbol{v}_3 = \langle c_2, c_2, 1\rangle$$

are mutually orthogonal; then find unit vectors, $\boldsymbol{u}_i$ in the direction of $\boldsymbol{v}_i$, for $i = 1,2,3$.

Solution: We need to find c_1 and c_2 solutions of the orthogonality conditions

$$\begin{matrix}\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0, & & \langle 1,2,5\rangle \cdot \langle -2,1,0\rangle = 0, & & 0 = 0,\\ \boldsymbol{v}_1 \cdot \boldsymbol{v}_3 = 0, & \Rightarrow & \langle 1,2,5\rangle \cdot \langle c_1,c_2,1\rangle = 0, & \Rightarrow & c_1 + 2c_2 + 5 = 0,\\ \boldsymbol{v}_2 \cdot \boldsymbol{v}_3 = 0 & & \langle -2,1,0\rangle \cdot \langle c_1,c_2,1\rangle = 0, & & -2c_1 + c_2 = 0.\end{matrix}$$

It is not hard to see that the solution is

$$c_1 = -1, \quad c_2 = -2,$$

which implies that $\boldsymbol{v}_3 = \langle -1,-2,1\rangle$. The length (magnitude) of these vectors is

$$\|\boldsymbol{v}_1\| = \sqrt{30}, \quad \|\boldsymbol{v}_2\| = \sqrt{5}, \quad \|\boldsymbol{v}_3\| = \sqrt{6}.$$

Therefore, the following vectors form an orthonormal set,

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{30}}\langle 1,2,5\rangle, \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{5}}\langle -2,1,0\rangle, \quad \boldsymbol{u}_3 = \frac{1}{\sqrt{6}}\langle -1,-2,1\rangle.$$

◁

Remark: Given two vectors, say $\boldsymbol{u}_1 = \langle 1,1,1\rangle$ and $\boldsymbol{u}_2 = \langle -2,1,1\rangle$, how to do you find a third vector, $\boldsymbol{u}_3$, perpendicular to these two vectors? With the cross product:

$$\boldsymbol{u}_3 = \boldsymbol{u}_1 \times \boldsymbol{u}_2 = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & 1 & 1 \\ -2 & 1 & 1 \end{vmatrix} = (1-1)\boldsymbol{i} - (1+2)\boldsymbol{j} + ((1+2)\boldsymbol{k} = \langle 0,-3,3\rangle.$$

Now it is simple to check that $\boldsymbol{u}_3$ is orthogonal to $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, since

$$\boldsymbol{u}_3 \cdot \boldsymbol{u}_1 = \langle 0,-3,3\rangle \cdot \langle 1,1,1\rangle = -3+3 = 0, \quad \boldsymbol{u}_3 \cdot \boldsymbol{u}_2 = \langle 0,-3,3\rangle \cdot \langle -2,1,1\rangle = -3+3 = 0.$$

4.1.3. Orthogonal Expansions. Sets of n mutually orthogonal vectors in the vector space $\mathbb{R}^n$ can be used to expand any other vector. The following theorem says precisely that, any vector in $\mathbb{R}^n$ can be decomposed as a linear combination of the these vectors. Furthermore, there is a simple formula for the vector components.

Theorem 4.1.8 (Orthogonal Expansion)**.** *Given an orthogonal set $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n\}$ in $\mathbb{R}^n$, every vector $\boldsymbol{v} \in \mathbb{R}^3$ can be decomposed as*

$$\boldsymbol{v} = v_1\, \boldsymbol{u}_1 + \cdots + v_n\, \boldsymbol{u}_n.$$

Furthermore, there is a formula for the vector components,

$$v_1 = \frac{(\boldsymbol{v} \cdot \boldsymbol{u}_1)}{(\boldsymbol{u}_1 \cdot \boldsymbol{u}_1)}, \quad \cdots, \quad v_n = \frac{(\boldsymbol{v} \cdot \boldsymbol{u}_n)}{(\boldsymbol{u}_n \cdot \boldsymbol{u}_n)}.$$

If the vectors are orthonormal, that is orthogonal and unit vectors, then the formula for the components reduces to

$$v_1 = \boldsymbol{v} \cdot \boldsymbol{u}_1, \quad \cdots, \quad v_n = \boldsymbol{v} \cdot \boldsymbol{u}_n.$$

Proof of Theorem 4.1.8: Since the vectors $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ are mutually perpendicular, that means they are linearly independent, so the set of all possible linear combinations of these vectors is the whole space $\mathbb{R}^n$. Therefore, given any vector $\boldsymbol{v} \in \mathbb{R}^n$, there exists constants $v_1, \cdots, v_n$ such that

$$\boldsymbol{v} = v_1\, \boldsymbol{u}_1 + \cdots + v_n\, \boldsymbol{u}_n.$$

Since the vectors $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ are mutually orthogonal, we can compute the dot product of the equation above with $\boldsymbol{u}_1$, and we get

$$\boldsymbol{u}_1 \cdot \boldsymbol{v} = v_1\, \boldsymbol{u}_1 \cdot \boldsymbol{u}_1 + \cdots + v_n\, \boldsymbol{u}_1 \cdot \boldsymbol{u}_n = v_1\, \boldsymbol{u}_1 \cdot \boldsymbol{u}_1 + 0 + \cdots + 0,$$

therefore, we get a formula for the component v_1,

$$v_1 = \frac{(\boldsymbol{v} \cdot \boldsymbol{u}_1)}{(\boldsymbol{u}_1 \cdot \boldsymbol{u}_1)}.$$

If the vector $\boldsymbol{u}_1$ is an unit vector, then $\boldsymbol{u}_1 \cdot \boldsymbol{u}_1 = 1$. A similar calculation provides the formulas for v_i, with $i = 2, \cdots n$. This establishes the Theorem. □

Example 4.1.9 (Orthogonal Expansion)**.** Find the expansion of the vector

$$\boldsymbol{w} = \langle 3, -2, 4 \rangle$$

on the orthogonal set

$$\{\boldsymbol{v}_1 = \langle 1, 1, 1 \rangle,\ \boldsymbol{v}_2 = \langle -2, 1, 1 \rangle,\ \boldsymbol{v}_3 = \langle 0, -3, 3 \rangle\}.$$

Solution: We know that the set above is orthogonal. So there are coefficients

$$w_1,\ w_2,\ w_3,$$

such that

$$\boldsymbol{w} = w_1\, \boldsymbol{v}_1 + w_2\, \boldsymbol{v}_2 + w_3\, \boldsymbol{v}_3,$$

where the coefficient w_i, with $i = 1, 2, 3$ are given by the formula

$$w_i = \frac{(\boldsymbol{w} \cdot \boldsymbol{v}_i)}{(\boldsymbol{v}_i \cdot \boldsymbol{v}_i)}.$$

So we get,

$$w_1 = \frac{\boldsymbol{w}\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1} = \frac{\langle 3,-2,4\rangle\cdot\langle 1,1,1\rangle}{\langle 1,1,1\rangle\cdot\langle 1,1,1\rangle} = \frac{5}{3},$$

$$w_2 = \frac{\boldsymbol{w}\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2} = \frac{\langle 3,-2,4\rangle\cdot\langle -2,1,1\rangle}{\langle -2,1,1\rangle\cdot\langle -2,1,1\rangle} = -\frac{4}{6} = -\frac{2}{3},$$

$$w_3 = \frac{\boldsymbol{w}\cdot\boldsymbol{v}_3}{\boldsymbol{v}_3\cdot\boldsymbol{v}_3} = \frac{\langle 3,-2,4\rangle\cdot\langle 0,-3,3\rangle}{\langle 0,-3,3\rangle\cdot\langle 0,-3,3\rangle} = \frac{18}{18} = 1.$$

Therefore, the vector $\boldsymbol{w} = \langle 3,-2,4\rangle$ is given by

$$\langle 3,-2,4\rangle = \frac{5}{3}\langle 1,1,1\rangle - \frac{2}{3}\langle -2,1,1\rangle + \langle 0,-3,3\rangle.$$

◁

Example 4.1.10 (Orthonormal Expansion)**.** Find the expansion of the vector

$$\boldsymbol{w} = \langle 3,-2,4\rangle$$

on the orthonormal set

$$\{\boldsymbol{u}_1 = \frac{1}{\sqrt{3}}\langle 1,1,1\rangle,\ \boldsymbol{u}_2 = \frac{1}{\sqrt{6}}\langle -2,1,1\rangle,\ \boldsymbol{u}_3 = \frac{1}{\sqrt{2}}\langle 0,-1,1\rangle\}.$$

Solution: We know that the set above is orthonormal. So there exist coefficients

$$\tilde{w}_1,\ \tilde{w}_2,\ \tilde{w}_3,$$

such that

$$\boldsymbol{w} = \tilde{w}_1\,\boldsymbol{u}_1 + \tilde{w}_2\,\boldsymbol{u}_2 + \tilde{w}_3\,\boldsymbol{u}_3,$$

where the coefficients $\tilde{w}_i$, with $i = 1,2,3$ are given by

$$\tilde{w}_i = (\boldsymbol{w}\cdot\boldsymbol{u}_i).$$

So we get,

$$\tilde{w}_1 = \boldsymbol{w}\cdot\boldsymbol{v}_1 = \langle 3,-2,4\rangle\cdot\frac{1}{\sqrt{3}}\langle 1,1,1\rangle = \frac{5}{\sqrt{3}},$$

$$\tilde{w}_2 = \boldsymbol{w}\cdot\boldsymbol{v}_2 = \langle 3,-2,4\rangle\cdot\frac{1}{\sqrt{6}}\langle -2,1,1\rangle = -\frac{4}{\sqrt{6}},$$

$$\tilde{w}_3 = \boldsymbol{w}\cdot\boldsymbol{v}_3 = \langle 3,-2,4\rangle\cdot\frac{1}{\sqrt{2}}\langle 0,-1,1\rangle = \frac{6}{\sqrt{2}}.$$

Therefore, the vector $\boldsymbol{w} = \langle 3,-2,4\rangle$ is given by

$$\langle 3,-2,4\rangle = \frac{5}{\sqrt{3}}\Big(\frac{1}{\sqrt{3}}\langle 1,1,1\rangle\Big) - \frac{4}{\sqrt{6}}\Big(\frac{1}{\sqrt{6}}\langle -2,1,1\rangle\Big) + \frac{6}{\sqrt{2}}\Big(\frac{1}{\sqrt{2}}\langle 0,-1,1\rangle\Big).$$

◁

4.1.4. Exercises.

4.1.1.- .

4.1.2.- .

4.2. Matrix Algebra

In this section we introduce an $m \times n$ matrix as a generalization of an m-vector from one to n columns. Then we show that $m \times n$ matrices are functions from the vector space of n-vectors to the vector space of m-vectors. This leads us to introduce operations on matrices just as we do it for regular functions on real numbers. We show how to compute linear combinations of matrices, the multiplication—which is just the composition—of matrices, the trace, determinant, and the inverse of a square matrix. These operations will be needed later when we study systems of differential equations.

4.2.1. Matrices and Linear Combinations.

We follow the lazy way to introduce matrices, which is as a generalization of column vectors from one column to multiple columns. While this approach is straightforward, it hides the fact that matrices are linear functions on vectors, which has to be discovered a little later. Before we start we review the notation introduced in the previous section. Let us recall that $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, which means $\mathbb{F}$ is either the set of real numbers $\mathbb{R}$ or the set of complex numbers $\mathbb{C}$. Then the space of all n-vectors, real or complex, is denoted as $\mathbb{F}^n$, while the space of all $m \times n$ matrices, again real or complex, will be $\mathbb{F}^{m,n}$.

Definition 4.2.1. *An $m \times n$ **matrix**, A, is an ordered array of numbers,*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}.$$

*The numbers $A_{i,j} \in \mathbb{F}$, for $i = 1, \cdots, m$ and $j = 1, \cdots n$, are called the matrix components. The space of all $m \times n$ matrices with components in $\mathbb{F}$ is called $\mathbb{F}^{m,n}$. The $n \times n$ matrices are called **square** matrices.*

Remark: We will use the *component notation* $A = [A_{ij}]$ to denote an $m \times n$ matrix, where A_{ij} are the matrix components. Analogously, an n-vector is denoted by $\boldsymbol{v} = [v_j]$, where v_j are the components of the vector $\boldsymbol{v}$.

Example 4.2.1. Examples of 2×2, 3×2, 2×3 and 3×3 matrices are, respectively,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

◁

We could right away define operation on matrices as generalization of operations on vectors. For example we can define the linear combination of matrices component wise in the same way we did it for vectors. But, before we go there it is useful to introduce one operation between matrices and vectors, called the matrix-vector product.

Definition 4.2.2. *The **matrix-vector product** of an $m \times n$ matrix $A = [A_{ij}]$ and an n-vector $\boldsymbol{u} = [u_j]$ is the m-vector $A\boldsymbol{u}$ given by*

$$A\boldsymbol{u} = \begin{bmatrix} A_{11}u_1 + \cdots + A_{1n}u_n \\ \vdots \\ A_{m1}u_1 + \cdots + A_{mn}u_n \end{bmatrix}.$$

Remark: If we use only components notation, the definition of the matrix-vector product above looks like

$$\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} A_{11}u_1 + \cdots + A_{1n}u_n \\ \vdots \\ A_{m1}u_1 + \cdots + A_{mn}u_n \end{bmatrix}.$$

Example 4.2.2. Below we compute a few matrix-vector products.

(a)

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ \mathbf{2} \end{bmatrix} = \begin{bmatrix} (1)(\mathbf{1}) + (2)(\mathbf{2}) \\ (3)(\mathbf{1}) + (4)(\mathbf{2}) \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \end{bmatrix}.$$

(b)

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 3 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \end{bmatrix} = \begin{bmatrix} (1)(\mathbf{1}) + (2)(\mathbf{2}) + (3)(\mathbf{3}) \\ (4)(\mathbf{1}) + (3)(\mathbf{2}) + (2)(\mathbf{3}) \end{bmatrix} = \begin{bmatrix} 14 \\ 16 \end{bmatrix}.$$

(c)

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ \mathbf{2} \end{bmatrix} = \begin{bmatrix} (1)(\mathbf{1}) + (2)(\mathbf{2}) \\ (3)(\mathbf{1}) + (4)(\mathbf{2}) \\ (5)(\mathbf{1}) + (6)(\mathbf{2}) \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \\ 17 \end{bmatrix}.$$

◁

A matrix together with the matrix-vector product imply that a matrix is a function on the space of vectors. Indeed, an $m \times n$ matrix A takes an n-vector $\boldsymbol{x}$ and creates a unique m-vector $\boldsymbol{y}$,

$$\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} A_{11}x_1 + \cdots + A_{1n}x_n \\ \vdots \\ A_{m1}x_1 + \cdots + A_{mn}x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \quad \Leftrightarrow \quad A\boldsymbol{x} = \boldsymbol{y}.$$

Therefore, matrices are actually functions on vectors. In fact, it is not difficult to see that matrices are linear functions on vectors, that is, for all n-vectors $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and all scalars a, b holds

$$A(a\boldsymbol{x}_1 + b\boldsymbol{x}_2) = a\,A\boldsymbol{x}_1 + b\,A\boldsymbol{x}_2.$$

It is also not difficult to see that the matrices can stretch or compress vectors, and they can change the vector direction.

This matrix-vector product gives us a way to generalize the linear combination of vectors into a linear combination of matrices. Recall that we defined the linear combination of two vectors, say $\boldsymbol{u} = [u_i]$ and $\boldsymbol{v} = [v_i]$, component wise as follows

$$a \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + b \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} (au_1 + bv_1) \\ \vdots \\ (au_n + bv_n) \end{bmatrix} \quad \Leftrightarrow \quad a[u_i] + b[v_i] = [au_i + bv_i],$$

where a, b are arbitrary scalars. We can do the same for matrices. The linear combination of two $m \times n$ matrices, A and B, is the following: for all scalar numbers a and b we define the $m \times n$ matrix $M = (aA + bB)$ as follows,

$$M\boldsymbol{x} = a\,A\boldsymbol{x} + b\,B\boldsymbol{x},$$

for all $\boldsymbol{x} \in \mathbb{F}^n$. Then it is simply to find a formula for the components of M in terms of the components of A and B,

$$M_{ij} = aA_{ij} + bB_{ij}, \qquad i = 1, \cdots, m, \quad j = 1, \cdots n.$$

In other words,

$$(aA + bB)_{i,j} = aA_{ij} + bB_{ij} \qquad i = 1, \cdots, m, \quad j = 1, \cdots n.$$

We can see how this definition works in the particular case of 2×2 matrices.

Example 4.2.3. Compute the components of the linear combination matrix

$$M = aA + bB$$

in terms of the components of the matrices A and B, for

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Solution: On the one hand we have

$$M\boldsymbol{x} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

On the other hand the definition of M says that

$$\begin{aligned} M\boldsymbol{x} &= a\,A\boldsymbol{x} + b\,B\boldsymbol{x} \\ &= a \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + b \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} aA_{11}x_1 + aA_{12}x_2 \\ aA_{21}x_1 + aA_{22}x_2 \end{bmatrix} + \begin{bmatrix} bB_{11}x_1 + bB_{12}x_2 \\ bB_{21}x_1 + bB_{22}x_2 \end{bmatrix} \\ &= \begin{bmatrix} (aA_{11} + bB_{11})x_1 + (aA_{12} + bB_{12})x_2 \\ (aA_{21} + bB_{21})x_1 + (aA_{22} + bB_{22})x_2 \end{bmatrix}. \end{aligned}$$

Therefore, we obtain

$$M\boldsymbol{x} = \begin{bmatrix} (aA_{11} + bB_{11}) & (aA_{12} + bB_{12}) \\ (aA_{21} + bB_{21}) & (aA_{22} + bB_{22}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Since the calculation above holds for all $\boldsymbol{x} \in \mathbb{R}^2$, we get a formula for the components of the matrix linear combination,

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} (aA_{11} + bB_{11}) & (aA_{12} + bB_{12}) \\ (aA_{21} + bB_{21}) & (aA_{22} + bB_{22}) \end{bmatrix},$$

which means that

$$M_{ij} = (aA_{ij} + bB_{ij}) \quad \Leftrightarrow \quad (aA + bB)_{ij} = (aA_{ij} + bB_{ij}), \quad i, j = 1, 2.$$

◁

The calculation done in the example above for 2×2 matrices can be easily generalized to $m \times n$ matrices. For these arbitrary matrices it is important to use index notation, to keep the equations small enough.

Definition 4.2.3. *Let $A = [A_{ij}]$ and $B = [B_{ij}]$ be $m \times n$ matrices in $\mathbb{F}^{m,n}$ and $a, b \in \mathbb{F}$ be scalars. The **linear combination** of A and B is also an $m \times n$ matrix in $\mathbb{F}^{m,n}$, denoted as $a\,A + b\,B$, and given by*

$$a\,A + b\,B = [a\,A_{ij} + b\,B_{ij}].$$

The particular case $b = 0$ corresponds to the multiplication of a matrix by a number, while the case $a = b = 1$ corresponds to the addition of two matrices, that is,

$$a\,A = [a\,A_{ij}], \qquad A + B = [A_{ij} + B_{ij}].$$

Example 4.2.4. Find the $A+B$, where $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 3 \\ 5 & 1 \end{bmatrix}$.

Solution: The addition of two equal size matrices is performed component-wise:

$$A + B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} (1+2) & (2+3) \\ (3+5) & (4+1) \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 8 & 5 \end{bmatrix}.$$

Therefore, we conclude

$$A + B = \begin{bmatrix} 3 & 5 \\ 8 & 5 \end{bmatrix}.$$

◁

Example 4.2.5. Find the $A+B$, where $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$.

Solution: The matrices have different sizes, so their addition is not defined. ◁

Example 4.2.6. Find $2A$ and $A/3$, where $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$.

Solution: The multiplication of a matrix by a number is done component-wise, therefore

$$2A = 2 \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 10 \\ 4 & 8 & 12 \end{bmatrix}, \qquad \frac{A}{3} = \frac{1}{3} \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & 1 & \frac{5}{3} \\ \frac{2}{3} & \frac{4}{3} & 2 \end{bmatrix}.$$

◁

4.2.2. Matrix Multiplication. The operation of matrix multiplication originates in the composition of functions. We call it matrix multiplication instead of matrix composition because it reduces to the multiplication of real numbers in the case of 1×1 real matrices. Unlike the multiplication of real numbers, the multiplication of general matrices is *not commutative*, that is, $AB \neq BA$ in the general case. This property comes from the fact that the composition of two functions is also not commutative.

Definition 4.2.4. *The **matrix multiplication** of an $m \times n$ matrix $A \in \mathbb{F}^{m,n}$ and an $n \times \ell$ matrix $B = \left[\boldsymbol{b}_1, \cdots, \boldsymbol{b}_\ell\right] \in \mathbb{F}^{n,\ell}$ is given by*

$$AB = \left[A\boldsymbol{b}_1, \cdots, A\boldsymbol{b}_\ell\right], \tag{4.2.1}$$

where $A\boldsymbol{b}_i$ are the matrix-vector product of matrix A and the n-vectors $\boldsymbol{b}_i$ for $i = 1, \cdots, \ell$.

The product is not defined for two arbitrary matrices, since the matrix-vector product $A\boldsymbol{b}_i$ must be well-defined. This means that the numbers of columns in the first matrix must match the numbers of rows in the second matrix.

$$\underset{m \times \boldsymbol{n}}{A} \quad \text{times} \quad \underset{\boldsymbol{n} \times \ell}{B} \quad \text{defines} \quad \underset{m \times \ell}{AB}$$

Example 4.2.7. Compute the product AB, where

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}.$$

Solution: We use Eq. (4.2.1) and we get

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \\ &= \left[\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} \right), \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} \right) \right] \\ &= \left[\begin{bmatrix} (5+14) \\ (15+28) \end{bmatrix}, \begin{bmatrix} (6+16) \\ (18+32) \end{bmatrix} \right] \\ &= \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}, \end{aligned}$$

therefore, we conclude that

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}.$$

Remark: Notice that there is a somewhat simpler way to obtain this product. The component $(AB)_{11} = 19$ is obtained from the first row in matrix A and the first column in matrix B as follows:

$$\begin{bmatrix} \mathbf{1} & \mathbf{2} \\ 3 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{5} & 6 \\ \mathbf{7} & 8 \end{bmatrix} = \begin{bmatrix} \mathbf{19} & 22 \\ 43 & 50 \end{bmatrix}, \qquad (1)(5) + (2)(7) = 19;$$

The component $(AB)_{12} = 22$ is obtained as follows:

$$\begin{bmatrix} \mathbf{1} & \mathbf{2} \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & \mathbf{6} \\ 7 & \mathbf{8} \end{bmatrix} = \begin{bmatrix} 19 & \mathbf{22} \\ 43 & 50 \end{bmatrix}, \qquad (1)(6) + (2)(8) = 22;$$

The component $(AB)_{21} = 43$ is obtained as follows:

$$\begin{bmatrix} 1 & 2 \\ \mathbf{3} & \mathbf{4} \end{bmatrix} \begin{bmatrix} \mathbf{5} & 6 \\ \mathbf{7} & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ \mathbf{43} & 22 \end{bmatrix}, \qquad (3)(5) + (4)(7) = 43;$$

And finally the component $(AB)_{22} = 50$ is obtained as follows:

$$\begin{bmatrix} 1 & 2 \\ \mathbf{3} & \mathbf{4} \end{bmatrix} \begin{bmatrix} 5 & \mathbf{6} \\ 7 & \mathbf{8} \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & \mathbf{50} \end{bmatrix}, \qquad (3)(6) + (4)(8) = 50.$$

$\lhd$

The alternative way to compute a matrix product mentioned in the example above can be generalized to all matrix products. This result is summarized in the following statement.

Theorem 4.2.5 (Components Formula)**.** *The matrix multiplication of an $m \times n$ matrix $A = [A_{ij}] \in \mathbb{F}^{m,n}$ and an $n \times \ell$ matrix $B = [B_{jk}] \in \mathbb{F}^{n,\ell}$, where $i = 1, \cdots, m$, $j = 1, \cdots, n$ and $k = 1, \cdots, \ell$, is the $m \times \ell$ matrix $AB \in \mathbb{F}^{m,\ell}$ given by*

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

Proof of Theorem 4.2.5: According to the Definition 4.2.4 we write the product AB as

$$AB = \left[A\boldsymbol{b}_1, \cdots, A\boldsymbol{b}_\ell \right].$$

The component $(AB)_{ij}$ of the matrix AB is the component i of the vector $A\boldsymbol{b}_j$, that is,

$$(AB)_{ij} = (A\boldsymbol{b}_j)_i,$$

where $i = 1, \cdots, m$ and $j = 1, \cdots, \ell$. Since the vectors $\boldsymbol{b}_j$ are the columns of matrix B, the vector components are given by

$$(\boldsymbol{b}_j)_k = B_{kj}, \qquad k = 1, \cdots n.$$

Then, we can write the components of the vectors $A\boldsymbol{b}_j$ as follows,

$$(A\boldsymbol{b}_j)_i = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

We then conclude that the components of the matrix product AB are given by

$$(AB)_{i,j} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

This establishes the Theorem. □

Example 4.2.8. Compute AB, where $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 0 \\ 2 & -1 \end{bmatrix}$.

Solution: The component $(AB)_{11} = 4$ is obtained from the first row in matrix A and the first column in matrix B as follows:

$$\begin{bmatrix} \mathbf{2} & \mathbf{-1} \\ -1 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{3} & 0 \\ \mathbf{2} & -1 \end{bmatrix} = \begin{bmatrix} \mathbf{4} & 1 \\ 1 & -2 \end{bmatrix}, \qquad (2)(3) + (-1)(2) = 4;$$

The component $(AB)_{12} = -1$ is obtained as follows:

$$\begin{bmatrix} \mathbf{2} & \mathbf{-1} \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 3 & \mathbf{0} \\ 2 & \mathbf{-1} \end{bmatrix} = \begin{bmatrix} 4 & \mathbf{1} \\ 1 & -2 \end{bmatrix}, \qquad (2)(0) + (-1)(1) = -1;$$

The component $(AB)_{21} = 1$ is obtained as follows:

$$\begin{bmatrix} 2 & -1 \\ \mathbf{-1} & \mathbf{2} \end{bmatrix} \begin{bmatrix} \mathbf{3} & 0 \\ \mathbf{2} & -1 \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ \mathbf{1} & -2 \end{bmatrix}, \qquad (-1)(3) + (2)(2) = 1;$$

And finally the component $(AB)_{22} = -2$ is obtained as follows:

$$\begin{bmatrix} 2 & -1 \\ \mathbf{-1} & \mathbf{2} \end{bmatrix} \begin{bmatrix} 3 & \mathbf{0} \\ 2 & \mathbf{-1} \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ 1 & \mathbf{-2} \end{bmatrix}, \qquad (-1)(0) + (2)(-1) = -2.$$

◁

Example 4.2.9. Compute BA, where $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 0 \\ 2 & -1 \end{bmatrix}$.

Solution: We find that $BA = \begin{bmatrix} 6 & -3 \\ 5 & -4 \end{bmatrix}$. Notice that in this case $AB \neq BA$. ◁

We can see from the previous two examples that the matrix product is not commutative, since we have computed both AB and BA and $AB \neq BA$. It is even more interesting to note that sometimes the matrix multiplication of matrices A and B may be defined for AB but it may not be even defined for BA.

Example 4.2.10. Compute AB and BA, where $A = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$.

Solution: The product AB is

$$AB = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \Rightarrow \quad AB = \begin{bmatrix} 16 & 23 & 30 \\ 6 & 9 & 12 \end{bmatrix}.$$

The product BA is not possible. ◁

Example 4.2.11. Compute AB and BA, where $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$.

Solution: We find that

$$AB = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix},$$

$$BA = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Remarks:

(a) Notice that in this case $AB \neq BA$.
(b) Also notice that $BA = 0$ but $A \neq 0$ and $B \neq 0$.
(c) Therefore, the product $AB = 0$ *does not imply* that either $A = 0$ or $B = 0$.

◁

4.2.3. Other Matrix Operations. Matrices can also be thought as generalizations of scalar-valued functions, in particular of the linear scalar-valued function $f(x) = ax$. The linear combination of matrices is then a generalization of the linear combination of scalar-valued functions. However, one can define operations on matrices that, unlike linear combinations or compositions, have no analogs on scalar-valued functions. One of such operations is the transpose of a matrix, which is a new matrix with the rows and columns interchanged.

Definition 4.2.6. *The* ***transpose*** *of an $m \times n$ matrix cb$A = [A_{ij}] \in \mathbb{F}^{m,n}$ is the $n \times m$ matrix denoted as $A^T = \big[(A^T)_{kl}\big] \in \mathbb{F}^{n,m}$, with its components given by*

$$\big(A^T\big)_{kl} = A_{lk}.$$

Remark: As we said above, if a matrix A is $m \times n$, then its transpose A^T is $n \times m$. We can compute A^T simply by interchanging rows and columns.

Example 4.2.12. Find the transpose of the matrices A, B, and C below,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}.$$

Solution: If a matrix M has components M_{ij}, then its transpose has components given by $(M^T)_{ji} = M_{ij}$. Therefore, the transpose of a 3×3 is also 3×3, the transpose of a 3×2 matrix is 2×3, and a transpose of a 2×3 matrix is 3×2. For the matrices A, B, and C above we get

$$A^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}, \quad B^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

◁

If a matrix has complex-valued coefficients, then the conjugate of a matrix can be defined as the conjugate of each component.

Definition 4.2.7. *The* ***complex conjugate*** *of a matrix* $A = [A_{ij}] \in \mathbb{F}^{m,n}$ *is the matrix*

$$\overline{A} = \big[\,\overline{A}_{ij}\big] \in \mathbb{F}^{m,n}.$$

Example 4.2.13. A matrix A and its conjugate is given below,

$$A = \begin{bmatrix} 1 & 2+i \\ -i & 3-4i \end{bmatrix}, \quad \Leftrightarrow \quad \overline{A} = \begin{bmatrix} 1 & 2-i \\ i & 3+4i \end{bmatrix}.$$

◁

Example 4.2.14. A matrix A has real coefficients iff $A = \overline{A}$; It has purely imaginary coefficients iff $A = -\overline{A}$. Here are examples of these two situations:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \Rightarrow \quad \overline{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = A;$$

$$A = \begin{bmatrix} i & 2i \\ 3i & 4i \end{bmatrix} \quad \Rightarrow \quad \overline{A} = \begin{bmatrix} -i & -2i \\ -3i & -4i \end{bmatrix} = -A.$$

◁

Definition 4.2.8. *The* ***adjoint*** *of a matrix* $A \in \mathbb{F}^{m,n}$ *is the matrix*

$$A^* = \overline{A}^{\,T} \in \mathbb{F}^{n,m}.$$

Remark: Since $\big(\overline{A}\big)^T = \overline{(A^T)}$, the order of the operations does not change the result, that is why there is no parenthesis in the definition of A^*.

Example 4.2.15. A matrix A and its adjoint is given below,

$$A = \begin{bmatrix} 1 & 2+i \\ -i & 3-4i \end{bmatrix}, \quad \Leftrightarrow \quad A^* = \begin{bmatrix} 1 & i \\ 2-i & 3+4i \end{bmatrix}.$$

◁

The transpose, conjugate and adjoint operations are useful to specify certain classes of matrices with particular symmetries. Here we introduce few of these classes.

Definition 4.2.9. *An* $n \times n$ *matrix* A *is called:*

(a) ***symmetric*** *iff* $A = A^T$*;*
(b) ***skew-symmetric*** *iff* $A = -A^T$*;*
(c) ***Hermitian*** *iff* $A = A^*$*;*
(d) ***skew-Hermitian*** *iff* $A = -A^*$*.*

Example 4.2.16. We present examples of each of the classes introduced in Def. 4.2.9.

Part (a): Matrices A and B are symmetric. Notice that A is also Hermitian, while B is not Hermitian,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 7 & 4 \\ 3 & 4 & 8 \end{bmatrix} = A^T, \qquad B = \begin{bmatrix} 1 & 2+3i & 3 \\ 2+3i & 7 & 4i \\ 3 & 4i & 8 \end{bmatrix} = B^T.$$

Part (b): Matrix C is skew-symmetric,

$$C = \begin{bmatrix} 0 & -2 & 3 \\ 2 & 0 & -4 \\ -3 & 4 & 0 \end{bmatrix} \quad \Rightarrow \quad C^T = \begin{bmatrix} 0 & 2 & -3 \\ -2 & 0 & 4 \\ 3 & -4 & 0 \end{bmatrix} = -C.$$

Notice that the diagonal elements in a skew-symmetric matrix must vanish, since the condition $C_{ij} = -C_{ji}$ in the case $i = j$ means $C_{ii} = -C_{ii}$, that is, $C_{ii} = 0$.

Part (c): Matrix D is Hermitian but is not symmetric:

$$D = \begin{bmatrix} 1 & 2+i & 3 \\ 2-i & 7 & 4+i \\ 3 & 4-i & 8 \end{bmatrix} \quad \Rightarrow \quad D^T = \begin{bmatrix} 1 & 2-i & 3 \\ 2+i & 7 & 4-i \\ 3 & 4+i & 8 \end{bmatrix} \neq D,$$

however,

$$D^* = \overline{D}^T = \begin{bmatrix} 1 & 2+i & 3 \\ 2-i & 7 & 4+i \\ 3 & 4-i & 8 \end{bmatrix} = D.$$

Notice that the diagonal elements in a Hermitian matrix must be real numbers, since the condition $A_{ij} = \overline{A}_{ji}$ in the case $i = j$ implies $A_{ii} = \overline{A}_{ii}$, that is,

$$2i\,\mathrm{Im}(A_{ii}) = A_{ii} - \overline{A}_{ii} = 0.$$

We can also verify what we said in part (a), matrix A is Hermitian since

$$A^* = \overline{A}^T = A^T = A.$$

Part (d): The following matrix E is skew-Hermitian:

$$E = \begin{bmatrix} i & 2+i & -3 \\ -2+i & 7i & 4+i \\ 3 & -4+i & 8i \end{bmatrix} \quad \Rightarrow \quad E^T = \begin{bmatrix} i & -2+i & 3 \\ 2+i & 7i & -4+i \\ -3 & 4+i & 8i \end{bmatrix}$$

therefore,

$$E^* = \overline{E}^T \begin{bmatrix} -i & -2-i & 3 \\ 2-i & -7i & -4-i \\ -3 & 4-i & -8i \end{bmatrix} = -E.$$

A skew-Hermitian matrix has purely imaginary elements in its diagonal, and the off diagonal elements have skew-symmetric real parts with symmetric imaginary parts. $\lhd$

Another oeration on matrices that has no analog on scalar functions is the trace of a square matrix. The trace is a number—the sum of all the diagonal elements of the matrix.

Definition 4.2.10. *The **trace** of a square matrix $A = \big[A_{ij}\big] \in \mathbb{F}^{n,n}$, denoted as $\mathrm{tr}\,(A) \in \mathbb{F}$, is the sum of its diagonal elements, that is, the scalar given by*

$$\mathrm{tr}\,(A) = A_{11} + \cdots + A_{nn}.$$

Example 4.2.17. Find the trace of the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

Solution: We only have to add up the diagonal elements:

$$\mathrm{tr}\,(A) = 1 + 5 + 9 \quad \Rightarrow \quad \mathrm{tr}\,(A) = 15.$$

$\lhd$

4.2.4. The Inverse Matrix. Since matrices are functions, one can try to find the inverse of a given matrix. We will show how this can be done only for 2×2 matrices. Before we start we need to introduce one important matrix.

Definition 4.2.11. *The* ***identity matrix*** *is the $n \times n$ given by*

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Remark: Sometimes we denote simply by I the $n \times n$ identity matrix I_n.

Example 4.2.18. The 2×2 and 3×3 identity matrices are

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Also, it is simple to see that the components of the $n \times n$ identity matrix are

$$I = [I_{ij}] \quad \text{with} \quad \begin{cases} I_{ii} = 1 \\ I_{ij} = 0 \quad i \neq j. \end{cases}$$

$\triangleleft$

These matrices I_n are called the identity matrices because of two reasons. First, their action on vectors is the same as the identity function, that is,

$$I_n \boldsymbol{x} = \boldsymbol{x} \qquad \forall \boldsymbol{x} \in \mathbb{F}^n,$$

where the symbol $\forall$ means "for all". It is also not difficult to prove that this property actually defines the identity matrix. Indeed, if a matrix B satisfies

$$B\boldsymbol{x} = \boldsymbol{x} \qquad \forall \boldsymbol{x} \in \mathbb{F}^n,$$

then $B = I_n$. The second reason these matrices are called identity matrices is that they act as the identity in the matrix multiplication, that is,

$$I_n A = A \qquad A I_n = A \qquad \forall A \in \mathbb{F}^{n,n}.$$

Again, this property defines the identity matrix. Indeed, if a matrix B satisfies

$$BA = A \qquad \forall A \in \mathbb{F}^{n,n},$$

then $B = I_n$. Same is true if we use the other product AB.

Now we are ready to introduce the concept of the inverse of a matrix.

Definition 4.2.12. *An $n \times n$ matrix $A \in \mathbb{F}^{n,n}$ is called* ***invertible*** *iff there exists another $n \times n$ matrix, denoted as A^{-1}, such that*

$$(A^{-1})A = I_n \quad \text{and} \quad A(A^{-1}) = I_n.$$

Remark: Notice we only need one of the equations in Definition 4.2.12, since one of the equations implies the other. For example, suppose we have a matrix A^{-1} such that

$$A^{-1}A = I_n.$$

Then, multiply this equation on the left by matrix A,

$$A\,(A^{-1}A) = A \quad \Rightarrow \quad (A\,A^{-1})\,A = A. \tag{4.2.2}$$

Now we need to show that matrix $A\,A^{-1}$ is the identity matrix I_n. To see that, first notice that the system of linear algebraic equations

$$A\boldsymbol{x} = \boldsymbol{y}$$

has a unique solution $\boldsymbol{x}$ for all $\boldsymbol{y} \in \mathbb{F}^n$. Indeed, multiplying this equation by A^{-1} on the left we get

$$A^{-1}A\boldsymbol{x} = A^{-1}\boldsymbol{y} \quad \Rightarrow \quad I_n\boldsymbol{x} = A^{-1}\boldsymbol{y} \quad \Rightarrow \quad \boldsymbol{x} = A^{-1}\boldsymbol{y},$$

where we used $A^{-1}A = I_n$. Since the vector $\boldsymbol{y}$ can be any vector in $\mathbb{F}^n$, then the vectors of the form $A\boldsymbol{x}$, must span the whole space $\mathbb{F}^n$ when $\boldsymbol{x}$ is any vector in $\mathbb{F}^n$. We can use this fact in Eq. (4.2.2) and we get

$$(A\,A^{-1})\,A\boldsymbol{x} = A\boldsymbol{x} \qquad \forall \boldsymbol{x} \in \mathbb{F}^n \quad \Leftrightarrow \quad (A\,A^{-1})\boldsymbol{y} = \boldsymbol{y} \qquad \forall \boldsymbol{y} \in \mathbb{F}^n,$$

where we recall that $\forall$ means "for all". Therefore, the matrix $A\,A^{-1}$ must be the identity matrix,

$$A\,A^{-1} = I_n.$$

This establishes the Remark.

Below we show an example of a matrix and its inverse in the case of 2×2 matrices.

Example 4.2.19. Verify that the matrix A^{-1} below is the inverse of matrix A, where

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}, \qquad A^{-1} = \frac{1}{4}\begin{bmatrix} 3 & -2 \\ -1 & 2 \end{bmatrix}.$$

Solution: We have to compute the products,

$$A\big(A^{-1}\big) = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} \frac{1}{4}\begin{bmatrix} 3 & -2 \\ -1 & 2 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}.$$

Therefore, we conclude that

$$A\big(A^{-1}\big) = I_2.$$

It is simple to check that the equation $\big(A^{-1}\big)A = I_2$ also holds Therefore, A^{-1} is indeed the inverse of A. ◁

Our next results applies only to 2×2 matrices. There is a simple condition on a 2×2 matrix to determine if this matrix is invertible and, in the case this condition is satisfied, there is a simple formula for the inverse matrix.

Theorem 4.2.13. *Given a 2×2 matrix A below introduce the number Δ as follows,*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad \Delta = ad - bc.$$

The matrix A is invertible iff $\Delta \neq 0$. Furthermore, if A is invertible, its inverse is

$$A^{-1} = \frac{1}{\Delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{4.2.3}$$

The number Δ is called the *determinant* of A, since it is the number that determines whether A is invertible or not.

Remark: In the case that the determinant $\Delta \neq 0$ we only have to check that the formula for the inverse matrix A^{-1} given in the theorem satisfies satisfies $A^{-1}A = I_2$. A straightforward calculation says that

$$A^{-1}\,A = \frac{1}{\Delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

It is also straightforward to check the other equation, (although we do not need to do it, check the Remark below Definition 4.2.12),

$$A\,A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\frac{1}{\Delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

However, it is not clear what happens in the case that $\Delta = 0$. In the proof below we construct A^{-1}. Then we can see that this construction works if and ony if $\Delta \neq 0$.

Proof of Theorem 4.2.13: let us think matrix A as function that maps every vector $\boldsymbol{x} = \langle x_1, x_2\rangle$ to a vector $\boldsymbol{y} = \langle y_1, y_2\rangle$, as follows

$$A\boldsymbol{x} = \boldsymbol{y} \quad \Rightarrow \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} a\,x_1 + b\,x_2 = y_1, \\ c\,x_1 + d\,x_2 = y_2. \end{cases}$$

Multiply by d the first equation on the far right above and by $-b$ the second equation on that far right and then add them up; we get

$$da\,x_1 + \cancel{db\,x_2} - bc\,x_1 - \cancel{bd\,x_2} = d\,y_1 - b\,y_2,$$

that is

$$\Delta\,x_1 = d\,y_1 - b\,y_2.$$

We see that we can solve for x_1 if and only if $\Delta \neq 0$. And in that case we get

$$x_1 = \frac{1}{\Delta}(d\,y_1 - b\,y_2).$$

Now multiply by c the first equation on the far right above and by $-a$ the second equation on that far right and then add them up; we get

$$\cancel{ca\,x_1} + cb\,x_2 - \cancel{ac\,x_1} - ad\,x_2 = c\,y_1 - a\,y_2,$$

that is

$$-\Delta\,x_2 = c\,y_1 - a\,y_2.$$

Again we see that we can solve for x_2 if and only if $\Delta \neq 0$. And in that case we get

$$x_2 = \frac{1}{\Delta}(-c\,y_1 + a\,y_2).$$

These formulas for x_1, x_2 say that

$$\left.\begin{aligned} x_1 &= \frac{1}{\Delta}(d\,y_1 - b\,y_2), \\ x_2 &= \frac{1}{\Delta}(-c\,y_1 + a\,y_2) \end{aligned}\right\} \quad \Rightarrow \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

In matrix notation we can summarize our calculations as follows. The system

$$A\boldsymbol{x} = \boldsymbol{y}, \quad \text{with} \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad \Delta = ad - bc \neq 0$$

has the solution

$$\boldsymbol{x} = A^{-1}\boldsymbol{y} \quad \text{with} \quad A^{-1} = \frac{1}{\Delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

The matrix A^{-1} exists if and only if the number $\Delta = ad - bc \neq 0$. We have already checked that this matrix satisfies $A^{-1}A = I_2$ and $A\,A^{-1} = I_2$, so this matrix A^{-1} is indeed the inverse of matrix A. This establishes the Theorem. □

Example 4.2.20. Compute the inverse of matrix $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$, given in Example 4.2.19.

Solution: Following Theorem 4.2.13 we first compute $\Delta = 6 - 4 = 4$. Since $\Delta \neq 0$, then A^{-1} exists and it is given by

$$A^{-1} = \frac{1}{4}\begin{bmatrix} 3 & -2 \\ -1 & 2 \end{bmatrix}.$$

◁

Example 4.2.21. Compute the inverse of matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$.

Solution: Following Theorem 4.2.13 we first compute $\Delta = 6 - 6 = 0$. Since $\Delta = 0$, then matrix A is not invertible. ◁

The matrix operations we have introduced are useful to solve matrix equations, where the unknown is a matrix. We now show an example of a matrix equation.

Example 4.2.22. Find a matrix X such that $AXB = I$, where

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Solution: There are many ways to solve a matrix equation. We choose to multiply the equation by the inverses of matrix A and B, if they exist. So first we check whether A is invertible. But

$$\det(A) = \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} = 1 - 6 = -5 \neq 0,$$

so A is indeed invertible. Regarding matrix B we get

$$\det(B) = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 4 - 1 = 3 \neq 0,$$

so B is also invertible. We then compute their inverses,

$$A^{-1} = \frac{1}{-5}\begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix}, \qquad B = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

We can now compute X,

$$AXB = I \quad \Rightarrow \quad A^{-1}(AXB)B^{-1} = A^{-1}IB^{-1} \quad \Rightarrow \quad X = A^{-1}B^{-1}.$$

Therefore,

$$X = \frac{1}{-5}\begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = -\frac{1}{15}\begin{bmatrix} 5 & -7 \\ -5 & 4 \end{bmatrix}$$

so we obtain

$$X = \begin{bmatrix} -\frac{1}{3} & \frac{7}{15} \\ \frac{1}{3} & -\frac{4}{15} \end{bmatrix}.$$

◁

Similarly to the 2×2 case, not every $n\times n$ matrix is invertible. It turns out one can define a number computed out of the matrix components that determines whether the matrix is invertible or not. That number is called the determinant, and we already introduced it for 2×2 matrices. We now give a formula for the determinant of a 3×3 matrix, and we mention that this number can be also defined on $n \times n$ matrices.

4.2.5. Overview of Determinants. A determinant is a number computed from a square matrix that gives some information about the matrix, for example if the matrix is invertible or not. We have already defined the determinant for 2×2 matrices. We now rewrite that definition using a more standard notation.

Definition 4.2.14. *The* ***determinant of a*** 2×2 *matrix* $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ *is given by*

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Remark: The absolute value of the determinant of a 2×2 matrix

$$A = \big[\boldsymbol{a}_1, \boldsymbol{a}_2\big]$$

has a geometrical meaning—it is the area of the parallelogram whose sides are given the vectors in the columns of the matrix A, as shown in Fig. 6.

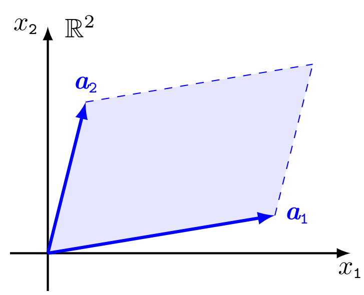


FIGURE 6. Geometrical meaning of the determinant of a 2×2 matrix.

It turns out that this geometrical meaning of the determinant is crucial to extend the definition of the determinant from 2×2 to 3×3 matrices. The result is given in the following definition.

Definition 4.2.15. *The* ***determinant of a*** 3×3 *matrix* $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ *is*

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Remarks:

(1) This is a recursive definition. The determinant of a 3×3 matrix is written in terms of three determinants of 2×2 matrices.

(2) The absolute value of the determinant of a 3×3 matrix

$$A = \big[\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\big]$$

is the volume of the parallelepiped formed by the columns of matrix A, as pictured in Fig. 7.

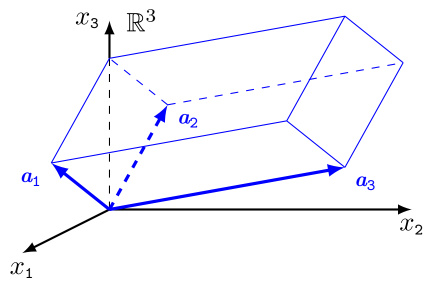


FIGURE 7. Geometrical meaning of the determinant of a 3×3 matrix.

Example 4.2.23. The following three examples show that the determinant can be a negative, zero or positive number.

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2, \qquad \begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix} = 8 - 3 = 5, \qquad \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 4 - 4 = 0.$$

The following is an example shows how to compute the determinant of a 3×3 matrix,

$$\begin{aligned} \begin{vmatrix} 1 & 3 & -1 \\ 2 & 1 & 1 \\ 3 & 2 & 1 \end{vmatrix} &= (1) \begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} - 3 \begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} + (-1) \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} \\ &= (1 - 2) - 3\,(2 - 3) - (4 - 3) \\ &= -1 + 3 - 1 \\ &= 1. \end{aligned}$$

$\triangleleft$

Remark: The determinant of upper or lower triangular matrices is the product of the diagonal coefficients.

Example 4.2.24. Compute the determinant of a 3×3 upper triangular matrix.

Solution:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ 0 & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} 0 & a_{23} \\ 0 & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} 0 & a_{22} \\ 0 & 0 \end{vmatrix} = a_{11} a_{22} a_{33}.$$

$\triangleleft$

Remark: Recall that one can prove that a 3×3 matrix is invertible if an only if it determinant is nonzero.

Example 4.2.25. Is matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 9 \end{bmatrix}$ invertible?

Solution: We only need to compute the determinant of A.

$$\det(A) = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 9 \end{vmatrix} = (1) \begin{vmatrix} 5 & 7 \\ 7 & 9 \end{vmatrix} - (2) \begin{vmatrix} 2 & 7 \\ 3 & 9 \end{vmatrix} + (3) \begin{vmatrix} 2 & 5 \\ 3 & 7 \end{vmatrix}.$$

Using the definition of determinant for 2×2 matrices we obtain

$$\det(A) = (45 - 49) - 2(18 - 21) + 3(14 - 15) = -4 + 6 - 3.$$

Since $\det(A) = -1$, that is, non-zero, matrix A is invertible. $\triangleleft$

The determinant of an $n \times n$ matrix can be defined generalizing the properties that areas of parallelograms have in two dimensions and volumes of parallelepipeds have in three dimensions. One can find a recursive formula for the determinant of an $n \times n$ matrix in terms of n determinants of $(n-1) \times (n-1)$ matrices. And the determinant so defined has its most important property—an $n \times n$ matrix is invertible if and only if its determinant is nonzero.

4.2.6. Exercises.

4.2.1.- .

4.2.2.- .

4.3. Eigenvalues and Eigenvectors

We know that a matrix is a function on vector spaces—a matrix acts on a vector and the result is another vector. In this section we see that, given an $n \times n$ matrix, real or complex, there may exist particular directions in $\mathbb{F}^n$ that are left invariant under the action of the matrix. This means that such a matrix acting on any vector along these particular directions results in a vector along these same directions. The vector is called an eigenvector and the proportionality factor is called an eigenvalue of that matrix.

These particular directions associated to a matrix have important applications. One of these applications is when we try to solve a system of linear differential equations. The eigenvalues and eigenvectors of the coefficient matrix in such system define particularly simple solutions to the differential equation. This was first shown by Augustine Cauchy (1789-1857) in 1840. David Hilbert (1862-1943) arrived at the concept of eigenvalues and eigenvectors when solving a different type of equations called integral equations. We use an adaptation of the names given by Hilbert, "eigenwert" (proper value) and "eigenvektor" (proper vector). In Chapter 5 we will study linear systems of differential equations and we will follow Cauchy to compute the solutions of these systems by finding eigenvalues and eigenvectors of their coefficient matrices.

4.3.1. Definition and Properties.

When a square matrix acts on a vector the result is another vector that, more often than not, points in a different direction from the original vector. However, there may exist vectors whose direction is not changed by the matrix. These will be important for us, so we give them a name.

Definition 4.3.1. *A number $\lambda \in \mathbb{F}$ and a nonzero n-vector $\boldsymbol{v} \in \mathbb{F}^n$ are an* ***eigenvalue*** *with corresponding* ***eigenvector*** *(eigenpair) of an $n \times n$ matrix $A \in \mathbb{F}^{n,n}$ iff they satisfy the equation*

$$A\boldsymbol{v} = \lambda \boldsymbol{v}.$$

Remarks:

(a) We still use the notation from previous sections, where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ means that $\mathbb{F}$ is either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$.

(b) We see that an eigenvector $\boldsymbol{v}$ determines a particular direction in the space $\mathbb{R}^n$, given by $(a\boldsymbol{v})$ for $a \in \mathbb{R}$, that remains invariant under the action of the matrix A. That is, the result of matrix A acting on any vector $(a\boldsymbol{v})$ on the line determined by $\boldsymbol{v}$ is again a vector on the same line, since

$$A(a\boldsymbol{v}) = a\,A\boldsymbol{v} = a\lambda\boldsymbol{v} = \lambda(a\boldsymbol{v}).$$

Example 4.3.1 (Verify Eigenvectors)**.** Verify that the pair λ_1, $\boldsymbol{v}_1$ and the pair λ_2, $\boldsymbol{v}_2$ are eigenpairs (eigenvalue and eigenvector pairs) of matrix A given below,

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}, \qquad \begin{cases} \lambda_1 = 4 & \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\[2ex] \lambda_2 = -2 & \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \end{cases}$$

Solution: We must verify the definition of eigenpairs given above. We start with the first pair,

$$A\boldsymbol{v}_1 = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda_1 \boldsymbol{v}_1 \quad \Rightarrow \quad A\boldsymbol{v}_1 = \lambda_1 \boldsymbol{v}_1.$$

A similar calculation for the second pair implies,

$$A\boldsymbol{v}_2 = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix} = -2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \lambda_2 \boldsymbol{v}_2 \quad \Rightarrow \quad A\boldsymbol{v}_2 = \lambda_2 \boldsymbol{v}_2.$$

◁

Example 4.3.2 (Geometrical Meaning)**.** Find the eigenpairs of the matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Solution: The action of this matrix on vectors on a plane is a reflection along the line $x_1 = x_2$. Therefore, this line $x_1 = x_2$, is left invariant under the action of this matrix. This property suggests that an eigenvector is any vector on that line, for example

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \lambda_1 = 1.$$

So, we have found one eigenvalue-eigenvector pair: $\lambda_1 = 1$, with $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. We remark that any nonzero vector proportional to $\boldsymbol{v}_1$ is also an eigenvector. Another choice for the eigenvector is $\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$.

It is not so easy to find a second eigenvector which does not belong to the line determined by $\boldsymbol{v}_1$. One way to find such eigenvector is noticing that the line perpendicular to the line $x_1 = x_2$ is also left invariant by matrix A. Therefore, any nonzero vector on that line must be an eigenvector. For example the vector $\boldsymbol{v}_2$ below, since

$$\boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = (-1) \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \lambda_2 = -1.$$

So, we have found a second eigenvalue-eigenvector pair: $\lambda_2 = -1$, with $\boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. We remark again that any nonzero vector proportional to $\boldsymbol{v}_2$ is also an eigenvector. Another choice for the eigenvector is $\boldsymbol{v}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$. The eigenvectors and eigenvalues of this matrix are displayed in Fig. 8. ◁

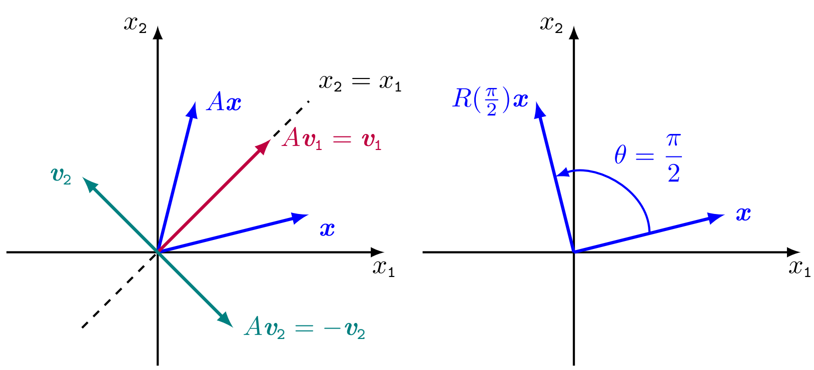


FIGURE 8. On the left we show the eigenpairs of the matrix in Example 4.3.2. On the right we show the matrix in Example 4.3.3 rotating a vector $\boldsymbol{x}$ counterclockwise by an angle $\theta = \pi/2$. When rotation is applied to every vector it shows that $R(\pi/2)$ does not have real eigenpairs.

There exist matrices that do not have eigenvalues and eigenvectors, as it is shown in the example below.

Example 4.3.3 (No Eigenvectors)**.** Fix a number $\theta \in (0, \pi)$ and define the matrix

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Show that this matrix $R(\theta)$ has no real eigenvalues.

Solution: One can compute the action of matrix $R(\theta)$ on several (real-valued) vectors and verify that the action of this matrix on the plane is a rotation counterclockwise by and angle θ, as shown in Fig. 8. In the particular case $\theta = \pi/2$ we have

$$R(\pi/2) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

and the action of this matrix on a given vector is shown in Figure 8. Since the eigenvectors of a matrix determine directions which are left invariant by the action of the matrix, and a rotation does not have such directions, we conclude that the matrix $R(\theta)$ above does not have real eigenvectors and so it does not have real eigenvalues either. Later on we show that this matrix $R(\theta)$, as a function on complex vectors, has complex-valued eigenpairs. ◁

4.3.2. Computing Eigenpairs. We now describe a method to find the eigenpairs of a matrix, if they exit. In other words, we are going to solve the eigenvalue-eigenvector problem: Given an $n \times n$ matrix A find, if possible, all its eigenvalues and eigenvectors, that is, all pairs λ and $\boldsymbol{v} \neq \boldsymbol{0}$ solutions of the equation

$$A\boldsymbol{v} = \lambda \boldsymbol{v}.$$

This problem is more complicated than finding the solution $\boldsymbol{x}$ to a linear system

$$A\boldsymbol{x} = \boldsymbol{b},$$

where A and $\boldsymbol{b}$ are known. In the eigenvalue-eigenvector problem above neither λ nor $\boldsymbol{v}$ are known. To solve the eigenvalue-eigenvector problem for a matrix A we proceed as follows:

(a) First, find the eigenvalues λ;
(b) Second, for each eigenvalue λ, find the corresponding eigenvectors $\boldsymbol{v}$.

The following result summarizes a way to solve the steps above.

Theorem 4.3.2 (Eigenvalues-Eigenvectors)**.**

(a) All the eigenvalues λ of an $n \times n$ matrix $A \in \mathbb{F}^{n,n}$ are the solutions of the scalar equation

$$\det(A - \lambda I) = 0. \tag{4.3.1}$$

(b) Given an eigenvalue λ of an $n \times n$ matrix A, the corresponding eigenvectors $\boldsymbol{v} \in \mathbb{F}^n$ are the nonzero solutions to the homogeneous linear system

$$(A - \lambda I)\boldsymbol{v} = \boldsymbol{0}. \tag{4.3.2}$$

Proof of Theorem 4.3.2: The number λ and the nonzero vector $\boldsymbol{v}$ are an eigenvalue-eigenvector pair of matrix A iff holds

$$A\boldsymbol{v} = \lambda \boldsymbol{v} \quad \Leftrightarrow \quad (A - \lambda I)\boldsymbol{v} = \boldsymbol{0},$$

where I is the $n \times n$ identity matrix. Since $\boldsymbol{v} \neq \boldsymbol{0}$, the last equation above says that the eigevalue λ is such that the columns of the matrix $(A - \lambda I)$ are linearly dependent. This means that λ is such that the matrix $(A - \lambda I)$ is not invertible. Since a matrix is not invertible iff its determinant vanishes, the eigenvalue λ must be solution of the equation

$$\det(A - \lambda I) = 0.$$

This is equation (4.3.1) and it determines the eigenvalues λ. Once this equation is solved, we substitute each solution λ back into the original eigenvalue-eigenvector equation

$$(A - \lambda I)\boldsymbol{v} = \mathbf{0}.$$

Since λ is known, this is a linear homogeneous system for the eigenvector components. It always has nonzero solutions, since λ is precisely the number that makes the coefficient matrix $(A - \lambda I)$ not invertible. This establishes the Theorem. □

Example 4.3.4 (Real Different)**.** Find the eigenvalues λ and eigenvectors $\boldsymbol{v}$ of the matrix

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}.$$

Solution: We first find the eigenvalues as the solutions of the Eq. (4.3.1). Compute

$$A - \lambda I = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} (1-\lambda) & 3 \\ 3 & (1-\lambda) \end{bmatrix}.$$

Then we compute its determinant,

$$0 = \det(A - \lambda I) = \begin{vmatrix} (1-\lambda) & 3 \\ 3 & (1-\lambda) \end{vmatrix} = (\lambda - 1)^2 - 9 \quad \Rightarrow \quad \begin{cases} \lambda_+ = 4, \\ \lambda_- = -2. \end{cases}$$

We have obtained two eigenvalues, so now we introduce $\lambda_+ = 4$ into Eq. (4.3.2), that is,

$$A - 4I = \begin{bmatrix} 1-4 & 3 \\ 3 & 1-4 \end{bmatrix} = \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix}.$$

Then we solve for $\boldsymbol{v}^+$ the equation

$$(A - 4I)\boldsymbol{v}^+ = \mathbf{0} \quad \Leftrightarrow \quad \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} v_1^+ \\ v_2^+ \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{aligned} -3v_1^+ + 3v_2^+ &= 0 \\ 3v_1^+ - 3v_2^+ &= 0. \end{aligned}$$

Notice that the second equation on the far right side is proportional to the first equation. Therefore, we only need to solve one equation,

$$-3v_1^+ + 3v_2^+ = 0 \quad \Rightarrow \quad v_1^+ = v_2^+$$

This means that the eigenvector equation determines only a relation between the components of the eigenvector, and not the whole eigenvector. In the equation above the component v_1^+ is fixed by v_2^+, and v_2^+ is free,

$$\boldsymbol{v}^+ = \begin{bmatrix} v_1^+ \\ v_2^+ \end{bmatrix} = \begin{bmatrix} v_2^+ \\ v_2^+ \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} v_2^+ \quad \Rightarrow \quad \boldsymbol{v}^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where we have chosen $v_2^+ = 1$. A similar calculation provides the eigenvector $\boldsymbol{v}^-$ associated with the eigenvalue $\lambda_- = -2$, that is, first compute the matrix

$$A + 2I = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}$$

then we solve for $\boldsymbol{v}^-$ the equation

$$(A + 2I)\boldsymbol{v}^- = \mathbf{0} \quad \Leftrightarrow \quad \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} v_1^- \\ v_2^- \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{aligned} 3v_1^+ + 3v_2^+ &= 0 \\ 3v_1^+ + 3v_2^+ &= 0. \end{aligned}$$

Notice that the second equation on the far right side is proportional to the first equation. Therefore, we only need to solve one equation,

$$3v_1^+ + 3v_2^+ = 0 \quad \Rightarrow \quad v_1^+ = -v_2^+$$

As above, the eigenvector equation determines only a relation between the components of the eigenvector, and not the whole eigenvector. In the equation above the component v_1^+ is fixed by v_2^+, and v_2^+ is free,

$$\boldsymbol{v}^- = \begin{bmatrix} v_1^- \\ v_2^- \end{bmatrix} = \begin{bmatrix} -v_2^- \\ v_2^- \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} v_2^- \quad \Rightarrow \quad \boldsymbol{v}^- = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

where we have chosen $v_2^- = 1$. We therefore conclude that the eigenvalues and eigenvectors of the matrix A above are given by

$$\lambda_+ = 4, \quad \boldsymbol{v}^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \lambda_- = -2, \quad \boldsymbol{v}^- = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

◁

Example 4.3.5 (Complex)**.** Find the eigenvalues λ and eigenvectors $\boldsymbol{v}$ of the matrix

$$A = \begin{bmatrix} -3 & 2 \\ -10 & 5 \end{bmatrix}.$$

Solution: The eigenvalues λ are the solutions of $\det(A - \lambda I) = 0$, that is,

$$0 = \det(A - \lambda I) = \begin{vmatrix} -3-\lambda & 2 \\ -10 & 5-\lambda \end{vmatrix} = (\lambda - 5)(\lambda + 3) + 20 = \lambda^2 - 2\lambda + 5.$$

The solutions of the equation above are

$$\lambda_\pm = \frac{1}{2}\big(2 \pm \sqrt{4-20}\big) = \frac{1}{2}(2 \pm 4i) \quad \Rightarrow \quad \lambda_\pm = 1 \pm 2i.$$

Notice that the eigenvalues are complex conjugate of each other, that is

$$\lambda_- = \overline{\lambda_+}$$

This happens for all real valued matrices with complex eigenvalues. A similar relation can be proven for their corresponding eigenvectors, $\boldsymbol{v}^- = \overline{\boldsymbol{v}^+}$.

Now we pick the eigenvalue λ_+ and we compute its corresponding eigenvector $\boldsymbol{v}^+$, as the solution of the system

$$(A - (1+2i)I)\boldsymbol{v}^+ = \mathbf{0} \quad \Leftrightarrow \quad \begin{bmatrix} -3-(1+2i) & 2 \\ -10 & 5-(1+2i) \end{bmatrix} \begin{bmatrix} v_1^+ \\ v_2^+ \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and from there we get

$$\begin{bmatrix} -4-2i & 2 \\ -10 & 4-2i \end{bmatrix} \begin{bmatrix} v_1^+ \\ v_2^+ \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} -(4+2i)\,v_1^+ + 2\,v_2^+ = 0 \\ -10\,v_1^+ + (4-2i)\,v_2^+ = 0. \end{cases}$$

The two equation on the far right are proportional to each other. indeed,

$$\big(-(4+2i)\,v_1^+ + 2\,v_2^+\big)(2-i) = -10\,v_1^+ + (4-2i)\,v_2^+,$$

so we have only one equation to solve, namely

$$-(4+2i)\,v_1^+ + 2\,v_2^+ = 0 \quad \Leftrightarrow \quad v_2^+ = (2+i)\,v_1^+.$$

All the eigenvectors $\boldsymbol{v}^+$ associated to the eigenvalue λ_+ are

$$\boldsymbol{v}^+ = \begin{bmatrix} v_1^+ \\ v_2^+ \end{bmatrix} = \begin{bmatrix} 1 \\ 2+i \end{bmatrix} v_1^+.$$

If we choose $v_1^+ = 1$, we get the eigenpair

$$\lambda_+ = 1 + 2i, \qquad \boldsymbol{v}^+ = \begin{bmatrix} 1 \\ 2+i \end{bmatrix}.$$

A similar calculation gives the eigenpair

$$\lambda_- = 1 - 2i, \qquad \boldsymbol{v}^- = \begin{bmatrix} 1 \\ 2-i \end{bmatrix}.$$

◁

Example 4.3.6 (Complex)**.** Fix a number $\theta \in (0, \pi)$ and define the matrix

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Find the eigenpairs of matrix $R(\theta)$.

Solution: We start computing the eigenvalues λ as the solution of $\det(R(\theta) - \lambda I) = 0$,

$$0 = \det(R(\theta) - \lambda I) = \begin{vmatrix} \cos(\theta) - \lambda & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) - \lambda \end{vmatrix} = (\cos(\theta) - \lambda)^2 + \sin^2(\theta),$$

$$\lambda^2 - 2\lambda\cos(\theta) + 1 = 0,$$

where we used that $\cos^2(\theta) + \sin^2(\theta) = 1$. The solutions of the equation above are

$$\lambda_\pm = 1 \pm \sqrt{\cos^2(\theta) - 1} \quad \Rightarrow \quad \lambda_\pm = 1 \pm i\sin(\theta),$$

where we used that $\theta \in (0, \pi)$, hence $\sin(\theta) > 0$. Now we pick the eigenvalue λ_+ and we compute its corresponding eigenvector $\boldsymbol{v}^+$, as the solution of the system

$$\big(R(\theta) - (1 + i\sin(\theta))I\big)\boldsymbol{v}^+ = \mathbf{0}$$

that is,

$$\begin{bmatrix} \cos(\theta) - (1 + i\sin(\theta)) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) - (1 + i\sin(\theta) \end{bmatrix} \begin{bmatrix} v_1^+ \\ v_2^+ \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

From here we get that

$$(-1 + \cos(\theta) - i\sin(\theta))\, v_1^+ = \sin(\theta)\, v_2^+.$$

Therefore, choosing $v_1^+ = \sin(\theta)$ and $v_2^+ = -1 + \cos(\theta) - i\sin(\theta)$ we get

$$\lambda_+ = 1 + i\sin(\theta), \qquad \boldsymbol{v}^+ = \begin{bmatrix} \sin(\theta) \\ -1 + \cos(\theta) - i\sin(\theta) \end{bmatrix},$$

and therefore,

$$\lambda_- = 1 - i\sin(\theta), \qquad \boldsymbol{v}^- = \begin{bmatrix} \sin(\theta) \\ -1 + \cos(\theta) + i\sin(\theta) \end{bmatrix}.$$

In the particular case $\theta = \pi/2$ the matrix is $R(\frac{\pi}{2}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ with eigenpairs

$$\lambda_\pm = 1 \pm i, \qquad \boldsymbol{v}^\pm = \begin{bmatrix} 1 \\ -1 \mp i \end{bmatrix}.$$

◁

4.3.3. Eigenvalue Multiplicity. It is useful to introduce few more concepts, that are common in the literature. We start with a function we have used already.

Definition 4.3.3. *The **characteristic polynomial** of an $n \times n$ matrix A is the function*

$$p(\lambda) = \det(A - \lambda I).$$

Example 4.3.7 (Characteristic Polynomial)**.** Find the characteristic polynomial of

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}.$$

Solution: We need to compute the determinant

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} (1-\lambda) & 3 \\ 3 & (1-\lambda) \end{vmatrix} = (1-\lambda)^2 - 9 = \lambda^2 - 2\lambda + 1 - 9.$$

We conclude that the characteristic polynomial is

$$p(\lambda) = \lambda^2 - 2\lambda - 8t.$$

$\lhd$

Since the matrix A in this example is 2×2, its characteristic polynomial has degree two. One can show that the characteristic polynomial of an $n \times n$ matrix has degree n. The eigenvalues of the matrix are the roots of the characteristic polynomial. Different matrices may have different types of roots, as it can be seen in the following result. We left the proof as an exercise.

Theorem 4.3.4 (Characteristic Polynomial)**.** *Given an $n \times n$ matrix A with real eigenvalues λ_i, where $i = 1, \cdots, k \leqslant n$, it is always possible to express the characteristic polynomial of A as*

$$p(\lambda) = (-1)^n \, (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_k)^{r_k}.$$

Remark: The number r_i is called the *algebraic multiplicity* of the eigenvalue λ_i. Furthermore, the *geometric multiplicity* of an eigenvalue λ_i, denoted as s_i, is the maximum number of eigenvectors corresponding to λ_i that form a linearly independent set.

Example 4.3.8 (Multiplicities)**.** Find the algebraic and geometric multiplicities of the the eigenvalues of the matrix

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$$

Solution: In order to find the algebraic multiplicity of the eigenvalues we need first to find the eigenvalues. We now that the characteristic polynomial of this matrix is given by

$$p(\lambda) = \begin{vmatrix} (1-\lambda) & 3 \\ 3 & (1-\lambda) \end{vmatrix} = (\lambda - 1)^2 - 9.$$

The roots of this polynomial are $\lambda_1 = 4$ and $\lambda_2 = -2$, so we know that $p(\lambda)$ can be rewritten in the following way,

$$p(\lambda) = (\lambda - 4)(\lambda + 2).$$

We conclude that the algebraic multiplicity of the eigenvalues are both one, that is,

$$\lambda_1 = 4, \quad r_1 = 1, \quad \text{and} \qquad \lambda_2 = -2, \quad r_2 = 1.$$

In order to find the geometric multiplicities of matrix eigenvalues we need first to find the matrix eigenvectors. This part of the work was already done in the Example **??** above and the result is

$$\lambda_1 = 4, \quad \boldsymbol{v}^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \lambda_2 = -2, \quad \boldsymbol{v}^{(2)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

From this expression we conclude that the geometric multiplicities for each eigenvalue are just one, that is,

$$\lambda_1 = 4, \quad s_1 = 1, \quad \text{and} \quad \lambda_2 = -2, \quad s_2 = 1.$$

◁

The following example shows that two matrices can have the same eigenvalues, and so the same algebraic multiplicities, but different eigenvectors with different geometric multiplicities.

Example 4.3.9 (Multiplicities)**.** Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution: We start finding the eigenvalues, the roots of the characteristic polynomial

$$p(\lambda) = \begin{vmatrix} (3-\lambda) & 0 & 1 \\ 0 & (3-\lambda) & 2 \\ 0 & 0 & (1-\lambda) \end{vmatrix} = -(\lambda-1)(\lambda-3)^2 \quad \Rightarrow \quad \begin{cases} \lambda_1 = 1, & r_1 = 1, \\ \lambda_2 = 3, & r_2 = 2. \end{cases}$$

We now compute the eigenvector associated with the eigenvalue $\lambda_1 = 1$, which is the solution of the linear system

$$(A - I)\boldsymbol{v}^{(1)} = \mathbf{0} \quad \Leftrightarrow \quad \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1^{(1)} \\ v_2^{(1)} \\ v_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{aligned} 2v_1^{(1)} + v_3^{(1)} &= 0 \\ 2v_2^{(1)} + 2v_3^{(1)} &= 0 \\ 0 &= 0. \end{aligned}$$

This system determines $v_1^{(1)}$ and $v_2^{(1)}$ in terms of $v_3^{(1)}$ and $v_3^{(1)}$ is free. It is simple to se that the solution of this system is

$$v_1^{(1)} = -\frac{v_3^{(1)}}{2}, \qquad v_2^{(1)} = -v_3^{(1)}, \qquad v_3^{(1)} \text{ free}.$$

Therefore, choosing $v_3^{(1)} = 2$ we obtain that

$$\boldsymbol{v}^{(1)} = \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix}, \qquad \lambda_1 = 1, \qquad r_1 = 1, \qquad s_1 = 1.$$

In a similar way we now compute the eigenvectors for the eigenvalue $\lambda_2 = 3$, which are all solutions of the linear system

$$(A - 3I)\boldsymbol{v}^{(2)} = \mathbf{0} \quad \Leftrightarrow \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} v_1^{(2)} \\ v_2^{(2)} \\ v_3^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{aligned} v_3^{(2)} &= 0 \\ 2v_3^{(2)} &= 0 \\ -2v_3^{(2)} &= 0. \end{aligned}$$

The obvious solution is $v_3^{(2)} = 0$. There is no condition on $v_1^{(2)}$ and $v_2^{(2)}$, which are free. Therefore, the eigenvectors are given by

$$\boldsymbol{v}^{(2)} = \begin{bmatrix} v_1^{(2)} \\ v_2^{(2)} \\ 0 \end{bmatrix} = \begin{bmatrix} v_1^{(2)} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ v_2^{(2)} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} v_1^{(2)} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} v_2^{(2)}.$$

Therefore, we obtain two linearly independent solutions, the first one $\boldsymbol{w}_1^{(2)}$ with the choice $v_1^{(2)} = 1$, $v_2^{(2)} = 0$, and the second one $\boldsymbol{w}_2^{(2)}$ with the choice $v_1^{(2)} = 0$, $v_2^{(2)} = 1$, that is

$$\boldsymbol{w}_1^{(2)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \boldsymbol{w}_2^{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \lambda_2 = 3, \qquad r_2 = 2, \qquad s_2 = 2.$$

Summarizing, the matrix in this example has three linearly independent eigenvectors. ◁

Example 4.3.10 (Multiplicities)**.** Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution: Notice that this matrix has only the coefficient a_{12} different from the previous example. Again, we start finding the eigenvalues, which are the roots of the characteristic polynomial

$$p(\lambda) = \begin{vmatrix} (3-\lambda) & 1 & 1 \\ 0 & (3-\lambda) & 2 \\ 0 & 0 & (1-\lambda) \end{vmatrix} = -(\lambda - 1)(\lambda - 3)^2 \quad \Rightarrow \quad \begin{cases} \lambda_1 = 1, & r_1 = 1, \\ \lambda_2 = 3, & r_2 = 2. \end{cases}$$

So this matrix has the same eigenvalues and algebraic multiplicities as the matrix in the previous example. We now compute the eigenvector associated with the eigenvalue $\lambda_1 = 1$, which is the solution of the linear system

$$(A - I)\boldsymbol{v}^{(1)} = \mathbf{0} \quad \Leftrightarrow \quad \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1^{(1)} \\ v_2^{(1)} \\ v_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{aligned} 2v_1^{(1)} + v_2^{(1)} + v_3^{(1)} &= 0 \\ 2v_2^{(1)} + 2v_3^{(1)} &= 0 \\ 0 &= 0. \end{aligned}$$

The second equation on the far right says $v_2^{(1)} = -v_3^{(1)}$ and this equation on the first equation on the far right implies

$$2v_1^{(1)} - v_3^{(1)} + v_3^{(1)} = 0 \quad \Rightarrow \quad v_1^{(1)} = 0.$$

Therefore, the solution of the system above is

$$v_1^{(1)} = 0, \qquad v_2^{(1)} = -v_3^{(1)}, \qquad v_3^{(1)} \text{ free}.$$

The eigenvector is given by

$$\boldsymbol{v}^{(1)} = \begin{bmatrix} v_1^{(1)} \\ v_2^{(1)} \\ v_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ -v_3^{(1)} \\ v_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} v_3^{(1)}.$$

Therefore, choosing $v_3^{(1)} = 1$ we obtain that

$$\boldsymbol{v}^{(1)} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \qquad \lambda_1 = 1, \qquad r_1 = 1, \qquad s_1 = 1.$$

In a similar way we now compute the eigenvectors for the eigenvalue $\lambda_2 = 3$. However, in this case we obtain only one solution, as this calculation shows,

$$(A - 3I)\boldsymbol{v}^{(2)} = \mathbf{0} \quad \Leftrightarrow \quad \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} v_1^{(2)} \\ v_2^{(2)} \\ v_3^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{aligned} v_2^{(2)} + v_3^{(2)} &= 0 \\ 2v_3^{(2)} &= 0 \\ 0 &= 0. \end{aligned}$$

From the system above we see that $v_3^{(2)} = 0$, which implies $v_2^{(2)} = 0$, while $v_1^{(2)}$ is free. The eigenvector is given by

$$\boldsymbol{v}^{(2)} = \begin{bmatrix} v_1^{(2)} \\ v_2^{(2)} \\ v_3^{(2)} \end{bmatrix} = \begin{bmatrix} -v_1^{(2)} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} v_1^{(2)}.$$

If we choose $v_1^{(2)} = 1$ we obtain that all eigenvectors are proportional to a single vector and we conclude that

$$\boldsymbol{v}^{(2)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \lambda_2 = 3, \qquad r_2 = 2, \qquad s_2 = 1.$$

Summarizing, the matrix in this example has only two linearly independent eigenvectors, and in the case of the eigenvalue $\lambda_2 = 3$ we have the strict inequality

$$1 = s_2 < r_2 = 2.$$

$\lhd$

4.3.4. Exercises.

4.3.1.- .

4.3.2.- .

4.4. Diagonalizable Matrices

We focus on two types of matrices, diagonal matrices and diagonalizable matrices. It is very simple to work with diagonal matrices, but they do not show in many applications. Diagonalizable matrices are simple enough so that certain computations to be performed exactly and are complicated enough so they can be used to describe several physical problems. We show that the eigenvectors of a matrix determine whether the matrix is diagonalizable or not.

4.4.1. Diagonal Matrices.

We first introduce the notion of a diagonal matrix. Later on we define a diagonalizable matrix as a matrix that can be transformed into a diagonal matrix by a simple transformation.

Definition 4.4.1. *An $n \times n$ matrix D is called **diagonal** iff* $D = \begin{bmatrix} d_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_{nn} \end{bmatrix}$.

The definition above says that a matrix is diagonal if and only if every non-diagonal coefficient vanishes. For example, the identity matrix is diagonal, but not all diagonal matrices are the identity matrix. We use either of the following two notations for a diagonal matrix D,

$$D = \operatorname{diag}\big[d_{11}, \cdots, d_{nn}\big] = \begin{bmatrix} d_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_{nn} \end{bmatrix}.$$

The notation in the middle equation says that the matrix D is diagonal and shows only the diagonal coefficients, which is all we need to determine a diagonal matrix, because we know for sure that any coefficient off the diagonal vanishes. The next result says that the eigenvalues of a diagonal matrix are the matrix diagonal elements, and it gives the corresponding eigenvectors.

Theorem 4.4.2 (Eigenvectors)**.** *If $D = \operatorname{diag}[d_{11}, \cdots, d_{nn}]$, then eigenpairs of D are*

$$\lambda_1 = d_{11}, \quad \boldsymbol{v}^{(1)} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \lambda_n = d_{nn}, \quad \boldsymbol{v}^{(n)} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Proof of Theorem 4.4.2: We just verify the formula in the Theorem. For example,

$$D\boldsymbol{v}^{(1)} = \begin{bmatrix} d_{11} & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = d_{11} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 \boldsymbol{v}^{(1)}.$$

A similar calculation with the rest of the eigenvectors establishes the Theorem. □

Diagonal matrices are simple to manipulate since they share many properties with numbers. For example *the product of two diagonal matrices is commutative.* Also, it is simple to compute power functions of a diagonal matrix.

Theorem 4.4.3 (Powers). *If* $D = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$, *then for any nonnegative integer* n *holds*

$$D^n = \begin{bmatrix} a^n & 0 \\ 0 & d^n \end{bmatrix}.$$

Proof of Theorem 4.4.3: We first computing D^2 as follows,

$$D^2 = DD = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} a^2 & 0 \\ 0 & d^2 \end{bmatrix}.$$

We now compute D^3,

$$D^3 = D^2 D = \begin{bmatrix} a^2 & 0 \\ 0 & d^2 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} a^3 & 0 \\ 0 & d^3 \end{bmatrix}.$$

The calculation above suggests we use induction. Let's assume the formula for the $n-1$ power is

$$D^{n-1} = \begin{bmatrix} a^{n-1} & 0 \\ 0 & d^{n-1} \end{bmatrix}.$$

Then, the next power is

$$D^n = D^{n-1}\, D = \begin{bmatrix} a^{n-1} & 0 \\ 0 & d^{n-1} \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} a^n & 0 \\ 0 & d^n \end{bmatrix}.$$

Therefore, we conclude that

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}^n = \begin{bmatrix} a^n & 0 \\ 0 & d^n \end{bmatrix}.$$

This establishes the Theorem. $\square$

Example 4.4.1. For every positive integer n find D^n, where $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$.

Solution: We start computing D^2 as follows,

$$D^2 = DD = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}.$$

We now compute D^3,

$$D^3 = D^2 D = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2^3 & 0 \\ 0 & 3^3 \end{bmatrix}.$$

Using induction, it is simple to see that

$$D^n = \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix}.$$

$\triangleleft$

A simple generalization of the calculation in the example above, which is left as an exercise, is the the proof of the following statement.

Theorem 4.4.4 (Powers). *If* $D = \operatorname{diag}\big[d_{11}, \cdots, d_{nn}\big]$, *then, for any integer* $k \geqslant 0$ *holds*

$$D^k = \operatorname{diag}\big[(d_{11})^k, \cdots, (d_{nn})^k\big].$$

4.4.2. Diagonalizable Matrices. A few properties of diagonal matrices are shared by diagonalizable matrices, which are matrices that can be converted into a diagonal matrix by a simple transformation.

Definition 4.4.5. *An $n \times n$ matrix A is called* ***diagonalizable*** *iff there exists an invertible matrix P and a diagonal matrix D such that*

$$A = PDP^{-1}.$$

Remarks:

(a) An equivalent characterization of a diagonalizable matrix is

$$A = PDP^{-1} \quad \Leftrightarrow \quad P^{-1}AP = D,$$

that is, a diagonalizable matrix can be transformed into a diagonal matrix simply by two matrix multiplications.

(b) The decomposition of a diagonalizable matrix,

$$A = PDP^{-1}$$

is not unique. Later on this section we see that there are infinitely many invertible matrices $\tilde{P}$ and diagonal matrices $\tilde{D}$ so that

$$A = \tilde{P}\tilde{D}\tilde{P}^{-1}.$$

(c) Systems of linear differential equations are simple to solve in the case that the coefficient matrix is diagonalizable. The multiplication by P and P^{-1} decouples the differential equations. We then solve the decoupled equations, one by one, and transform the solutions back to the original unknowns.

(d) Not every square matrix is diagonalizable. For example, matrix A below is diagonalizable while B is not,

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}, \qquad B = \frac{1}{2}\begin{bmatrix} 3 & 1 \\ -1 & 5 \end{bmatrix}.$$

Later on we study how to find out whether a matrix is diagonalizable or not. It turns out, the eigenvectors of a matrix determine whether the matrix is or is not diagonalizable.

Example 4.4.2 (Diagonalizable)**.** Show that matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ is diagonalizable, with

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}.$$

Solution: That matrix P is invertible can be verified by computing its determinant,

$$\det(P) = 1 - (-1) = 2.$$

Since the determinant is nonzero, P is invertible. Using linear algebra methods one can find out that the inverse matrix is

$$P^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Now we only need to verify that PDP^{-1} is indeed A, which can be done by a straightforward calculation,

$$\begin{aligned} PDP^{-1} &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 2 \\ 4 & -2 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \quad \Rightarrow \quad PDP^{-1} = A. \end{aligned}$$

Equivalently, we can verify that $P^{-1}AP$ is diagonal, since

$$\begin{aligned} P^{-1}AP &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 4 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 2 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix} \quad \Rightarrow \quad P^{-1}AP = D. \end{aligned}$$

◁

4.4.3. Eigenvectors and Diagonalizable Matrices. In the example above we showed that $A = PDP^{-1}$, where

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

We also know that the matrix A has eigenvalues and eigenvectors given by

$$\lambda_1 = 4, \quad \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad \lambda_2 = -2, \quad \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Notice that matrix D contains the eigenvalues of A and matrix P contains the eigenvectors of A, in the same order, that is,

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \qquad P = \begin{bmatrix} \boldsymbol{v}_1, \boldsymbol{v}_2 \end{bmatrix}.$$

We see in the example above that there is a relation between the eigenvalues and eigenvectors of a matrix (λ_i, $\boldsymbol{v}_i$ for $i = 1, 2$), and the diagonal and invertible matrices (D, P), that factorize a diagonalizable matrix. It turns out that this relation is not a coincidence, instead, this is a particular case of a general result.

Theorem 4.4.6 (Eigenvectors and Diagonalizability)**.** *A 2×2 matrix, A, has two eigenvectors, $\boldsymbol{v}_1, \boldsymbol{v}_2$ not proportional to each other iff A is diagonalizable, $A = PDP^{-1}$, with*

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad P = [\boldsymbol{v}_1, \boldsymbol{v}_2],$$

where λ_i is the eigenvalue of the eigenvector $\boldsymbol{v}_i$, for $i = 1, 2$.

Before we prove Theorem 4.4.6 it is convenient we show a result that we will need in the proof of Theorem 4.4.6.

Theorem 4.4.7 (Proportionallity)**.** *Given vectors* $\boldsymbol{v}_1$*,* $\boldsymbol{v}_2$ *and a constant* c*, then*

$$\boldsymbol{v}_1 = c\,\boldsymbol{v}_2 \quad \Leftrightarrow \quad \det\big([\boldsymbol{v}_1, \boldsymbol{v}_2]\big) = 0.$$

Proof of Theorem 4.4.7: Denoting $P = \big[\boldsymbol{v}_1, \boldsymbol{v}_2\big]$, then we need to show that

$$\boldsymbol{v}_1 = c\,\boldsymbol{v}_2 \quad \Leftrightarrow \quad \det(P) = 0.$$

($\Rightarrow$) Let's write the vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ in components,

$$\boldsymbol{v}_1 = \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix}, \qquad \boldsymbol{v}_2 = \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} \quad \Rightarrow \quad P = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}.$$

Since these vectors are proportional to each other,

$$v_{11} = c\,v_{12}, \quad v_{21} = c\,v_{22} \quad \Rightarrow \quad \det(P) = \begin{vmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{vmatrix} = \begin{vmatrix} c\,v_{12} & v_{12} \\ c\,v_{22} & v_{22} \end{vmatrix} = c\,v_{12}\,v_{22} - c\,v_{22}\,v_{12} = 0.$$

We have shown that $\det(P) = 0$.

($\Leftarrow$) If $\boldsymbol{v}_1 = \boldsymbol{0}$, then $\boldsymbol{v}_1 = c\boldsymbol{v}_2$ with $c = 0$. So, we consider the case $\boldsymbol{v}_1 \neq \boldsymbol{0}$. Let's write the vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ in components,

$$\boldsymbol{v}_1 = \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix}, \qquad \boldsymbol{v}_2 = \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix}.$$

Since $\boldsymbol{v}_1 \neq \boldsymbol{0}$, then either $v_{11} \neq 0$ or $v_{21} \neq 0$. We now assume that $v_{11} \neq 0$. (The proof for $v_{21} \neq 0$ is similar and left as an exercise.) Then, matrix P is given by

$$P = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}.$$

We know that $\det(P) = 0$, and recall $v_{11} \neq 0$, then

$$0 = \begin{vmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{vmatrix} = v_{11}v_{22} - v_{12}v_{21} \quad \Rightarrow \quad v_{22} = \frac{v_{12}v_{21}}{v_{11}}.$$

Therefore we get that

$$\boldsymbol{v}_2 = \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} = \begin{bmatrix} v_{12} \\ \dfrac{v_{12}v_{21}}{v_{11}} \end{bmatrix} = \frac{v_{12}}{v_{11}} \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} = c\boldsymbol{v}_1, \quad \text{with} \quad c = \frac{v_{12}}{v_{11}}.$$

Therefore, $\det(P) = 0$ implies $\boldsymbol{v}_1 = c\,\boldsymbol{v}_2$. This establishes the Theorem. □

Recall two things: first, the contrapositive of a statement $B \Rightarrow C$ is No $C \Rightarrow$ No B; second, the contrapositive and the original statement are equivalent. If we compute the contrapositive on each implication in Theorem 4.4.7 we obtain the following equivalent statement.

Theorem 4.4.8 (Nonproportionallity)**.** *Given vectors* $\boldsymbol{v}_1$*,* $\boldsymbol{v}_2$*, then for every* $c \in \mathbb{R}$

$$\boldsymbol{v}_1 \neq c\,\boldsymbol{v}_2 \quad \Leftrightarrow \quad P = \big[\boldsymbol{v}_1, \boldsymbol{v}_2\big] \quad \text{is invertible.}$$

Now we are ready to prove Theorem 4.4.6.

Proof of Theorem 4.4.6:
($\Rightarrow$) Since λ_i, $\boldsymbol{v}_i$, for $i = 1, 2$ are eigenvalues and eigenvectors of matrix A, then

$$A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1, \qquad A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2.$$

Define $P = \big[\boldsymbol{v}_1, \boldsymbol{v}_2\big]$ and $D = \operatorname{diag}\big[\lambda_1, \lambda_2\big]$. Now we show that $AP = PD$. We start computing the left-hand side, AP,

$$AP = A\,\big[\boldsymbol{v}_1, \boldsymbol{v}_2\big] = \big[A\boldsymbol{v}_1, A\boldsymbol{v}_2\big] = \big[\lambda_1\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2\big].$$

Now we compute PD,

$$PD = \big[\boldsymbol{v}_1, \boldsymbol{v}_2\big]\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} = \big[\lambda_1\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2\big].$$

We conclude that $AP = PD$. Finally, we know that $\boldsymbol{v}_1 \neq c\,\boldsymbol{v}_2$ for any constant c, then Theorem 4.4.8 shows that matrix P is invertible. Multiply $AP = PD$ by P^{-1} from the right,

$$APP^{-1} = PDP^{-1} \quad \Rightarrow \quad A = PDP^{-1}.$$

This proves the ($\Rightarrow$) part of the Theorem.

($\Leftarrow$) We know that $A = PDP^{-1}$, for some matrices P and D, where P is invertible and D is diagonal. Let's write them as

$$D = \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}, \qquad P = [\boldsymbol{v}_1, \boldsymbol{v}_2],$$

We need to show that

$$A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1, \qquad A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2,$$

and that $\boldsymbol{v}_1$ is not proportional to $\boldsymbol{v}_2$, that is, for every constant c holds $\boldsymbol{v}_1 \neq c\boldsymbol{v}_2$. The first part is simple, since $A = PDP^{-1}$ implies that

$$AP = PD \quad \Rightarrow \quad A\,\big[\boldsymbol{v}_1, \boldsymbol{v}_2\big] = \big[\boldsymbol{v}_1, \boldsymbol{v}_2\big]\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} \quad \Rightarrow \quad \big[A\boldsymbol{v}_1, A\boldsymbol{v}_2\big] = \big[\lambda_1\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2\big].$$

This last equation implies

$$A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1, \qquad A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2,$$

so λ_i, $\boldsymbol{v}_i$, for $i = 1, 2$ are eigenvalues and eigenvectors of matrix A. Since matrix P is invertible, then Theorem 4.4.8 shows that the eigenvectors are not proportional to each other. This establishes the ($\Leftarrow$) part of the Theorem and then, the Theorem. □

Example 4.4.3 (Diagonalizable)**.** Show that matrix $A = \begin{bmatrix}1 & 3\\ 3 & 1\end{bmatrix}$ is diagonalizable.

Solution: We know that the eigenvalues and eigenvectors of matrix A are given by

$$\lambda_1 = 4, \quad \boldsymbol{v}_1 = \begin{bmatrix}1\\ 1\end{bmatrix} \quad \text{and} \quad \lambda_2 = -2, \quad \boldsymbol{v}_2 = \begin{bmatrix}-1\\ 1\end{bmatrix}.$$

Recall that matrix P must contain nonproportional eigenvectors of A in any order, while matrix D must contain in the diagonal the corresponding eigenvalues, in the same order as the eigenvectors in P. Therefore, introduce P and D as follows,

$$P = \big[\boldsymbol{v}_1, \boldsymbol{v}_2\big], \qquad D = \operatorname{diag}\big[\lambda_1, \lambda_2\big],$$

which means

$$P = \begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}, \qquad D = \begin{bmatrix}4 & 0\\ 0 & -2\end{bmatrix}. \quad \text{Therefore} \quad P^{-1} = \frac{1}{2}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}.$$

Now we can compute P^{-1},

$$P^{-1} = \frac{1}{2}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}.$$

We must show that $A = PDP^{-1}$. This is indeed the case, since

$$\begin{aligned} PDP^{-1} &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 2 \\ 4 & -2 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \\ &= A. \end{aligned}$$

Therefore, we have obtained that $PDP^{-1} = A$, that is, A is diagonalizable.

Remark: Matrices P and D are not unique. An equivalent choice is

$$\tilde{P} = \left[\boldsymbol{v}_2, \boldsymbol{v}_1\right], \qquad \tilde{D} = \operatorname{diag}\left[\lambda_2, \lambda_1\right],$$

where we switched the order of the columns in P and also in D, which means,

$$\tilde{P} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \quad \tilde{D} = \begin{bmatrix} -2 & 0 \\ 0 & 4 \end{bmatrix}.$$

Another equivalent choice is to use nonzero multiples of the eigenvectors,

$$\hat{P} = \left[a\boldsymbol{v}_2, b\boldsymbol{v}_1\right], \qquad \hat{D} = \operatorname{diag}\left[\lambda_2, \lambda_1\right],$$

where a and b are nonzero but otherwise arbitrary scalars. For example, for $a = -2$ and $b = 3$ we get

$$\hat{P} = \left[-2\boldsymbol{v}_2, 3\boldsymbol{v}_1\right], \qquad \hat{D} = \operatorname{diag}\left[\lambda_2, \lambda_1\right],$$

which means

$$\hat{P} = \begin{bmatrix} 2 & 3 \\ -2 & 3 \end{bmatrix} \quad \hat{D} = \begin{bmatrix} -2 & 0 \\ 0 & 4 \end{bmatrix}.$$

As an exercise show that

$$A = \tilde{P}\tilde{D}\tilde{P}^{-1}, \qquad A = \hat{P}\hat{D}\hat{P}^{-1}.$$

The most general choices for matrices P and D are either

$$P = \left[a\boldsymbol{v}_1, b\boldsymbol{v}_2\right], \qquad D = \operatorname{diag}\left[\lambda_1, \lambda_2\right],$$

or

$$P = \left[c\boldsymbol{v}_2, d\boldsymbol{v}_1\right], \qquad D = \operatorname{diag}\left[\lambda_2, \lambda_1\right],$$

where a, b, c, d are nonzero but otherwise arbitrary scalars. As an extra exercise show

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

for either of the two general choices made above for matrices P and D, where $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ are the eigenvectors found in this example and λ_1, λ_2 are the eigenvalues found in this example. ◁

With Theorem 4.4.9 we can show that a matrix *is not* diagonalizable.

Example 4.4.4 (Not Diagonalizable)**.** Show that matrix below is not diagonalizable,

$$A = \frac{1}{2} \begin{bmatrix} 3 & 1 \\ -1 & 5 \end{bmatrix}.$$

Solution: We first compute the matrix eigenvalues. The characteristic polynomial is

$$p(\lambda) = \begin{vmatrix} \left(\frac{3}{2} - \lambda\right) & \frac{1}{2} \\ -\frac{1}{2} & \left(\frac{5}{2} - \lambda\right) \end{vmatrix} = \left(\frac{3}{2} - \lambda\right)\left(\frac{5}{2} - \lambda\right) + \frac{1}{4} = \lambda^2 - 4\lambda + 4.$$

The roots of the characteristic polynomial are computed in the usual way,

$$\lambda = \frac{1}{2}\left[4 \pm \sqrt{16 - 16}\right] \quad \Rightarrow \quad \lambda = 2, \qquad r = 2.$$

We have obtained a single eigenvalue with algebraic multiplicity $r = 2$. The associated eigenvectors are computed as the solutions to the equation $(A - 2I)\boldsymbol{v} = \mathbf{0}$. Then,

$$(A - 2I)\boldsymbol{v} = \mathbf{0} \quad \Rightarrow \quad \begin{bmatrix} \left(\frac{3}{2} - 2\right) & \frac{1}{2} \\ -\frac{1}{2} & \left(\frac{5}{2} - 2\right) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

From here we get

$$\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad -v_1 + v_2 = 0.$$

Therefore, $v_1 = v_2$ and the eigenvectors have the form

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} v_1.$$

We choose $v_1 = 1$, then we get

$$\boldsymbol{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad s = 1.$$

We conclude that the largest linearly independent set of eigenvectors for the 2×2 matrix A contains only one vector, instead of two. Therefore, matrix A is not diagonalizable. $\lhd$

The result in Theorem 4.4.6 above can be generalized to $n \times n$ matrices.

Theorem 4.4.9 (Eigenvectors and Diagonalizability)**.** *An $n \times n$ matrix A has n linearly independent eigenvectors, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ iff matrix A is diagonalizable, $A = PDP^{-1}$, and*

$$D = \operatorname{diag}\left[\lambda_1, \cdots, \lambda_n\right], \quad P = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n],$$

where λ_i is the eigenvalue of the eigenvector $\boldsymbol{v}_i$, for $i = 1, \cdots, n$.

In Theorem 4.4.6 we had the condition that the vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ be not proportional to each other, $\boldsymbol{v}_1 \neq c\,\boldsymbol{v}_2$ for any c. In Theorem 4.4.9 the analogous condition is that the vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ be *linearly independent*. Recall we defined linearly independent vectors in Section 4.1 saying that no vector among the $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ is a linear combination of the other $n-1$ vectors. We also said that the vectors are *linearly dependent* when they are not linearly independent.

For example, in the case of three vectors in $\mathbb{R}^3$, say $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$, these vectors are linearly independent if $\boldsymbol{v}_1$ is not a linear combination of $\boldsymbol{v}_2$, $\boldsymbol{v}_3$, and $\boldsymbol{v}_2$ is not a linear combination of $\boldsymbol{v}_1$, $\boldsymbol{v}_3$, and $\boldsymbol{v}_3$ is not a linear combination of $\boldsymbol{v}_1$, $\boldsymbol{v}_2$. Linearly independent vectors in $\mathbb{R}^3$ span a volume in $\mathbb{R}^3$. The three vectors are linearly dependent when one of the vectors is a linear combination of the other two. The vectors are linearly dependent when they span a plane or a line in $\mathbb{R}^3$ but not a volume.

The difficult part in the proof of Theorem 4.4.9 is the relation between the linear independence of the vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ and the invertibility of matrix P. It turns out that Theorem 4.4.8 can be generalized from two vectors to not proportional to each other to n vectors linearly independent.

Theorem 4.4.10 (Linearly Independent)**.** *The vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ are linearly independent iff the matrix $P = \big[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\big]$ is invertible.*

The proof is a bit too long for our textbook, but the interested reader can find it in Tom Apostol's book, Calculus, Volume II, Section 3.9, [2]. Once Theorem 4.4.10 is proven, then it is not too hard to prove Theorem 4.4.9, since we only need to generalize the idea in the proof of Theorem 4.4.6.

Proof of Theorem 4.4.9:
$(\Rightarrow)$ Let $\boldsymbol{v}^{(i)}$,, for $i = 1, \cdots, n$, be eigenvector of matrix A with eigenvalue λ_i, that is,

$$A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i.$$

Now use the eigenvectors to construct matrix $P = \big[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\big]$. Since the vectors in the columns of P are linearly independent, then Theorem 4.4.10 implies that matrix P is invertible. We now show $AP = PD$. We start computing the product

$$AP = A\big[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\big] = \big[A\boldsymbol{v}_1, \cdots, A\boldsymbol{v}_n\big], = \big[\lambda_1 \boldsymbol{v}_1 \cdots, \lambda_n \boldsymbol{v}_n\big].$$

Now we compute the other product,

$$PD = \big[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\big] \operatorname{diag}\big[\lambda_1, \cdots, \lambda_n\big] = \big[\lambda_1 \boldsymbol{v}_1, \cdots, \lambda_n \boldsymbol{v}_n\big].$$

Therefore, $AP = PD$. We already showed that matrix P is invertible, then multiply the equation $AP = PD$ by P^{-1} from the right,

$$APP^{-1} = PDP^{-1} \quad \Rightarrow \quad A = PDP^{-1}.$$

This means that A is diagonalizable. This establishes the $(\Rightarrow)$ part of the Theorem.

$(\Leftarrow)$ Since matrix A is diagonalizable, there exist an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$. Let's denote

$$D = \operatorname{diag}\big[\lambda_1, \cdots, \lambda_n\big], \quad P = \big[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\big],$$

where λ_i denote the diagonal elements of D and $\boldsymbol{v}_i$ denote the column vectors of P, where $i = 1, \cdots, n$. Since P is invertible, Theorem 4.4.10 implies that the vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ are linearly independent. Now, multiply the equation $A = PDP^{-1}$ by P on the right, then

$$AP = PD.$$

The left side of this equation is

$$AP = A\big[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\big] = \big[A\boldsymbol{v}_1, \cdots, A\boldsymbol{v}_n\big],$$

while the right side is

$$PD = \big[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\big] \operatorname{diag}\big[\lambda_1, \cdots, \lambda_n\big] = \big[\lambda_1 \boldsymbol{v}_1, \cdots, \lambda_n \boldsymbol{v}_n\big].$$

The equation $AP = DP$ implies

$$A\boldsymbol{v}_1 = \lambda_1 \boldsymbol{v}_1 \quad \cdots \quad A\boldsymbol{v}_n = \lambda_n \boldsymbol{v}_n.$$

This establishes the $(\Leftarrow)$ part of the Theorem and with that, the Theorem. $\square$

Example 4.4.5 (Diagonalizable)**.** Show that matrix A below is diagonalizable,

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution: We have seen in Example 4.3.9 that this matrix has an eigenvalues and eigenvectors

$$\lambda_1 = 1, \quad \boldsymbol{v}^{(1)} = \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix} \quad \text{and} \quad \lambda_2 = 3, \quad \boldsymbol{w}_1^{(2)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{w}_2^{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Since matrix A is 3×3 and has three linearly independent eigenvectors, then Theorem 4.4.9 says that this matrix is diagonalizable,

$$A = PDP^{-1}$$

and a choice for matrices P and D are

$$P = \begin{bmatrix} -1 & 1 & 0 \\ -2 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Another choice for these matrices is

$$P = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -2 & 1 \\ 0 & 2 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

$\triangleleft$

Example 4.4.6 (Nondiagonalizable)**.** Show that matrix A below is not diagonalizable,

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution: We have seen in Example 4.3.10 that this matrix has an eigenvalues and eigenvectors

$$\lambda_1 = 1, \quad \boldsymbol{v}^{(1)} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \quad \text{and} \quad \lambda_2 = 3, \quad \boldsymbol{w}^{(2)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Since this matrix A is 3×3 and has at most two linearly independent eigenvectors, Theorem 4.4.9 says this matrix is not diagonalizable. $\triangleleft$

4.4.4. The Case of Different Eigenvalues.

Theorem 4.4.9 shows the importance of knowing whether an $n \times n$ matrix has a linearly independent set of n eigenvectors. More often than not, there is no simple way to check this property other than to compute all the matrix eigenvectors. However, when an $n \times n$ matrix has n *different* eigenvalues we do not need to compute the eigenvectors. The following result says that such matrix always have a linearly independent set of n eigenvectors, and then Theorem 4.4.9 says this matrix is diagonalizable.

Theorem 4.4.11 (Different Eigenvalues)**.** *If an $n \times n$ matrix has n different eigenvalues, then this matrix has a linearly independent set of n eigenvectors.*

Proof of Theorem 4.4.11: Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of an $n \times n$ matrix A, all different from each other. Let $\boldsymbol{v}^{(1)}, \cdots, \boldsymbol{v}^{(n)}$ the corresponding eigenvectors, that is, $A\boldsymbol{v}^{(i)} = \lambda_i \boldsymbol{v}^{(i)}$, with $i = 1, \cdots, n$. We have to show that the set $\{\boldsymbol{v}^{(1)}, \cdots, \boldsymbol{v}^{(n)}\}$ is linearly independent. We assume that the opposite is true and we obtain a contradiction. Let us assume that the set above is linearly dependent, that is, there are constants $c_1, \cdots, c_n$, not all zero, such that,

$$c_1 \boldsymbol{v}^{(1)} + \cdots + c_n \boldsymbol{v}^{(n)} = \boldsymbol{0}. \tag{4.4.1}$$

Let us name the eigenvalues and eigenvectors such that $c_1 \neq 0$. Now, multiply the equation above by the matrix A, the result is,

$$c_1 \lambda_1 \boldsymbol{v}^{(1)} + \cdots + c_n \lambda_n \boldsymbol{v}^{(n)} = \boldsymbol{0}.$$

Multiply Eq. (4.4.1) by the eigenvalue λ_n, the result is,

$$c_1 \lambda_n \boldsymbol{v}^{(1)} + \cdots + c_n \lambda_n \boldsymbol{v}^{(n)} = \boldsymbol{0}.$$

Subtract the second from the first of the equations above, then the last term on the right-hand sides cancels out, and we obtain,

$$c_1(\lambda_1 - \lambda_n)\boldsymbol{v}^{(1)} + \cdots + c_{n-1}(\lambda_{n-1} - \lambda_n)\boldsymbol{v}^{(n-1)} = \boldsymbol{0}. \tag{4.4.2}$$

Repeat the whole procedure starting with Eq. (4.4.2), that is, multiply this later equation by matrix A and also by λ_{n-1}, then subtract the second from the first, the result is,

$$c_1(\lambda_1 - \lambda_n)(\lambda_1 - \lambda_{n-1})\boldsymbol{v}^{(1)} + \cdots + c_{n-2}(\lambda_{n-2} - \lambda_n)(\lambda_{n-2} - \lambda_{n-1})\boldsymbol{v}^{(n-2)} = \boldsymbol{0}.$$

Repeat the whole procedure a total of $n - 1$ times, in the last step we obtain the equation

$$c_1(\lambda_1 - \lambda_n)(\lambda_1 - \lambda_{n-1}) \cdots (\lambda_1 - \lambda_3)(\lambda_1 - \lambda_2)\boldsymbol{v}^{(1)} = \boldsymbol{0}.$$

Since all the eigenvalues are different, we conclude that $c_1 = 0$, however this contradicts our assumption that $c_1 \neq 0$. Therefore, the set of n eigenvectors must be linearly independent. This establishes the Theorem. □

Example 4.4.7 (Different Eigenvalues)**.** Is matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ diagonalizable?

Solution: The characteristic polynomial of matrix A is

$$p(\lambda) = \begin{vmatrix} (1-\lambda) & 1 \\ 1 & (1-\lambda) \end{vmatrix} = (1-\lambda)^2 - 1 = \lambda^2 - 2\lambda \quad \Rightarrow \quad p(\lambda) = \lambda(\lambda - 2).$$

The roots of the characteristic polynomial are the matrix eigenvalues,

$$\lambda_1 = 0, \qquad \lambda_2 = 2.$$

The eigenvalues are different, then Theorem 4.4.11 says that matrix A is diagonalizable.◁

4.4.5. Exercises.

4.4.1.- .

4.4.2.- .

4.5. The Matrix Exponential

When we multiply two square matrices the result is another square matrix. This property allow us to define power functions and polynomials of a square matrix. In this section we go one step further and define the exponential of a square matrix. We will show that the derivative of the exponential function on matrices, as the one defined on real numbers, is proportional to itself.

4.5.1. The Scalar Exponential. The exponential function defined on real numbers, $f(x) = e^{ax}$, where a is a constant and $x \in \mathbb{R}$, satisfies $f'(x) = a\, f(x)$. We want to find a function of a square matrix with a similar property. Since the exponential on real numbers can be defined in several equivalent ways, we start with a short review of three of ways to define the exponential e^x.

(a) The exponential function can be defined as a generalization of the power function from the positive integers to the real numbers. One starts with positive integers n, defining

$$e^n = e \cdots e, \quad n\text{-times}.$$

Then one defines $e^0 = 1$, and for negative integers $-n$

$$e^{-n} = \frac{1}{e^n}.$$

The next step is to define the exponential for rational numbers, $\dfrac{m}{n}$, with m, n integers,

$$e^{\frac{m}{n}} = \sqrt[n]{e^m}.$$

The difficult part in this definition of the exponential is the generalization to irrational numbers, x, which is done by a limit procedure,

$$e^x = \lim_{\frac{m}{n} \to x} e^{\frac{m}{n}}.$$

It is nontrivial to define that limit precisely, which is why many calculus textbooks do not show it. Because all this, it is not clear how to generalize this definition from real numbers, x, to square matrices, X.

(b) The exponential function can be defined as the inverse of the natural logarithm function $g(x) = \ln(x)$, which in turns is defined as the area under the graph of the function $h(x) = \dfrac{1}{x}$ from 1 to x, that is,

$$\ln(x) = \int_1^x \frac{1}{y}\, dy, \qquad x \in (0, \infty).$$

Again, it is not clear how to extend to matrices this definition of the exponential function on real numbers.

(c) The exponential function can be defined also by its Taylor series expansion,

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

Most calculus textbooks show this series expansion, a Taylor expansion, as a result from the exponential definition, not as a definition itself. But one can define the exponential using this series and prove that the function so defined satisfies the properties in (a) and (b). It turns out, this series expression can be generalized square matrices.

4.5.2. The Matrix Exponential. We now use the idea in (c) to define the exponential function on square matrices. We start with the power function of a square matrix, $f(X) = X^n = X \cdots X$, n-times, for X a square matrix and n a positive integer. Then we define a polynomial of a square matrix,

$$p(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0 I.$$

Now we are ready to define the exponential of a square matrix.

Definition 4.5.1. *The **exponential** of a square matrix A, denoted as e^A, is given by*

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}. \tag{4.5.1}$$

This definition makes sense because the infinite sum in Eq. (4.5.1) converges. This is shown, in Section 7.5 of Apostol's Calculus, [2]. A different proof uses the Spectral Theorem—therefore this proof works in the particular case of normal matrices—and can be found in Section 2.1.2 and 4.5 in Hassani's Mathematical Physics, [6]. Another proof uses the Cayley-Hamilton Theorem, which reduces the infinite sume to a finite sum, and then computing the exponential is then equivalent to solving a linear system of equations for the exponential components.

Usually it is difficult to compute the exponential of a general square matrix. But, when the matrix is diagonal the exponential is remarkably simple.

Theorem 4.5.2 (Exponential of Diagonal Matrices)**.** *If $D = \operatorname{diag}\big[d_{11}, \cdots, d_{nn}\big]$, then*

$$e^D = \operatorname{diag}\big[e^{d_{11}}, \cdots, e^{d_{nn}}\big].$$

Proof of Theorem 4.5.2: We start from the definition of the exponential,

$$e^D = \sum_{k=0}^{\infty} \frac{1}{k!} \big(\operatorname{diag}\big[d_{11}, \cdots, d_{nn}\big]\big)^k.$$

We know that for diagonal matrices holds

$$\operatorname{diag}\big[d_{11}, \cdots, d_{nn}\big]^k = \operatorname{diag}\big[(d_{11})^k, \cdots, (d_{nn})^k\big],$$

therefore we get

$$e^D = \sum_{k=0}^{\infty} \frac{1}{k!} \operatorname{diag}\big[(d_{11})^k, \cdots, (d_{nn})^k\big].$$

Then,

$$e^D = \sum_{k=0}^{\infty} \operatorname{diag}\Big[\frac{(d_{11})^k}{k!}, \cdots, \frac{(d_{nn})^k}{k!}\Big] = \operatorname{diag}\Big[\sum_{k=0}^{\infty} \frac{(d_{11})^k}{k!}, \cdots, \sum_{k=0}^{\infty} \frac{(d_{nn})^k}{k!}\Big].$$

Each infinite sum in the diagonal of matrix above converges to an exponential,

$$\sum_{k=0}^{\infty} \frac{(d_{ii})^k}{k!} = e^{d_{ii}}.$$

So, we arrive to the equation

$$e^{\operatorname{diag}\big[d_{11}, \cdots, d_{nn}\big]} = \operatorname{diag}\big[e^{d_{11}}, \cdots, e^{d_{nn}}\big].$$

This establishes the Theorem. □

Example 4.5.1 (Exponential of Diagonal Matrix)**.** Compute e^D, where $D = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$.

Solution: We follow the proof of Theorem 4.5.2. From the definition of the exponential,

$$e^D = \sum_{n=0}^{\infty} \frac{D^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}^n.$$

Since the matrix D is diagonal, we have that

$$\begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}^n = \begin{bmatrix} 2^n & 0 \\ 0 & 7^n \end{bmatrix},$$

then,

$$e^D = \sum_{n=0}^{\infty} \frac{1}{n!} \begin{bmatrix} 2^n & 0 \\ 0 & 7^n \end{bmatrix} = \sum_{n=0}^{\infty} \begin{bmatrix} \frac{2^n}{n!} & 0 \\ 0 & \frac{7^n}{n!} \end{bmatrix} = \begin{bmatrix} \sum_{n=0}^{\infty} \frac{2^n}{n!} & 0 \\ 0 & \sum_{n=0}^{\infty} \frac{7^n}{n!} \end{bmatrix}.$$

Since

$$\sum_{n=0}^{\infty} \frac{a^n}{n!} = e^a,$$

for $a = 2$, 7, we obtain that

$$e^{\begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}} = \begin{bmatrix} e^2 & 0 \\ 0 & e^7 \end{bmatrix}.$$

$\lhd$

Remark: In the particular case of of 2×2 matrices, Theorem 4.5.2 implies that

$$e^{\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}} = \begin{bmatrix} e^a & 0 \\ 0 & e^d \end{bmatrix}.$$

However, in the case of a general 2×2 matrix we have that

$$e^{\begin{bmatrix} a & b \\ c & d \end{bmatrix}} \neq \begin{bmatrix} e^a & e^b \\ e^c & e^d \end{bmatrix}.$$

4.5.3. Formula for Diagonalizable Matrices. The exponential of a diagonalizable matrix is simple to compute, although not as simple as for diagonal matrices. The infinite sum in the exponential of a diagonalizable matrix reduces to a product of three matrices. We start with the following result, the nth-power of a diagonalizable matrix.

Theorem 4.5.3 (Powers of Diagonalizable Matrices)**.** *If an $n \times n$ matrix A is diagonalizable, with invertible matrix P and diagonal matrix D satisfying $A = PDP^{-1}$, then for every integer $k \geqslant 1$ holds*

$$A^k = PD^kP^{-1}. \tag{4.5.2}$$

Proof of Theorem 4.5.3: Since the case $n = 1$ is trivially true, we start computing the case $n = 2$. We get

$$A^2 = \left(PDP^{-1}\right)^2 = \left(PDP^{-1}\right)\left(PDP^{-1}\right) = PDDP^{-1} \quad \Rightarrow \quad A^2 = PD^2P^{-1},$$

that is, Eq. (4.5.2) holds for $k = 2$. Now assume that Eq. (4.5.2) is true for k. This equation also holds for $k + 1$, since

$$A^{(k+1)} = A^kA = \left(PD^kP^{-1}\right)\left(PDP^{-1}\right) = PD^kP^{-1}PDP^{-1} = PD^kDP^{-1}.$$

We conclude that $A^{(k+1)} = PD^{(k+1)}P^{-1}$. This establishes the Theorem. $\square$

We are ready to compute the exponential of a diagonalizable matrix.

Theorem 4.5.4 (Exponential of Diagonalizable Matrices)**.** *If an $n \times n$ matrix A is diagonalizable, with invertible matrix P and diagonal matrix D satisfying $A = PDP^{-1}$, then the exponential of matrix A is given by*

$$e^A = Pe^D P^{-1}. \tag{4.5.3}$$

Remark: Theorem 4.5.4 says that the infinite sum in the definition of e^A reduces to a product of three matrices when the matrix A is diagonalizable. This Theorem also says that *to compute the exponential of a diagonalizable matrix we need to compute the eigenvalues and eigenvectors of that matrix.*

Proof of Theorem 4.5.4: We start with the definition of the exponential,

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!}\, A^n = \sum_{k=0}^{\infty} \frac{1}{k!}\, (PDP^{-1})^n = \sum_{k=0}^{\infty} \frac{1}{k!}\, (PD^n P^{-1}),$$

where the last step comes from Theorem 4.5.3. Now, in the expression on the far right we can take common factor P on the left and P^{-1} on the right, that is,

$$e^A = P\Big(\sum_{k=0}^{\infty} \frac{1}{k!}\, D^n\Big)\, P^{-1}.$$

The sum in between parenthesis is the exponential of the diagonal matrix D, which we computed in Theorem 4.5.2,

$$e^A = Pe^D P^{-1}.$$

This establishes the Theorem. □

Remark: We have defined the exponential function

$$\tilde{F}(A) = e^A : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n},$$

which is a function from the space of square matrices into the space of square matrices. However, when one studies solutions to linear systems of differential equations, one needs a slightly different type of functions. One needs functions of the form

$$F(t) = e^{At} : \mathbb{R} \to \mathbb{R}^{n\times n},$$

where A is a constant square matrix and the independent variable is $t \in \mathbb{R}$. That is, one needs to generalize the real constant a in the function $f(t) = e^{at}$ to an $n \times n$ matrix A.

In the case that the matrix A is diagonalizable, with $A = PDP^{-1}$, so is matrix At, and $At = P(Dt)P^{-1}$. Therefore, the formula for the exponential of At is simply

$$e^{At} = Pe^{Dt}P^{-1}.$$

We use this formula in the following example.

Example 4.5.2. Compute e^{At}, where $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ and $t \in \mathbb{R}$.

Solution: To compute e^{At} we need the decomposition $A = PDP^{-1}$, which in turns implies that $At = P(Dt)P^{-1}$. Matrices P and D are constructed with the eigenvectors and eigenvalues of matrix A. We computed them in Example 4.3.4,

$$\lambda_1 = 4, \quad \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \lambda_2 = -2, \quad \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Introduce P and D as follows,

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \Rightarrow \quad P^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}.$$

Then, the exponential function is given by

$$e^{At} = Pe^{Dt}P^{-1} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{4t} & 0 \\ 0 & e^{-2t} \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Usually one leaves the function in this form. If we multiply the three matrices out we get

$$e^{At} = \frac{1}{2} \begin{bmatrix} (e^{4t} + e^{-2t}) & (e^{4t} - e^{-2t}) \\ (e^{4t} - e^{-2t}) & (e^{4t} + e^{-2t}) \end{bmatrix}.$$

◁

4.5.4. Properties of the Exponential. We summarize some simple properties of the exponential function in the following result. We leave the proof as an exercise.

Theorem 4.5.5 (Algebraic Properties)**.** *If A is an $n \times n$ matrix, then*

(a) If 0 is the $n \times n$ zero matrix, then $e^0 = I$.
(b) $\left(e^A\right)^T = e^{(A^T)}$, where T means transpose.
(c) For all nonnegative integers k holds $A^k e^A = e^A A^k$.
(d) If $AB = BA$, then $A e^B = e^B A$ and $e^A e^B = e^B e^A$.

An important property of the exponential on real numbers is not true for the exponential on matrices. We know that $e^a e^b = e^{a+b}$ for all real numbers a, b. However, there exist $n \times n$ matrices A, B such that $e^A e^B \neq e^{A+B}$. We now prove a weaker property.

Theorem 4.5.6 (Group Property)**.** *For any $n \times n$ matrix A and s, t real numbers holds*

$$e^{As} e^{At} = e^{A(s+t)}.$$

Proof of Theorem 4.5.6: We start with the definition of the exponential function

$$e^{As} e^{At} = \Big(\sum_{j=0}^{\infty} \frac{A^j s^j}{j!}\Big) \Big(\sum_{k=0}^{\infty} \frac{A^k t^k}{k!}\Big) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{A^{j+k} s^j t^k}{j!\, k!}.$$

We now introduce the new label $n = j + k$, then $j = n - k$, and we reorder the terms,

$$e^{As} e^{At} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{A^n s^{n-k} t^k}{(n-k)!\, k!} = \sum_{n=0}^{\infty} \frac{A^n}{n!} \Big(\sum_{k=0}^{n} \frac{n!}{(n-k)!\, k!} s^{n-k} t^k\Big).$$

If we recall the binomial theorem, $(s+t)^n = \sum_{k=0}^{n} \frac{n!}{(n-k)!\, k!} s^{n-k} t^k$, we get

$$e^{As} e^{At} = \sum_{n=0}^{\infty} \frac{A^n}{n!} (s+t)^n = e^{A(s+t)}.$$

This establishes the Theorem. □

If we set $s = 1$ and $t = -1$ in the Theorem 4.5.6 we get that

$$e^A e^{-A} = e^{A(1-1)} = e^0 = I,$$

so we have a formula for the inverse of the exponential. We write this result as its own theorem.

Theorem 4.5.7 (Inverse Exponential)**.** *If A is an $n \times n$ matrix, then*

$$\left(e^{A}\right)^{-1} = e^{-A}.$$

Example 4.5.3. Verify Theorem 4.5.7 for e^{At}, where $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ and $t \in \mathbb{R}$.

Solution: In Example 4.5.2 we found that

$$e^{At} = \frac{1}{2} \begin{bmatrix} (e^{4t} + e^{-2t}) & (e^{4t} - e^{-2t}) \\ (e^{4t} - e^{-2t}) & (e^{4t} + e^{-2t}) \end{bmatrix}.$$

We know that a 2×2 matrix is invertible iff its determinant is nonzero. In our case,

$$\det\left(e^{At}\right) = \frac{1}{2}\left(e^{4t} + e^{-2t}\right) \frac{1}{2}\left(e^{4t} + e^{-2t}\right) - \frac{1}{2}\left(e^{4t} - e^{-2t}\right) \frac{1}{2}\left(e^{4t} - e^{-2t}\right)$$

which gives us

$$\det\left(e^{At}\right) = e^{2t},$$

hence e^{At} is invertible. The inverse is

$$\left(e^{At}\right)^{-1} = \frac{1}{e^{2t}} \frac{1}{2} \begin{bmatrix} (e^{4t} + e^{-2t}) & (-e^{4t} + e^{-2t}) \\ (-e^{4t} + e^{-2t}) & (e^{4t} + e^{-2t}) \end{bmatrix},$$

that is

$$\left(e^{At}\right)^{-1} = \frac{1}{2} \begin{bmatrix} (e^{2t} + e^{-4t}) & (-e^{2t} + e^{-4t}) \\ (-e^{2t} + e^{-4t}) & (e^{2t} + e^{-4t}) \end{bmatrix}.$$

We now compute e^{-At}, which is pretty simple,

$$e^{-At} = \frac{1}{2} \begin{bmatrix} (e^{-4t} + e^{2t}) & (e^{-4t} - e^{2t}) \\ (e^{-4t} - e^{2t}) & (e^{-4t} + e^{2t}) \end{bmatrix}.$$

We see that

$$\left(e^{At}\right)^{-1} = e^{-At}.$$

$\lhd$

We now want to compute the derivative of the function $F(t) = e^{At}$, where A is a constant $n \times n$ matrix and $t \in \mathbb{R}$. It is not difficult to show the following result.

Theorem 4.5.8 (Derivative of the Exponential)**.** *If A is an $n \times n$ matrix, and $t \in \mathbb{R}$, then*

$$\frac{d}{dt} e^{At} = A\, e^{At}.$$

Remark: Recall that Theorem 4.5.5 says that $A\, e^{A} = e^{A}\, A$, so we have that

$$\frac{d}{dt} e^{At} = A\, e^{At} = e^{A}\, A.$$

First Proof of Theorem 4.5.8: We use the definition of the exponential,

$$\frac{d}{dt} e^{At} = \frac{d}{dt} \sum_{n=0}^{\infty} \frac{A^n t^n}{n!} = \sum_{n=0}^{\infty} \frac{A^n}{n!} \frac{d}{dt}(t^n) = \sum_{n=1}^{\infty} \frac{A^n t^{n-1}}{(n-1)!} = A \sum_{n=1}^{\infty} \frac{A^{(n-1)} t^{n-1}}{(n-1)!},$$

therefore we get

$$\frac{d}{dt} e^{At} = A\, e^{At}.$$

This establishes the Theorem. $\square$

Second Proof of Theorem 4.5.8: We use the definition of derivative and Theorem 4.5.6,

$$F'(t) = \lim_{h\to 0} \frac{e^{A(t+h)} - e^{At}}{h} = \lim_{h\to 0} \frac{e^{At}\, e^{Ah} - e^{At}}{h} = e^{At} \Big(\lim_{h\to 0} \frac{e^{Ah} - I}{h} \Big),$$

and using now the power series definition of the exponential we get

$$F'(t) = e^{At} \Big[\lim_{h\to 0} \frac{1}{h} \Big(Ah + \frac{A^2 h^2}{2!} + \cdots \Big) \Big] = e^{At}\, A.$$

This establishes the Theorem. □

Example 4.5.4. Verify Theorem 4.5.8 for $F(t) = e^{At}$, where $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ and $t \in \mathbb{R}$.

Solution: In Example 4.5.2 we found that

$$e^{At} = \frac{1}{2} \begin{bmatrix} (e^{4t} + e^{-2t}) & (e^{4t} - e^{-2t}) \\ (e^{4t} - e^{-2t}) & (e^{4t} + e^{-2t}) \end{bmatrix}.$$

Therefore, if we derivate component by component we get

$$\frac{d}{dt} e^{At} = \frac{1}{2} \begin{bmatrix} (4\,e^{4t} - 2\,e^{-2t}) & (4\,e^{4t} + 2\,e^{-2t}) \\ (4\,e^{4t} + 2\,e^{-2t}) & (4\,e^{4t} - 2\,e^{-2t}) \end{bmatrix}.$$

On the other hand, if we compute

$$\begin{aligned} A\, e^{At} &= \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} (e^{4t} + e^{-2t}) & (e^{4t} - e^{-2t}) \\ (e^{4t} - e^{-2t}) & (e^{4t} + e^{-2t}) \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} (4\,e^{4t} - 2\,e^{-2t}) & (4\,e^{4t} + 2\,e^{-2t}) \\ (4\,e^{4t} + 2\,e^{-2t}) & (4\,e^{4t} - 2\,e^{-2t}) \end{bmatrix} \end{aligned}$$

Therefore, $\dfrac{d}{dt} e^{At} = A\, e^{At}$. The relation $\dfrac{d}{dt} e^{At} = e^{At}\, A$ is shown in a similar way. ◁

We end this brief summary of the matrix exponential showing that not all the properties of the exponential of scalar numbers hold for the exponential of a matrix. The formula

$$e^{a+b} = e^{A} a\, e^{b},$$

which holds for all scalars a, b, does not hold for all matrices.

Theorem 4.5.9 (Exponent Rule). *If A, B are $n \times n$ matrices such that $AB = BA$, then*

$$e^{A+B} = e^{A}\, e^{B}.$$

Proof of Theorem 4.5.9: Introduce the function

$$F(t) = e^{(A+B)t}\, e^{-Bt}\, e^{-At},$$

where $t \in \mathbb{R}$. Compute the derivative of $F(t)$,

$$F'(t) = (A+B)\, e^{(A+B)t}\, e^{-Bt}\, e^{-At} + e^{(A+B)t}\, (-B)\, e^{-Bt}\, e^{-At} + e^{(A+B)t}\, e^{-Bt}\, (-A)\, e^{-At}.$$

Since $AB = BA$, we know that $e^{-Bt} A = A\, e^{-Bt}$, so we get

$$F'(t) = (A+B)\, e^{(A+B)t}\, e^{-Bt}\, e^{-At} - e^{(A+B)t}\, B\, e^{-Bt}\, e^{-At} - e^{(A+B)t}\, A\, e^{-Bt}\, e^{-At}.$$

Now $AB = BA$ also implies that $(A+B)\, B = B\, (A+B)$, therefore Theorem 4.5.5 implies

$$e^{(A+B)t}\, B = B\, e^{(A+B)t}.$$

Analogously, we have that $(A+B)\, A = A\, (A+B)$, therefore Theorem 4.5.5 implies that

$$e^{(A+B)t}\, A = A\, e^{(A+B)t}.$$

Using these equations in $F'(t)$ we get

$$F'(t) = (A + B)\,F(t) - B\,F(t) - A\,F(t) \quad \Rightarrow \quad F'(t) = 0.$$

Therefore, $F(t)$ is a constant matrix, $F(t) = F(0) = I$. So we get

$$e^{(A+B)t}\,e^{-Bt}\,e^{-At} = I \quad \Rightarrow \quad e^{(A+B)t} = e^{At}\,e^{Bt}.$$

This establishes the Theorem. □

4.5.5. Exercises.

4.5.1.- Use the definition of the matrix exponential to prove Theorem **??**. Do not use any other theorems in this Section.

4.5.2.- If $A^2 = A$, find a formula for e^A which does not contain an infinite sum.

4.5.3.- Compute e^A for the following matrices:

(a) $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

(b) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

(c) $A = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}$.

4.5.4.- Show that, if A is diagonalizable,

$$\det\big(e^A\big) = e^{\operatorname{tr}(A)}.$$

Remark: This result is true for all square matrices, but it is hard to prove for nondiagonalizable matrices.

4.5.5.- Compute e^A for the following matrices:

(a) $A = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$.

(b) $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$.

4.5.6.- If $A^2 = I$, show that

$$2\,e^A = \Big(e + \frac{1}{e}\Big)\, I + \Big(e - \frac{1}{e}\Big)\, A.$$

4.5.7.- If λ and $\boldsymbol{v}$ are an eigenvalue and eigenvector of A, then show that

$$e^A \boldsymbol{v} = e^{\lambda}\, \boldsymbol{v}.$$

4.5.8.- By direct computation show that $e^{(A+B)} \neq e^A\, e^B$ for

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

CHAPTER 5

Systems of Differential Equations

In these chapter we find formulas for solutions to systems of linear differential equations. We start with homogeneous linear systems with constant coefficients. We then use these solutions to characterize the solutions of systems of nonlinear differential equations. First find their constant solutions, called critical points. Then we student the behavior of solutions to nonlinear systems near their critical points.

5.1. Two-Dimensional Linear Systems

In this section we introduce 2×2 systems of first order linear differential equations. We then focus on homogeneous systems with constant coefficients. These systems are simple enough so their solutions can be computed and classified. But they are non-trivial enough so their solutions describe several situations including exponential decays and oscillations. In later sections we will use these systems as approximations of more complicated nonlinear systems.

5.1.1. 2×2 Linear Systems.

The simplest systems of differential equations are the linear systems. The simplest linear systems are the 2×2 linear systems.

Definition 5.1.1. *An 2×2* ***first order linear differential system*** *is the equation*

$$\boldsymbol{x}'(t) = A(t)\, \boldsymbol{x}(t) + \boldsymbol{b}(t), \tag{5.1.1}$$

where the 2×2 coefficient matrix A, the source n-vector $\boldsymbol{b}$, and the unknown n-vector $\boldsymbol{x}$ are given in components by

$$A(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{bmatrix}, \qquad \boldsymbol{b}(t) = \begin{bmatrix} b_1(t) \\ b_2(t) \end{bmatrix}, \qquad \boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.$$

(i) The system in 5.1.1 is ***homogeneous*** *iff the source vector $\boldsymbol{b} = \boldsymbol{0}$.*
(ii) The system in 5.1.1 has ***constant coefficients*** *iff the matrix A is constant.*
(iii) The system in 5.1.1 is ***diagonalizable*** *iff the matrix A is diagonalizable.*

Remarks:

(a) The derivative of a a vector valued function is defined as $\boldsymbol{x}'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix}$.

(b) By the definition of the matrix-vector product, Eq. (5.1.1) can be written as

$$x_1'(t) = a_{11}(t)\, x_1(t) + a_{12}(t)\, x_2(t) + b_1(t),$$
$$x_2'(t) = a_{21}(t)\, x_1(t) + a_{22}(t)\, x_2(t) + b_2(t).$$

(c) We recall that in § 4.4 we say that a square matrix A is diagonalizable iff there exists an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

A *solution* of an 2×2 linear differential system is a 2-vector valued function $\boldsymbol{x}(t)$, that is, a 2-vector with components $x_1(t)$, $x_2(t)$, that satisfy both differential equations in the system. When we write the equations we usually write $\boldsymbol{x}$ instead of $\boldsymbol{x}(t)$.

Example 5.1.1. A single linear differential equation for a function $y(t)$,

$$y' = a(t)\, y + b(t),$$

can be written using a notation similar to the one used above, that is, find a function $x_1(t)$ solution of

$$x_1' = a_{11}(t)\, x_1 + b_1(t).$$

So, the single first order linear differential equation would be the case of a 1×1 system. Since this is a linear first order equation, its solutions can be found with the integrating factor method described in § 1.2. ◁

Example 5.1.2. Use matrix notation to write down the 2×2 system given by

$$x_1' = x_1 - x_2 + e^{2t},$$
$$x_2' = -x_1 + x_2 - t^2\, e^{-t}.$$

Solution: We can rewrite the system above using vectors as follows

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ -x_1 + x_2 \end{bmatrix} + \begin{bmatrix} e^{2t} \\ -t^2 \, e^{-t} \end{bmatrix}$$

The first term on the right-hand side above is a matrix-vector product,

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} e^{2t} \\ -t^2 \, e^{-t} \end{bmatrix}$$

Therefore, the matrix of coefficients, the source vector, and unknown vector are

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \qquad \boldsymbol{b}(t) = \begin{bmatrix} e^{2t} \\ -t^2 \, e^{-t} \end{bmatrix}, \qquad \boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.$$

With these definitions the differential equation can be written as follows,

$$\boldsymbol{x}' = A\boldsymbol{x} + \boldsymbol{b}.$$

◁

Example 5.1.3. Find the explicit expression for the linear system $\boldsymbol{x}' = A\boldsymbol{x} + \boldsymbol{b}$, where

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}, \qquad \boldsymbol{b}(t) = \begin{bmatrix} e^{t} \\ 2e^{3t} \end{bmatrix}, \qquad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Solution: The 2×2 linear system is given by

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} e^{t} \\ 2e^{3t} \end{bmatrix} \quad \Leftrightarrow \quad \begin{aligned} x_1' &= x_1 + 3x_2 + e^{t}, \\ x_2' &= 3x_1 + x_2 + 2e^{3t}. \end{aligned}$$

◁

Example 5.1.4. Show that the vector valued functions

$$\boldsymbol{x}^1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{2t}, \qquad \boldsymbol{x}^2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-t}$$

are solutions to the 2×2 linear system $\boldsymbol{x}' = A\boldsymbol{x}$, where $A = \begin{bmatrix} 3 & -2 \\ 2 & -2 \end{bmatrix}$.

Solution: We compute the left-hand side and the right-hand side of the differential equation above for the function $\boldsymbol{x}^1$ and we see that both side match, that is,

$$A\boldsymbol{x}^1 = \begin{bmatrix} 3 & -2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{2t} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} e^{2t} = 2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{2t}; \quad \boldsymbol{x}^{1\,\prime} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \left(e^{2t}\right)' = \begin{bmatrix} 2 \\ 1 \end{bmatrix} 2\, e^{2t},$$

so we conclude that $\boldsymbol{x}^{1\,\prime} = A\boldsymbol{x}^1$. Analogously,

$$A\boldsymbol{x}^2 = \begin{bmatrix} 3 & -2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-t} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} e^{-t} = - \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-t}; \quad \boldsymbol{x}^{2\,\prime} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \left(e^{-t}\right)' = - \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-t},$$

so we conclude that $\boldsymbol{x}^{2\,\prime} = A\boldsymbol{x}^2$. ◁

5.1.2. Order Transformations. There is a relation between solutions to 2×2 systems of linear differential equations and the solutions of 2nd order linear scalar differential equations. We study two of these relations, the first order reduction and the second order reduction. All we do in this section can be generalized to $n \times n$ systems.

It is useful to have a correspondence between solutions of an 2×2 linear system and an 2nd order scalar equation. One reason is that concepts developed for one of the equations can be translated to the other equation. For example, we have introduced several concepts when we studied 2nd order scalar linear equations in § 2.1, concepts such as the superposition property, fundamental solutions, general solutions, the Wronskian, and Abel's theorem. It turns out that these concepts can be translated to 2×2 (and in general to $n \times n$) linear differential systems.

We now introduce the first order reduction—any second order linear differential equation can be written as a 2×2 system of first order linear differential equations.

Theorem 5.1.2 (First Order Reduction)**.** *A function y solves the second order equation*

$$y'' + a_1(t)\, y' + a_0(t)\, y = b(t), \tag{5.1.2}$$

iff the functions $x_1 = y$ and $x_2 = y'$ are solutions to the 2×2 first order differential system

$$x_1' = x_2, \tag{5.1.3}$$

$$x_2' = -a_0(t)\, x_1 - a_1(t)\, x_2 + b(t). \tag{5.1.4}$$

Proof of Theorem 5.1.2:
($\Rightarrow$) Given a solution y of Eq. (5.1.2), introduce the functions $x_1 = y$ and $x_2 = y'$. Therefore Eq. (5.1.3) holds, due to the relation

$$x_1' = y' = x_2,$$

Also Eq. (5.1.4) holds, because of the equation

$$x_2' = y'' = -a_0(t)\, y - a_1(t)\, y' + b(t) \quad \Rightarrow \quad x_2'' = -a_0(t)\, x_1 - a_1(t)\, x_2 + b(t).$$

($\Leftarrow$) Differentiate Eq. (5.1.3) and introduce the result into Eq. (5.1.4), that is,

$$x_1'' = x_2' \quad \Rightarrow \quad x_1'' = -a_0(t)\, x_1 - a_1(t)\, x_1' + b(t).$$

Denoting $y = x_1$, we obtain,

$$y'' + a_1(t)\, y' + a_0(t)\, y = b(t).$$

This establishes the Theorem. □

Example 5.1.5. Express as a first order system the second order equation

$$y'' + 2y' + 2y = \sin(at).$$

Solution: Introduce the new unknowns

$$x_1 = y, \quad x_2 = y' \quad \Rightarrow \quad x_1' = x_2.$$

Then, the differential equation can be written as

$$x_2' + 2\, x_2 + 2\, x_1 = \sin(at).$$

We conclude that

$$x_1' = x_2,$$
$$x_2' = -2\, x_1 - 2\, x_2 + \sin(at).$$

◁

We now introduce a second reductions—any 2×2 system of first order linear differential equations which is homogeneous and has constant coefficients can be written as a second order linear differential equation for each variable in the system.

Theorem 5.1.3 (Second Order Reduction)**.** *Any* 2×2 *first order linear homogeneous system with constant coefficients*

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

can be written as the second order equation

$$\boldsymbol{x}'' - \operatorname{tr}(A)\,\boldsymbol{x}' + \det(A)\,\boldsymbol{x} = \boldsymbol{0}. \tag{5.1.5}$$

Furthermore, the solution to the initial value problem

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0,$$

also solves the initial value problem given by Eq. (5.1.5) with initial conditions

$$\boldsymbol{x}(0) = \boldsymbol{x}_0, \qquad \boldsymbol{x}'(0) = A\boldsymbol{x}_0. \tag{5.1.6}$$

Remark: If we write $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, then Eq. (5.1.5) in components has the form

$$x_1'' - \operatorname{tr}(A)\,x_1' + \det(A)\,x_1 = 0, \tag{5.1.7}$$

$$x_2'' - \operatorname{tr}(A)\,x_2' + \det(A)\,x_2 = 0. \tag{5.1.8}$$

Also, recall that given a 2×2 matrix $a = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ the trace and the determinant of such matrix are

$$\operatorname{tr}(A) = a + d, \qquad \det(A) = ad - bc.$$

First Proof of Theorem 5.1.3: We start with the following identity, which is satisfied by every 2×2 matrix A, (exercise: prove it on 2×2 matrices by a straightforward calculation)

$$A^2 - \operatorname{tr}(A)\,A + \det(A)\,I = 0.$$

This identity is the particular case $n = 2$ of the Cayley-Hamilton Theorem, which holds for every $n \times n$ matrix. If we use this identity on the equation for $\boldsymbol{x}''$ we get the equation in Theorem 5.1.3, because

$$\boldsymbol{x}'' = \big(A\,\boldsymbol{x}\big)' = A\,\boldsymbol{x}' = A^2\,\boldsymbol{x} = \operatorname{tr}(A)\,A\boldsymbol{x} - \det(A) I\boldsymbol{x}.$$

Recalling that $A\,\boldsymbol{x} = \boldsymbol{x}'$, and $I\boldsymbol{x} = \boldsymbol{x}$, we get the vector equation

$$\boldsymbol{x}'' - \operatorname{tr}(A)\,\boldsymbol{x}' + \det(A)\,\boldsymbol{x} = \boldsymbol{0}.$$

The initial conditions for a second order differential equation are $\boldsymbol{x}(0)$ and $\boldsymbol{x}'(0)$. The first condition is given by hypothesis, $\boldsymbol{x}(0) = \boldsymbol{x}_0$. The second condition comes from the original first order system evaluated at $t = 0$, that is $\boldsymbol{x}'(0) = A\boldsymbol{x}(0) = A\boldsymbol{x}_0$. This establishes the Theorem. □

Second Proof of Theorem 5.1.3: This proof is based on a straightforward computation. Denote $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, then the system has the form

$$x_1' = a_{11}\,x_1 + a_{12}\,x_2 \tag{5.1.9}$$

$$x_2' = a_{21}\,x_1 + a_{22}\,x_2. \tag{5.1.10}$$

We start considering the case $a_{12} \neq 0$. Compute the derivative of the first equation,

$$x_1'' = a_{11}\,x_1' + a_{12}\,x_2'.$$

Use Eq. (5.1.10) to replace x_2' on the right-hand side above,

$$x_1'' = a_{11}\, x_1' + a_{12}\big(a_{21}\, x_1 + a_{22}\, x_2\big).$$

Since we are assuming that $a_{12} \neq 0$, we can replace the term with x_2 above using Eq. (5.1.9),

$$x_1'' = a_{11}\, x_1' + a_{12}a_{21}\, x_1 + a_{12}a_{22}\, \frac{\big(x_1' - a_{11}\, x_1\big)}{a_{12}}.$$

A simple cancellation and reorganization of terms gives the equation,

$$x_1'' = (a_{11} + a_{22})\, x_1' + (a_{12}a_{21} - a_{11}a_{22})\, x_1.$$

Recalling that $\operatorname{tr}(A) = a_{11} + a_{22}$, and $\det(A) = a_{11}a_{22} - a_{12}a_{21}$, we get

$$x_1'' - \operatorname{tr}(A)\, x_1' + \det(A)\, x_1 = 0.$$

The initial conditions for x_1 are $x_1(0)$ and $x_1'(0)$. The first one comes from the first component of $\boldsymbol{x}(0) = \boldsymbol{x}_0$, that is

$$x_1(0) = x_{01}. \tag{5.1.11}$$

The second condition comes from the first component of the first order differential equation evaluated at $t = 0$, that is $\boldsymbol{x}'(0) = A\boldsymbol{x}(0) = A\boldsymbol{x}_0$. The first component is

$$x_1'(0) = a_{11}\, x_{01} + a_{12}\, x_{02}. \tag{5.1.12}$$

Consider now the case $a_{12} = 0$. In this case the system is

$$x_1' = a_{11}\, x_1$$
$$x_2' = a_{21}\, x_1 + a_{22}\, x_2.$$

In this case compute one more derivative in the first equation above,

$$x_1'' = a_{11}\, x_1'.$$

Now rewrite the first equation in the system as follows

$$a_{22}\,(-x_1' + a_{11}\, x_1) = 0.$$

Adding these last two equations for x_1 we get

$$x_1'' - a_{11}\, x_1' + a_{22}\,(-x_1' + a_{11}\, x_1) = 0,$$

So we get the equation

$$x_1'' - (a_{11} + a_{22})\, x_1' + (a_{11}a_{22})\, x_1 = 0.$$

Recalling that in the case $a_{12} = 0$ we have $\operatorname{tr}(A) = a_{11} + a_{22}$, and $\det(A) = a_{11}a_{22}$, we get

$$x_1'' - \operatorname{tr}(A)\, x_1' + \det(A)\, x_1 = 0.$$

The initial conditions are the same as in the case $a_{12} \neq 0$. A similar calculation gives x_2 and its initial conditions. This establishes the Theorem. □

Example 5.1.6. Express as a single second order equation the 2×2 system and solve it,

$$x_1' = -x_1 + 3x_2,$$
$$x_2' = x_1 - x_2.$$

Solution: Instead of using the result from Theorem 5.1.3, we solve this problem following the second proof of that theorem. But instead of working with x_1, we work with x_2. We start computing x_1 from the second equation: $x_1 = x_2' + x_2$. We then introduce this expression into the first equation,

$$(x_2' + x_2)' = -(x_2' + x_2) + 3x_2 \quad \Rightarrow \quad x_2'' + x_2' = -x_2' - x_2 + 3x_2,$$

so we obtain the second order equation

$$x_2'' + 2x_2' - 2x_2 = 0.$$

We solve this equation with the methods studied in Chapter 2, that is, we look for solutions of the form $x_2(t) = e^{rt}$, with r solution of the characteristic equation

$$r^2 + 2r - 2 = 0 \quad \Rightarrow \quad r_\pm = \frac{1}{2}\left[-2 \pm \sqrt{4+8}\right] \quad \Rightarrow \quad r_\pm = -1 \pm \sqrt{3}.$$

Therefore, the general solution to the second order equation above is

$$x_2 = c_+ \, e^{(1+\sqrt{3})t} + c_- \, e^{(1-\sqrt{3})t}, \qquad c_+, \; c_- \in \mathbb{R}.$$

Since x_1 satisfies the same equation as x_2, we obtain the same general solution

$$x_1 = \tilde{c}_+ \, e^{(1+\sqrt{3})t} + \tilde{c}_- \, e^{(1-\sqrt{3})t}, \qquad \tilde{c}_+, \; \tilde{c}_- \in \mathbb{R}.$$

◁

Example 5.1.7. Write the first order initial value problem

$$\boldsymbol{x}' = A\,\boldsymbol{x}, \qquad A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad \boldsymbol{x}(0) = \begin{bmatrix} 5 \\ 6 \end{bmatrix},$$

as a second order initial value problem for x_1. Repeat the calculations for x_2.

Solution: From Theorem 5.1.3 we know that both x_1 and x_2 satisfy the same differential equation. Since $\operatorname{tr}(A) = 1 + 4 = 5$ and $\det(A) = 4 - 6 = -2$, the differential equations are

$$x_1'' - 5\,x_1' - 2\,x_1 = 0, \qquad x_2'' - 5\,x_2' - 2\,x_2 = 0.$$

From the same Theorem we know that the initial conditions for the second order differential equations above are $\boldsymbol{x}(0) = \boldsymbol{x}_0$ and $\boldsymbol{x}'(0) = A\,\boldsymbol{x}_0$, that is

$$\boldsymbol{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}, \qquad \boldsymbol{x}'(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 17 \\ 39 \end{bmatrix},$$

therefore, the initial conditions for x_1 and x_2 are

$$x_1(0) = 5, \quad x_1'(0) = 17, \qquad \text{and} \qquad x_2(0) = 6, \quad x_2'(0) = 39.$$

◁

5.1.3. Diagonalizable Systems.

Linear systems are the simplest systems to solve, but do not think they are simple to solve—that would be a mistake. In systems, the differential equations are coupled. To understand what coupled means, consider a 2×2, constant coefficients, homogeneous, linear system

$$\begin{aligned} x_1' &= a_{11}\,x_1 + a_{12}\,x_2 \\ x_2' &= a_{21}\,x_1 + a_{22}\,x_2. \end{aligned}$$

We cannot simply integrate the first equation to obtain x_1, because x_2 is on the right-hand side, and we do not know what this function is. Analogously, we cannot simply integrate the second equation to obtain x_2, because x_1 is on the right-hand side, and we do not know what this function is either. This is what we mean by the system to be coupled—one cannot solve for one variable at a time, one must solve it for both variables together.

In the particular case that the coefficient matrix is *diagonalizable*, it is possible to *decouple the system*. In the following example we show how this can be done.

Example 5.1.8. Find functions x_1, x_2 solutions of the first order, 2×2, constant coefficients, homogeneous differential system

$$x_1' = x_1 + 3x_2,$$
$$x_2' = 3x_1 + x_2.$$

Solution: If we write this system in matrix form we get

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Notice that the coefficient matrix is

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix},$$

which is diagonalizable. We know that this matrix A has eignpairs

$$\lambda_+ = 4, \quad \boldsymbol{v}^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \lambda_- = -2, \quad \boldsymbol{v}^- = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

This means that A can be written as

$$A = PDP^{-1}, \quad P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}, \quad P^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

If we add this information into the differential equation we get

$$\boldsymbol{x}' = PDP^{-1}\boldsymbol{x}.$$

We now want to do linear combinations among the equations in the system. One way to do that in an efficient way is to multiply the whole system by a matrix. We choose to multiply the system by P^{-1}, that is,

$$P^{-1}\boldsymbol{x}' = P^{-1}PDP^{-1}\boldsymbol{x} \quad \Rightarrow \quad (P^{-1}\boldsymbol{x})' = D\,(P^{-1}\boldsymbol{x}),$$

where we used that P^{-1} is a constant matrix, so its t-derivative is zero, hence we get $P^{-1}\boldsymbol{x}' = (P^{-1}\boldsymbol{x})'$. If we introduce the new variable $\boldsymbol{y} = P^{-1}\boldsymbol{x}$, we got the system

$$\boldsymbol{y}' = D\boldsymbol{y},$$

which is a diagonal system. We now *repeat* the steps above, but a bit slower, so we can show explicitly what is the effect on the system when we multiply it by P^{-1}. What we did is

$$P^{-1}\boldsymbol{x}' = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \frac{1}{2} \begin{bmatrix} (x_1 + x_2)' \\ (-x_1 + x_2)' \end{bmatrix}.$$

So, multiplying the system by P^{-1} means, in this case, to add and subtract the two equations in the original system,

$$(x_1 + x_2)' = 4x_1 + 4x_2 \quad \Rightarrow \quad (x_1 + x_2)' = 4(x_1 + x_2).$$
$$(x_2 - x_1)' = 2x_1 - 2x_2 \quad \Rightarrow \quad (x_2 - x_1)' = -2(x_2 - x_1).$$

Introduce the new variables $y_1 = (x_2 + x_1)/2$, and $y_2 = (x_2 - x_1)/2$, which written in matrix notation is

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} (x_1 + x_2) \\ (-x_1 + x_2) \end{bmatrix} \quad \Leftrightarrow \quad \boldsymbol{y} = P^{-1}\boldsymbol{x}.$$

In terms of this new variable $\boldsymbol{y}$, the system is

$$\left.\begin{aligned} y_1' &= 4y_1, \\ y_2' &= -2y_2 \end{aligned}\right\} \quad \Leftrightarrow \quad \boldsymbol{y}' = D\boldsymbol{y}.$$

We have decoupled the original system. The original system for $\boldsymbol{x}$ is coupled, but the new system for $\boldsymbol{y}$ is diagonal, so decoupled. The solution is

$$\left.\begin{array}{l} y_1' = 4y_1 \quad \Rightarrow \quad y_1 = c_1\, e^{4t}, \\ y_2' = 2y_2 \quad \Rightarrow \quad y_2 = c_2\, e^{-2t}, \end{array}\right\} \quad \Rightarrow \quad \boldsymbol{y}(t) = \begin{bmatrix} c_1\, e^{4t} \\ c_2\, e^{-2t} \end{bmatrix},$$

with c_1, $c_2 \in \mathbb{R}$. Now we go back to the original variables. Since $\boldsymbol{y} = P^{-1}\boldsymbol{x}$, then

$$\boldsymbol{x} = P\boldsymbol{y} \quad \Rightarrow \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} x_1 = y_1 - y_2, \\ x_2 = y_1 + y_2. \end{cases}$$

So the general solution is

$$x_1(t) = c_1\, e^{4t} - c_2\, e^{-2t}, \qquad x_2(t) = c_1\, e^{4t} + c_2\, e^{-2t}.$$

In vector notation we get

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (c_1\, e^{4t} - c_2\, e^{-2t}) \\ (c_1\, e^{4t} + c_2\, e^{-2t}) \end{bmatrix} = \begin{bmatrix} c_1\, e^{4t} \\ c_1\, e^{4t} \end{bmatrix} + \begin{bmatrix} -c_2\, e^{-2t} \\ c_2\, e^{-2t} \end{bmatrix}.$$

Therefore, we get all the solutions for the 2×2 linear differential system,

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

$\lhd$

In the example above we solved the differential equation

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix},$$

and all the solutions to that equation are

$$\boldsymbol{x}(t) = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{4t} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t},$$

where $c_1, c_2 \in \mathbb{R}$ are arbitrary constants. So we can write all the solutions above as arbitrary linear combinations of two solutions,

$$\boldsymbol{x}_1(t) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{4t}, \qquad \boldsymbol{x}_2(t) = \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}.$$

Notice that these solutions are constructed with the eigenpairs of the coefficient matrix A. We have found in previous sections that the eigenpairs of $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ are

$$\lambda_1 = 4, \quad \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad \lambda_2 = -2, \quad \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

So, we can write the solutions above as

$$\boldsymbol{x}_1(t) = \boldsymbol{v}_1\, e^{\lambda_1 t}, \qquad \boldsymbol{x}_2(t) = \boldsymbol{v}_2\, e^{\lambda_2 t}.$$

We now show that this is not a coincidence. We have actually discovered a fundamental property of diagonalizable systems of linear differential equations.

Theorem 5.1.4 (Diagonalizable Systems)**.** *If the 2×2 constant matrix A is diagonalizable with eigenpairs λ_1, $\boldsymbol{v}_1$, and λ_2, $\boldsymbol{v}_2$, then **all solutions** of $\boldsymbol{x}' = A\,\boldsymbol{x}$ are given by*

$$\boldsymbol{x}(t) = c_1\, e^{\lambda_1 t}\, \boldsymbol{v}_1 + c_2\, e^{\lambda_2 t}\, \boldsymbol{v}_2. \tag{5.1.13}$$

Remark: It is simple to check that the functions given in Eq. (5.1.13) are in fact solutions of the differential equation. The reason is that each function $\boldsymbol{x}_i = e^{\lambda_i t}\, \boldsymbol{v}_i$, for $i = 1, 2$, is solution of the system $\boldsymbol{x}' = A\,\boldsymbol{x}$. Indeed

$$\boldsymbol{x}_i' = \lambda_i\, e^{\lambda_i t}\, \boldsymbol{v}_i, \quad \text{and} \quad A\boldsymbol{x}_i = A\left(e^{\lambda_i t}\, \boldsymbol{v}_i\right) = e^{\lambda_i t}\, A\, \boldsymbol{v}_i = \lambda_i\, e^{\lambda_i t}\, \boldsymbol{v}_i,$$

hence $\boldsymbol{x}_i' = A\,\boldsymbol{x}_i$. What it is not so easy to prove is that Eq. (5.1.13) contains all the solutions.

Proof of Theorem 5.1.4: Since the coefficient matrix A is diagonalizable, there exist an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$. Introduce this expression into the differential equation and multiplying the whole equation by P^{-1},

$$P^{-1}\boldsymbol{x}'(t) = P^{-1}\big(PDP^{-1}\big)\, \boldsymbol{x}(t).$$

Notice that to multiply the differential system by the matrix P^{-1} means to perform a very particular type of linear combinations among the equations in the system. This is the linear combination that *decouples* the system. Indeed, since matrix A is constant, so is P and D. In particular $P^{-1}\boldsymbol{x}' = \big(P^{-1}\boldsymbol{x}\big)'$, hence

$$(P^{-1}\boldsymbol{x})' = D\left(P^{-1}\,\boldsymbol{x}\right).$$

Define the new variable $\boldsymbol{y} = \big(P^{-1}\boldsymbol{x}\big)$. The differential equation is now given by

$$\boldsymbol{y}'(t) = D\,\boldsymbol{y}(t).$$

Since matrix D is diagonal, the system above is a *decoupled* for the variable $\boldsymbol{y}$. Solve the decoupled initial value problem $\boldsymbol{y}'(t) = D\,\boldsymbol{y}(t)$,

$$\left.\begin{aligned} y_1'(t) &= \lambda_1\, y_1(t), \\ y_2'(t) &= \lambda_2\, y_2(t), \end{aligned}\right\} \quad \Rightarrow \quad \left\{\begin{aligned} y_1(t) &= c_1\, e^{\lambda_1 t}, \\ y_2(t) &= c_2\, e^{\lambda_2 t}, \end{aligned}\right\} \quad \Rightarrow \quad \boldsymbol{y}(t) = \begin{bmatrix} c_1\, e^{\lambda_1 t} \\ c_2\, e^{\lambda_2 t} \end{bmatrix}.$$

Once $\boldsymbol{y}$ is found, we transform back to $\boldsymbol{x}$,

$$\boldsymbol{x}(t) = P\,\boldsymbol{y}(t) = \big[\boldsymbol{v}_1, \boldsymbol{v}_2\big] \begin{bmatrix} c_1\, e^{\lambda_1 t} \\ c_2\, e^{\lambda_2 t} \end{bmatrix} = c_1\, e^{\lambda_1 t}\, \boldsymbol{v}_1 + c_2\, e^{\lambda_2 t}\, \boldsymbol{v}_2.$$

This establishes the Theorem. □

Remark: Since the eigenvalues are the roots of the characteristic polynomial, we will often use the notation $\lambda_{\pm}$ to denote these roots, where $\lambda_+ \geqslant \lambda_-$, for real eigenvalues.

We classify the diagonalizable 2×2 linear differential systems by the eigenvalues of their coefficient matrix.

(i) The eigenvalues λ_+, λ_- are real and distinct;
(ii) The eigenvalues $\lambda_{\pm} = \alpha \pm \beta i$ are distinct and complex, with $\lambda_+ = \overline{\lambda_-}$;
(iii) The eigenvalues $\lambda_+ = \lambda_- = \lambda_0$ is repeated and real.

We now provide a few examples of systems on each of the cases above, starting with an example of case *(i)*.

Example 5.1.9. Find the solution of the initial value problem

$$\boldsymbol{x}' = A\,\boldsymbol{x}, \qquad A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \qquad \boldsymbol{x}(0) = \begin{bmatrix} a \\ b \end{bmatrix},$$

where a, b are arbitrary constants.

Solution: It is not difficult to show that the eigenpairs of the coefficient matrix are

$$\lambda_+ = 3, \quad \boldsymbol{v}_+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad \lambda_- = -1, \quad \boldsymbol{v}_- = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

This coefficient matrix has distinct real eigenvalues, so the general solution of the differential equation is

$$\boldsymbol{x}(t) = c_+ \, e^{3t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_- \, e^{-t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Now, the initial condition is

$$\begin{bmatrix} a \\ b \end{bmatrix} = \boldsymbol{x}(0) = c_+ \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_- \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_+ \\ c_- \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Therefore, we get

$$\begin{bmatrix} c_+ \\ c_- \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{2} \begin{bmatrix} (a+b) \\ (-a+b) \end{bmatrix} \quad \Rightarrow \quad \begin{cases} c_+ = \dfrac{1}{2}(a+b) \\ c_- = \dfrac{1}{2}(-a+b). \end{cases}$$

Then, the solution of the initial value problem is

$$\boldsymbol{x}(t) = \frac{1}{2}(a+b) \, e^{3t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}(-a+b) \, e^{-t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

◁

5.1.4. The Case of Complex Eigenvalues. We now focus on case *(ii)*. The coefficient matrix is real-valued with the complex-valued eigenvalues. In this case each eigenvalue is the complex conjugate of the other. A similar result is true for $n \times n$ real-valued matrices. When such $n \times n$ matrix has a complex eigenvalue λ, there is another eigenvalue $\overline{\lambda}$. A similar result holds for the respective eigenvectors.

Theorem 5.1.5 (Conjugate Pairs)**.** *If an $n \times n$ real-valued matrix A has a complex eigenpair λ, $\boldsymbol{v}$, then the complex conjugate pair $\overline{\lambda}$, $\overline{\boldsymbol{v}}$ is also an eigenpair of matrix A.*

Proof of Theorem 5.1.5: Complex conjugate the eigenvalue eigenvector equation for λ and $\boldsymbol{v}$, and recall that matrix A is real-valued, hence $\overline{A} = A$. We obtain,

$$A\boldsymbol{v} = \lambda \boldsymbol{v} \quad \Leftrightarrow \quad A\overline{\boldsymbol{v}} = \overline{\lambda}\,\overline{\boldsymbol{v}},$$

This establishes the Theorem. □

Complex eigenvalues of a matrix with real coefficients are always complex conjugate pairs. Same it's true for their respective eigenvectors. So they can be written in terms of their real and imaginary parts as follows,

$$\lambda_{\pm} = \alpha \pm i\beta, \qquad \boldsymbol{v}_{\pm} = \boldsymbol{a} \pm i\boldsymbol{b}, \tag{5.1.14}$$

where $\alpha, \beta \in \mathbb{R}$ and $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$.

The general solution formula in Eq. (5.1.13) still holds in the case that A has complex eigenvalues and eigenvectors. The main drawback of this formula is similar to what we found in Chapter 2. It is difficult to separate real-valued from complex-valued solutions. The fix to that problem is also similar to the one found in Chapter 2—find a real-valued fundamental set of solutions.

Theorem 5.1.6 (Complex and Real Solutions)**.** *If $\lambda_{\pm} = \alpha \pm i\beta$ are eigenvalues of an 2×2 constant matrix A with eigenvectors $\boldsymbol{v}_{\pm} = \boldsymbol{a} \pm i\boldsymbol{b}$, where $\alpha, \beta \in \mathbb{R}$ and $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2$, then a linearly independent set of two complex-valued solutions to $\boldsymbol{x}' = A\,\boldsymbol{x}$ is*

$$\left\{ \boldsymbol{x}_+(t) = e^{\lambda_+ t}\, \boldsymbol{v}_+,\ \boldsymbol{x}_-(t) = e^{\lambda_- t}\, \boldsymbol{v}_-, \right\}. \tag{5.1.15}$$

Furthermore, a linearly independent set of two real-valued solutions to $\boldsymbol{x}' = A\,\boldsymbol{x}$ is given by

$$\Big\{\boldsymbol{x}_1(t) = \big(\boldsymbol{a}\,\cos(\beta t) - \boldsymbol{b}\,\sin(\beta t)\big)\,e^{\alpha t},\ \boldsymbol{x}_2(t) = \big(\boldsymbol{a}\,\sin(\beta t) + \boldsymbol{b}\,\cos(\beta t)\big)\,e^{\alpha t}\Big\}. \tag{5.1.16}$$

Proof of Theorem 5.1.6: Theorem 5.1.4 implies the set in (5.1.15) is a linearly independent set. The new information in Theorem 5.1.6 above is the real-valued solutions in Eq. (5.1.16). They are obtained from Eq. (5.1.15) as follows:

$$\begin{aligned}\boldsymbol{x}_\pm &= (\boldsymbol{a} \pm i\boldsymbol{b})\,e^{(\alpha\pm i\beta)t}\\ &= e^{\alpha t}(\boldsymbol{a} \pm i\boldsymbol{b})\,e^{\pm i\beta t}\\ &= e^{\alpha t}(\boldsymbol{a} \pm i\boldsymbol{b})\,\big(\cos(\beta t) \pm i\sin(\beta t)\big)\\ &= e^{\alpha t}\big(\boldsymbol{a}\,\cos(\beta t) - \boldsymbol{b}\,\sin(\beta t)\big) \pm ie^{\alpha t}\big(\boldsymbol{a}\,\sin(\beta t) + \boldsymbol{b}\,\cos(\beta t)\big).\end{aligned}$$

Since the differential equation $\boldsymbol{x}' = A\boldsymbol{x}$ is linear, the functions below are also solutions,

$$\begin{aligned}\boldsymbol{x}_1 &= \frac{1}{2}\big(\boldsymbol{x}_+ + \boldsymbol{x}_-\big) = \big(\boldsymbol{a}\,\cos(\beta t) - \boldsymbol{b}\,\sin(\beta t)\big)\,e^{\alpha t},\\ \boldsymbol{x}_2 &= \frac{1}{2i}\big(\boldsymbol{x}_+ - \boldsymbol{x}_-\big) = \big(\boldsymbol{a}\,\sin(\beta t) + \boldsymbol{b}\,\cos(\beta t)\big)\,e^{\alpha t}.\end{aligned}$$

This establishes the Theorem. □

Example 5.1.10. Find a real-valued set of fundamental solutions to the differential equation

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad A = \begin{bmatrix} 2 & 3\\ -3 & 2\end{bmatrix}. \tag{5.1.17}$$

Solution: Fist find the eigenvalues of matrix A above,

$$0 = \begin{vmatrix} (2-\lambda) & 3\\ -3 & (2-\lambda)\end{vmatrix} = (\lambda-2)^2 + 9 \quad \Rightarrow \quad \lambda_\pm = 2 \pm 3i.$$

Then find the respective eigenvectors. The one corresponding to λ_+ is the solution of the homogeneous linear system with coefficients given by

$$\begin{bmatrix} 2-(2+3i) & 3\\ -3 & 2-(2+3i)\end{bmatrix} = \begin{bmatrix} -3i & 3\\ -3 & -3i\end{bmatrix} \to \begin{bmatrix} -i & 1\\ -1 & -i\end{bmatrix} \to \begin{bmatrix} 1 & i\\ -1 & -i\end{bmatrix} \to \begin{bmatrix} 1 & i\\ 0 & 0\end{bmatrix}.$$

Therefore the eigenvector $\boldsymbol{v}_+ = \begin{bmatrix} v_{+1}\\ v_{+2}\end{bmatrix}$ is given by

$$v_{+1} = -iv_{+2} \quad \Rightarrow \quad v_{+2} = 1, \quad v_{+1} = -i, \quad \Rightarrow \quad \boldsymbol{v}_+ = \begin{bmatrix} -i\\ 1\end{bmatrix}, \quad \lambda_+ = 2+3i.$$

The second eigenvector is the complex conjugate of the eigenvector found above, that is,

$$\boldsymbol{v}_- = \begin{bmatrix} i\\ 1\end{bmatrix}, \quad \lambda_- = 2-3i.$$

Notice that

$$\boldsymbol{v}_\pm = \begin{bmatrix} 0\\ 1\end{bmatrix} \pm \begin{bmatrix} -1\\ 0\end{bmatrix} i.$$

Then, the real and imaginary parts of the eigenvalues and of the eigenvectors are given by

$$\alpha = 2, \qquad \beta = 3, \qquad \boldsymbol{a} = \begin{bmatrix} 0\\ 1\end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} -1\\ 0\end{bmatrix}.$$

So a real-valued expression for a fundamental set of solutions is given by

$$\boldsymbol{x}_1 = \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} \cos(3t) - \begin{bmatrix} -1 \\ 0 \end{bmatrix} \sin(3t) \right) e^{2t} \quad \Rightarrow \quad \boldsymbol{x}_1 = \begin{bmatrix} \sin(3t) \\ \cos(3t) \end{bmatrix} e^{2t},$$

$$\boldsymbol{x}_2 = \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} \sin(3t) + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \cos(3t) \right) e^{2t} \quad \Rightarrow \quad \boldsymbol{x}_2 = \begin{bmatrix} -\cos(3t) \\ \sin(3t) \end{bmatrix} e^{2t}.$$

◁

Example 5.1.11. Find the solution of the initial value problem

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad \boldsymbol{x}(0) = \begin{bmatrix} a \\ b \end{bmatrix},$$

where a and b are arbitrary constants and A is a 2×2 matrix with eigenvalues and eigenvectors

$$\lambda_1 = -2 + 3i, \quad \boldsymbol{v}_1 = \begin{bmatrix} 5 - 7i \\ -2 + 3i \end{bmatrix} \quad \text{and} \quad \lambda_2 = -2 - 3i, \quad \boldsymbol{v}_2 = \begin{bmatrix} 5 + 7i \\ -2 - 3i \end{bmatrix}.$$

Solution: We know that for a system $\boldsymbol{x}' = A\boldsymbol{x}$ with coefficient matrix A having eigenvalues $\lambda_{\pm} = \alpha \pm \beta i$ (the convention is that λ_+ is the eigenvalue with positive imaginary part, that is $\beta > 0$) and corresponding eigenvectors $\boldsymbol{v}_{\pm} = \boldsymbol{a} \pm \boldsymbol{b}i$, the fundamental solutions are

$$\boldsymbol{x}_1(t) = \big(\boldsymbol{a} \cos(\beta t) - \boldsymbol{b} \sin(\beta t)\big)\, e^{\alpha t},$$

$$\boldsymbol{x}_2(t) = \big(\boldsymbol{a} \sin(\beta t) + \boldsymbol{b} \cos(\beta t)\big)\, e^{\alpha t}.$$

To find α, β, $\boldsymbol{a}$, and $\boldsymbol{b}$ in our case we only need to focus on the eigenvalue with positive imaginary part and call it λ_+, that is $\lambda_+ = \lambda_1 = -2 + 3i$. This implies that

$$\alpha = -2, \qquad \beta = 3.$$

If $\lambda_1 = \lambda_+$, then $\boldsymbol{v}_1 = \boldsymbol{v}_+$. Since $\boldsymbol{v}_+ = \boldsymbol{a} + \boldsymbol{b}i$, we get

$$\boldsymbol{a} + \boldsymbol{b}i = \boldsymbol{v}_+ = \boldsymbol{v}_1 = \begin{bmatrix} 5 - 7i \\ -2 + 3i \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix} + \begin{bmatrix} -7 \\ 3 \end{bmatrix} i \quad \Rightarrow \quad \boldsymbol{a} = \begin{bmatrix} 5 \\ -2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} -7 \\ 3 \end{bmatrix}.$$

With all that information we can write the first real-valued fundamental solution

$$\boldsymbol{x}_1(t) = \big(\boldsymbol{a} \cos(\beta t) - \boldsymbol{b} \sin(\beta t)\big)\, e^{\alpha t} = \left(\begin{bmatrix} 5 \\ -2 \end{bmatrix} \cos(3t) - \begin{bmatrix} -7 \\ 3 \end{bmatrix} \sin(3t) \right) e^{-2t},$$

which gives us the solution

$$\boldsymbol{x}_1(t) = \begin{bmatrix} 5\cos(3t) + 7\sin(3t) \\ -2\cos(3t) - 3\sin(3t) \end{bmatrix} e^{-2t}.$$

The other real-valued fundamental solution is given by

$$\boldsymbol{x}_2(t) = \big(\boldsymbol{a} \sin(\beta t) + \boldsymbol{b} \cos(\beta t)\big)\, e^{\alpha t} = \left(\begin{bmatrix} 5 \\ -2 \end{bmatrix} \sin(3t) + \begin{bmatrix} -7 \\ 3 \end{bmatrix} \cos(3t) \right) e^{-2t},$$

which gives us the solution

$$\boldsymbol{x}_2(t) = \begin{bmatrix} -7\cos(3t) + 5\sin(3t) \\ 3\cos(3t) - 2\sin(3t) \end{bmatrix} e^{-2t}.$$

With the real-valued fundamental solutions we can write the general solution of the differential equation, $\boldsymbol{x}(t) = c_1 \boldsymbol{x}_1(t) + c_2 \boldsymbol{x}_2(t)$, as follows

$$\boldsymbol{x}(t) = \left(c_1 \begin{bmatrix} 5\cos(3t) + 7\sin(3t) \\ -2\cos(3t) - 3\sin(3t) \end{bmatrix} + c_2 \begin{bmatrix} -7\cos(3t) + 5\sin(3t) \\ 3\cos(3t) - 2\sin(3t) \end{bmatrix} \right) e^{-2t}.$$

The initial condition implies

$$\begin{bmatrix} a \\ b \end{bmatrix} = \boldsymbol{x}(0) = c_1 \begin{bmatrix} 5 \\ -2 \end{bmatrix} + c_2 \begin{bmatrix} -7 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 & -7 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

The determinant of the matrix on the far right is one, nonzero, so the matrix is invertible,

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{1} \begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3a + 7b \\ 2a + 5b \end{bmatrix},$$

meaning $c_1 = 3a + 7b$ and $c_2 = 2a + 5b$. Then, the solution of the initial value problem is

$$\boldsymbol{x}(t) = \Big((3a + 7b) \begin{bmatrix} 5\cos(3t) + 7\sin(3t) \\ -2\cos(3t) - 3\sin(3t) \end{bmatrix} + (2a + 5b) \begin{bmatrix} -7\cos(3t) + 5\sin(3t) \\ 3\cos(3t) - 2\sin(3t) \end{bmatrix} \Big) e^{-2t}.$$

◁

We end with case (iii). There are no many possibilities left for a 2×2 real matrix that both is diagonalizable and has a repeated eigenvalue. Such matrix must be proportional to the identity matrix.

Theorem 5.1.7. *Every 2×2 diagonalizable matrix with repeated eigenvalue λ_0 has the form*

$$A = \lambda_0 I.$$

Proof of Theorem 5.1.7: Since matrix A diagonalizable, there exists a matrix P invertible such that $A = PDP^{-1}$. Since A is 2×2 with a repeated eigenvalue λ_0, then

$$D = \begin{bmatrix} \lambda_0 & 0 \\ 0 & \lambda_0 \end{bmatrix} = \lambda_0 \, I.$$

Put these two facts together,

$$A = PDP^{-1} = P\lambda_0 I P^{-1} = \lambda_0 P \, P^{-1} = \lambda_0 I \quad \Rightarrow \quad A = \begin{bmatrix} \lambda_0 & 0 \\ 0 & \lambda_0 \end{bmatrix}.$$

□

Remark: The general solution $\boldsymbol{x}$ for $\boldsymbol{x}' = \lambda I \, \boldsymbol{x}$ is simple to write. Since any non-zero 2-vector is an eigenvector of $\lambda_0 I$, we choose the linearly independent set

$$\Big\{ \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Big\}.$$

Using these eigenvectors we can write the general solution,

$$\boldsymbol{x}(t) = c_1 \, e^{\lambda_0 t} \, \boldsymbol{v}_1 + c_2 \, e^{\lambda_0 t} \, \boldsymbol{v}_2 = c_1 \, e^{\lambda_0 t} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \, e^{\lambda_0 t} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \boldsymbol{x}(t) = e^{\lambda t} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

5.1.5. Non-Diagonalizable Systems.

A 2×2 linear systems might not be diagonalizable. This can happen only when the coefficient matrix has a repeated eigenvalue and all eigenvectors are proportional to each other. If we denote by λ the repeated eigenvalue of a 2×2 matrix A, and by $\boldsymbol{v}$ an associated eigenvector, then one solution to the differential system $\boldsymbol{x}' = A \, \boldsymbol{x}$ is

$$\boldsymbol{x}_1(t) = e^{\lambda t} \, \boldsymbol{v}.$$

Every other eigenvector $\tilde{\boldsymbol{v}}$ associated with λ is proportional to $\boldsymbol{v}$. So any solution of the form $\tilde{\boldsymbol{v}} e^{\lambda t}$ is proportional to the solution above. The next result provides a linearly independent set of two solutions to the system $\boldsymbol{x}' = A\boldsymbol{x}$ associated with the repeated eigenvalue λ.

Theorem 5.1.8 (Repeated Eigenvalue)**.** *If an 2×2 matrix A has a repeated eigenvalue λ with only one associated eigen-direction, given by the eigenvector $\boldsymbol{v}$, then the differential system $\boldsymbol{x}'(t) = A\,\boldsymbol{x}(t)$ has a linearly independent set of solutions*

$$\big\{\boldsymbol{x}_1(t) = e^{\lambda t}\,\boldsymbol{v}, \quad \boldsymbol{x}_2(t) = e^{\lambda t}\,\big(\boldsymbol{v}\,t + \boldsymbol{w}\big)\big\},$$

where the vector $\boldsymbol{w}$ is one of infinitely many solutions of the algebraic linear system

$$(A - \lambda I)\boldsymbol{w} = \boldsymbol{v}. \tag{5.1.18}$$

Remark: The eigenvalue λ is the precise number that makes matrix $(A-\lambda I)$ not invertible, that is, $\det(A-\lambda I) = 0$. This implies that an algebraic linear system with coefficient matrix $(A - \lambda I)$ may or may not have solutions, depending on what the source vector is. The Theorem above says that Eq. (5.1.18) has solutions when the source vector is $\boldsymbol{v}$. The reason that this system has solutions is that $\boldsymbol{v}$ is an eigenvector of A.

We give two proofs of this theorem. We start with a verification proof, that is, we show that the two functions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, given in the theorem are in fact fundamental solutions of the differential equation.

Proof of Theorem 5.1.8: We already know that $\boldsymbol{x}_1(t) = e^{\lambda t}\boldsymbol{v}$ is solution of $\boldsymbol{x}' = A\boldsymbol{x}$, because of the eigenpair equation $A\boldsymbol{v} = \lambda\boldsymbol{v}$. Indeed,

$$\boldsymbol{x}_1' = (e^{\lambda t})'\,\boldsymbol{v} = \lambda\,e^{\lambda t}\boldsymbol{v} = \lambda\boldsymbol{x}_1, \quad \text{and} \quad A\boldsymbol{x}_1 = e^{\lambda t}A\boldsymbol{v} = e^{\lambda t}\lambda\boldsymbol{v} = \lambda\boldsymbol{x}_1,$$

which says that $\boldsymbol{x}_1' = A\boldsymbol{x}_1$. Now we do a similar calculation for $\boldsymbol{x}_2$. On the one hand we have

$$\boldsymbol{x}_2' = \big(e^{\lambda t}(t\boldsymbol{v} + \boldsymbol{w})\big)' = (e^{\lambda t}t)'\boldsymbol{v} + (e^{\lambda t})'\boldsymbol{w} = e^{\lambda t}\boldsymbol{v} + \lambda e^{\lambda t}t\boldsymbol{v} + \lambda e^{\lambda t}\boldsymbol{w} = e^{\lambda t}\boldsymbol{v} + \lambda\boldsymbol{x}_2.$$

On the other hand,

$$A\boldsymbol{x}_2 = e^{\lambda t}tA\boldsymbol{v} + e^{\lambda t}A\boldsymbol{w} = e^{\lambda t}t\lambda\boldsymbol{v} + e^{\lambda t}A\boldsymbol{w} = \lambda\big(e^{\lambda t}t\boldsymbol{v} + e^{\lambda t}\boldsymbol{w} - e^{\lambda t}\boldsymbol{w}\big) + e^{\lambda t}A\boldsymbol{w},$$

which means,

$$A\boldsymbol{x}_2 = \lambda\boldsymbol{x}_2 + e^{\lambda t}(A\boldsymbol{w} - \lambda\boldsymbol{w}).$$

Therefore, $\boldsymbol{x}_2' = A\boldsymbol{x}_2$ if and only if

$$e^{\lambda t}\boldsymbol{v} + \lambda\boldsymbol{x}_2 = \lambda\boldsymbol{x}_2 + e^{\lambda t}(A\boldsymbol{w} - \lambda\boldsymbol{w}) \quad \Leftrightarrow \quad (A - \lambda I)\boldsymbol{w} = \boldsymbol{v}.$$

So, $\boldsymbol{x}_2$ is solution of $\boldsymbol{x}' = A\boldsymbol{x}$ if and only if the vector $\boldsymbol{w}$ is solution of $(A - \lambda I)\boldsymbol{w} = \boldsymbol{v}$. This establishes the Theorem. □

We now give a constructive proof of the same theorem, that is, we find the formula for the second fundamental solution x_2 using a generalization of the reduction of order method.

Proof of Theorem 5.1.8: One solution to the differential system is $\boldsymbol{x}_1(t) = e^{\lambda t}\,\boldsymbol{v}$. Inspired by the reduction order method we look for a second solution of the form

$$\boldsymbol{x}_2(t) = e^{\lambda t}\,\boldsymbol{u}(t).$$

Inserting this function into the differential equation $\boldsymbol{x}' = A\,\boldsymbol{x}$ we get

$$\boldsymbol{u}' + \lambda\,\boldsymbol{u} = A\,\boldsymbol{u} \quad \Rightarrow \quad (A - \lambda I)\,\boldsymbol{u} = \boldsymbol{u}'.$$

We now introduce a power series expansion of the vector-valued function $\boldsymbol{u}$,

$$\boldsymbol{u}(t) = \boldsymbol{u}_0 + \boldsymbol{u}_1 t + \boldsymbol{u}_2 t^2 + \cdots,$$

into the differential equation above,

$$(A - \lambda I)(\boldsymbol{u}_0 + \boldsymbol{u}_1 t + \boldsymbol{u}_2 t^2 + \cdots) = (\boldsymbol{u}_1 + 2\boldsymbol{u}_2 t + \cdots).$$

If we evaluate the equation above at $t = 0$, and then its derivative at $t = 0$, and so on, we get the following infinite set of linear algebraic equations

$$\begin{aligned}(A - \lambda I)\boldsymbol{u}_0 &= \boldsymbol{u}_1,\\ (A - \lambda I)\boldsymbol{u}_1 &= 2\boldsymbol{u}_2,\\ (A - \lambda I)\boldsymbol{u}_2 &= 3\boldsymbol{u}_3\\ &\vdots\end{aligned}$$

Here is where we use Cayley-Hamilton's Theorem. Recall that the characteristic polynomial $p(\tilde{\lambda}) = \det(A - \tilde{\lambda} I)$ has the form

$$p(\tilde{\lambda}) = \tilde{\lambda}^2 - \operatorname{tr}(A)\,\tilde{\lambda} + \det(A).$$

Cayley-Hamilton Theorem says that the matrix-valued polynomial $p(A) = 0$, that is,

$$A^2 - \operatorname{tr}(A)\,A + \det(A)\,I = 0.$$

Since in the case we are interested in matrix A has a repeated root λ, then

$$p(\tilde{\lambda}) = (\tilde{\lambda} - \lambda)^2 = \tilde{\lambda}^2 - 2\lambda\,\tilde{\lambda} + \lambda^2.$$

Therefore, Cayley-Hamilton Theorem for the matrix in this Theorem has the form

$$0 = A^2 - 2\lambda\,A + \lambda^2\,I \quad\Rightarrow\quad (A - \lambda I)^2 = 0.$$

This last equation is the one we need to solve the system for the vector-valued $\boldsymbol{u}$. Multiply the first equation in the system by $(A - \lambda I)$ and use that $(A - \lambda I)^2 = 0$, then we get

$$\boldsymbol{0} = (A - \lambda I)^2\boldsymbol{u}_0 = (A - \lambda I)\,\boldsymbol{u}_1 \quad\Rightarrow\quad (A - \lambda I)\boldsymbol{u}_1 = \boldsymbol{0}.$$

This implies that $\boldsymbol{u}_1$ is an eigenvector of A with eigenvalue λ. We can denote it as $\boldsymbol{u}_1 = \boldsymbol{v}$. Using this information in the rest of the system we get

$$\begin{aligned}(A - \lambda I)\boldsymbol{u}_0 &= \boldsymbol{v},\\ (A - \lambda I)\boldsymbol{v} &= 2\boldsymbol{u}_2 \quad\Rightarrow\quad \boldsymbol{u}_2 = \boldsymbol{0},\\ (A - \lambda I)\boldsymbol{u}_2 &= 3\boldsymbol{u}_3 \quad\Rightarrow\quad \boldsymbol{u}_3 = \boldsymbol{0},\\ &\vdots\end{aligned}$$

We conclude that all terms $\boldsymbol{u}_2 = \boldsymbol{u}_3 = \cdots = \boldsymbol{0}$. Denoting $\boldsymbol{u}_0 = \boldsymbol{w}$ we obtain the following system of algebraic equations,

$$\begin{aligned}(A - \lambda I)\boldsymbol{w} &= \boldsymbol{v},\\ (A - \lambda I)\boldsymbol{v} &= \boldsymbol{0}.\end{aligned}$$

For vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ solution of the system above we get $\boldsymbol{u}(t) = \boldsymbol{w} + t\boldsymbol{v}$. This means that the second solution to the differential equation is

$$\boldsymbol{x}_2(t) = e^{\lambda t}\,(t\boldsymbol{v} + \boldsymbol{w}).$$

This establishes the Theorem. □

Example 5.1.12. Find the fundamental solutions of the differential equation

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad A = \frac{1}{4}\begin{bmatrix} -6 & 4\\ -1 & -2\end{bmatrix}.$$

Solution: As usual, we start finding the eigenvalues and eigenvectors of matrix A. The former are the solutions of the characteristic equation

$$0 = \begin{vmatrix} \left(-\frac{3}{2} - \lambda\right) & 1 \\ -\frac{1}{4} & \left(-\frac{1}{2} - \lambda\right) \end{vmatrix} = \Big(\lambda + \frac{3}{2}\Big)\Big(\lambda + \frac{1}{2}\Big) + \frac{1}{4} = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2.$$

Therefore, there solution is the repeated eigenvalue $\lambda = -1$. The associated eigenvectors are the vectors $\boldsymbol{v}$ solution to the linear system $(A + I)\boldsymbol{v} = \mathbf{0}$,

$$\begin{bmatrix} \left(-\frac{3}{2} + 1\right) & 1 \\ -\frac{1}{4} & \left(-\frac{1}{2} + 1\right) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & 1 \\ -\frac{1}{4} & \frac{1}{2} \end{bmatrix} \to \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} \quad \Rightarrow \quad v_1 = 2v_2.$$

Choosing $v_2 = 1$, then $v_1 = 2$, and we obtain

$$\lambda = -1, \qquad \boldsymbol{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Any other eigenvector associated to $\lambda = -1$ is proportional to the eigenvector above. The matrix A above is not diagonalizable. So. we follow Theorem 5.1.8 and we solve for a vector $\boldsymbol{w}$ the linear system $(A + I)\boldsymbol{w} = \boldsymbol{v}$. The augmented matrix for this system is given by,

$$\left[\begin{array}{cc|c} -\frac{1}{2} & 1 & 2 \\ -\frac{1}{4} & \frac{1}{2} & 1 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & -2 & -4 \\ 1 & -2 & -4 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & -2 & -4 \\ 0 & 0 & 0 \end{array}\right] \quad \Rightarrow \quad w_1 = 2w_2 - 4.$$

We have obtained infinitely many solutions given by

$$\boldsymbol{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} w_2 + \begin{bmatrix} -4 \\ 0 \end{bmatrix}.$$

As one could have imagined, given any solution $\boldsymbol{w}$, the $c\boldsymbol{v} + \boldsymbol{w}$ is also a solution for any $c \in \mathbb{R}$. We choose the simplest solution given by

$$\boldsymbol{w} = \begin{bmatrix} -4 \\ 0 \end{bmatrix}.$$

Therefore, a fundamental set of solutions to the differential equation above is formed by

$$\boldsymbol{x}_1(t) = e^{-t} \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \qquad \boldsymbol{x}_2(t) = e^{-t} \Big(t \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} -4 \\ 0 \end{bmatrix} \Big). \tag{5.1.19}$$

◁

5.1.6. Exercises.

5.1.1.- .

5.1.2.- .

5.2. Two-Dimensional Phase portraits

Figures are easier to understand than words. Words are easier to understand than equations. The qualitative behavior of a function is often simpler to visualize from a graph than from an explicit or implicit expression of the function. In this section we show graphical representation of the solutions found in the previous section , which are solutions of 2×2 linear systems of differential equations. We focus on phase portraits of these solutions and we characterize the trivial solution $\boldsymbol{x} = \boldsymbol{0}$ as stable, unstable, saddle, or center.

5.2.1. Review of Solutions Formulas.

In the previous section we found explicit formulas for the solutions of 2×2 linear homogeneous systems of differential equations with constant coefficients. We summarize the results from the previous section in the following theorem.

Theorem 5.2.1. *A pair of fundamental solutions of a 2×2 system $\boldsymbol{x}' = A\,\boldsymbol{x}$, where A is a constant matrix, depend on the eigenpairs of A, say $\lambda_\pm$, $\boldsymbol{v}_\pm$, as follows.*

(a) *If $\lambda_+ \neq \lambda_-$ and real, then A is diagonalizable and a pair of fundamental solutions is*

$$\boldsymbol{x}_+ = \boldsymbol{v}_+\, e^{\lambda_+ t}, \qquad \boldsymbol{x}_- = \boldsymbol{v}_-\, e^{\lambda_- t},$$

(b) *If $\lambda_{pm} = \alpha \pm \beta i$ and $\boldsymbol{v}_\pm = \boldsymbol{a} \pm \boldsymbol{b} i$, then A is diagonalizable and fundamental solutions is*

$$\boldsymbol{x}_1 = \big(\boldsymbol{a}\cos(\beta t) - \boldsymbol{b}\sin(\beta t)\big)\, e^{\alpha t},$$
$$\boldsymbol{x}_2 = \big(\boldsymbol{a}\sin(\beta t) + \boldsymbol{b}\cos(\beta t)\big)\, e^{\alpha t}.$$

(c) *If $\lambda_+ = \lambda_- = \lambda_0$ and A is diagonalizable, then $A = \lambda_0 I$ and a pair of fundamental solutions is*

$$\boldsymbol{x}_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{\lambda_0 t}, \qquad \boldsymbol{x}_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{\lambda_0 t},$$

(d) *If $\lambda_+ = \lambda_- = \lambda_0$ and A is **not** diagonalizable, then a pair of fundamental solutions is*

$$\boldsymbol{x}_+ = \boldsymbol{v}\, e^{\lambda_0 t}, \qquad \boldsymbol{x}_- = (t\,\boldsymbol{v} + \boldsymbol{w})\, e^{\lambda_0 t},$$

where

$$(A - \lambda_0 I)\boldsymbol{v} = \boldsymbol{0}, \qquad (A - \lambda_0 I)\boldsymbol{w} = \boldsymbol{v}.$$

This theorem has been proven in the previous section. We see that the formulas for fundamental solutions of a 2×2 system $\boldsymbol{x}' = A\boldsymbol{x}$ change considerably depending on the eigenpairs of the coefficient matrix A. The main reason for this property of the solutions is that—unlike the eigenvalues, which depend continuously on the matrix coefficients—the eigenvectors do not depend continuously on the matrix coefficients. This means that two matrices with very similar coefficients will have very similar eigenvalues but they might have very different eigenvectors.

We are interested in graphical representations of solutions to 2×2 systems of differential equations. One possible graphical representation of a solution vector

$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

is to graph each component, x_1 and x_2, as functions of t. We have seen another way to graph a 2-vector-valued function. We plot the whole vector $\boldsymbol{x}(t)$ at t on the plane x_1, x_2. Each vector $\boldsymbol{x}(t)$ is represented by its end point, while the whole solution $\boldsymbol{x}$ is represented by a line with arrows pointing in the direction of increasing t. Such a figure is called a *phase portrait* or *phase diagram* of a solution.

Phase portraits have a simple physical interpretation in the case that the solution vector $\boldsymbol{x}(t)$ is the position function of a particle moving in a plane at the time t. In this case the phase portrait is the trajectory of the particle. The arrows added to this trajectory indicate the motion of the particle as time increases.

5.2.2. Real Distinct Eigenvalues. We now focus on solutions to the system $\boldsymbol{x}' = A\boldsymbol{x}$, in the case that matrix A has two real eigenvalues $\lambda_+ \neq \lambda_-$. We study the case where both eigenvalues are non-zero. (The case where one eigenvalue vanishes is left as an exercise.) So, the eigenvalues belong to one of the following cases:

(i) $\lambda_+ > \lambda_- > 0$, both eigenvalues positive;
(ii) $\lambda_+ > 0 > \lambda_-$, one eigenvalue negative and the other positive;
(iii) $0 > \lambda_+ > \lambda_-$, both eigenvalues negative.

The phase portrait of several solutions $\boldsymbol{x}(t)$ can be displayed in the same picture. If the fundamental solutions are $\boldsymbol{x}_+$ and $\boldsymbol{x}_-$, the any solution is given by

$$\boldsymbol{x} = c_+\,\boldsymbol{x}_+ + c_-\,\boldsymbol{x}_-.$$

We indicate what solution we are plotting by specifying the values for the constants c_+ and c_-. A phase diagram can be sketched by following these few steps.

(a) Plot the eigenvectors $\boldsymbol{v}^+$ and $\boldsymbol{v}^-$ corresponding to the eigenvalues λ_+ and λ_-.
(b) Draw the whole lines parallel to these vectors and passing through the origin. These straight lines correspond to solutions with either c_+ or c_- zero.
(c) Draw arrows on these lines to indicate how the solution changes as the variable t increases. If t is interpreted as time, the arrows indicate how the solution changes into the future. The arrows point towards the origin if the corresponding eigenvalue λ is negative, and they point away form the origin if the eigenvalue is positive.
(d) Find the non-straight curves correspond to solutions with both coefficient c_+ and c_- non-zero. Again, arrows on these curves indicate the how the solution moves into the future.

Case $\lambda_+ > \lambda_- > 0$, Source, (Unstable Point).

Example 5.2.1. Sketch the phase diagram of the solutions to the differential equation

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad A = \frac{1}{4}\begin{bmatrix} 11 & 3 \\ 1 & 9 \end{bmatrix}. \tag{5.2.1}$$

Solution: The characteristic equation for this matrix A is given by

$$\det(A - \lambda I) = \lambda^2 - 5\lambda + 6 = 0 \quad \Rightarrow \quad \begin{cases} \lambda_+ = 3, \\ \lambda_- = 2. \end{cases}$$

One can show that the corresponding eigenvectors are given by

$$\boldsymbol{v}^+ = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \qquad \boldsymbol{v}^- = \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

So the general solution to the differential equation above is given by

$$\boldsymbol{x}(t) = c_+\,\boldsymbol{v}^+ e^{\lambda_+ t} + c_-\,\boldsymbol{v}^- e^{\lambda_- t} \quad \Leftrightarrow \quad \boldsymbol{x}(t) = c_+ \begin{bmatrix} 3 \\ 1 \end{bmatrix} e^{3t} + c_- \begin{bmatrix} -2 \\ 2 \end{bmatrix} e^{2t}.$$

In Fig. 1 we have sketched four curves, each representing a solution $\boldsymbol{x}$ corresponding to a particular choice of the constants c_+ and c_-. These curves actually represent eight different solutions, for eight different choices of the constants c_+ and c_-, as is described below. The arrows on these curves represent the change in the solution as the variable t grows. Since both eigenvalues are positive, the length of the solution vector always increases as t increases. The straight lines correspond to the following four solutions:

$$c_+ = 1,\ c_- = 0; \quad c_+ = 0,\ c_- = 1; \quad c_+ = -1,\ c_- = 0; \quad c_+ = 0,\ c_- = -1.$$

The curved lines on each quadrant correspond to the following four solutions:

$$c_+ = 1,\ c_- = 1; \quad c_+ = 1,\ c_- = -1; \quad c_+ = -1,\ c_- = 1; \quad c_+ = -1,\ c_- = -1.$$

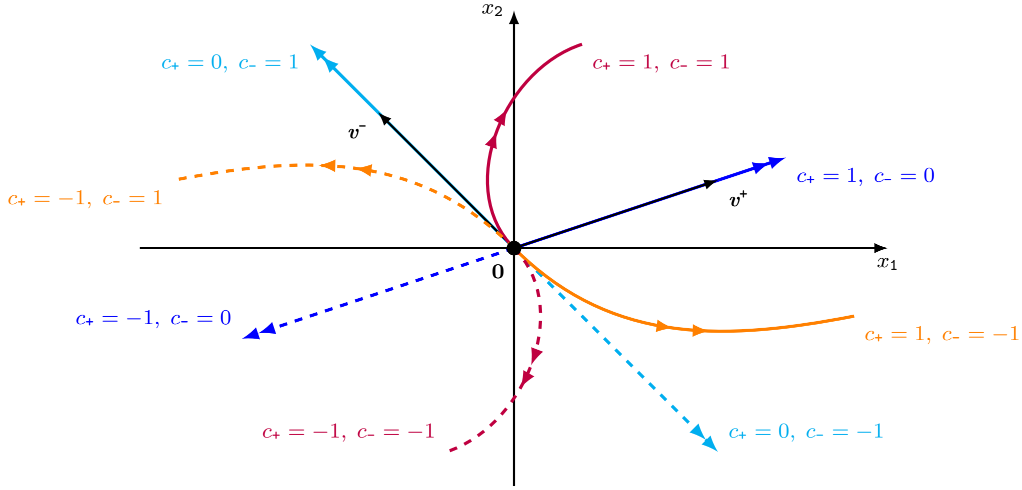


FIGURE 1. Eight solutions to Eq. (5.2.1), where $\lambda_+ > \lambda_- > 0$. The trivial solution $\boldsymbol{x} = \mathbf{0}$ is called a **source**, or an **unstable point**.

◁

Case $\lambda_+ > 0 > \lambda_-$, Saddle, (Unstable Point).

Example 5.2.2. Sketch the phase diagram of the solutions to the differential equation

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}. \tag{5.2.2}$$

Solution: In a previous section we have computed the eigenvalues and eigenvectors of this coefficient matrix, and the result is

$$\lambda_+ = 4, \quad \boldsymbol{v}^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \lambda_- = -2, \quad \boldsymbol{v}^- = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Therefore, all solutions of the differential equation above are

$$\boldsymbol{x}(t) = c_+ \boldsymbol{v}^+ e^{\lambda_+ t} + c_- \boldsymbol{v}^- e^{\lambda_- t} \quad \Leftrightarrow \quad \boldsymbol{x}(t) = c_+ \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{4t} + c_- \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t},$$

In Fig. 2 we have sketched four curves, each representing a solution $\boldsymbol{x}$ corresponding to a particular choice of the constants c_+ and c_-. These curves actually represent eight different solutions, for eight different choices of the constants c_+ and c_-, as is described below. The arrows on these curves represent the change in the solution as the variable t grows. The part of the solution with positive eigenvalue increases exponentially when t grows, while the part of the solution with negative eigenvalue decreases exponentially when t grows. The straight lines correspond to the following four solutions:

$$c_+ = 1,\ c_- = 0; \quad c_+ = 0,\ c_- = 1; \quad c_+ = -1,\ c_- = 0; \quad c_+ = 0,\ c_- = -1.$$

The curved lines on each quadrant correspond to the following four solutions:

$$c_+ = 1,\ c_- = 1; \quad c_+ = 1,\ c_- = -1; \quad c_+ = -1,\ c_- = 1; \quad c_+ = -1,\ c_- = -1.$$

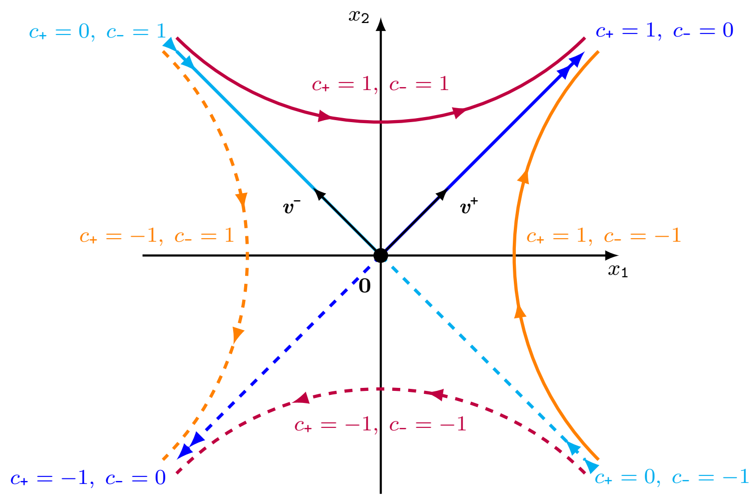


FIGURE 2. Several solutions to Eq. (5.2.2), $\lambda_+ > 0 > \lambda_-$. The trivial solution $\boldsymbol{x} = \mathbf{0}$ is called a **saddle** or an **unstable point**.

◁

Case $0 > \lambda_+ > \lambda_-$**, Sink, (Stable Point).**

Example 5.2.3. Sketch the phase diagram of the solutions to the differential equation

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad A = \frac{1}{4}\begin{bmatrix} -9 & 3 \\ 1 & -11 \end{bmatrix}. \tag{5.2.3}$$

Solution: The characteristic equation for this matrix A is given by

$$\det(A - \lambda I) = \lambda^2 + 5\lambda + 6 = 0 \quad \Rightarrow \quad \begin{cases} \lambda_+ = -2, \\ \lambda_- = -3. \end{cases}$$

One can show that the corresponding eigenvectors are given by

$$\boldsymbol{v}^+ = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \qquad \boldsymbol{v}^- = \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

So the general solution to the differential equation above is given by

$$\boldsymbol{x}(t) = c_+\, \boldsymbol{v}^+ e^{\lambda_+ t} + c_-\, \boldsymbol{v}^- e^{\lambda_- t} \quad \Leftrightarrow \quad \boldsymbol{x}(t) = c_+ \begin{bmatrix} 3 \\ 1 \end{bmatrix} e^{-2t} + c_- \begin{bmatrix} -2 \\ 2 \end{bmatrix} e^{-3t}.$$

In Fig. 3 we have sketched four curves, each representing a solution $\boldsymbol{x}$ corresponding to a particular choice of the constants c_+ and c_-. These curves actually represent eight different solutions, for eight different choices of the constants c_+ and c_-, as is described below. The arrows on these curves represent the change in the solution as the variable t grows.

Since both eigenvalues are negative, the length of the solution vector always decreases as t grows and the solution vector always approaches zero. The straight lines correspond to the following four solutions:

$$c_+ = 1,\ c_- = 0; \quad c_+ = 0,\ c_- = 1; \quad c_+ = -1,\ c_- = 0; \quad c_+ = 0,\ c_- = -1.$$

The curved lines on each quadrant correspond to the following four solutions:

$$c_+ = 1,\ c_- = 1; \quad c_+ = 1,\ c_- = -1; \quad c_+ = -1,\ c_- = 1; \quad c_+ = -1,\ c_- = -1.$$

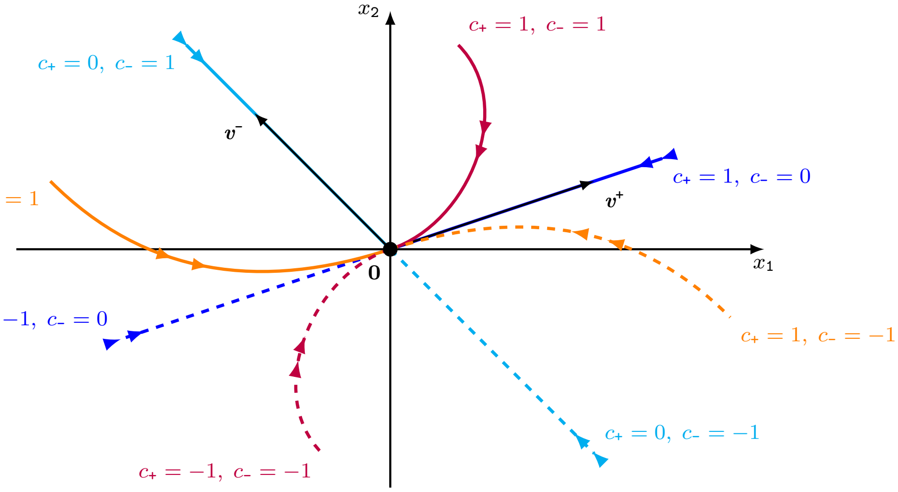


FIGURE 3. Several solutions to Eq. (5.2.3), where $0 > \lambda_+ > \lambda_-$. The trivial solution $\boldsymbol{x} = \boldsymbol{0}$ is called a **sink** or a **stable point**.

◁

5.2.3. Complex Eigenvalues. A real-valued matrix may have complex-valued eigenvalues. These complex eigenvalues come in pairs, because the matrix is real-valued. If λ is one of these complex eigenvalues, then $\overline{\lambda}$ is also an eigenvalue. A usual notation is $\lambda_\pm = \alpha \pm i\beta$, with $\alpha,\ \beta \in \mathbb{R}$. The same happens with their eigenvectors, which are written as $\boldsymbol{v}_\pm = \boldsymbol{a} \pm i\boldsymbol{b}$, with $\boldsymbol{a},\ \boldsymbol{b} \in \mathbb{R}^2$. When the matrix is the coefficient matrix of a differential equation,

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

the solutions $\boldsymbol{x}_+(t) = \boldsymbol{v}^+ e^{\lambda_+ t}$ and $\boldsymbol{x}_-(t) = \boldsymbol{v}^- e^{\lambda_- t}$ are complex-valued. We know that real-valued fundamental solutions can be constructed with the real part and the imaginary part of the solution $\boldsymbol{x}_+$. The resulting formulas are

$$\boldsymbol{x}_1(t) = \big(\boldsymbol{a}\,\cos(\beta t) - \boldsymbol{b}\,\sin(\beta t)\big)\,e^{\alpha t}, \qquad \boldsymbol{x}_2(t) = \big(\boldsymbol{a}\,\sin(\beta t) + \boldsymbol{b}\,\cos(\beta t)\big)\,e^{\alpha t}. \tag{5.2.4}$$

These real-valued solutions are used to draw the phase portraits. We recall how to obtain real-valued fundamental solutions in the following example.

Example 5.2.4. Find a real-valued set of fundamental solutions to the differential equation below and sketch a phase portrait, where

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad A = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}.$$

Solution: We know from a previous section that the eigenpairs of the coefficient matrix are

$$\lambda_{\pm} = 2 \pm 3i, \quad \boldsymbol{v}^{\pm} = \begin{bmatrix} \mp i \\ 1 \end{bmatrix}.$$

Writing them in real and imaginary parts, $\lambda_{\pm} = \alpha \pm i\beta$ and $\boldsymbol{v}_{\pm} = \boldsymbol{a} \pm i\boldsymbol{b}$, we get

$$\alpha = 2, \qquad \beta = 3, \qquad \boldsymbol{a} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

These eigenvalues and eigenvectors imply the following real-valued fundamental solutions,

$$\Big\{ \boldsymbol{x}_1(t) = \begin{bmatrix} \sin(3t) \\ \cos(3t) \end{bmatrix} e^{2t}, \ \boldsymbol{x}_2(t) = \begin{bmatrix} -\cos(3t) \\ \sin(3t) \end{bmatrix} e^{2t} \Big\}. \tag{5.2.5}$$

Before plotting the phase portrait of these solutions it is useful to plot the phase portrait of the following functions

$$\tilde{\boldsymbol{x}}_1(t) = \begin{bmatrix} \sin(3t) \\ \cos(3t) \end{bmatrix} \qquad \tilde{\boldsymbol{x}}_2(t) = \begin{bmatrix} -\cos(3t) \\ \sin(3t) \end{bmatrix}.$$

These functions $\tilde{\boldsymbol{x}}_1$ and $\tilde{\boldsymbol{x}}_2$ are **not** solutions of the differential equation. But they are simple to plot, because both vector functions have length one for all t,

$$\|\tilde{\boldsymbol{x}}_1(t)\| = \sqrt{\sin^2(3t) + \cos^2(3t)} = 1, \quad \|\tilde{\boldsymbol{x}}_2(t)\| = \sqrt{\cos^2(3t) + \sin^2(3t)} = 1.$$

Therefore, one can see that these functions describe a circle in the plane x_1x_2, represented by the dashed line in Fig. 4. Once we know that, we realize that the fundamental solutions of the differential equation can be written as

$$\boldsymbol{x}_1(t) = \tilde{\boldsymbol{x}}_1(t)\, e^{2t}, \quad \boldsymbol{x}_2(t) = \tilde{\boldsymbol{x}}_2(t)\, e^{2t}.$$

Therefore, the solutions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ must be spirals going away from the origin as t increases, see Fig. 4. ◁

In general, when a coefficient matrix A has complex eigenpairs,

$$\lambda_{\pm} = \alpha \pm i\beta, \qquad \boldsymbol{v}^{\pm} = \boldsymbol{a} \pm i\boldsymbol{b},$$

real-valued fundamental solutions of the differential equation $\boldsymbol{x}]' = A\boldsymbol{x}$ are given by

$$\boldsymbol{x}_1(t) = \big(\boldsymbol{a}\cos(\beta t) - \boldsymbol{b}\sin(\beta t)\big)\, e^{\alpha t}, \qquad \boldsymbol{x}_2(t) = \big(\boldsymbol{a}\sin(\beta t) + \boldsymbol{b}\cos(\beta t)\big)\, e^{\alpha t}.$$

The phase portrait of these solutions can be done as follows. First plot the vectors $\boldsymbol{a}$, $\boldsymbol{b}$. Then plot the auxiliary functions

$$\tilde{\boldsymbol{x}}_1(t) = \big(\boldsymbol{a}\cos(\beta t) - \boldsymbol{b}\sin(\beta t)\big), \qquad \tilde{\boldsymbol{x}}_2(t) = \big(\boldsymbol{a}\sin(\beta t) + \boldsymbol{b}\cos(\beta t)\big).$$

It can be shown that the phase portrait of both auxiliary functions is the same ellipse. In the case the vector $\boldsymbol{a}$ is perpendicular to $\boldsymbol{b}$ this ellipse has main axes given by the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. These ellipses are plotted below in dashed lines. Since

$$\boldsymbol{x}_1(t) = \tilde{\boldsymbol{x}}_1(t)\, e^{\alpha t}, \quad \boldsymbol{x}_2(t) = \tilde{\boldsymbol{x}}_2(t)\, e^{\alpha t},$$

the phase portrait of the solutions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are going to be spirals. As t increases we see that the solution spirals out for $\alpha > 0$, stays in the ellipse for $\alpha = 0$, and spirals into the origin for $\alpha < 0$.

We now choose arbitrary vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ and we sketch phase portraits of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ for a few choices of α. The result is given in Fig. 5.

We now make a different choice for the vector $\boldsymbol{b}$, and we repeat the three phase portraits given above; for $\alpha > 0$, $\alpha = 0$, and $\alpha < 0$. The result is given in Fig. 6. Comparing Figs. 5 and 6 shows that the relative directions of the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ determines the rotation direction of the solutions as t increases.

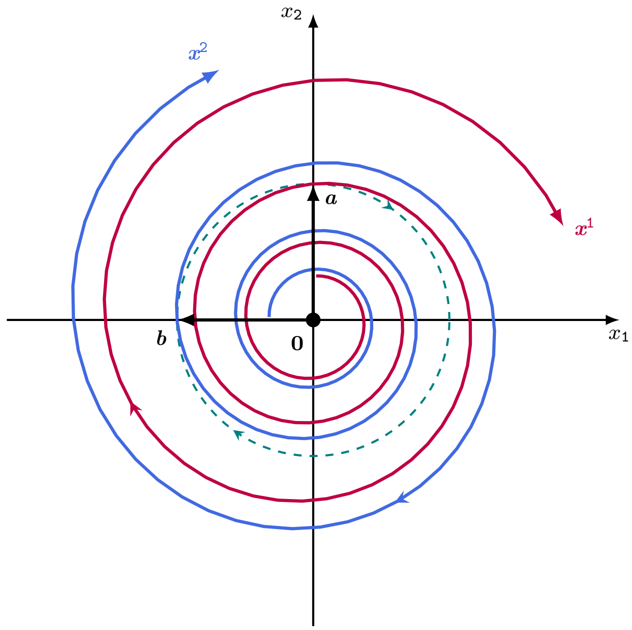


FIGURE 4. The graph of the fundamental solutions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in Eq. (5.2.5). The trivial solution $\boldsymbol{x} = \boldsymbol{0}$ is an **unstable spiral**.

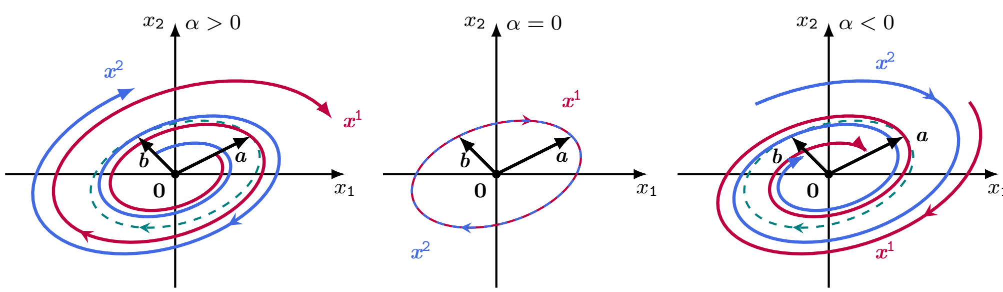


FIGURE 5. Fundamental solutions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in Eq. (5.2.4) for $\alpha > 0$, $\alpha = 0$, and $\alpha < 0$. The relative positions of $\boldsymbol{a}$ and $\boldsymbol{b}$ determines the rotation direction. Compare with Fig. 6. The trivial solution $\boldsymbol{x} = \boldsymbol{0}$ is called an **unstable spiral** for $\alpha > 0$, a **center** for $\alpha = 0$, and a **stable spiral** for $\alpha < 0$.

5.2.4. Repeated Eigenvalues. A matrix with repeated eigenvalues may or may not be diagonalizable. If a 2×2 matrix A is diagonalizable with repeated eigenvalues, then we know that this matrix is proportional to the identity matrix, $A = \lambda_0 I$, with λ_0 the repeated eigenvalue. We saw in a previous section that the general solution of a differential system with such coefficient matrix is

$$\boldsymbol{x}(t) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} e^{\lambda_0 t}.$$

Phase portraits of these solutions are just straight lines, starting from the origin for $\lambda_0 > 0$, or ending at the origin for $\lambda_0 < 0$.

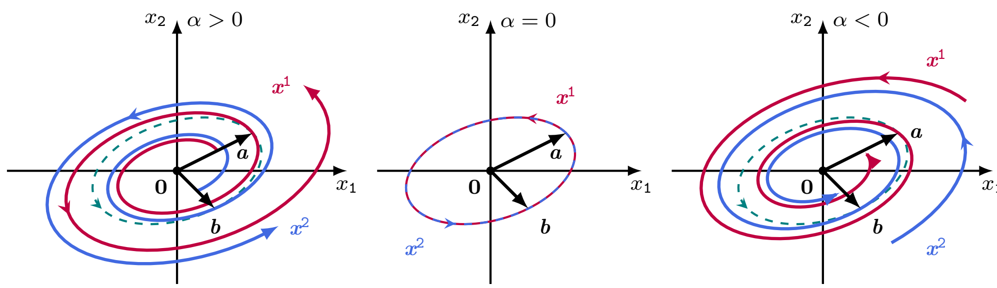


FIGURE 6. Fundamental solutions $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ in Eq. (5.2.4) for $\alpha > 0$, $\alpha = 0$, and $\alpha < 0$. The relative positions of $\boldsymbol{a}$ and $\boldsymbol{b}$ determines the rotation direction. Compare with Fig. 5. The trivial solution $\boldsymbol{x} = \boldsymbol{0}$ is called an **unstable spiral** for $\alpha > 0$, a **center** for $\alpha = 0$, and a **stable spiral** for $\alpha < 0$.

Non-diagonalizable 2×2 differential systems are more interesting. If $\boldsymbol{x}' = A\,\boldsymbol{x}$ is such a system, it has fundamental solutions

$$\boldsymbol{x}_1(t) = \boldsymbol{v}\, e^{\lambda_0 t}, \qquad \boldsymbol{x}_2(t) = (\boldsymbol{v}\, t + \boldsymbol{w})\, e^{\lambda_0 t}, \tag{5.2.6}$$

where λ_0 is the repeated eigenvalue of A with eigenvector $\boldsymbol{v}$, and vector $\boldsymbol{w}$ is any solution of the linear algebraic system

$$(A - \lambda_0 I)\boldsymbol{w} = \boldsymbol{v}.$$

The phase portrait of these fundamental solutions is given in Fig 7. To construct this figure start drawing the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$. The solution $\boldsymbol{x}_1$ is simpler to draw than $\boldsymbol{x}_2$, since the former is a straight semi-line starting at the origin and parallel to $\boldsymbol{v}$.

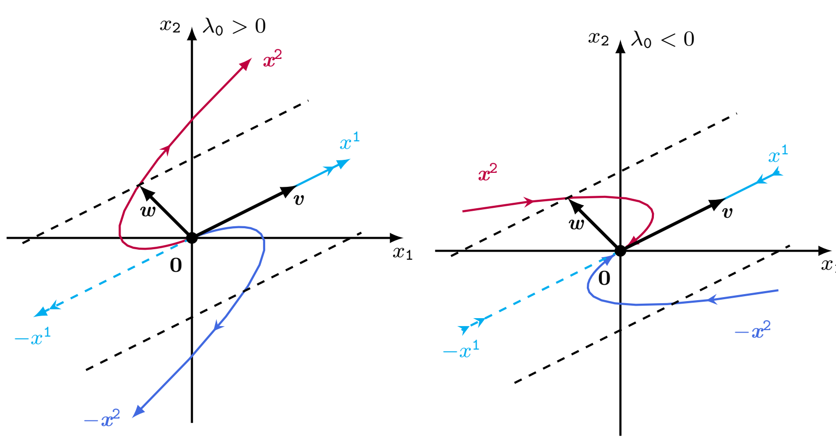


FIGURE 7. Functions $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ in Eq. (5.2.6) for the cases $\lambda_0 > 0$ and $\lambda_0 < 0$. The trivial solution $\boldsymbol{x}{=}\boldsymbol{0}$ is a **source** or an **unstable point** for $\lambda_0 > 0$, and a **sink** or a **stable point** for $\lambda_0 < 0$.

The solution $\boldsymbol{x}_2$ is more difficult to draw. One way is to first draw the trajectory of the auxiliary function

$$\tilde{\boldsymbol{x}}_2 = \boldsymbol{v}\, t + \boldsymbol{w}.$$

This is a straight line parallel to $\boldsymbol{v}$ passing through $\boldsymbol{w}$, one of the black dashed lines in Fig. 7, the one passing through $\boldsymbol{w}$. The solution $\boldsymbol{x}_2$ can be written as

$$\boldsymbol{x}_2(t) = \tilde{\boldsymbol{x}}_2(t)\, e^{\lambda_0 t}.$$

Consider the case $\lambda_0 > 0$. For $t > 0$ we have $\boldsymbol{x}_2(t) > \tilde{\boldsymbol{x}}_2(t)$, and the opposite happens for $t < 0$. In the limit $t \to -\infty$ the solution values $\boldsymbol{x}_2(t)$ approach the origin, since the exponential factor $e^{\lambda_0 t}$ decreases faster than the linear factor t increases. The result is the purple line in the first picture of Fig. 7. The other picture, for $\lambda_0 < 0$ can be constructed following similar ideas.

5.2.5. The Stability of Linear Systems. In the last part of this section we focus on a particular solution of the system $\boldsymbol{x}' = A\boldsymbol{x}$, the trivial solution $\boldsymbol{x}_0 = \mathbf{0}$. First notice that $\boldsymbol{x}_0 = \mathbf{0}$ is indeed a solution of the differential equation, since $\boldsymbol{x}_0' = \mathbf{0} = A\mathbf{0} = A\boldsymbol{x}_0$. Second, this is a solution that is t-independent, that is, a constant solution. Third, this solution is not very interesting in itself, that's why we have not mentioned it at all till now.

However, we are now interested in the behavior of solutions with initial data nearby the trivial solution $\boldsymbol{x}_0 = \mathbf{0}$. We introduce two main characterizations of the behavior of solutions with initial data nearby the trivial solution. The first characterization is not so precise as the second characterization. Because of that, the first characterization can be extended to more systems than the second characterization.

The first characterization is very general and can be extended to a great variety of systems, including $n \times n$ systems, and nonlinear systems.

Definition 5.2.2. *The solution $\boldsymbol{x}_0 = \mathbf{0}$ of a 2×2 linear system $\boldsymbol{x}' = A\boldsymbol{x}$ is:*

(1) ***stable*** *iff all solutions $\boldsymbol{x}(t)$ with $\boldsymbol{x}(0)$ sufficiently close to $\boldsymbol{x}_0 = \mathbf{0}$ remain close to it for all $t > 0$.*

(2) ***asymptotically stable*** *iff all solutions $\boldsymbol{x}(t)$ with $\boldsymbol{x}(0)$ sufficiently close to $\boldsymbol{x}_0 = \mathbf{0}$ satisfy that $\boldsymbol{x}(t) \to \boldsymbol{x}_0$ as $t \to \infty$.*

Otherwise, the trivial solution $\boldsymbol{x}_0$ is called ***unstable****.*

This characterization of the trivial solution is usually called the stability of the trivial solution. It is not difficult to relate the stability of the trivial solution with the sign of the eigenvalues of the coefficient matrix of the linear system.

Theorem 5.2.3 (Stability I)**.** *Let $\boldsymbol{x}(t)$ be the solution of a 2×2 linear system $\boldsymbol{x}' = A\boldsymbol{x}$, with $\det(A) \neq 0$ and initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_1$. Then the trivial solution $\mathbf{x}_0 = \mathbf{0}$ is:*

(1) stable if both eigenvalues of A have non-positive real parts.

(2) asymptotically stable if both eigenvalues of A have negative real parts.

Remark: If the trivial solution is asymptotically stable, then it is also stable. The converse statement is not true.

Proof of Theorem 5.2.3: We start with part (2). If the eigenvalues of the coefficient matrix A have non-positive real parts, then we have the following possibilities.

(a) If the eigenvalues are real and $\lambda_- < \lambda_+ < 0$, then the general solution is

$$\boldsymbol{x}(t) = c_+ \boldsymbol{v}_+ e^{\lambda_+ t} + c_- \boldsymbol{v}_- e^{\lambda_- t}.$$

Therefore, the analysis done in this section says that $\boldsymbol{x}(t) \to \boldsymbol{x}_0 = \mathbf{0}$ as $t \to \infty$.

(b) If the eigenvalues are real and $\lambda_+ = \lambda_- = \lambda_0 < 0$, then the general solution is

$$\boldsymbol{x}(t) = c_1 \boldsymbol{v}_0 e^{\lambda_0 t} + c_- (t\boldsymbol{v}_0 + \boldsymbol{w})\, e^{\lambda_0 t}.$$

Therefore, the analysis done in this section says that $\boldsymbol{x}(t) \to \boldsymbol{x}_0 = \mathbf{0}$ as $t \to \infty$.

(c) If the eigenvalues are complex, $\lambda_\pm = \alpha \pm i\beta$, with $\alpha < 0$, then

$$\boldsymbol{x}_1(t) = \big(\boldsymbol{a} \cos(\beta t) - \boldsymbol{b} \sin(\beta t)\big)\, e^{\alpha t},$$
$$\boldsymbol{x}_2(t) = \big(\boldsymbol{a} \sin(\beta t) + \boldsymbol{b} \cos(\beta t)\big)\, e^{\alpha t},$$

are fundamental solutions of the differential equation. Since $\alpha < 0$, it is simple to see that $\boldsymbol{x}(t) \to \boldsymbol{x}_0 = \mathbf{0}$ as $t \to \infty$.

Regarding part (1), it includes all cases in part (2). But part (1) also includes the following two cases.

(a) Suppose that one eigenvalue of A is $\lambda_1 = 0$, with eigenvector $\boldsymbol{v}_1$. In this case one fundamental solution is $\boldsymbol{x}_1 = \boldsymbol{v}_1$, a constant, nonzero solution. Then a picture helps to understand that for initial data $\boldsymbol{x}(0)$ close enough to $\boldsymbol{x}_0 = \mathbf{0}$, the solution $\boldsymbol{x}(t)$ remains close to $\boldsymbol{x}_0$ for $t > 0$.

(b) The last case is when the coefficient matrix A has pure imaginary eigenvalues, $\lambda_\pm = \pm\beta i$. In this case the solutions are ellipses with $\boldsymbol{x}_0 = \mathbf{0}$ in its interior.

□

The second characterization of the trivial solution $\boldsymbol{x}_0 = \mathbf{0}$ is more precise, but some parts of this characterization cannot be extended to more general systems. For example, the saddle point characterization below applies only to 2×2 systems.

Definition 5.2.4. *The solution $\boldsymbol{x}_0 = \mathbf{0}$ of a 2×2 linear system $\boldsymbol{x}' = A\boldsymbol{x}$ is:*

*(a) a **source** iff for any initial condition $\boldsymbol{x}(0)$ the corresponding solution $\boldsymbol{x}(t)$ satisfies*

$$\lim_{t\to\infty} |\boldsymbol{x}(t)| = \infty.$$

*(b) a **sink** iff for any initial condition $\boldsymbol{x}(0)$ the corresponding solution $\boldsymbol{x}(t)$ satisfies*

$$\lim_{t\to\infty} \boldsymbol{x}(t) = \mathbf{0}.$$

*(c) a **saddle** iff for some initial data $\boldsymbol{x}(0)$ the corresponding solution $\boldsymbol{x}(t)$ behaves as in (a) and for other initial data the solution behaves as in (b).*

*(d) a **center** iff for any initial data $\boldsymbol{x}(0)$ the corresponding solution $\boldsymbol{x}(t)$ describes a closed periodic trajectory around $\boldsymbol{x}_0 = \mathbf{0}$.*

Recall that in Theorem 5.2.1 we have explicit formulas for all solutions of the 2×2 system $\boldsymbol{x}' = A\boldsymbol{x}$, for any matrix A. From those formulas it is not difficult to prove the following result.

Theorem 5.2.5 (Stability II)**.** *Let $\boldsymbol{x}(t)$ be the solution of a 2×2 linear system $\boldsymbol{x}' = A\boldsymbol{x}$, with $\det(A) \neq 0$ and initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_1$. Then the trivial solution $\mathrm{x}_0 = \mathbf{0}$ is:*

*(i) a **source** if both eigenvalues of A have positive real parts.*

*(ii) a **sink** if both eigenvalues of A have negative real parts.*

*(iii) a **saddle** if one eigenvalues of A is positive and the other is negative.*

*(iv) a **center** if both eigenvalues of A are purely imaginary.*

Proof of Theorem 5.2.5: Parts (i) and (ii)are simple to prove. From the solution formulas given above in this section one can see the following. If the coefficient matrix A has both eigenvalues with positive real parts, then for all initial data $\boldsymbol{x}(0) \neq \mathbf{0}$ the corresponding solutions satisfy that $\|\boldsymbol{x}(t)\| \to \infty$ as $t \to \infty$. This shows part (i). If the eigenvalues have negative real parts, then $\boldsymbol{x}(t) \to \mathbf{0}$ as $t \to \infty$.

Part (iii) is also simple to prove using the solution formulas from the beginning of this section. If the initial data $\boldsymbol{x}(0)$ lies on the eigenspace of A with negative eigenvalue, then the corresponding solution satisfies $\boldsymbol{x}(t) \to \mathbf{0}$ as $t \to \infty$. For any orhter initial data, the corresponding solution satisfies that $\|\boldsymbol{x}(t)\| \to \infty$ as $t \to \infty$.

Finally, part (iv) is simple to see from the solutions formulas at the beginning of this section. If A has pure imaginary eigenvalues, then the solutions describe ellipses around $\boldsymbol{x}_0 = \mathbf{0}$, which are closed periodic trajectories. □

5.2.6. Exercises.

5.2.1.- .

5.2.2.- .

5.3. Qualitative Analysis of Nonlinear Systems

We consider 2×2 systems of differential equations which are nonlinear. We focus on autonomous systems—systems where the independent variable does not appear explicitly. We study the behavior of certain solutions that are close to equilibrium solutions. We find that such solutions to nonlinear systems behave in a similar way as solutions of appropriately chosen linear systems. Therefore, the qualitative behavior of solutions to certain nonlinear systems can be obtained by studying solutions of several linear systems.

5.3.1. 2×2 Autonomous Systems. We study a more complicated system described by two functions, solutions of two differential equations involving only first derivatives of these functions. This is called a 2×2 system of first order differential equations.

Definition 5.3.1. *A 2×2* ***System of First Order Differential Equations*** *(SFODE) for the variables $x_1(t)$, $x_2(t)$ is*

$$x_1' = f_1(t, x_1, x_2), \tag{5.3.1}$$

$$x_2' = f_2(t, x_1, x_2). \tag{5.3.2}$$

The system above is called ***autonomous*** *when the functions f_1, f_2 do not depend explicitly on the independent variable t, that is,*

$$x_1' = f_1(x_1, x_2), \tag{5.3.3}$$

$$x_2' = f_2(x_1, x_2). \tag{5.3.4}$$

Let us introduce the variable $x = (x_1, x_2)$ and the vector

$$\boldsymbol{F} = \langle f_1, f_2 \rangle = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix},$$

where we use both notations, rows and columns, to denote vector components. Then, a *solution* of the system of differential equations in (5.3.1), (5.3.2) is a curve

$$x(t) = (x_1(t), x_2(t))$$

in the x_1x_2-plane, called the *phase space*, where the independent variable t is the parameter of the curve. The vector tangent to the solution curve $x(t)$ is its t-derivative, which in components is given by

$$\boldsymbol{x}' = \langle x_1', x_2' \rangle = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}.$$

Then, the system of differential equations in (5.3.1), (5.3.2) can be written in a more compact vector notation as

$$\boldsymbol{x}' = \boldsymbol{F}(t, x). \tag{5.3.5}$$

The system is autonomous if $\boldsymbol{x}' = \boldsymbol{F}(x)$.

Systems fo First Order Differential Equations can be linear or nonlinear.

Example 5.3.1 (Linear Systems)**.** Recall that a 2×2 system of first order linear differential equations (SFOLDE), homogeneous, and with constant coefficients, have the form

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

These equations can be written as

$$x_1' = a\,x_1 + b\,x_2,$$

$$x_2' = c\,x_1 + d\,x_2.$$

which means that the vector $\boldsymbol{F}$ for linear systems is given by the linear functions

$$\boldsymbol{F}(x_1, x_2) = \begin{bmatrix} a\,x_1 + b\,x_2 \\ c\,x_1 + d\,x_2 \end{bmatrix}.$$

$\triangleleft$

Examples of 2×2 *nonlinear* systems include the competing species system, the predator prey system, and the first order reduction of the equation for the movement of a pendulum.

Example 5.3.2 (Competing Species)**.** The physical system consists of two species that compete on the same food resources. For example, rabbits and sheep, which compete on the grass on a particular piece of land. If x_1 and x_2 are the competing species popultions, the the differential equations, also called Lotka-Volterra equations for competition, are

$$x_1' = r_1\,x_1\Big(1 - \frac{x_1}{K_1} - \alpha\,x_2\Big),$$
$$x_2' = r_2\,x_2\Big(1 - \frac{x_2}{K_2} - \beta\,x_1\Big).$$

The constants r_1, r_2, α, β are all nonnegative, and K_1, K_2 are positive. Note that in the case of absence of one species, say $x_2 = 0$, the population of the other species, x_1 is described by a logistic equation. The terms $-\alpha\,x_1x_2$ and $-\beta\,x_1x_2$ say that the competition between the two species is proportional to the number of competitive pairs x_1x_2. $\triangleleft$

Example 5.3.3 (Predator-Prey)**.** The physical system consists of two biological species where one preys on the other. For example cats prey on mice, foxes prey on rabbits. If we call x_1 the predator population, and x_2 the prey population, then predator-prey equations, also known as Lotka-Volterra equations for predator prey, are

$$x_1' = -a\,x_1 + b\,x_1x_2,$$
$$x_2' = -c\,x_1x_2 + d\,x_2.$$

The constants a, b, c, d are all nonnegative. Notice that in the case of absence of predators, $x_1 = 0$, the prey population grows without bounds, since $x_2' = d\,x_2$. In the case of absence of prey, $x_2 = 0$, the predator population becames extinct, since $x_1' = -a\,x_1$. The term $-c\,x_1x_2$ represents the prey death rate due to predation, which is porportional to the number of encounters, x_1x_2, between predators and prey. These encounters have a positive contribution $b\,x_1x_2$ to the predator population. $\triangleleft$

Example 5.3.4 (The Nonlinear Pendulum)**.** A pendulum of mass m, length ℓ, oscillating under the gravity acceleration g, moves according to Newton's second law of motion

$$m(\ell\theta)'' = -mg\sin(\theta),$$

where the angle θ depends on time t. If we rearrange terms we get a second order scalar equation

$$\theta'' + \frac{g}{\ell}\sin(\theta) = 0.$$

This second order equation can be written as a first order system. If we introduce the new variables $x_1 = \theta$ and $x_2 = \theta'$, then

$$x_1' = x_2$$
$$x_2' = -\frac{g}{\ell}\sin(x_1).$$

$\triangleleft$

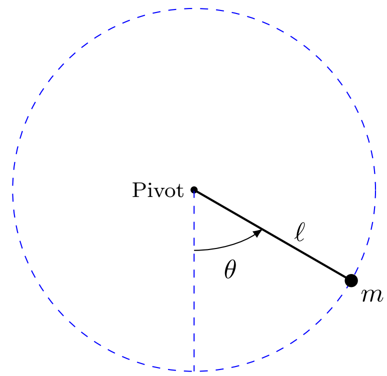


FIGURE 8. Pendulum of mass m, length ℓ, oscillating around the pivot.

It is hard to find formulas for solutions of nonlinear systems of differential equations. And even in the case that one can find such formulas, they often are too complicated to provide much insight into the solution behavior. Instead, we will try to get a qualitative understanding of the solutions of nonlinear systems. In other words, we will try to get an idea of the phase portrait of solutions without solving the differential equation.

We first find the constant solutions, called equilibrium solutions, or equilibrium points, or critical points. Then we study the behavior of solutions near the equilibrium solutions. It turns out that, sometimes, this behavior can be obtained from studying solutions of *linear systems* of differential equations. In this section we study when this is possible and how to obtain the appropriate linear system to study.

5.3.2. Equilibrium Solutions.

We start with the definition of equilibrium solutions.

Definition 5.3.2. *An **equilibrium solution** of the autonomous system $\boldsymbol{x}' = \boldsymbol{F}(x)$ is a constant solution, that is, a point x^0 in the phase space satisfying*

$$\boldsymbol{F}(x^0) = \boldsymbol{0}.$$

Remarks:

- Equilibrium solutions in components have the form $x^0 = (x_1^0, x_2^0)$, which is a point in the x_1x_2-plane, called phase space. This is why equilibrium solutions are also called *equilibrium points* or *critical points*.
- If we write the field is $\boldsymbol{F} = \langle f_1, f_2 \rangle$, the equilibrium solutions must satisfy the equations

$$f_1(x_1^0, x_2^0) = 0$$
$$f_2(x_1^0, x_2^0) = 0.$$

Example 5.3.5. Find all the equilibrium solutions of the 2×2 linear system

$$\boldsymbol{x}' = A\boldsymbol{x}, \qquad \det(A) \neq 0.$$

Solution: The equilibrium solutions are the constant vectors $\boldsymbol{x}$ solutions of $A\boldsymbol{x} = \boldsymbol{0}$. Since the coefficient matrix A is invertible,

$$A^{-1}A\boldsymbol{x} = A^{-1}\boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{x} = \boldsymbol{0}.$$

This linear system has only one equilibrium solution and it is the zero solution $\boldsymbol{x} = \boldsymbol{0}$. ◁

Example 5.3.6. Find all the equilibrium solutions of the competing species system

$$x_1' = 3\,x_1(1 - x_1) - 3\,x_1x_2$$
$$x_2' = 4\,x_2\Big(1 - \frac{x_2}{2}\Big) - 6\,x_1x_2.$$

Solution: We need to find all constants $x = (x_1, x_2)$ solutions of

$$3\,x_1(1 - x_1) - 3\,x_1x_2 = 0$$
$$4\,x_2\Big(1 - \frac{x_2}{2}\Big) - 6\,x_1x_2 = 0.$$

It is convenient to write the left sides above a products,

$$3\,x_1(1 - x_1 - x_2) = 0$$
$$2\,x_2(2 - x_2 - 3\,x_1) = 0.$$

From that expression we see that we have four possible solutions. HTe first three of them are

$$x_1 = 0, \quad x_2 = 0 \quad \text{and} x_1 = 0, \quad x_2 = 2 \quad \text{and} x_1 = 1, \quad x_2 = 0.$$

The last equilibrium solution is the solution of

$$\left.\begin{aligned} 1 - x_1 - x_2 = 0 \\ 2 - x_2 - 3\,x_1 = 0. \end{aligned}\right\} \quad \Rightarrow \quad x_1 = \frac{1}{2}, \quad x_2 = frac12.$$

So we have found four equilibrium solutions (or critical points), and they are given by

$$x^0 = (0, 0), \quad x^1 = (0, 2), \quad x^2 = (1, 0), \quad x^3 = \Big(\frac{1}{2}, \frac{1}{2}\Big).$$

◁

Example 5.3.7. Find all the critical points of the two-dimensional (decoupled) system

$$x_1' = -x_1 + (x_1)^3$$
$$x_2' = -2\,x_2.$$

Solution: We need to find all constants $x = (x_1, x_2)$ solutions of

$$-x_1 + (x_1)^3 = 0,$$
$$-2\,x_2 = 0.$$

From the second equation we get $x_2 = 0$. From the first equation we get

$$x_1\big((x_1)^2 - 1\big) = 0 \quad \Rightarrow \quad x_1 = 0, \quad \text{or} \quad x_1 = \pm 1.$$

Therefore, we got three critical points, $x^0 = (0, 0)$, $x^1 = (1, 0)$, $x^2 = (-1, 0)$. ◁

5.3.3. Linearizations. Once we know the constant solutions of a nonlinear differential system we study the equation itself near these constant solutions. Consider the two-dimensional system

$$x_1' = f_1(x_1, x_2),$$
$$x_2' = f_2(x_1, x_2),$$

Assume that f_1, f_2 have Taylor expansions at $x^0 = (x_1^0, x_2^0)$. We denote

$$u_1 = (x_1 - x_1^0), \qquad u_2 = (x_2 - x_2^0),$$

and

$$f_1^0 = f_1(x_1^0, x_2^0), \qquad f_2^0 = f_2(x_1^0, x_2^0).$$

Then, by the Taylor expansion theorem,

$$f_1(x_1, x_2) = f_1^0 + \frac{\partial f_1}{\partial x_1}\Big|_{x^0} u_1 + \frac{\partial f_1}{\partial x_2}\Big|_{x^0} u_2 + O\big((u_1)^2, (u_2)^2, u_1 u_2\big),$$

$$f_2(x_1, x_2) = f_2^0 + \frac{\partial f_2}{\partial x_1}\Big|_{x^0} u_1 + \frac{\partial f_2}{\partial x_2}\Big|_{x^0} u_2 + O\big((u_1)^2, (u_2)^2, u_1 u_2\big),$$

where $O\big((u_1)^2, (u_2)^2, u_1 u_2\big)$ denotes quadratic terms in u_1 and u_2. Let us simplify the notation a bit further. Let us denote

$$\partial_1 f_1 = \frac{\partial f_1}{\partial x_1}\Big|_{x^0}, \qquad \partial_2 f_1 = \frac{\partial f_1}{\partial x_2}\Big|_{x^0},$$

$$\partial_1 f_2 = \frac{\partial f_2}{\partial x_1}\Big|_{x^0}, \qquad \partial_2 f_2 = \frac{\partial f_2}{\partial x_2}\Big|_{x^0}.$$

then the Taylor expansion of $\boldsymbol{F}$ has the form

$$f_1(x_1, x_2) = f_1^0 + (\partial_1 f_1)\, u_1 + (\partial_2 f_1)\, u_2 + O\big((u_1)^2, (u_2)^2, u_1 u_2\big),$$

$$f_2(x_1, x_2) = f_2^0 + (\partial_1 f_2)\, u_1 + (\partial_2 f_2)\, u_2 + O\big((u_1)^2, (u_2)^2, u_1 u_2\big).$$

We now use this Taylor expansion of the field $\boldsymbol{F}$ into the differential equation $\boldsymbol{x}' = \boldsymbol{F}$. Recall that

$$x_1 = x_1^0 + u_1, \qquad x_2 = x_2^0 + u_2,$$

and that x_1^0 and x_2^0 are constants, then

$$u_1' = f_1^0 + (\partial_1 f_1)\, u_1 + (\partial_2 f_1)\, u_2 + O\big((u_1)^2, (u_2)^2, u_1 u_2\big),$$

$$u_2' = f_2^0 + (\partial_1 f_2)\, u_1 + (\partial_2 f_2)\, u_2 + O\big((u_1)^2, (u_2)^2, u_1 u_2\big).$$

Let us write this differential equation using vector notation. If we introduce the vectors and the matrix

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \qquad \boldsymbol{F}_0 = \begin{bmatrix} f_1^0 \\ f_2^0 \end{bmatrix}, \qquad DF_0 = \begin{bmatrix} \partial_1 f_1 & \partial_2 f_1 \\ \partial_1 f_2 & \partial_2 f_2 \end{bmatrix},$$

then, we have that

$$\boldsymbol{x}' = \boldsymbol{F}(x) \quad \Leftrightarrow \quad \boldsymbol{u}' = \boldsymbol{F}_0 + (DF_0)\, \boldsymbol{u} + O\big((u_1)^2, (u_2)^2, u_1 u_2\big).$$

In the case that x^0 is a critical point, then $\boldsymbol{F}_0 = \boldsymbol{0}$. In this case we have that

$$\boldsymbol{x}' = \boldsymbol{F}(x) \quad \Leftrightarrow \quad \boldsymbol{u}' = (DF_0)\, \boldsymbol{u} + O\big((u_1)^2, (u_2)^2, u_1 u_2\big).$$

The relation above says that, when x is close to x^0, the equation coefficients of $\boldsymbol{x}' = \boldsymbol{F}(x)$ are close to the coefficients of the linear differential equation $\boldsymbol{u}' = (DF_0)\, \boldsymbol{u}$. For this reason, we give this linear differential equation a name.

Definition 5.3.3. *The **linearization** of a nonlinear 2×2 system for the function $\boldsymbol{x}(t)$,*

$$\boldsymbol{x}' = \boldsymbol{F}(x)$$

at a critical point x^0 is the 2×2 linear system for a function $\boldsymbol{u}(t)$ given by

$$\boldsymbol{u}' = (DF_0)\, \boldsymbol{u},$$

*where we have introduced the **Jacobian**, or derivative, matrix at x^0,*

$$DF_0 = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}\Big|_{x^0} & \frac{\partial f_1}{\partial x_2}\Big|_{x^0} \\ \frac{\partial f_2}{\partial x_1}\Big|_{x^0} & \frac{\partial f_2}{\partial x_2}\Big|_{x^0} \end{bmatrix} = \begin{bmatrix} \partial_1 f_1 & \partial_2 f_1 \\ \partial_1 f_2 & \partial_2 f_2 \end{bmatrix}.$$

Remark: In components, the nonlinear system is

$$x_1' = f_1(x_1, x_2),$$
$$x_2' = f_2(x_1, x_2),$$

and the linearization at x^0 is

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}' = \begin{bmatrix} \partial_1 f_1 & \partial_2 f_1 \\ \partial_1 f_2 & \partial_2 f_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Once we have these definitions, we can summarize the calculation we did above.

Theorem 5.3.4. *If a* 2×2 *nonlinear autonomous system*

$$\boldsymbol{x}' = \boldsymbol{F}(\boldsymbol{x})$$

has a critical point x^0*, then in a neighborhood of* x^0 *the equation coefficients of this nonlinear system are close to the equation coefficients of its linearization at* x^0*, given by*

$$\boldsymbol{u}' = (DF_0)\, \boldsymbol{u},$$

where $\boldsymbol{u}(t) = \boldsymbol{x}(t) - \boldsymbol{x}^0$.

Remark: The proof is the calculation given above the Definition 5.3.3. This result says that near a critical point, the equation coefficients of the nonlinear system are close to the equation coefficients of its linearization at the critical point. This is a result about the equations, not their solutions. Any relation between solutions is hard to prove. This is the main subject of the Hartman-Grobman Theorem 5.3.6, which can be established for only a particular type of nonlinear systems.

Example 5.3.8. Find the linearization at every critical point of the nonlinear system

$$x_1' = -x_1 + (x_1)^3$$
$$x_2' = -2\, x_2.$$

Solution: We found earlier that this system has three critial points,

$$x^0 = (0,0), \qquad x^1 = (1,0), \qquad x^2 = (-1,0).$$

This means we need to compute three linearizations, one for each critical point. We start computing the derivative matrix at an arbitrary point x,

$$DF(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1}(-x_1 + x_1^3) & \frac{\partial}{\partial x_2}(-x_1 + x_1^3) \\ \frac{\partial}{\partial x_1}(-2x_2) & \frac{\partial}{\partial x_2}(-2x_2) \end{bmatrix},$$

so we get that

$$DF(x) = \begin{bmatrix} -1 + 3\, x_1^2 & 0 \\ 0 & -2 \end{bmatrix}.$$

We only need to evaluate this matrix Df at the critical points. We start with $\boldsymbol{x}_0$,

$$x^0 = (0,0) \quad \Rightarrow \quad DF_0 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}' = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

The Jacobian at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is the same, so we get the same linearization at these points,

$$x^1 = (1,0) \quad \Rightarrow \quad DF_1 = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}' = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$x^2 = (-1, 0) \quad \Rightarrow \quad DF_2 = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}' = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

◁

We classify critical points of nonlinear systems by the eigenvalues of their linearizations.

Definition 5.3.5. *A critical point x^0 of a two-dimensional system $\boldsymbol{x}' = \boldsymbol{F}(x)$ is a:*

(a) ***source node*** *iff both eigenvalues of DF_0 are real and positive;*
(b) ***source spiral*** *iff both eigenvalues of DF_0 are complex with positive real parts;*
(c) ***sink node*** *iff both eigenvalues of DF_0 are real and negative;*
(d) ***sink spiral*** *iff both eigenvalues of DF_0 are complex with negative real part;*
(e) ***saddle node****, iff both eigenvalues of DF_0 are real, one positive, the other negative;*
(f) ***center****, iff both eigenvalues of DF_0 are pure imaginary;*
(g) ***higher order*** *critical point iff at least one eigenvalue of DF_0 is zero.*

A critical point x^0 is called ***hyperbolic*** *iff it belongs to cases (a-e), that is, the real part of all eigenvalues of DF_0 are nonzero.*

Example 5.3.9. Classify all the critical points of the nonlinear system

$$x_1' = -x_1 + (x_1)^3$$
$$x_2' = -2\, x_2.$$

Solution: We already know that this system has three critical points,

$$x^0 = (0, 0), \qquad x^1 = (1, 0), \qquad x^2 = (-1, 0).$$

We have already computed the linearizations at these critical points too.

$$DF_0 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}, \qquad DF_1 = DF_2 = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

We now need to compute the eigenvalues of the Jacobian matrices above.

- For x^0 we have $\lambda_+ = -1$, $\lambda_- = -2$, so x^0 is an attractor.
- For x^1 and x^2 we have $\lambda_+ = 2$, $\lambda_- = -2$, so x^1 and x^2 are saddle points.

◁

We have seen that in a neighborhood of a critical point the equation coefficients are close to the equation coefficients of its linearization at that critical point. This is a relation between equation coefficients. Now we relate their corresponding solutions. It turns out that such relation between solutions can be established when the Jacobian matrix at a critical point has eigenvalues with nonzero real part. In this case, the solutions of the nonlinear system near that critical point are close to the solutions of the linearization at that critical point. We summarize this result in the following statement.

Theorem 5.3.6 (Hartman-Grobman). *Consider a two-dimensional nonlinear autonomous system with a continuously differentiable field $\boldsymbol{F}$,*

$$\boldsymbol{x}' = \boldsymbol{F}(x),$$

and consider its linearization at a hyperbolic critical point x^0,

$$\boldsymbol{u}' = (DF_0)\, \boldsymbol{u}.$$

Then, there is a neighborhood of the hyperbolic critical point x^0 where all the solutions of the linear system can be transformed into solutions of the nonlinear system by a continuous, invertible, transformation.

Remark: The Hartman-Grobman theorem implies that the phase portrait of the linear system in a neighborhood of a hyperbolic critical point can be transformed into the phase portrait of the nonlinear system by a continuous, invertible, transformation. When that happens we say that the two phase portraits are *topologically equivalent*.

Remark: This theorem says that, for hyperbolic critical points, the phase portrait of the linearization at the critical point is enough to determine the phase portrait of the nonlinear system near that critical point.

Example 5.3.10. Use the Hartman-Grobman theorem to sketch the phase portrait of

$$x_1' = -x_1 + (x_1)^3$$
$$x_2' = -2\,x_2.$$

Solution: We have found before that the critical points are

$$x^0 = (0,0), \qquad x^1 = (1,0), \qquad x^2 = (-1,0),$$

where x^0 is a sink node and x^1, x^2 are saddle nodes.

The phase portrait of the linearized systems at the critical points is given in Fig 5.3.3. These critical points have all linearizations with eigenvalues having nonzero real parts. This means that the critical points are hyperbolic, so we can use the Hartman-Grobman theorem. This theorem says that the phase portrait in Fig. 5.3.3 is precisely the phase portrait of the nonlinear system in this example.

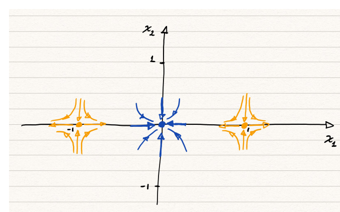


FIGURE 9. Phase portraits of the linear systems at x^0, x^1, and x^2.

Since we now know that Fig 5.3.3 is also the phase portrait of the nonlinear, we only need to fill in the gaps in that phase portrait. In this example, a decoupled system, we can complete the phase portrait from the symmetries of the solution. Indeed, in the x_2 direction all trajectories must decay to exponentially to the $x_2 = 0$ line. In the x_1 direction, all trajectories are attracted to $x_1 = 0$ and repelled from $x_1 = \pm 1$. The vertical lines $x_1 = 0$ and $x_1 = \pm 1$ are invariant, since $x_1' = 0$ on these lines; hence any trajectory that start on these lines stays on these lines. Similarly, $x_2 = 0$ is an invariant horizontal line. We also note that the phase portrait must be symmetric in both x_1 and x_2 axes, since the equations are invariant under the transformations $x_1 \to -x_1$ and $x_2 \to -x_2$. Putting all this extra information together we arrive to the phase portrait in Fig. 10.

◁

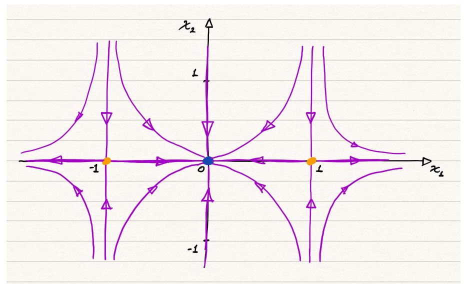


FIGURE 10. Phase portraits of the nonlinear systems in the Example 5.3.10

5.3.4. Exercises.

5.3.1.- .

5.3.2.- .

5.4. Applications of the Qualitative Analysis

In this section we carry out the analysis outlined in § 5.3 in a few physical systems. These systems include the competing species, the predator-prey, and the nonlinear pendulum. We use the analysis in § 5.3 to find qualitative properties of solutions to nonlinear autonomous systems

5.4.1. Competing Species.

The physical system consists of two species that compete on the same, limited, food resources. For example, rabbits and sheep, which compete on the grass on a particular piece of land. If x_1 and x_2 are the competing species populations, the differential equations, also called Lotka-Volterra equations for competition, are

$$x_1' = r_1\, x_1 \Big(1 - \frac{x_1}{K_1} - \alpha\, x_2\Big),$$
$$x_2' = r_2\, x_2 \Big(1 - \frac{x_2}{K_2} - \beta\, x_1\Big).$$

The constants r_1, r_2, α, β are all nonnegative, and K_1, K_2 are positive. The constants r_1, r_2 are the growth rate per capita, the constants K_1, K_2 are the carrying capacities, and α, β the competition constants for the respective species.

Example 5.4.1 (Competing Species: Extinction)**.** Sketch in the phase space all the critical points and several solution curves tin order to get a qualitative understanding of the behavior of all solutions to the competing species system (found in Strogatz book [9]),

$$x_1' = x_1\,(3 - x_1 - 2\,x_2), \tag{5.4.1}$$
$$x_2' = x_2\,(2 - x_2 - x_1), \tag{5.4.2}$$

where $x_1(t)$ is the population of one of the species, say rabbits, and $x_2(t)$ is the population of the other species, say sheep, at the time t. We restrict our study to nonnegative functions x_1, x_2.

Solution: We start finding all the critical points of the rabbits-sheep system. We need to find all constants (x_1, x_2) solutions of

$$x_1\,(3 - x_1 - 2\,x_2) = 0, \tag{5.4.3}$$
$$x_2\,(2 - x_2 - x_1) = 0. \tag{5.4.4}$$

- One solution is
$$x_1 = 0 \quad \text{and} \quad x_2 = 0,$$
which gives the critical point $x^0 = (0, 0)$.
- A second solution is
$$x_1 = 0 \quad \text{and} \quad (2 - x_2 - x_1) = 0,$$
but since $x_1 = 0$, then $x_2 = 2$, which gives the critical point $x^1 = (0, 2)$.
- A third solution is
$$(3 - x_1 - 2\,x_2) = 0 \quad \text{and} \quad x_2 = 0,$$
but since $x_2 = 0$, then $x_1 = 3$, which gives the critical point $x^2 = (3, 0)$.
- The fourth solution is
$$(2 - x_2 - x_1) = 0 \quad \text{and} \quad (3 - x_1 - 2\,x_2) = 0$$
which gives $x_1 = 1$ and $x_2 = 1$, so we get the critical point $x^3 = (1, 1)$.

We now compute the linearization of the rabbits-sheep system in Eqs.(5.4.1)-(5.4.2). We first compute the derivative of the field $\boldsymbol{F}$, where

$$\boldsymbol{F}(x) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} x_1\,(3 - x_1 - 2\,x_2) \\ x_2\,(2 - x_2 - x_1) \end{bmatrix}.$$

The derivative of $\boldsymbol{F}$ at an arbitrary point x is

$$DF(x) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} (3 - 2\,x_1 - 2\,x_2) & -2\,x_1 \\ -x_2 & (2 - x_1 - 2\,x_2) \end{bmatrix}.$$

We now evaluate the matrix $DF(x)$ at each of the critical points we found.

$$\text{At} \quad x^0 = (0,0) \quad \text{we get} \quad (DF_0) = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}.$$

This coefficient matrix has eigenvalues $\lambda_{0+} = 3$ and $\lambda_{0-} = 2$, both positive, which means that the critical point x^0 is a Source Node. To sketch the phase portrait we will need the corresponding eigenvectors, $\boldsymbol{v}_0^+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{v}_0^- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

$$\text{At} \quad x^1 = (0,2) \quad \text{we get} \quad (DF_1) = \begin{bmatrix} -1 & 0 \\ -2 & -2 \end{bmatrix}.$$

This coefficient matrix has eigenvalues $\lambda_{1+} = -1$ and $\lambda_{1-} = -2$, both negative, which means that the critical point x^1 is a Sink Node. One can check that the corresponding eigenvectors are $\boldsymbol{v}_1^+ = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\boldsymbol{v}_1^- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

$$\text{At} \quad x^2 = (3,0) \quad \text{we get} \quad (DF_2) = \begin{bmatrix} -3 & -6 \\ 0 & -1 \end{bmatrix}.$$

This coefficient matrix has eigenvalues $\lambda_{2+} = -1$ and $\lambda_{2-} = -3$, both negative, which means that the critical point x^2 is a Sink Node. One can check that the corresponding eigenvectors are $\boldsymbol{v}_2^+ = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ and $\boldsymbol{v}_2^- = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

$$\text{At} \quad x^3 = (1,1) \quad \text{we get} \quad (DF_3) = \begin{bmatrix} -1 & -2 \\ -1 & -1 \end{bmatrix}.$$

It is not difficult to check that this coefficient matrix has eigenvalues $\lambda_{3+} = -1 + \sqrt{2}$ and $\lambda_{3-} = -1 - \sqrt{2}$, which means that the critical point x^3 is a Saddle Node. One can check that the corresponding eigenvectors are $\boldsymbol{v}_3^+ = \begin{bmatrix} -\sqrt{2} \\ 1 \end{bmatrix}$ and $\boldsymbol{v}_3^- = \begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix}$. We summarize this information about the linearized systems in Fig. 11.

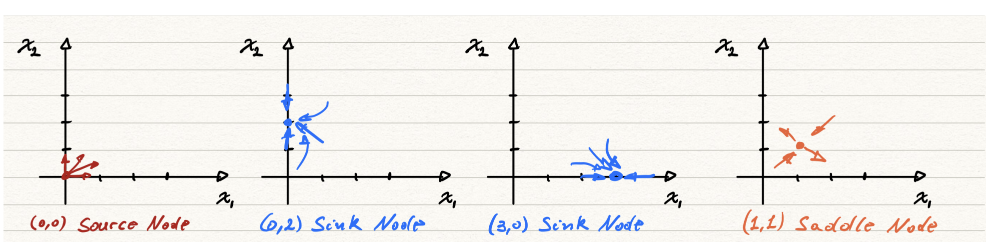


FIGURE 11. The linearizations of the rabbits-sheeps system in Eqs. (5.4.1)-(5.4.2).

We now notice that all these critical points have nonzero real part, that means they are hyperbolic critical points. Then we can use Hartman-Grobman Theorem 5.3.6 to construct the phase portrait of the nonlinear system in (5.4.1)-(5.4.2) around these critical points. The Hartman-Grobman theorem says that the qualitative structure of the phase portrait for the linearized system is the same for the phase portrait of the nonlinear system around the critical point. So we get the picture in Fig. 12.

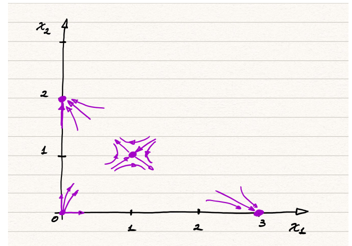


FIGURE 12. Phase Portrait for Eqs. (5.4.1)-(5.4.2).

We would like to have the complete phase portrait for the nonlinear system, that is, we would like to fill the gaps in Fig. 12. This is difficult to do analytically in this example as well as in general nonlinear autonomous systems. At this point is where we need to turn to computer generated solutions to fill the gaps in Fig. 12. The result is in Fig. 13.

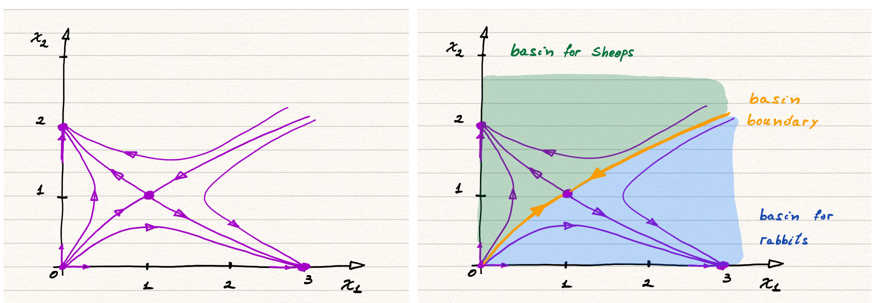


FIGURE 13. The phase portrait of the rabbits-sheeps system in Eqs. (5.4.1)-(5.4.2).

We can now study the phase portrait in Fig. 13 to obtain some biological insight on the rabbits-sheep system. The picture on the right says that most of the time one species drives the other to extinction. If the initial data for the system is a point on the blue region, called the *rabbit basin*, then the solution evolves in time toward the critical point $x^2 = (3, 0)$. This means that the sheep become extinct. If the initial data for the system is a point on the green region, called the *sheep basin*, then the solution evolves in time toward the critical point $x^1 = (0, 2)$. This means that the rabbits become extinct.

The two basins of attractions are separated by a curve, called the *basin boundary*. Only when the initial data lies on that curve the rabbits and sheep coexist with neither becoming extinct. The solution moves towards the critical point $x^3 = (1, 1)$. Therefore, the populations of rabbits and sheep become equal to each other as the time goes to infinity. But, if we pick an initial data outside this basin boundary, no matter how close this boundary, one of the species becomes extinct. ◁

Our next example is also a competing species system. The coefficient in this model are slightly different from the previous example, but the behavior of the species population predicted by this new model is very different from the previous example. The main prediction of the previous example is that one species goes extinct. We will see that the main prediction for the next example is that both species can coexist.

Example 5.4.2 (Competing Species: Coexistence)**.** Sketch in the phase space all the critical points and several solution curves tin order to get a qualitative understanding of the behavior of all solutions to the competing species system (found in Strogatz book [9]),

$$x_1' = x_1\,(1 - x_1 - x_2), \tag{5.4.5}$$

$$x_2' = x_2\,\Big(\frac{3}{4} - x_2 - \frac{1}{2}\,x_1\Big), \tag{5.4.6}$$

where $x_1(t)$ is the population of one of the species, say rabbits, and $x_2(t)$ is the population of the other species, say sheep, at the time t. We restrict our study to nonnegative functions x_1, x_2.

Solution: We start computing the critical points. We need to find all constants (x_1, x_2) solutions of

$$x_1\,(1 - x_1 - x_2) = 0, \tag{5.4.7}$$

$$x_2\,(3 - 4\,x_2 - 2\,x_1) = 0. \tag{5.4.8}$$

- One solution is
$$x_1 = 0 \quad \text{and} \quad x_2 = 0,$$
which gives the critical point $x^0 = (0, 0)$.
- A second solution is
$$x_1 = 0 \quad \text{and} \quad (3 - 4\,x_2 - 2\,x_1) = 0,$$
but since $x_1 = 0$, then $x_2 = 3/4$, which gives the critical point $x^1 = (0, 3/4)$.
- A third solution is
$$x_1\,(1 - x_1 - x_2) = 0 \quad \text{and} \quad x_2 = 0,$$
but since $x_2 = 0$, then $x_1 = 1$, which gives the critical point $x^2 = (1, 0)$.
- The fourth solution is
$$(1 - x_1 - x_2) = 0 \quad \text{and} \quad (3 - 4\,x_2 - 2\,x_1) = 0$$
which gives $x_1 = 1/2$ and $x_2 = 1/2$, so we get the critical point $x^3 = (1/2, 1/2)$.

We now compute the linearization of Eqs.(5.4.5)-(5.4.6). We first compute the derivative of the field $\boldsymbol{F}$, where

$$\boldsymbol{F}(x) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} x_1\,(1 - x_1 - x_2) \\ x_2\,\big(\frac{3}{4} - x_2 - \frac{1}{2}\,x_1\big) \end{bmatrix}.$$

The derivative of $\boldsymbol{F}$ at an arbitrary point x is

$$DF(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} (1 - 2\,x_1 - x_2) & -x_1 \\ -\frac{1}{2}\,x_2 & (\frac{3}{4} - 2x_2 - \frac{1}{2}\,x_1) \end{bmatrix}.$$

We now evaluate the matrix $DF(x)$ at each of the critical points we find the following.

At $\quad x^0 = (0, 0), \qquad (DF_0) = \begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{\frac{3}{4}} \end{bmatrix}, \qquad \boldsymbol{u} = \boldsymbol{0} \qquad$ is a Source Node.

At $\quad x^1 = (0, 3/4), \qquad (DF_1) = \begin{bmatrix} \mathbf{\frac{1}{3}} & 0 \\ -\frac{3}{8} & \mathbf{-\frac{3}{4}} \end{bmatrix}, \qquad \boldsymbol{u} = \boldsymbol{0} \qquad$ is a Saddle Node.

At $\quad x^2 = (1, 0), \qquad (DF_2) = \begin{bmatrix} \mathbf{-1} & -1 \\ 0 & \mathbf{\frac{1}{4}} \end{bmatrix}, \qquad \boldsymbol{u} = \boldsymbol{0} \qquad$ is a Saddle Node.

At $\quad x^3 = (1/2, 1/2), \qquad (DF_3) = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{4} & -\frac{1}{2} \end{bmatrix}, \qquad \boldsymbol{u} = \boldsymbol{0} \qquad$ is a Sink Node.

In the expressions above, we highlighted in red boldface the eigenvalues of the linearizations. For the last linearization, the eigenvalues are

$$\lambda_{3\pm} = \frac{1}{4}(-2 \pm \sqrt{2}) < 0.$$

If we put all this information together in a phase diagram, and we use the Hartman-Grobman Theorem 5.3.6, we obtain the picture in Fig. 14.

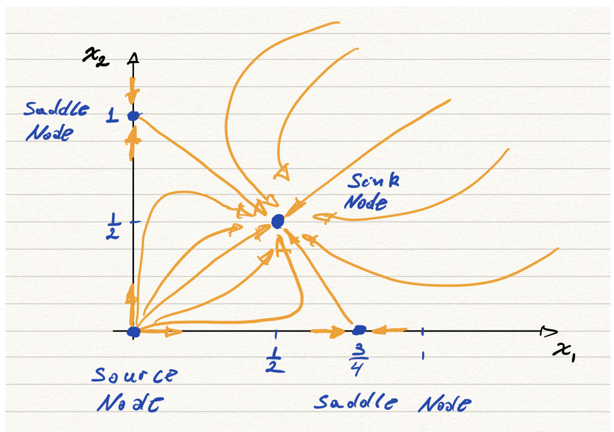


FIGURE 14. The phase portrait of the rabbits-sheep system in Eqs. (5.4.5)-(5.4.6).

We can see that this competing species system predicts that both species will coexist, in this case with population values given by the critical point $x^3 = (1/2, 1/2)$. No matter what the initial conditions are, the solution curve will move towards the critical point x^3.

◁

5.4.2. Predator-Prey. In this section we construct the phase diagram of predator prey systems, first introduced in § **??**. These are systems of equations that model physical systems consisting of two biological species where one species preys on the other. For example cats prey on mice, foxes prey on rabbits. If we call x_1 the predator population, and x_2 the prey population, then predator-prey equations, also known as Lotka-Volterra equations for predator prey, are

$$x_1' = -a\,x_1 + \alpha\,x_1 x_2, \tag{5.4.9}$$

$$x_2' = b\,x_2 - \beta\,x_1 x_2. \tag{5.4.10}$$

The constants a, b, α, and β are all positive. The vector field, $\boldsymbol{F}$, of the equation is

$$\boldsymbol{F}(x) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} -a\,x_1 + \alpha\,x_1 x_2 \\ b\,x_2 - \beta\,x_1 x_2 \end{bmatrix}.$$

The equilibrium solutions, or critical points, are the solutions of $\boldsymbol{F}(x) = \boldsymbol{0}$, which in components we get

$$x_1(-a + \alpha\,x_2) = 0 \tag{5.4.11}$$

$$x_2(b - \beta\,x_1) = 0. \tag{5.4.12}$$

There are two solutions of the equations above, which give us the critical points

$$x^0 = (0, 0), \qquad x^1 = \Big(\frac{b}{\beta}, \frac{a}{\alpha}\Big).$$

The derivative matrix for the predator-prey system is

$$DF(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} (-a + \alpha\, x_2) & \alpha\, x_1 \\ -\beta\, x_2 & (b - \beta\, x_1) \end{bmatrix}.$$

At the critical point $x^0 = (0,0)$ we get

$$(DF_0) = \begin{bmatrix} -a & 0 \\ 0 & b \end{bmatrix}$$

which has eigenpairs

$$\lambda_{0+} = b, \quad \boldsymbol{v}_{0+} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad \lambda_{0-} = -a, \quad \boldsymbol{v}_{0-} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

which means that $x^0 = (0,0)$ is a Saddle Node. At the critical point $x^1 = (b/\beta, a/\alpha)$ we get

$$(DF_1) = \begin{bmatrix} 0 & \frac{\alpha}{\beta}\, b \\ -\frac{\beta}{\alpha}\, a & 0 \end{bmatrix}$$

which has eigenpairs

$$\lambda_{1\pm} = \pm\sqrt{ab}\, i, \qquad \boldsymbol{v}_{1\pm} = \begin{bmatrix} \alpha/\beta \\ \pm\sqrt{a/b}\, i \end{bmatrix},$$

which means that $x^0 = (b/\beta, a/\alpha)$ is a Center. In particular, notice that we can write the eigenvector $\boldsymbol{v}_{1+}$ above as

$$\boldsymbol{v}_{1+} = \begin{bmatrix} \alpha/\beta \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \sqrt{a/b} \end{bmatrix} i = \boldsymbol{a} + \boldsymbol{b}\, i \quad \Rightarrow \quad \boldsymbol{a} = \begin{bmatrix} \alpha/\beta \\ 0 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 0 \\ \sqrt{a/b} \end{bmatrix}.$$

These vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ determine the direction a point in a solution curve moves as time increases on the x_1x_2-plane, which in this system is clockwise.

It could be useful to write an approximate expression for the solutions $\boldsymbol{x}(t)$ of the nonlinear predator-prey system near the critical point x^1. From the eigenpairs of the linearization matrix DF_1 we can compute the fundamental solutions of the linearization,

$$\boldsymbol{u}' = DF_1\, \boldsymbol{u},$$

at the critical point $x^1 = (b/\beta, a/\alpha)$. If we recall the formulas from § **??**, them we can see that these fundamental solutions are given by

$$\boldsymbol{u}_1(t) = \cos(\sqrt{ab}\, t) \begin{bmatrix} \alpha/\beta \\ 0 \end{bmatrix} - \sin(\sqrt{ab}\, t) \begin{bmatrix} 0 \\ \sqrt{a/b} \end{bmatrix} = \begin{bmatrix} \frac{\alpha}{\beta} \cos(\sqrt{ab}\, t) \\ -\sqrt{\frac{a}{b}} \sin(\sqrt{ab}\, t) \end{bmatrix},$$

$$\boldsymbol{u}_2(t) = \sin(\sqrt{ab}\, t) \begin{bmatrix} \alpha/\beta \\ 0 \end{bmatrix} + \cos(\sqrt{ab}\, t) \begin{bmatrix} 0 \\ \sqrt{a/b} \end{bmatrix} = \begin{bmatrix} \frac{\alpha}{\beta} \sin(\sqrt{ab}\, t) \\ \sqrt{\frac{a}{b}} \cos(\sqrt{ab}\, t) \end{bmatrix}.$$

Therefore, the general solution of the linearization is

$$\boldsymbol{u}(t) = c_1\, \boldsymbol{u}_1(t) + c_1\, \boldsymbol{u}_2(t),$$

where c_1, c_2 are arbitrary constants that could be determined by appropriate initial conditions. Then, the solutions of the nonlinear system near the critical point x^1 are given by $\boldsymbol{x}(t) \simeq \boldsymbol{x}^1 + \boldsymbol{u}(t)$, that is,

$$\boldsymbol{x}(t) \simeq \begin{bmatrix} b/\beta \\ a/\alpha \end{bmatrix} + c_1 \begin{bmatrix} \frac{\alpha}{\beta} \cos(\sqrt{ab}\, t) \\ -\sqrt{\frac{a}{b}} \sin(\sqrt{ab}\, t) \end{bmatrix} + c_2 \begin{bmatrix} \frac{\alpha}{\beta} \sin(\sqrt{ab}\, t) \\ \sqrt{\frac{a}{b}} \cos(\sqrt{ab}\, t) \end{bmatrix}.$$

Remarks:

- Near the critical point x^1 the populations of predators and preys oscillate in time with a frequency $\omega = \sqrt{ab}$, that is, period $T = 2\pi/\sqrt{ab}$. This period is the same for all solutions near the equilibrium solution x^1, hence independent of the initial conditions of the solutions.
- The populations of predators and preys are out of phase by $-T/4$.
- The amplitude of the oscillations in the predator and prey populations do depend on the initial conditions of the solutions.
- It can be shown that the average populations of predators and prey are, respectively, b/β and a/α, the same as the equilibrium populations in x^1.

We can now put together all the information we found in the discussion and sketch a phase portrait for the solutions of the Predator-Prey system in Equations (5.4.9)-(5.4.10). The phase space is the x_1x_2-plane, where in the horizontal axis we put the predator and in the vertical axis the prey. We only consider the region $x_1 \geqslant 0$, $x_2 \geqslant 0$ in the phase space, since populations cannot be negative. We choose some arbitrary length for the vectors $\boldsymbol{a}$, $\boldsymbol{b}$, to be able to sketch the phase portrait, although we faithfully represent their directions, horizontal to the right for $\boldsymbol{a}$ and vertical upwards for $\boldsymbol{b}$. The result is given in Fig. 15.

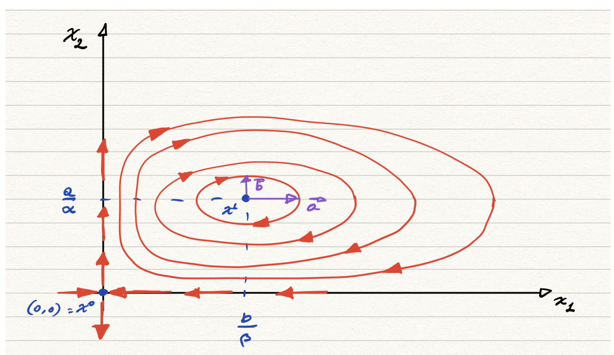


FIGURE 15. The phase portrait of a predator-prey system.

Example 5.4.3 (Predator-Prey: Infinite Food)**.** Sketch the phase diagram of the predator-prey system

$$x_1' = -x_1 + \frac{1}{2}\,x_1x_2,$$
$$x_2' = \frac{3}{4}\,x_2 - \frac{1}{4}\,x_1x_2.$$

Solution: We start computing the critical points, the constants (x_1, x_2) solutions of

$$x_1\Big(-1 + \frac{1}{2}\,x_2\Big) = 0, \tag{5.4.13}$$
$$x_2\,\Big(\frac{3}{4} - \frac{1}{4}\,x_1\Big) = 0. \tag{5.4.14}$$

- One solution is

$$x_1 = 0 \quad \text{and} \quad x_2 = 0,$$

which gives the critical point $x^0 = (0, 0)$.

- The only other solution is when
$$\Big(-1 + \frac{1}{2}\,x_2\Big) = 0 \quad \text{and} \quad \Big(\frac{3}{4} - \frac{1}{4}\,x_1\Big) = 0,$$
which means $x_1 = 3$ and $x_2 = 2$. This gives us the critical point $x^1 = (3, 2)$.

We now compute the linearization of Eqs.(5.4.13)-(5.4.14). We first compute the derivative of the field $\boldsymbol{F}$, where
$$\boldsymbol{F}(x) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} x_1\big(-1 + \frac{1}{2}\,x_2\big) \\ x_2\,\big(\frac{3}{4} - \frac{1}{4}\,x_1\big) \end{bmatrix}.$$
The derivative of $\boldsymbol{F}$ at an arbitrary point x is
$$DF(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} (-1 + \frac{1}{2}\,x_2) & \frac{1}{2}\,x_1 \\ -\frac{1}{4}\,x_2 & (\frac{3}{4} - \frac{1}{4}\,x_1) \end{bmatrix}.$$
We now evaluate the matrix $DF(x)$ at each of the critical points we find the following.

At $\quad x^0 = (0, 0), \quad (DF_0) = \begin{bmatrix} \mathbf{-1} & 0 \\ 0 & \mathbf{\frac{3}{4}} \end{bmatrix}, \quad x^0$ is a Saddle Node.

At $\quad x^1 = (3, 2), \quad (DF_1) = \begin{bmatrix} 0 & \frac{3}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}, \quad x^1$ is a Center.

The critical point x^0 is a Saddle Node, because the eigenvalues of the linearization, which are the coefficients in red boldface, satisfy $\lambda_- = -1 < 0 < \lambda_+ = 3/4$. The corresponding eigenvectors are
$$\boldsymbol{v}_- = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \boldsymbol{v}_+ = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$
The critical point x^1 is a Center, because the eigenpairs of the linearization are
$$\lambda_\pm = \pm\sqrt{\frac{3}{4}}\,i, \qquad \boldsymbol{v}_\pm = \begin{bmatrix} \sqrt{3} \\ \pm i \end{bmatrix}.$$
The eigenvectors above implies that as time increases the solution moves clockwise around the center critical point. If we put all this information together, we obtain the phase diagram similar to the one given in Fig. 15 ◁

It is simple to modify the Predator-Prey system above to consider the case where the prey has access to *finite* food resources.

Example 5.4.4 (Predator-Prey: Finite Food)**.** Characterize the critical points the predator-prey system
$$x_1' = -\frac{3}{4}\,x_1 + \frac{1}{4}\,x_1x_2, \tag{5.4.15}$$
$$x_2' = x_2 - \sigma\,x_2^2 - \frac{1}{2}\,x_1x_2, \tag{5.4.16}$$
where σ is a constant satisfying $0 < \sigma < 1/3$.

Solution: We start computing the critical points, the constants (x_1, x_2) solutions of
$$x_1\Big(-\frac{3}{4} + \frac{1}{4}\,x_2\Big) = 0,$$
$$x_2\,\Big(1 - \sigma\,x_2 - \frac{1}{2}\,x_1\Big) = 0.$$

- One solution is
$$x_1 = 0 \quad \text{and} \quad x_2 = 0,$$
which gives the critical point $x^0 = (0,0)$.
- Another solutions is
$$x_1 = 0 \quad \text{and} \quad x_2 = \frac{1}{\sigma},$$
which gives the critical point $x^1 = (0, 1/\sigma)$.
- The last solution is when
$$\Big(-\frac{3}{4} + \frac{1}{4}\,x_2\Big) = 0 \quad \text{and} \quad \Big(1 - \sigma\,x_2 - \frac{1}{2}\,x_1\Big) = 0,$$
which gives us the critical point $x^2 = (2(1-3\sigma), 3)$.

We now compute the linearization of Eqs. (5.4.15)-(5.4.16). We first compute the derivative of the field $\boldsymbol{F}$, where
$$\boldsymbol{F}(x) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} -\frac{3}{4}\,x_1 + \frac{1}{4}\,x_1x_2 \\ x_2 - \sigma\,x_2^2 - \frac{1}{2}\,x_1x_2 \end{bmatrix}.$$
The derivative of $\boldsymbol{F}$ at an arbitrary point x is
$$DF(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} (-\frac{3}{4} + \frac{1}{4}\,x_2) & \frac{1}{4}\,x_1 \\ -\frac{1}{2}\,x_2 & (1 - 2\sigma\,x_2 - \frac{1}{2}\,x_1) \end{bmatrix}.$$
We now evaluate the matrix $DF(x)$ at each of the critical points we find the following.

$$\text{At} \quad x^0 = (0,0), \qquad (DF_0) = \begin{bmatrix} \boldsymbol{-\frac{3}{4}} & 0 \\ 0 & \boldsymbol{1} \end{bmatrix}, \qquad x^0 \text{ is a Saddle Node.}$$

$$\text{At} \quad x^1 = (0,1/\sigma), \qquad (DF_1) = \begin{bmatrix} \boldsymbol{\frac{1}{4}(\frac{1}{\sigma} - 3)} & 0 \\ -\frac{1}{2\sigma} & \boldsymbol{-1} \end{bmatrix}, \qquad x^1 \text{ is a Saddle Node.}$$

$$\text{At} \quad x^2 = (2(1-3\sigma),3), \qquad (DF_2) = \begin{bmatrix} 0 & \frac{1}{2}(1-3\sigma) \\ -\frac{3}{2} & -3\sigma \end{bmatrix}, \qquad x^2 \text{ depends on } \sigma.$$

The critical point x^0 is a Saddle Node, because the eigenvalues of the linearization, which are the coefficients in **red boldface**, satisfy $\lambda_- = -1 < 0 < \lambda_+ = 3/4$. The corresponding eigenvectors are
$$\boldsymbol{v}_- = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \boldsymbol{v}_+ = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$
The critical point x^1 is also a Saddle Node, because the eigenvalues of the linearization, which are the coefficients in **red boldface**, satisfy $\lambda_- = -3 < 0$ and $\lambda_+ = (1/\sigma - 3)/4 > 0$, since we assumed $\sigma < 1/3$. The corresponding eigenvectors are
$$\boldsymbol{v}_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \boldsymbol{v}_+ = \begin{bmatrix} (1+\sigma) \\ -2 \end{bmatrix}.$$
Finally, we focus on the critical point x^2. It takes some algebra to compute the eigenvalues of the linearization, which are
$$\lambda_\pm = \frac{3}{2}\big(-\sigma \pm \sqrt{\sigma^2 + \sigma - 1/3}\big).$$
Then, one can see that for $0 < \sigma < 1/3$ we have two cases:

- For $\sigma \in (0, \sigma_0)$, with $\sigma_0 = (-1 + \sqrt{7/3})/2 \simeq 0.2638$, the radicand $(\sigma^2 + \sigma - 1/3) < 0$, therefore the critical point x^2 is a Sink Spiral.
- For $\sigma \in (\sigma_0, 1/3)$, with σ_0 as above, the radicand $(\sigma^2 + \sigma - 1/3) > 0$, and also $\lambda_+ < 0$, therefore the critical point x^2 is a Sink Node.

◁

5.4.3. Nonlinear Pendulum. Consider a pendulum formed by a ball of mass m and a rod of length ℓ oscillating under the gravity acceleration g around a pivot point, in a medium with damping constant d, as given in Fig. 16. The forces acting on the ball are its weight

$$\boldsymbol{f}_g = -mg\,\boldsymbol{j},$$

the friction force which opposes the velocity of the ball,

$$\boldsymbol{f}_d = -d\,\boldsymbol{v},$$

where $\boldsymbol{v}$ is the velocity of the ball computed later on, and the tension $\boldsymbol{T}$ of the rod connecting the ball to the pivot. The position of the ball in the coordinate system xy given also in Fig. 16 is

$$\boldsymbol{r}(t) = \ell\,(\sin(\theta(t))\,\boldsymbol{i} - \cos(\theta(t))\,\boldsymbol{j}).$$

where $\boldsymbol{i}$ and $\boldsymbol{j}$ are the unit vectors along the axis x and y, respectively and $\theta(t)$ is the angle given in Fig. 16. It is convenient to introduce the radial unit vector

$$\boldsymbol{\rho}(\theta) = \sin(\theta)\,\boldsymbol{i} - \cos(\theta)\,\boldsymbol{j},$$

and the tangential unit vector $\boldsymbol{\tau}(\theta) = d\boldsymbol{\rho}/d\theta$, which is then given by

$$\boldsymbol{\tau}(\theta) = \cos(\theta)\,\boldsymbol{i} + \sin(\theta)\,\boldsymbol{j}.$$

Using this notation the position of the ball as function of time is

$$\boldsymbol{r}(t) = \ell\,\boldsymbol{\rho}(\theta(t)),$$

which we will write simply as $\boldsymbol{r} = \ell\,\boldsymbol{\rho}$. The velocity of the ball is

$$\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt} = \ell\Big(\frac{d\boldsymbol{\rho}}{d\theta}\Big)\,\theta'(t) = \ell\,\boldsymbol{\tau}\,\theta' \quad\Rightarrow\quad \boldsymbol{v} = \ell\,\theta'\,\boldsymbol{\tau}.$$

The acceleration of the ball is

$$\boldsymbol{a}(t) = \frac{d\boldsymbol{v}}{dt} = \ell\,\theta''\,\boldsymbol{\tau} + \ell\,(\theta')^2\,\frac{d\boldsymbol{\tau}}{d\theta}.$$

It is simple to see that $d\boldsymbol{\tau}/d\theta = -\boldsymbol{\rho}$, therefore

$$\boldsymbol{a} = \ell\,\theta''\,\boldsymbol{\tau} - \ell\,(\theta')^2\,\boldsymbol{\rho}.$$

The second Newton's law of motion for the ball says

$$m\,\boldsymbol{a} = \boldsymbol{f}_g + \boldsymbol{f}_d + \boldsymbol{T},$$

that is

$$m\ell\,\theta''\,\boldsymbol{\tau} - m\ell\,(\theta')^2\,\boldsymbol{\rho} = mg\,(\cos(\theta)\,\boldsymbol{\rho} - \sin(\theta)\,\boldsymbol{\tau}) - d\ell\,\theta'\,\boldsymbol{\tau} - T\,\boldsymbol{\rho}$$

where we used that $\boldsymbol{j} = -\cos(\theta)\,\boldsymbol{\rho} + \sin(\theta)\,\boldsymbol{\tau}$ and the tension force is radial $\boldsymbol{T} = T\,\boldsymbol{\rho}$. The equation above are actually two equations, one along the radial direction $\boldsymbol{\rho}$ and the other along the tangential direction $\boldsymbol{\tau}$. These equations are, respectively

$$T = mg\cos(\theta) + m\ell\,(\theta')^2,$$

which defines the tension force $\boldsymbol{T}$, and

$$m\ell\,\theta'' = -mg\sin(\theta) - d\ell\,\theta',$$

which determines the movement of the ball in the pendulum. This is a second order scalar equation for the angular position of the ball, $\theta(t)$, which can be rewritten as

$$\theta'' + \frac{d}{m}\,\theta' + \frac{g}{\ell}\,\sin(\theta) = 0. \tag{5.4.17}$$

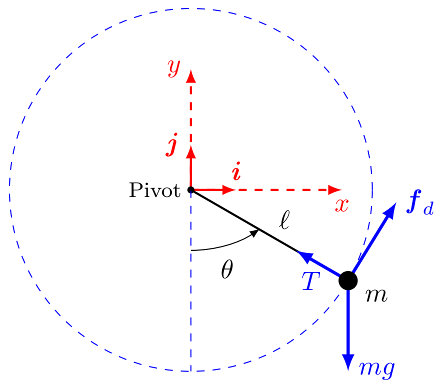


FIGURE 16. Pendulum with a ball of mass m, rod with length ℓ, oscillating around the pivot. We also show the tension force $\boldsymbol{T}$, the weight force $\boldsymbol{f}_g$ and the friction force $\boldsymbol{f}_d$. We assumed that the ball is moving downwards to graph the friction force $\boldsymbol{f}_d = -d\,\boldsymbol{v}$.

We want to construct a phase diagram for the pendulum. The first step is to write Eq. (5.4.17) as a 2×2 first order system. For this reason we introduce $x_1 = \theta$ and $x_2 = \theta'$, then the second order Eq. (5.4.17) can be written as

$$x_1' = x_2 \tag{5.4.18}$$

$$x_2' = -\frac{g}{\ell}\sin(x_1) - \frac{d}{m}\,x_2. \tag{5.4.19}$$

Remark: This is a nonlinear system. In the case of small oscillations we have the approximation $\sin(x_1) \sim x_1$, and we obtain the linear system.

$$x_1' = x_2$$

$$x_2' = -\frac{g}{\ell}\,x_1 - \frac{d}{m}\,x_2,$$

which we have studied in previous sections.

The phase space for this system is the plane x_1x_2, where $x_1 = \theta$ and $x_2 = \theta'$.

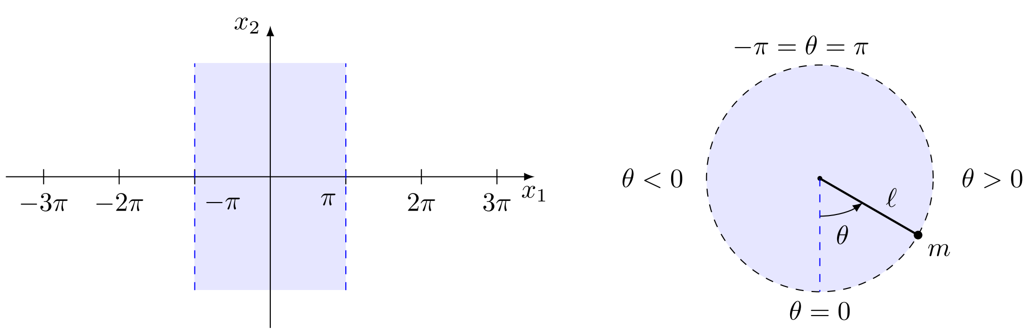


FIGURE 17. On the left we have the phase space of a pendulum.

If the pendulum does one complete turn counterclockwise and then stops at the downwards vertical position, this position is not $\theta = 0$, but $\theta = 2\pi$. Similarly, if the pendulum does one complete turn clockwise and then stops at the downwards vertical position, this position is not $\theta = 0$, but $\theta = -2\pi$.

Example 5.4.5. Sketch a phase diagram for a pendulum. Consider the particular case of $g/\ell = 1$, and mass $m = 1$. Make separate diagrams for the case of no damping, $d = 0$, and for the case with damping, $d > 0$.

Solution: The first order differential equations for the pendulum in this case are

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= -\sin(x_1) - d\,x_2. \end{aligned}$$

We start finding the critical points, which are the constants (x_1, x_2) solutions of

$$\left.\begin{aligned} x_2 &= 0 \\ -\sin(x_1) - d\,x_2 &= 0. \end{aligned}\right\} \quad \Rightarrow \quad \sin(x_1) = 0, \quad x_2 = 0.$$

Therefore, we get infinitely many critical point of the form

$$x^n = (n\pi, 0), \qquad n = 0, \pm 1, \pm 2, \cdots .$$

We now compute the linearization of Eqs.(5.4.18)-(5.4.19). We first compute the derivative of the field $\boldsymbol{F}$, where

$$\boldsymbol{F}(x) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\sin(x_1) - d\,x_2 \end{bmatrix}.$$

The derivative of $\boldsymbol{F}$ at an arbitrary point x is

$$DF(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\cos(x_1) & -d \end{bmatrix}.$$

We now evaluate the matrix $DF(x)$ at each of the critical points we find that

$$DF(x^n) = \begin{bmatrix} 0 & 1 \\ (-1)^{n+1} & -d \end{bmatrix}, \qquad n = 0, \pm 1, \pm 2, \cdots .$$

From here we see that we have two types of critical points, when n is even, $n = 2k$, and when n is odd, $n = 2k + 1$. The corresponding linearizations are

$$DF(x^{2k}) = \begin{bmatrix} 0 & 1 \\ -1 & -d \end{bmatrix}, \qquad DF(x^{2k+1}) = \begin{bmatrix} 0 & 1 \\ 1 & -d \end{bmatrix}. \tag{5.4.20}$$

In order to sketch the phase diagram, it is useful to consider three cases: (1) No damping, $d = 0$, (2) small damping, and (3) large damping.

In the case of no damping, $d = 0$, we have,

$$x^{2k} = (2k\pi, 0), \qquad DF_{2k} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \lambda_\pm = \pm i, \quad \boldsymbol{v}_\pm = \begin{bmatrix} \mp i \\ 1 \end{bmatrix}, \quad x^{2k} \text{ are Centers.}$$

$$x^{2k+1} = ((2k+1)\pi, 0), \quad DF_{2k+1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \lambda_\pm = \pm 1, \quad \boldsymbol{v}_\pm = \begin{bmatrix} \pm 1 \\ 0 \end{bmatrix}, \quad x^{2k+1} \text{ are Saddles.}$$

Notice that from the center critical points, x^{2k}, we can get the direction of the solution curves. Since the eigenvectors for these center critical points are

$$\boldsymbol{v}_+ = \begin{bmatrix} \mp i \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} i = \boldsymbol{a} + \boldsymbol{b}\,i \quad \Rightarrow \quad \boldsymbol{a} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

and the direction of increasing time is $\boldsymbol{a} \to -\boldsymbol{b}$, we get that around the centers the solution moves clockwise as time increases. If we put all this information together, we get the phase diagram in Fig. 18.

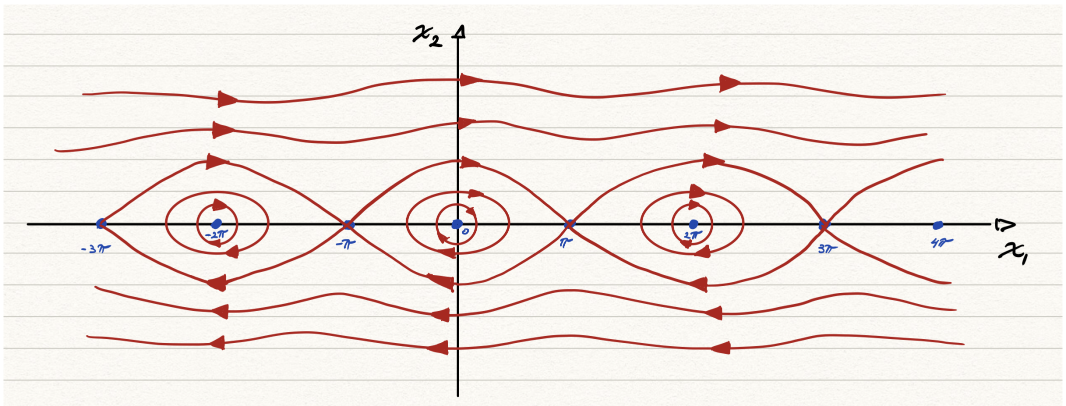


FIGURE 18. Phase diagram of a pendulum without damping, $d = 0$.

In the case that there is damping, either small or large, we need to go back to the linearizations in Eq. (5.4.20), and compute the eigenvalues and eigenvectors for the even critical points x^{2k} and the odd critical points x^{2k+1}. In the case of the odd critical points, x^{2k+1}, we get that they are Saddle Nodes, just as in the case $d = 0$, because the eigenvalues are

$$\lambda_{\pm} = \frac{1}{2}\big(-d \pm \sqrt{d^2 + 4}\big) \quad \Rightarrow \quad \lambda_- < 0 < \lambda_+ \quad \Rightarrow \quad \text{Saddle Nodes.}$$

So, the odd critical points are saddle nodes, no matter the value of of the damping constant d. The even critical points behave in a different way. In the case of even critical points, x^{2k}, the eigenvalues of the linearizations are

$$\lambda_{\pm} = \frac{1}{2}\big(-d \pm \sqrt{d^2 - 4}\big).$$

We see that we have two different types of eigenvalues, depending whether $d^2 - 4 < 0$ or $d^2 - 4 \geqslant 0$. Since we assume here that $d > 0$, these two cases are, respectively,

$$0 < d < 2, \quad \text{or} \quad d \geqslant 2.$$

The first case above is called *small damping*, and in this case the eigenvalues of the linearization are complex valued,

$$\lambda_{\pm} = \frac{1}{2}\big(-d \pm \sqrt{4 - d^2}\, i\big), \quad 0 < d < 2 \quad \Rightarrow \quad \text{Sink Spirals.}$$

The second case above is called *large damping*, and in this case the eigenvalues of the linearizations are real and negative,

$$\lambda_{\pm} = \frac{1}{2}\big(-d \pm \sqrt{d^2 - 4}\big), \quad d \geqslant 2 \quad \Rightarrow \quad \text{Sink Nodes.}$$

In Fig. 19 we sketch the phase diagram of a pendulum with small damping.

◁

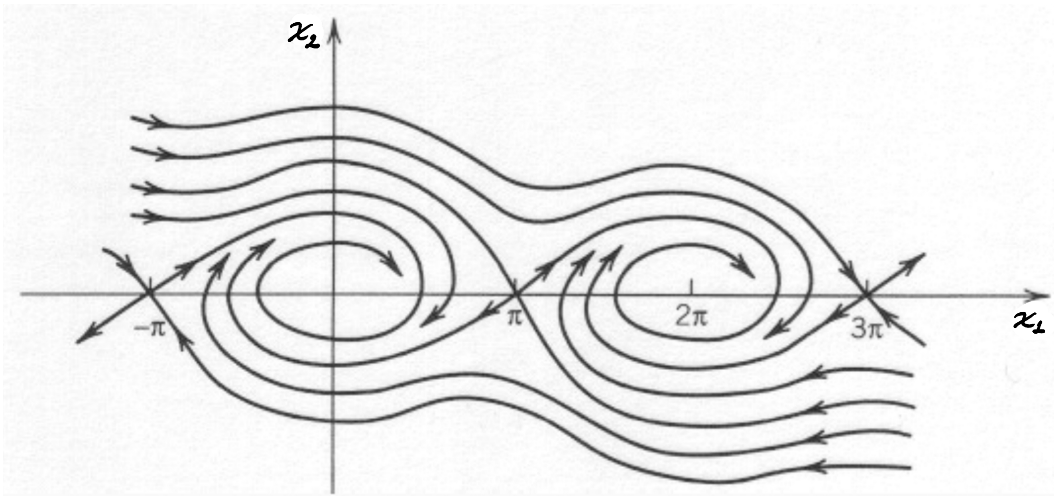


FIGURE 19. Phase diagram of a pendulum with small damping, $0 < d < 2$.

5.4.4. Exercises.

5.4.1.- .

5.4.2.- .

CHAPTER 6

Boundary Value Problems

In this chapter we focus on boundary value problems both for ordinary differential equations and for a particular partial differential equation, the heat equation. We start introducing boundary value problems for ordinary differential equations. We then introduce a particular type of boundary value problem, eigenfunction problems. Later on we introduce more general eigenfunction problems, the Sturm-Liouville Problem. We show that solutions to Sturm-Liouville problems have two important properties, completeness and orthognonality. Wwe use these results to introduce the Fourier series expansion of continuous and discontinuous functions. We end this chapter introducing the separation of variables method to find solutions of a partial differential equation, the heat equation.

6.1. Eigenfunction Problems

In this Section we consider second order, linear, ordinary differential equations. In the first half of the Section we study boundary value problems for these equations and in the second half we focus on a particular type of boundary value problems, called the eigenvalue-eigenfunction problem for these equations.

6.1.1. Two-Point Boundary Value Problems.

We start with the definition of a two-point boundary value problem.

Definition 6.1.1. *A* ***two-point boundary value problem*** *(BVP) is the following: Find solutions to the differential equation*

$$y'' + a_1(x)\, y' + a_0(x)\, y = f(x)$$

satisfying the boundary conditions (BC)

$$b_1\, y(x_1) + b_2\, y'(x_1) = y_1,$$
$$\tilde{b}_1\, y(x_2) + \tilde{b}_2\, y'(x_2) = y_2,$$

where b_1, b_2, $\tilde{b}_1$, $\tilde{b}_2$, x_1, x_2, y_1, and y_2 are given and $x_1 \neq x_2$. The boundary conditions are ***homogeneous*** *iff $y_1 = 0$ and $y_2 = 0$.*

Remarks:

(a) The two boundary conditions are held at *different* points, $x_1 \neq x_2$.
(b) Both y and y' may appear in the boundary condition.

Example 6.1.1 (The Four Main Problems). We now show four examples of boundary value problems that differ only on the boundary conditions: Solve the different equation

$$y'' + a_1(x)\, y' + a_0(x)\, y = f(x)$$

with the boundary conditions at $x_1 = 0$ and $x_2 = L$ given below.

(a)

$$\text{Dirichlet BC:} \quad \begin{Bmatrix} y(0) = y_1, \\ y(L) = y_2, \end{Bmatrix} \quad \text{which is the case} \quad \begin{Bmatrix} b_1 = 1, & b_2 = 0, \\ \tilde{b}_1 = 1, & \tilde{b}_2 = 0. \end{Bmatrix}$$

(b)

$$\text{Neumann BC:} \quad \begin{Bmatrix} y'(0) = y_1, \\ y'(L) = y_2, \end{Bmatrix} \quad \text{which is the case} \quad \begin{Bmatrix} b_1 = 0, & b_2 = 1, \\ \tilde{b}_1 = 0, & \tilde{b}_2 = 1. \end{Bmatrix}$$

(c)

$$\text{Mixed (1) BC:} \quad \begin{Bmatrix} y(0) = y_1, \\ y'(L) = y_2, \end{Bmatrix} \quad \text{which is the case} \quad \begin{Bmatrix} b_1 = 1, & b_2 = 0, \\ \tilde{b}_1 = 0, & \tilde{b}_2 = 1. \end{Bmatrix}$$

(d)

$$\text{Mixed (2) BC:} \quad \begin{Bmatrix} y'(0) = y_1, \\ y(L) = y_2, \end{Bmatrix} \quad \text{which is the case} \quad \begin{Bmatrix} b_1 = 0, & b_2 = 1, \\ \tilde{b}_1 = 1, & \tilde{b}_2 = 0. \end{Bmatrix}$$

The most general case is when each boundary condition contain both y and y'. Below is one example of this case.

$$\text{Steklov BC:} \quad \begin{Bmatrix} 2\, y(0) + y'(0) = y_1, \\ y'(L) + 3\, y'(L) = y_2, \end{Bmatrix} \quad \text{which is the case} \quad \begin{Bmatrix} b_1 = 2, & b_2 = 1, \\ \tilde{b}_1 = 1, & \tilde{b}_2 = 3. \end{Bmatrix}$$

◁

6.1.2. Comparison: IVP and BVP. We now review the initial boundary value problem for the equation above, which was discussed in Sect. 2.1, where we showed in Theorem 2.1.3 that this initial value problem always has a unique solution.

Definition 6.1.2 (IVP)**.** *Find all solutions of the differential equation $y'' + a_1\, y' + a_0\, y = 0$ satisfying the initial condition (IC)*

$$y(t_0) = y_0, \qquad y'(t_0) = y_1. \tag{6.1.1}$$

Remarks: In an initial value problem we usually the following happens.

- The variable t represents time.
- The variable y represents position.
- The IC are position and velocity at the initial time.

A typical boundary value problem that appears in many applications is the following.

Definition 6.1.3 (BVP)**.** *Find all solutions of the differential equation $y'' + a_1\, y' + a_0\, y = 0$ satisfying the boundary condition (BC)*

$$y(0) = y_0, \qquad y(L) = y_1, \qquad L \neq 0. \tag{6.1.2}$$

Remarks: In a boundary value problem we usually the following happens.

- The variable x represents position.
- The variable y may represents a physical quantity such us temperature.
- The BC are the temperature at two different positions.

The names "initial value problem" and "boundary value problem" come from physics. An example of the former is to solve Newton's equations of motion for the position function of a point particle that starts at a given initial position and velocity. An example of the latter is to find the equilibrium temperature of a cylindrical bar with thermal insulation on the round surface and held at constant temperatures at the top and bottom sides.

Let's recall an important result we saw in § 2.1 about solutions to initial value problems.

Theorem 6.1.4 (IVP)**.** *The equation $y'' + a_1\, y' + a_0\, y = 0$ with IC $y(t_0) = y_0$ and $y'(t_0) = y_1$ has a unique solution y for each choice of the IC.*

The solutions to boundary value problems are more complicated to describe. A boundary value problem may have a unique solution, or may have infinitely many solutions, or may have no solution, depending on the boundary conditions. In the case of the boundary value problem in Def. 6.1.3 we get the following.

Theorem 6.1.5 (BVP)**.** *Consider the differential equation*

$$y'' + a_1\, y' + a_0\, y = 0$$

with Dirichlet boundary conditions

$$y(0) = y_0, \qquad y(L) = y_1,$$

where $L \neq 0$. Let $r_\pm$ be the solutions of the characteristic equation

$$r^2 + a_1 r + a_0 = 0.$$

Then we have the following:

(A) If r_+, r_- are real, then the BVP above has a unique solution for all y_0, $y_1 \in \mathbb{R}$.

(B) If $r_\pm = \alpha \pm i\,\beta$ are complex, with α, $\beta \in \mathbb{R}$, then the solution of the BVP above belongs to one of the following three possibilities:

(i) There exists a unique solution;

(ii) There exists infinitely many solutions;

(iii) There exists no solution.

Proof of Theorem 6.1.5:

Part (A): Let r_+, r_- be solutions of the characteristic equation

$$r^2 + a_1\,r + a_0 = 0.$$

Let us assume that $r_+ \neq r_-$. Then, the general solution of the differential equation is

$$y(x) = c_+\,e^{r_+ x} + c_-\,e^{r_- x}.$$

The boundary conditions are

$$\left.\begin{aligned} y_0 &= y(0) = c_+ + c_- \\ y_1 &= y(L) = c_+\,e^{c_+ L} + c_-\,e^{c_- L} \end{aligned}\right\} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 1 \\ e^{r_+ L} & e^{r_- L} \end{bmatrix} \begin{bmatrix} c_+ \\ c_- \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}.$$

Let's denote the coefficient matrix in this equation as

$$P = \begin{bmatrix} 1 & 1 \\ e^{r_+ L} & e^{r_- L} \end{bmatrix}.$$

This system of algebraic equations for the coefficients c_+, c_- has a unique solution iff the coefficient matrix P is invertible. When the inverse matrix exists the solution for the coefficients is

$$\begin{bmatrix} c_+ \\ c_- \end{bmatrix} = P^{-1} \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}.$$

To find out whether P is invertible we compute its determinant,

$$\det(P) = \begin{vmatrix} 1 & 1 \\ e^{r_+ L} & e^{r_- L} \end{vmatrix} = e^{r_- L} - e^{r_+ L}.$$

If the roots r_+, r_- are real and $r_+ \neq r_-$, then $e^{r_- L} \neq e^{r_+ L}$. This implies that $\det(P) \neq 0$, then P^{-1} exists, and that gives us a unique solution c_+, c_-. We conclude that there is a unique solution y of the BVP.

Now, let us assume that $r_+ = r_- = r_0$. Then, the general solution of the differential equation is

$$y(x) = (c_1 + c_2\,x)\,e^{r_0 x}, \qquad c_1, c_2 \in \mathbb{R}.$$

Again, the boundary conditions in Eq. (6.1.2) determine the values of the constants c_1 and c_2 as follows:

$$\left.\begin{aligned} y_0 &= y(0) = c_1 \\ y_1 &= y(L) = c_1 e^{r_0 L} + c_2 L\,e^{r_0 L} \end{aligned}\right\} \quad \Rightarrow \quad \begin{bmatrix} 1 & 0 \\ e^{r_0 L} & L\,e^{r_0 L} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}.$$

Let's denote the coefficient matrix in this equation as

$$Q = \begin{bmatrix} 1 & 0 \\ e^{r_0 L} & L\,e^{r_0 L} \end{bmatrix}.$$

To find out whether Q is invertible we compute its determinant,

$$\det(Q) = \begin{vmatrix} 1 & 0 \\ e^{r_0 L} & L\,e^{r_0 L} \end{vmatrix} = L\,e^{r_0 L}.$$

We see that for $L \neq 0$ we get $\det(Q) \neq 0$, meaning that Q^{-1} exists, which says that there is a unique solution c_1, c_2. We conclude that there is a unique solution y of the BVP.

Part (B): Let r_+, r_- be solutions of the characteristic equation

$$r^2 + a_1\, r + a_0 = 0.$$

Let us assume that $r_\pm - \alpha \pm \beta i$. Then, the general solution of the differential equation is

$$y(x) = c_+\, e^{r_+ x} + c_-\, e^{r_- x}.$$

The same calculation we did in part (A) implies that c_+, c_- must be solution of

$$\begin{bmatrix} 1 & 1 \\ e^{r_+ L} & e^{r_- L} \end{bmatrix} \begin{bmatrix} c_+ \\ c_- \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}.$$

Once again, let's denote the coefficient matrix in this equation as

$$P = \begin{bmatrix} 1 & 1 \\ e^{r_+ L} & e^{r_- L} \end{bmatrix}.$$

This system of algebraic equations for the coefficients c_+, c_- has a unique solution iff the coefficient matrix P is invertible. To find out whether P is invertible we compute its determinant,

$$\det(P) = \begin{vmatrix} 1 & 1 \\ e^{r_+ L} & e^{r_- L} \end{vmatrix} = e^{r_- L} - e^{r_+ L}.$$

In this case the roots are complex valued, so we have that

$$e^{r_\pm L} = e^{(\alpha \pm i\beta)L} = e^{\alpha L}(\cos(\beta L) \pm i \sin(\beta L)),$$

therefore

$$\begin{aligned}
\det(P) &= e^{r_- L} - e^{r_+ L} \\
&= e^{\alpha L}\big(\cos(\beta L) - i\sin(\beta L)\big) - e^{\alpha L}\big(\cos(\beta L) + i\sin(\beta L)\big) \\
&= -2i\, e^{\alpha L}\, \sin(\beta L).
\end{aligned}$$

We conclude that

$$\det(P) = -2i\, e^{\alpha L}\, \sin(\beta L) = 0 \quad \Leftrightarrow \quad \beta L = n\pi.$$

So for $\beta L \neq n\pi$ the BVP has a unique solution, case (Bi). But for $\beta L = n\pi$ the BVP has either no solution or infinitely many solutions, cases (Bii) and (Biii). This establishes the Theorem. □

Example 6.1.2. Find all solutions to the BVPs $y'' + y = 0$ with the BCs:

$$\textbf{BC (a)} \quad \begin{cases} y(0) = 1, \\ y(\pi) = 0. \end{cases} \qquad \textbf{BC (b)} \quad \begin{cases} y(0) = 1, \\ y(\pi/2) = 1. \end{cases} \qquad \textbf{BC (c)} \quad \begin{cases} y(0) = 1, \\ y(\pi) = -1. \end{cases}$$

Solution: We first find the roots of the characteristic polynomial $r^2 + 1 = 0$, that is, $r_\pm = \pm i$. So the general solution of the differential equation is

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

BC (a):

$$\begin{aligned}
1 = y(0) = c_1 \quad &\Rightarrow \quad c_1 = 1. \\
0 = y(\pi) = -c_1 \quad &\Rightarrow \quad c_1 = 0.
\end{aligned}$$

Therefore, there is no solution.

BC (b):

$$1 = y(0) = c_1 \quad \Rightarrow \quad c_1 = 1.$$

$$1 = y(\pi/2) = c_2 \quad \Rightarrow \quad c_2 = 1.$$

So there is a unique solution $y(x) = \cos(x) + \sin(x)$.
BC (c):

$$1 = y(0) = c_1 \quad \Rightarrow \quad c_1 = 1.$$

$$-1 = y(\pi) = -c_1 \quad \Rightarrow \quad c_1 = 1.$$

Therefore, c_2 is arbitrary, so we have infinitely many solutions

$$y(x) = \cos(x) + c_2 \sin(x), \qquad c_2 \in \mathbb{R}.$$

$\lhd$

Example 6.1.3. Find all solutions to the BVPs $y'' + 4\,y = 0$ with the BCs:

$$\textbf{BC (a)} \quad \begin{cases} y(0) = 1, \\ y(\pi/4) = -1. \end{cases} \qquad \textbf{BC (b)} \quad \begin{cases} y(0) = 1, \\ y(\pi/2) = -1. \end{cases} \qquad \textbf{BC (c)} \quad \begin{cases} y(0) = 1, \\ y(\pi/2) = 1. \end{cases}$$

Solution: We first find the roots of the characteristic polynomial $r^2 + 4 = 0$, that is, $r_\pm = \pm 2i$. So the general solution of the differential equation is

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

BC (a):

$$1 = y(0) = c_1 \quad \Rightarrow \quad c_1 = 1.$$

$$-1 = y(\pi/4) = c_2 \quad \Rightarrow \quad c_2 = -1.$$

Therefore, there is a unique solution $y(x) = \cos(2x) - \sin(2x)$.
BC (b):

$$1 = y(0) = c_1 \quad \Rightarrow \quad c_1 = 1.$$

$$-1 = y(\pi/2) = -c_1 \quad \Rightarrow \quad c_1 = 1.$$

So, c_2 is arbitrary and we have infinitely many solutions

$$y(x) = \cos(2x) + c_2 \sin(2x), \qquad c_2 \in \mathbb{R}.$$

BC (c):

$$1 = y(0) = c_1 \quad \Rightarrow \quad c_1 = 1.$$

$$1 = y(\pi/2) = -c_1 \quad \Rightarrow \quad c_1 = -1.$$

Therefore, we have no solution. $\lhd$

6.1.3. Simple Eigenfunction Problems. We now focus on boundary value problems that have infinitely many solutions. A particular type of these problems are called an eigenfunction problems. They are similar to the eigenvector problems we studied in § 4.3. Recall that the eigenvector problem is the following: Given an $n \times n$ matrix A, find all numbers λ and nonzero vectors $\boldsymbol{v}$ solution of the algebraic linear system

$$A\boldsymbol{v} = \lambda \boldsymbol{v}.$$

We saw that for each λ there are infinitely many solutions $\boldsymbol{v}$, because if $\boldsymbol{v}$ is a solution so is any multiple $a\boldsymbol{v}$. An eigenfunction problem is something similar.

Definition 6.1.6. *An* ***eigenfunction problem*** *is the following: Given a linear operator*

$$L(y) = a_2\, y'' + a_1\, y' + a_0\, y,$$

find all numbers λ *and a nonzero functions* y *solution of the differential equation*

$$L(y) = \lambda\, y,$$

with homogeneous boundary conditions

$$b_1\, y(x_1) + b_2\, y'(x_1) = 0,$$
$$\tilde{b}_1\, y(x_2) + \tilde{b}_2\, y'(x_2) = 0.$$

Remarks:

- Notice that $y = 0$ is always a solution of the BVP above.
- Eigenfunctions are the nonzero solutions of the BVP above.
- Hence, the eigenfunction problem is a BVP with infinitely many solutions.
- In the case that $L(y) = -y''$ with Dirichlet boundary conditions, that is, the operator in Theorem 6.1.5, we look for λ such that the operator $L(y) - \lambda\, y$ has characteristic polynomial with complex roots, that is, $L(y) - \lambda\, y$ has oscillatory solutions.
- Most of our examples focus on the linear operator $L(y) = -y''$, which is the one that appears when we solve the heat equation.

Example 6.1.4. Find the eigenvalues and eigenfunctions of the operator $L(y) = -y''$, with $y(0) = 0$ and $y(L) = 0$. This is equivalent to finding all numbers λ and nonzero functions y solutions of the BVP

$$-y'' = \lambda\, y, \quad \text{with (homogeneous) Dirichlet BC} \quad y(0) = 0, \quad y(L) = 0, \qquad L > 0.$$

Solution: We divide the problem in three cases: **(a)** $\lambda < 0$, **(b)** $\lambda = 0$, and **(c)** $\lambda > 0$.

Case (a): $\lambda = -\mu^2 < 0$, so the differential equation is

$$y'' - \mu^2 y = 0.$$

The characteristic equation is

$$r^2 - \mu^2 = 0 \quad \Rightarrow \quad r_\pm = \pm\mu.$$

The general solution is $y = c_+\, e^{\mu x} + c_-\, e^{-\mu x}$. The BC imply

$$0 = y(0) = c_+ + c_-, \qquad 0 = y(L) = c_+\, e^{\mu L} + c_-\, e^{-\mu L}.$$

So from the first equation we get $c_+ = -c_-$, so

$$0 = -c_-\, e^{\mu L} + c_-\, e^{-\mu L} \quad \Rightarrow \quad -c_-(e^{\mu L} - e^{-\mu L}) = 0 \quad \Rightarrow \quad c_- = 0, \quad c_+ = 0.$$

So the only the solution is $y = 0$, then there are no eigenfunctions with negative eigenvalues.

Case (b): $\lambda = 0$, so the differential equation is

$$y'' = 0 \quad \Rightarrow \quad y = c_0 + c_1 x.$$

The BC imply

$$0 = y(0) = c_0, \qquad 0 = y(L) = c_1 L \quad \Rightarrow \quad c_1 = 0.$$

So the only solution is $y = 0$, then there are no eigenfunctions with eigenvalue $\lambda = 0$.

Case (c): $\lambda = \mu^2 > 0$, so the differential equation is

$$y'' + \mu^2 y = 0.$$

The characteristic equation is

$$r^2 + \mu^2 = 0 \quad \Rightarrow \quad r_\pm = \pm \mu i.$$

The general solution is $y = c_+ \cos(\mu x) + c_- \sin(\mu x)$. The BC imply

$$0 = y(0) = c_+, \qquad 0 = y(L) = c_+ \cos(\mu L) + c_- \sin(\mu L).$$

Since $c_+ = 0$, the second equation above is

$$c_- \sin(\mu L) = 0, \quad c_- \neq 0 \quad \Rightarrow \quad \sin(\mu L) = 0 \quad \Rightarrow \quad \mu_n L = n\pi, \quad n = 1, 2, 3, \cdots.$$

So we get $\mu_n = n\pi/L$, hence the eigenvalue eigenfunction pairs are

$$\lambda_n = \Big(\frac{n\pi}{L}\Big)^2, \qquad y_n(x) = c_n \sin\Big(\frac{n\pi x}{L}\Big).$$

Since we need only one eigenfunction for each eigenvalue, we choose $c_n = 1$, and we get

$$\lambda_n = \Big(\frac{n\pi}{L}\Big)^2, \qquad y_n(x) = \sin\Big(\frac{n\pi x}{L}\Big), \qquad n = 1, 2, 3, \cdots.$$

◁

Example 6.1.5. Find the numbers λ and the nonzero functions y solutions of the BVP

$$-y'' = \lambda y, \qquad y(0) = 0, \qquad y'(L) = 0, \qquad L > 0.$$

Solution: We divide the problem in three cases: **(a)** $\lambda < 0$, **(b)** $\lambda = 0$, and **(c)** $\lambda > 0$.

Case (a): Let $\lambda = -\mu^2$, with $\mu > 0$, so the differential equation is

$$y'' - \mu^2 y = 0.$$

The characteristic equation is

$$r^2 - \mu^2 = 0 \quad \Rightarrow \quad r_\pm = \pm \mu,$$

The general solution is $y(x) = c_1 e^{-\mu x} + c_2 e^{\mu x}$. The BC imply

$$\left.\begin{aligned} 0 = y(0) &= c_1 + c_2, \\ 0 = y'(L) &= -\mu c_1 e^{-\mu L} + \mu c_2 e^{\mu L} \end{aligned}\right\} \quad \Rightarrow \quad \begin{bmatrix} 1 & 1 \\ -\mu e^{-\mu L} & \mu e^{\mu L} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The matrix above is invertible, because

$$\begin{vmatrix} 1 & 1 \\ -\mu e^{-\mu L} & \mu e^{\mu L} \end{vmatrix} = \mu\big(e^{\mu L} + e^{-\mu L}\big) \neq 0.$$

So, the linear system above for c_1, c_2 has a unique solution $c_1 = c_2 = 0$. Hence, we get the only solution $y = 0$. This means there are no eigenfunctions with negative eigenvalues.

Case (b): Let $\lambda = 0$, so the differential equation is

$$y'' = 0 \quad \Rightarrow \quad y(x) = c_1 + c_2 x, \qquad c_1, c_2 \in \mathbb{R}.$$

The boundary conditions imply the following conditions on c_1 and c_2,

$$0 = y(0) = c_1, \qquad 0 = y'(L) = c_2.$$

So the only solution is $y = 0$. This means there are no eigenfunctions with eigenvalue $\lambda = 0$.

Case (c): Let $\lambda = \mu^2$, with $\mu > 0$, so the differential equation is

$$y'' + \mu^2 y = 0.$$

The characteristic equation is

$$r^2 + \mu^2 = 0 \quad \Rightarrow \quad r_\pm = \pm\mu\, i.$$

The general solution is $y(x) = c_1 \cos(\mu x) + c_2 \sin(\mu x)$. The BC imply

$$\left.\begin{aligned} 0 &= y(0) = c_1, \\ 0 &= y'(L) = -\mu c_1 \sin(\mu L) + \mu c_2 \cos(\mu L) \end{aligned}\right\} \quad \Rightarrow \quad c_2 \cos(\mu L) = 0.$$

Since we are interested in non-zero solutions y, we look for solutions with $c_2 \neq 0$. This implies that μ cannot be arbitrary but must satisfy the equation

$$\cos(\mu L) = 0 \quad \Leftrightarrow \quad \mu_n L = (2n-1)\frac{\pi}{2}, \qquad n = 1, 2, 3, \cdots.$$

We therefore conclude that the eigenvalues and eigenfunctions are given by

$$\lambda_n = \frac{(2n-1)^2\pi^2}{4L^2}, \qquad y_n(x) = c_n \sin\Big(\frac{(2n-1)\pi x}{2L}\Big), \qquad n = 1, 2, 3, \cdots.$$

Since we only need one eigenfunction for each eigenvalue, we choose $c_n = 1$, and we get

$$\lambda_n = \frac{(2n-1)^2\pi^2}{4L^2}, \qquad y_n(x) = \sin\Big(\frac{(2n-1)\pi x}{2L}\Big), \qquad n = 1, 2, 3, \cdots.$$

◁

In the next example we see how these calculations of eigenvalues and eigenfunctions can be generalized to more complicated operators, such as $L(y) = -x^2\, y'' + x\, y$.

Example 6.1.6. Find the numbers λ and the nonzero functions y solutions of the BVP

$$-x^2\, y'' + x\, y' = \lambda\, y, \qquad y(1) = 0, \quad y(L) = 0, \quad L > 1.$$

Solution: Let us rewrite the equation as

$$x^2\, y'' - x\, y' + \lambda y = 0.$$

This equation is called an Euler equidimensional equation, and by trial and error one can find that solutions are given by

$$y_1(x) = x^{r_+}, \qquad y_2(x) = x^{r_-},$$

where the power $r_\pm$ are the solutions of the indicial polynomial

$$r(r-1) - r + \lambda = 0 \quad \Rightarrow \quad r^2 - 2r + \lambda = 0 \quad \Rightarrow \quad r_\pm = 1 \pm \sqrt{1-\lambda}.$$

In the case that $r_\pm = r_0$, the root is repeated and one can check that the fundamental solutions are given by

$$y_1(x) = x^{r_0}, \qquad y_2(x) = x^{r_0} \ln(x),$$

Case (a): Let $1 - \lambda = 0$, so we have a repeated root $r_+ = r_- = 1$. In this case the general solution to the differential equation is

$$y(x) = \big(c_1 + c_2 \ln(x)\big)\, x.$$

The boundary conditions imply the following conditions on c_1 and c_2,

$$\left.\begin{aligned} 0 &= y(1) = c_1, \\ 0 &= y(L) = \big(c_1 + c_2 \ln(L)\big)\, L \end{aligned}\right\} \quad \Rightarrow \quad c_2\, L\, \ln(L) = 0 \quad \Rightarrow \quad c_2 = 0.$$

Therefore, we see that the only solution is $y = 0$. This means there are no eigenfunctions with eigenvalue $\lambda = 1$.

Case (b): Let $1-\lambda > 0$, so we can rewrite it as $1-\lambda = \mu^2$, with $\mu > 0$. Recalling that $r_\pm = 1 \pm \sqrt{1-\lambda}$ we get that

$$r_\pm = 1 \pm \mu,$$

hence the general solution to the differential equation is given by

$$y(x) = c_1 x^{(1-\mu)} + c_2 x^{(1+\mu)},$$

The boundary conditions imply the following conditions on c_1 and c_2,

$$\left.\begin{aligned} 0 &= y(1) = c_1 + c_2, \\ 0 &= y(L) = c_1 L^{(1-\mu)} + c_2 L^{(1+\mu)} \end{aligned}\right\} \quad \Rightarrow \quad \begin{bmatrix} 1 & 1 \\ L^{(1-\mu)} & L^{(1+\mu)} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The matrix above is invertible, because

$$\begin{vmatrix} 1 & 1 \\ L^{(1-\mu)} & L^{(1+\mu)} \end{vmatrix} = L\left(L^{\mu} - L^{-\mu}\right) \neq 0 \quad \Leftrightarrow \quad L \neq \pm 1.$$

Since we assumed that $L > 1$, the matrix above is invertible, and the linear system for the constants c_1, c_2 has a unique solution given by

$$c_1 = c_2 = 0.$$

Hence we get the only solution is $y = 0$. This means there are no eigenfunctions with eigenvalues $\lambda < 1$.

Case (c): Let $1-\lambda < 0$, so we can rewrite it as $1-\lambda = -\mu^2$, with $\mu > 0$. Recalling that $r_\pm = 1 \pm \sqrt{1-\lambda}$, we get that

$$r_\pm = 1 \pm i\mu.$$

In this cae it can be shown that the general solution of the differential equation can be written as

$$y(x) = x\big[c_1 \cos\big(\mu \ln(x)\big) + c_2 \sin\big(\mu \ln(x)\big)\big].$$

The boundary conditions imply the following conditions on c_1 and c_2,

$$\left.\begin{aligned} 0 &= y(1) = c_1, \\ 0 &= y(L) = c_1 L \cos\big(\mu \ln(L)\big) + c_2 L \sin\big((\mu \ln(L)\big) \end{aligned}\right\} \quad \Rightarrow \quad c_2 L \sin\big(\mu \ln(L)\big) = 0.$$

Since we are interested in nonzero solutions y, we look for solutions with $c_2 \neq 0$. This implies that μ cannot be arbitrary but must satisfy the equation

$$\sin\big(\mu \ln(L)\big) = 0 \quad \Leftrightarrow \quad \mu_n \ln(L) = n\pi, \qquad n = 1, 2, 3, \cdots.$$

Recalling that $1 - \lambda_n = -\mu_n^2$, we get $\lambda_n = 1 + \mu_n^2$, hence,

$$\lambda_n = 1 + \frac{n^2\pi^2}{\ln^2(L)}, \qquad y_n(x) = c_n x \sin\Big(\frac{n\pi \ln(x)}{\ln(L)}\Big), \qquad n = 1, 2, 3, \cdots.$$

Since we only need one eigenfunction for each eigenvalue, we choose $c_n = 1$, and we get

$$\lambda_n = 1 + \frac{n^2\pi^2}{\ln^2(L)}, \qquad y_n(x) = x \sin\Big(\frac{n\pi \ln(x)}{\ln(L)}\Big), \qquad n = 1, 2, 3, \cdots.$$

◁

6.1.4. Exercises.

6.1.1.- .

6.1.2.- .

6.2. Sturm-Liouville Problems

In this section we introduce orthogonal functions and then we solve Sturm-Liouville problems, which are eigenfunction problems for Sturm-Liouville operators. The latter are second order linear operators with a very interesting property: the eigenvalues of the operator are real valued and eigenfunctions of different eigenvalues are orthogonal.

6.2.1. Orthogonal Functions.

The idea of orthogonal functions originates with orthogonal vectors in $\mathbb{R}^n$, in particular, orthogonal vectors in $\mathbb{R}^3$. We saw that the geometric properties of vector projections in three-dimensional space can be captured by the dot product of vectors in $\mathbb{R}^3$, more precisely, by the following three analytic properties of the dot product: positivity, symmetry, and linearity. Then, inspired by the dot product of vectors, we introduce a dot product of functions—called inner product—such that it has these same three analytic properties: positivity, symmetry, and linearity.

One important difference between $\mathbb{R}^3$ and the space of functions is that in three-dimensional space we have a geometric intuition of perpendicular vectors but in the space of functions we do not have such geometric intuition of perpendicular functions. So, instead of geometric intuition, we use the inner product of functions to *define* orthogonal functions.

Notice that there are infinitely many products of two vectors in $\mathbb{R}^3$ such that the result is a number and the product has the positivity, symmetry, and linearity properties. Our geometric intuition leads us to select one, the one we called the dot product. In the space of functions there are also infinitely many products of two functions such that the result is a number and the product has the positivity, symmetry, and linearity properties. The lack of geometric intuition in the space of functions leads us to choose the simplest looking product, and here it is.

Definition 6.2.1. *An **inner product** of two functions f, g on a non-empty $[a,b] \subset \mathbb{R}$, with weight function r, is*

$$f \cdot g = \int_a^b r(x)\, f(x)\, g(x)\, dx,$$

where the weight function is positive, that is $r(x) > 0$ for all $x \in [a,b]$.

Remarks:

(a) We have a different inner product for each choice of the weight function r.

(b) The case $r(x) = 1$ for all $x \in [a,b]$ gives a simple generalization of the dot product for vectors. The dot product of two vectors $\boldsymbol{u} = \langle u_1, u_2, u_3 \rangle$ and $\boldsymbol{v} = \langle v_1, v_2, v_3 \rangle$ is

$$\boldsymbol{u} \cdot \boldsymbol{v} = \sum_{i=1}^{3} u_i\, v_i.$$

The inner product of two functions f and g on the interval $[a,b]$, with weight function $r = 1$ is

$$f \cdot g = \int_a^b f(x)\, g(x)\, dx.$$

The values $f(x)$ and $g(x)$ for every $x \in [a,b]$ play the role of the vector components, while the integral on the interval $[a,b]$ is the continuum analog of the sum over all components.

The dot product in Def. 6.2.1 above takes two functions and produces a number. And one can verify that the product has the following properties.

Theorem 6.2.2. *For every functions f, g, h and every a, $b \in \mathbb{R}$ holds,*

(a) *Positivity:* $f \cdot f = 0$ *iff* $f = 0$*; and* $f \cdot f > 0$ *for* $f \neq 0$.
(b) *Symmetry:* $f \cdot g = g \cdot f$.
(c) *Linearity:* $(a\,f + b\,g) \cdot h = a\,(f \cdot h) + b\,(g \cdot h)$.

Remark: The proof is not difficult and it is left as an exercise.

The *magnitude* of a function f is the nonnegative number

$$\|f\| = \sqrt{f \cdot f} = \Big(\int_a^b r(x)\,\big(f(x)\big)^2\,dx \Big)^{1/2}.$$

We use a double bar to denote the magnitude so we do not confuse it with $|f|$, which means the absolute value. A function f is a *unit function* iff $f \cdot f = 1$. Since we do not have a geometric intuition of perpendicular functions, we *define* perpendicular functions using the dot product.

Definition 6.2.3. *Two functions f, g are **orthogonal** (perpendicular) iff* $f \cdot g = 0$.

A set of functions is an *orthogonal set* if all the functions in the set are mutually perpendicular. An *orthonormal set* is an orthogonal set where all the functions are unit functions.

Example 6.2.1 (Legendre Polynomials)**.** Show that the the first three Legendre polynomials below from an orthogonal set under the inner product in Def. 6.2.1 on the interval $[-1, 1]$ with weight function $r = 1$. Are these polynomials unit functions?

$$p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = \frac{1}{2}\big(3x^2 - 1\big).$$

Solution: We need to show that these functions are mutually perpendicular, that is,

$$p_0 \cdot p_1 = 0, \quad p_0 \cdot p_2 = 0, \quad p_1 \cdot p_2 = 0.$$

We start showing that p_0 is orthogonal to p_1, since

$$p_0 \cdot p_1 = \int_{-1}^1 x\,dx = \frac{x^2}{2}\Big|_{-1}^1 = 0.$$

Now we show that p_0 is orthogonal to p_2, since

$$p_0 \cdot p_2 = \int_{-1}^1 \frac{1}{2}\big(3x^2 - 1\big)dx = \frac{1}{2}\big(x^3 - x\big)\Big|_{-1}^1 = \frac{1}{2}\big((1-1) - (-1+1)\big) = 0.$$

Finally, we show that p_1 is orthogonal to p_2, since

$$p_1 \cdot p_2 = \int_{-1}^1 x\,\frac{1}{2}\big(3x^2 - 1\big)dx = \frac{1}{2}\int_{-1}^1 \big(3x^3 - x\big)\,dx = \frac{1}{2}\big(\frac{3}{4}x^4 - \frac{x^2}{2}\big)\Big|_{-1}^1 = 0.$$

We now compute the length (or magnitude) of these polynomials.

$$p_0 \cdot p_0 = \int_{-1}^1 dx = x\Big|_{-1}^1 = 1 - (-1) = 2,$$

$$p_1 \cdot p_1 = \int_{-1}^1 x^2 dx = \frac{x^3}{3}\Big|_{-1}^1 = \frac{1}{3} - \frac{(-1)}{3} = \frac{2}{3}.$$

The length of p_3 is a bit more complicated to compute,

$$\begin{aligned}
p_2 \cdot p_2 &= \frac{1}{4}\int_{-1}^{1} \big(3x^2 - 1\big)^2\, dx \\
&= \frac{1}{4}\int_{-1}^{1} \big(9x^4 - 6x^2 + 1\big)^2\, dx \\
&= \frac{1}{4}\Big(\frac{9}{5}x^5 - 2x^3 + x\Big)\Big|_{-1}^{1} \\
&= \frac{1}{2}\Big(\frac{9}{5} - 2 + 1\Big) \\
&= \frac{2}{5}.
\end{aligned}$$

Therefore, these Legendre polynomials are not unit functions. ◁

6.2.2. Sturm-Liouville Problem. We now introduce a differential operator that has very interesting properties and that shows up very often when we describe natural processes. In this section we will be interested in the eigenvalues and eigenfunctions of this operator.

Definition 6.2.4. *A* ***Sturm-Liouville operator*** *is the differential operator*

$$L(y) = -\big(p\, y'\big)' + q\, y, \tag{6.2.1}$$

where y is a twice continuously differentiable function on $[a,b] \subset \mathbb{R}$, with $b > a$, the function p is differentiable, the function q is continuous, and p non-negative, meaning that $p(x) \geqslant 0$ for all $x \in [a,b]$. The Sturm-Liouville operator is called ***regular*** *when p is positive, that is $p(x) > 0$ for all $x \in [a,b]$.*

Remarks:

(a) Sturm-Liouville operators are a particular case of second order linear operators

$$L(y) = a_2\, y'' + a_1\, y' + a_0\, y$$

where $a_2 < 0$ and $a_1 = a_2'$. This can be sen by writing the Sturm-Liouville operator as

$$L(y) = -p\, y'' - p'\, y' + q\, y.$$

(b) The Sturm-Liouville operator, just as any other differential operator, is determined not only by the functions p and q, but also by the type of functions y it is applied. For example, let's fix functions p and q, then the Sturm-Liouville operator given by the expression $L(y) = -(p\, y')' + q\, y$ acting on functions y twice continuously differentiable satisfying

$$y(a) = 0, \quad y(b) = 0,$$

is different from the Sturm-Liouville operator $L(y) = -(p\, y')' + q\, y$ acting on functions y twice continuously differentiable satisfying

$$y'(a) = 0, \quad y'(b) = 0.$$

The set of functions on which these two operators operate are different. In other words, these two operators have different domains.

The Sturm-Liouville operators have interesting properties, and one of them is the Lagrange identity.

Theorem 6.2.5 (Lagrange Identity)**.** *If L is a Sturm-Liouville operator as in (6.2.1), then the following equation holds for every twice-continuously differentiable functions u, v,*

$$v\,L(u) - u\,L(v) = \big(p\,(u\,v' - u'\,v)\big)'. \tag{6.2.2}$$

Proof of Theorem 6.2.5: From the definition of the Sturm-Liouville operator we get

$$v\,L(u) = -v\,(p\,u')' + q\,v\,u,$$
$$u\,L(v) = -u\,(p\,v')' + q\,u\,v.$$

Then, the difference of the equations above is

$$v\,L(u) - u\,L(v) = -v\,(p\,u')' + u\,(p\,v')'.$$

But the product rule of derivatives implies

$$v\,(p\,u')' = (v\,p\,u')' - v'\,p\,u'.$$

Analogously,

$$u\,(p\,v')' = (u\,p\,v')' - u'\,p\,v'.$$

Therefore,

$$\begin{aligned} v\,L(u) - u\,L(v) &= -(v\,p\,u')' + v'\,p\,u' + (u\,p\,v')' - u'\,p\,v' \\ &= \big(p\,(u\,v' - u'\,v)\big)'. \end{aligned}$$

This establishes the Theorem. □

The Lagrange identity will be important when we prove properties of the solutions to eigenfunction problems for the Sturm-Liouville operator. Let's now introduce these eigenfunction problems.

Definition 6.2.6. *A* ***regular Sturm-Liouville system*** *is the eigenfunction problem*

$$-\big(p\,y'\big)' + q\,y = \lambda\,r\,y, \tag{6.2.3}$$

on $(a,b) \subset \mathbb{R}$, with $b > a$, with function p differentiable, functions q and r continuous, functions p and r are positive, meaning $p > 0$, $r > 0$ for all $x \in [a,b]$, and boundary conditions

$$a_1\,y(a) + a_2\,y'(a) = 0, \tag{6.2.4}$$
$$b_1\,y(b) + b_2\,y'(b) = 0, \tag{6.2.5}$$

where a_1, a_2, b_1, b_2 are given constants with $|a_1| + |a_2| > 0$ and $|b_1| + |b_2| > 0$.

Remarks:

(a) The condition on the constants a_1 and a_2 simply says that both constants cannot be zero at the same time. We assume the same for the constants b_1 and b_2.

(b) When the function r, called the weight function, is $r(x) = 1$ for all $x \in [a,b]$, the regular Sturm-Liouville system reduces to the problem of finding a number λ and a nonzero function y solutions of

$$L(y) = \lambda\,y$$

with boundary conditions as in Eqs. (6.2.4)-(6.2.5). This is a particular case of the eigenfunction problems we introduced in the previous section. Indeed, all the eigenfunction problems we solved in the previous section are regular Sturm-Liouville systems.

Example 6.2.2 (Dirichlet)**.** Find the eigenvalues and unit eigenfunctions of the regular Sturm-Liouville system on the interval $[0, L]$, with $L > 0$,

$$-y'' = \lambda\, y, \qquad y(0) = 0, \quad y(L) = 0.$$

Solution: This is a regular Sturm-Liouville systems, where $p = 1$, $q = 0$, $r = 1$, $a = 0$, and $b = L$, while the boundary condition coefficients $a_1 = 1$, $a_2 = 0$, $b_1 = 1$, and $b_2 = 0$. The differential equation is

$$y'' + \lambda\, y = 0.$$

From § 6.1 we know that the boundary value problem can have nonzero solutions (eigenfunctions) only when $\lambda > 0$. So, we consider only the case $\lambda > 0$. In this case, the characteristic equation is

$$r^2 + \lambda = 0 \quad \Rightarrow \quad r_{\pm} = \pm\sqrt{\lambda}\, i.$$

The general solution of the differential equation is

$$y(x) = c_1 \cos(\sqrt{\lambda}\, x) + c_2 \sin(\sqrt{\lambda}\, x).$$

The first boundary condition, $y(0) = 0$ implies that $c_1 = 0$. Then the solution so far is

$$y(x) = c_2 \sin(\sqrt{\lambda}\, x).$$

The second boundary condition, $y(L) = 0$ implies

$$c_2 \sin(\sqrt{\lambda}\, L) = 0, \quad c_2 \neq 0 \quad \Rightarrow \quad \sin(\sqrt{\lambda}\, L) = 0 \quad \Rightarrow \quad \sqrt{\lambda_n}\, L = n\pi,$$

for $n = 1, 2, 3, \cdots$. Therefore, the eigenvalues and eigenfunctions are

$$\lambda_n = \Big(\frac{n\pi}{L}\Big)^2, \quad y_n(x) = c_n \sin\Big(\frac{n\pi x}{L}\Big), \quad n = 1, 2, 3, \cdots,$$

for any nonzero constants c_n. Since we want unit eigenfunctions, we compute the length of the eigenfunction above,

$$\|y_n\|^2 = y_n \cdot y_n = (c_n)^2 \int_0^L \sin^2\Big(\frac{n\pi x}{L}\Big)\, dx.$$

Recalling the identity,

$$\sin^2(x) = \frac{1}{2}(1 - \cos(2x))$$

we get

$$\begin{aligned}
\|y_n\|^2 &= \frac{(c_n)^2}{2} \int_0^L \Big(1 - \cos\Big(\frac{2n\pi x}{L}\Big)\Big)\, dx \\
&= \frac{(c_n)^2}{2} \Big(x - \frac{L}{2n\pi} \sin\Big(\frac{2n\pi x}{L}\Big)\Big)\Big|_0^L \\
&= \frac{(c_n)^2}{2} \Big(L - \frac{L}{n\pi} \sin(2n\pi) - 0 + 0\Big) \\
&= \frac{(c_n)^2 L}{2}.
\end{aligned}$$

Therefore, $\|y_n\| = 1$ iff $c_n = \sqrt{2/L}$. Then, unit eigenfunctions are given by

$$y_n(x) = \sqrt{\frac{2}{L}}\, \sin\Big(\frac{n\pi x}{L}\Big).$$

◁

Example 6.2.3 (Mixed)**.** Find the eigenvalues and eigenfunctions of the regular Sturm-Liouville system on the interval $[0, L]$, with $L > 0$,

$$-y'' = \lambda\, y, \qquad y'(0) = 0, \quad y(L) = 0.$$

Solution: This is a regular Sturm-Liouville systems, where $p = 1$, $q = 0$, $r = 1$, $a = 0$, and $b = L$, while the boundary condition coefficients $a_1 = 0$, $a_2 = 1$, $b_1 = 1$, and $b_2 = 0$. The differential equation is

$$y'' + \lambda\, y = 0.$$

From § 6.1 we know that the boundary value problem can have nonzero solutions (eigenfunctions) only when $\lambda > 0$. So, we consider only the case $\lambda > 0$. In this case, the characteristic equation is

$$r^2 + \lambda = 0 \quad \Rightarrow \quad r_{\pm} = \pm\sqrt{\lambda}\, i.$$

The general solution of the differential equation is

$$y(x) = c_1 \cos(\sqrt{\lambda}\, x) + c_2 \sin(\sqrt{\lambda}\, x).$$

The derivative is

$$y'(x) = -\sqrt{\lambda}\, c_1 \sin(\sqrt{\lambda}\, x) + \sqrt{\lambda}\, c_2 \cos(\sqrt{\lambda}\, x).$$

The first boundary condition, $y'(0) = 0$ implies that $c_2 = 0$. Then the solution so far is

$$y(x) = c_1 \cos(\sqrt{\lambda}\, x).$$

The second boundary condition, $y(L) = 0$ implies

$$c_1 \cos(\sqrt{\lambda}\, L) = 0, \quad c_1 \neq 0 \quad \Rightarrow \quad \cos(\sqrt{\lambda}\, L) = 0 \quad \Rightarrow \quad \sqrt{\lambda_n}\, L = (2n-1)\frac{\pi}{2},$$

for $n = 1, 2, 3, \cdots$. Therefore, the eigenvalues and eigenfunctions are

$$\lambda_n = \Big(\frac{(2n-1)\pi}{2L}\Big)^2, \quad y_n(x) = c_n \cos\Big(\frac{(2n-1)\pi x}{2L}\Big), \quad n = 1, 2, 3, \cdots,$$

for any nonzero constants c_n. Since we want unit eigenfunctions, we compute the length of the eigenfunction above,

$$\|y_n\|^2 = y_n \cdot y_n = (c_n)^2 \int_0^L \cos^2\Big(\frac{(2n-1)\pi x}{2L}\Big)\, dx.$$

Recalling the identity,

$$\cos^2(x) = \frac{1}{2}(1 + \cos(2x))$$

we get

$$\begin{aligned}
\|y_n\|^2 &= \frac{(c_n)^2}{2} \int_0^L \Big(1 + \cos\Big(\frac{2(2n-1)\pi x}{2L}\Big)\Big)\, dx \\
&= \frac{(c_n)^2}{2} \Big(x + \frac{L}{(2n-1)\pi} \sin\Big(\frac{(2n-1)\pi x}{L}\Big)\Big)\Big|_0^L \\
&= \frac{(c_n)^2}{2} \Big(L + \frac{L}{(2n-1)\pi} \sin((2n-1)\pi) - 0 - 0\Big) \\
&= \frac{(c_n)^2 L}{2}.
\end{aligned}$$

Therefore, $\|y_n\| = 1$ iff $c_n = \sqrt{2/L}$. Then, unit eigenfunctions are given by

$$y_n(x) = \sqrt{\frac{2}{L}} \cos\Big(\frac{(2n-1)\pi x}{2L}\Big).$$

◁

Example 6.2.4 (Mixed)**.** Find the eigenvalues and eigenfunctions of the regular Sturm-Liouville system on the interval $[0, L]$, with $L > 0$,

$$-y'' = \lambda\, y, \qquad y(0) = 0, \quad y'(L) = 0.$$

Solution: This is a regular Sturm-Liouville systems, where $p = 1$, $q = 0$, $r = 1$, $a = 0$, and $b = L$, while the boundary condition coefficients $a_1 = 1$, $a_2 = 0$, $b_1 = 0$, and $b_2 = 1$. The differential equation is

$$y'' + \lambda\, y = 0.$$

From § 6.1 we know that the boundary value problem can have nonzero solutions (eigenfunctions) only when $\lambda > 0$. So, we consider only the case $\lambda > 0$. In this case, the characteristic equation is

$$r^2 + \lambda = 0 \quad \Rightarrow \quad r_{\pm} = \pm\sqrt{\lambda}\, i.$$

The general solution of the differential equation is

$$y(x) = c_1 \cos(\sqrt{\lambda}\, x) + c_2 \sin(\sqrt{\lambda}\, x).$$

The derivative is

$$y'(x) = -\sqrt{\lambda}\, c_1 \sin(\sqrt{\lambda}\, x) + \sqrt{\lambda}\, c_2 \cos(\sqrt{\lambda}\, x).$$

The first boundary condition, $y(0) = 0$ implies that $c_1 = 0$. Then the solution so far is

$$y(x) = c_2 \sin(\sqrt{\lambda}\, x).$$

The second boundary condition, $y'(L) = 0$ implies

$$\sqrt{\lambda}\, c_2 \cos(\sqrt{\lambda}\, L) = 0, \quad c_2 \neq 0 \quad \Rightarrow \quad \cos(\sqrt{\lambda}\, L) = 0 \quad \Rightarrow \quad \sqrt{\lambda_n}\, L = (2n-1)\frac{\pi}{2},$$

for $n = 1, 2, 3, \cdots$. Therefore, the eigenvalues and eigenfunctions are

$$\lambda_n = \Big(\frac{(2n-1)\pi}{2L}\Big)^2, \quad y_n(x) = c_n \sin\Big(\frac{(2n-1)\pi x}{2L}\Big), \quad n = 1, 2, 3, \cdots,$$

for any nonzero constants c_n. Since we want unit eigenfunctions, we compute the length of the eigenfunction above,

$$\|y_n\|^2 = y_n \cdot y_n = (c_n)^2 \int_0^L \sin^2\Big(\frac{(2n-1)\pi x}{2L}\Big)\, dx.$$

Recalling the identity,

$$\sin^2(x) = \frac{1}{2}(1 - \cos(2x))$$

we get

$$\begin{aligned}
\|y_n\|^2 &= \frac{(c_n)^2}{2} \int_0^L \Big(1 - \cos\Big(\frac{2(2n-1)\pi x}{2L}\Big)\Big)\, dx \\
&= \frac{(c_n)^2}{2} \Big(x - \frac{L}{(2n-1)\pi} \sin\Big(\frac{(2n-1)\pi x}{L}\Big)\Big)\Big|_0^L \\
&= \frac{(c_n)^2}{2} \Big(L - \frac{L}{(2n-1)\pi} \sin((2n-1)\pi) - 0 + 0\Big) \\
&= \frac{(c_n)^2 L}{2}.
\end{aligned}$$

Therefore, $\|y_n\| = 1$ iff $c_n = \sqrt{2/L}$. Then, unit eigenfunctions are given by

$$y_n(x) = \sqrt{\frac{2}{L}} \sin\Big(\frac{(2n-1)\pi x}{2L}\Big).$$

◁

Example 6.2.5 (Neumann)**.** Find the eigenvalues and eigenfunctions of the regular Sturm-Liouville system on the interval $[0, L]$, with $L > 0$,

$$-y'' = \lambda\, y, \qquad y'(0) = 0, \quad y'(L) = 0.$$

Solution: This is a regular Sturm-Liouville systems, where $p = 1$, $q = 0$, $r = 1$, $a = 0$, and $b = L$, while the boundary condition coefficients $a_1 = 0$, $a_2 = 1$, $b_1 = 0$, and $b_2 = 1$. The differential equation is

$$y'' + \lambda\, y = 0.$$

From § 6.1 we know that the boundary value problem can has unique solutions for $\lambda < 0$. For $\lambda = 0$ we have

$$y'' = 0 \quad \Rightarrow \quad y(x) = c_1 + c_2\, x.$$

The derivative is

$$y'(x) = c_2.$$

Both boundary conditions imply $c_2 = 0$. So we got that eigenvalue eigenfunction

$$\lambda_0 = 0, \quad y_0(x) = c_0 \neq 0.$$

Now we consider the case $\lambda > 0$. In this case, the characteristic equation is

$$r^2 + \lambda = 0 \quad \Rightarrow \quad r_\pm = \pm\sqrt{\lambda}\, i.$$

The general solution of the differential equation is

$$y(x) = c_1 \cos(\sqrt{\lambda}\, x) + c_2 \sin(\sqrt{\lambda}\, x).$$

The derivative is

$$y'(x) = -\sqrt{\lambda}\, c_1 \sin(\sqrt{\lambda}\, x) + \sqrt{\lambda}\, c_2 \cos(\sqrt{\lambda}\, x).$$

The first boundary condition, $y'(0) = 0$ implies that $c_2 = 0$. Then the solution so far is

$$y(x) = c_1 \cos(\sqrt{\lambda}\, x).$$

The second boundary condition, $y'(L) = 0$ implies

$$-\sqrt{\lambda}\, c_1 \sin(\sqrt{\lambda}\, L) = 0, \quad c_1 \neq 0 \quad \Rightarrow \quad \sin(\sqrt{\lambda}\, L) = 0 \quad \Rightarrow \quad \sqrt{\lambda_n}\, L = n\pi,$$

for $n = 1, 2, 3, \cdots$. Therefore, if we include the case of $\lambda = 0$ computed above, then the eigenvalues and eigenfunctions solutions of the Sturm-Liouville problem in this example are given by

$$\lambda_n = \Big(\frac{n\pi}{L}\Big)^2, \quad y_n(x) = c_n \cos\Big(\frac{n\pi x}{L}\Big), \quad n = 0, 1, 2, 3, \cdots,$$

for any nonzero constants c_n. Since we want unit eigenfunctions, we compute the length of the eigenfunction above,

$$\|y_n\|^2 = y_n \cdot y_n = (c_n)^2 \int_0^L \cos^2\Big(\frac{n\pi x}{L}\Big)\, dx.$$

The case $n = 0$ is $\|y_0\|^2 = (c_0)^2 L = 1$, which implies $c_0 = \sqrt{1/L}$. For the case $n > 0$ we recall the identity,

$$\cos^2(x) = \frac{1}{2}(1 + \cos(2x))$$

then, we get

$$\begin{aligned}\|y_n\|^2 &= \frac{(c_n)^2}{2}\int_0^L \Big(1+\cos\Big(\frac{2n\pi x}{L}\Big)\Big)\,dx\\ &= \frac{(c_n)^2}{2}\Big(x+\frac{L}{2n\pi}\sin\Big(\frac{2n\pi x}{L}\Big)\Big)\Big|_0^L\\ &= \frac{(c_n)^2}{2}\Big(L+\frac{L}{2n\pi}\sin(2n\pi)-0-0\Big)\\ &= \frac{(c_n)^2 L}{2}.\end{aligned}$$

Therefore, $\|y_n\| = 1$ iff $c_n = \sqrt{2/L}$. Then, unit eigenfunctions are given by

$$y_0 = \sqrt{\frac{1}{L}}, \qquad y_n(x) = \sqrt{\frac{2}{L}}\cos\Big(\frac{n\pi x}{L}\Big), \quad n = 1, 2, 3, \cdots .$$

$\triangleleft$

In all the examples above we see that the eigenvalues are real-valued, they can be ordered as

$$\lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots ,$$

and for each eigenvalue there is one eigenfunction (except multiplicative factors). These are actually properties of the eigenvalues and eigenvectors of all regular Sturm-Liouville systems. We summarize these properties in the following result.

Theorem 6.2.7 (Regular Sturm-Liouville)**.** *The solutions of a regular Sturm-Liouville system on an interval $[a, b]$ given in equations (6.2.3)-(6.2.5) has the following properties:*

(a) *All eigenvalues are real.*
(b) *The eigenfunctions of different eigenvalues are orthogonal.*
(c) *There are infinite many eigenvalues, which can be arranged in increasing order,*

$$\lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots \quad \textit{with} \quad \lambda_n \to \infty \quad \textit{with} \quad n \to \infty.$$

(d) *If the eigenvalues are labeled as λ_n, for $n = 0, 1, 2, \cdots$, then the corresponding eigenfunctions y_n have exaclty n zeros in the interval (a, b) and they are uniquely determined up to a constant factor.*

Remark: The proof of **(c)**, **(d)** can be found in many textbooks on differential equations. Here we highlight a few of them, for example Section 6.2 of Pinchover and Rubinstein [7], Section 5.3 in Teschl [10], and chapter 10 in Birkhoff [3].

Proof of (a), (b) in Theorem 6.2.7:
Part (a): Suppose that λ and y are solutions of the regular Sturm-Liouville system in (6.2.3)-(6.2.5). We want to show that $\lambda = \overline{\lambda}$, where $\overline{\lambda}$ denotes the complex conjugate of λ. Since the coefficients p, q, and r are real-valued, then

$$\overline{-\big(p\,y'\big)' + q\,y} = -\big(p\,\overline{y}'\big)' + q\,\overline{y} \quad \Rightarrow \quad \overline{L(y)} = L\big(\overline{y}\big).$$

Since λ and y are solutions of the regular Sturm-Liouville system,

$$L(y) = \lambda\, r\, y. \tag{6.2.6}$$

Now complex conjugate this equation and recalling $\overline{L(y)} = L\big(\overline{y}\big)$ we get

$$L\big(\overline{y}\big) = \overline{\lambda}\, r\, \overline{y}. \tag{6.2.7}$$

The Lagrange identity for y and $\overline{y}$ implies

$$\overline{y}\, L(y) - y\, L(\overline{y}) = \big(p\,(y\,\overline{y}' - y'\,\overline{y})\big)'.$$

If we use equations (6.2.6), (6.2.7) we get

$$\lambda\, r\, \overline{y}\, y - \overline{\lambda}\, r y\, \overline{y} = \big(p\,(y\,\overline{y}' - y'\,\overline{y})\big)',$$

that is

$$(\lambda - \overline{\lambda})\, r\, |y|^2 = \big(p\,(y\,\overline{y}' - y'\,\overline{y})\big)'.$$

Let's integrate the equation above on the interval $[a, b]$,

$$\begin{aligned}(\lambda - \overline{\lambda}) \int_a^b r(x)\, |y(x)|^2\, dx &= \int_a^b \big(p(x)\,(y(x)\,\overline{y}(x)' - y(x)'\,\overline{y}(x))\big)'\, dx \\ &= p(x)\,\big(y(x)\,\overline{y}(x)' - y(x)'\,\overline{y}(x)\big)\Big|_a^b. \qquad (6.2.8)\end{aligned}$$

Recall that the function y satisfies the boundary condition

$$b_1\, y(b) + b_2\, y'(b) = 0 \quad \Rightarrow \quad b_1\, \overline{y}(b) + b_2\, \overline{y}'(b) = 0.$$

Since b_1 and b_2 cannot be both zero, assume the $b_1 \neq 0$, and then

$$y(b) = -\frac{b_2}{b_1}\, y'(b), \quad \Rightarrow \quad \overline{y}(b) = -\frac{b_2}{b_1}\, \overline{y}'(b).$$

If we use this boundary condition in the right-hand side of Eq. (6.2.8) we get

$$y(b)\,\overline{y}'(b) - y'(b)\,\overline{y}(b) = -\frac{b_2}{b_1}\, y'(b)\,\overline{y}'(b) + y'(b)\,\frac{b_2}{b_1}\,\overline{y}'(b) = 0.$$

A similar calculation holds if we assume $b_2 \neq 0$. Also, something similar happens at $x = a$, that is, the function y satisfies the boundary condition

$$a_1\, y(a) + a_2\, y'(a) = 0 \quad \Rightarrow \quad a_1\, \overline{y}(a) + a_2\, \overline{y}'(a) = 0.$$

Since a_1 and a_2 cannot be both zero, assume the $a_1 \neq 0$, and then

$$y(a) = -\frac{a_2}{a_1}\, y'(a), \quad \Rightarrow \quad \overline{y}(a) = -\frac{a_2}{a_1}\, \overline{y}'(a).$$

If we use this boundary condition in the right-hand side of Eq. (6.2.8) we get

$$y(a)\,\overline{y}'(a) - y'(a)\,\overline{y}(a) = -\frac{a_2}{a_1}\, y'(a)\,\overline{y}'(a) + y'(a)\,\frac{a_2}{a_1}\,\overline{y}'(a) = 0.$$

A similar calculation holds if we assume $a_2 \neq 0$. Therefore, we have shown that

$$(\lambda - \overline{\lambda}) \int_a^b r(x)\, |y(x)|^2\, dx = 0,$$

and since the integral is strictly positive, we conclude that

$$\lambda = \overline{\lambda}.$$

This establishes part (a) of the Theorem.

Part (b): Suppose that λ_1, y_1 and λ_2, y_2 are solutions of the regular Sturm-Liouville system in (6.2.3)-(6.2.5) and that $\lambda_1 \neq \lambda_2$. This means

$$L(y_1) = \lambda_1\, r\, y_1, \qquad L(y_2) = \lambda_2\, r\, y_2.$$

Multiply the first equation by y_2 and the second equation by y_1 and then substract the second from the first equation,

$$\begin{aligned}y_2\, L(y_1) - y_1\, L(y_2) &= \lambda_1\, r\, y_2\, y_1 - \lambda_2\, r\, y_1\, y_2 \\ &= (\lambda_1 - \lambda_2)\, r\, y_2\, y_1.\end{aligned}$$

The Lagrange identity is

$$y_2\,L(y_1) - y_1\,L(y_2) = \big(p\,(y_1\,y_2' - y_1'\,y_2)\big)'.$$

These two equations imply

$$(\lambda_1 - \lambda_2)\,r\,y_2\,y_1 = \big(p\,(y_1\,y_2' - y_1'\,y_2)\big)'.$$

If we integrate on both sides on the interval $[a,b]$ and we use the definition of the inner product of two functions we get that

$$(\lambda_1 - \lambda_2)\,(y_1 \cdot y_2) = p(x)\,\big(y_1(x)\,y_2'(x) - y_1'(x)\,y_2(x)\big)\Big|_a^b.$$

Since y_1 and y_2 satisfy the boundary conditions (6.2.4), (6.2.5), we have shown in the proof of part (a) that the right-hand side above vanishes, that is

$$(\lambda_1 - \lambda_2)\,(y_1 \cdot y_2) = 0.$$

But the eigenvalues are different, so we conclude that

$$y_1 \cdot y_2 = 0,$$

which means the eigenfunctions are orthogonal. This establishes part (b) of the Theorem. □

6.2.3. Eigenfunction Expansions. Theorem 6.2.7 summarizes the main properties of the solutions of regular Sturm-Liouville systems, which are the eigenvalues and eigenfunctions of the Sturm-Liouville operator. There is one more property of these solutions of a Sturm-Liouville system, which is important to us and we decided to state it in a separated statement.

Theorem 6.2.8 (Eigenfunction Expansions)**.** *Let y_n, for $n = 0, 1, 2, \cdots$, be eigenfunctions solutions of a regular Sturm-Liouville system.*

(a) *If a function f is continuous and its derivative, f', is piecewise continuous on $[a,b]$, then the function f can be written as*

$$f(x) = F(x)$$

for all $x \in [a,b]$, where

$$F(x) = \sum_{n=0}^{\infty} c_n\,y_n(x), \tag{6.2.9}$$

and the coefficients c_n are given by

$$c_n = \frac{f \cdot y_n}{y_n \cdot y_n}.$$

(b) *If a function f and its derivative f' are piecewise continuous on $[a,b]$ and $F(x)$ is the function given in Eq. (6.2.9) above, then*

$$F(x) = f(x)$$

for all $x \in [a,b]$ where f is continuous, and

$$F(x_0) = \frac{1}{2}\Big(\lim_{x\to x_0^+} f(x) + \lim_{x\to x_0^-} f(x)\Big)$$

for all $x_0 \in [a,b]$ where f is discontinuous.

Remarks:

(a) When we say that $f(x) = F(x)$, with F being an infinite sum, we mean that for a fixed $x \in [a, b]$ we have

$$F_N(x) = \sum_{n=0}^{N} c_n\, y_n(x) \to f(x) \quad \text{as} \quad N \to \infty,$$

in other words, $F_N(x)$ converges to $f(x)$ pointwise in $[a, b]$.

(b) The proof of the convergence statements in Theorem 6.2.8 can be found in Section 6.2 of Pinchover and Rubinstein [7] and references therein. In this notes we only prove the formula for the coefficients c_n.

(c) If the eigenfunctions are unit functions, that is, the eigenfunctions have length one, then the coefficients c_n in the expansion in (6.2.9) are given by the simpler formula

$$c_n = f \cdot y_n.$$

Proof of the Coefficients Formula in Theorem 6.2.8: Theorem 6.2.7 says that any continuous function f on the interval $[a, b]$ can be written in terms of the eigenfunctions of the regular Sturm-liouville system, that is,

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

Now multiply both sides of the equation by $r\, y_m$ and integrate on $[a, b]$,

$$\begin{aligned}\int_a^b r(x)\, y_m(x)\, f(x)\, dx &= \int_a^b r(x)\, y_m(x) \sum_{n=0}^{\infty} c_n\, y_n(x)\, dx \\ &= \sum_{n=0}^{\infty} c_n \int_a^b r(x)\, y_m(x)\, y_n(x)\, dx.\end{aligned}$$

The equation above can be written using the inner product notation,

$$f \cdot y_m = \sum_{n=0}^{\infty} c_n\, (y_m \cdot y_n).$$

But the eigenfunctions are orthogonal, then $y_m \cdot y_n = 0$ for $m \neq n$, which means

$$f \cdot y_m = c_m\, (y_m \cdot y_m) \quad \Rightarrow \quad c_m = \frac{f \cdot y_m}{y_m \cdot y_m}.$$

Since the eigenfunctions are unit functions, so $y_m \cdot y_m = 1$ and we get the formula (renaming the index from m back to n),

$$c_n = f \cdot y_n.$$

This establishes the Theorem. □

Example 6.2.6. Find the expansion of the polynomial

$$f(x) = 1 + x + x^2, \qquad x \in [-1, 1],$$

in terms of the first three Legendre's polynomials

$$p_0(x) = 1, \qquad p_1(x) = x, \qquad p_2(x) = \frac{1}{2}(3\, x^2 - 1).$$

Solution: We know that the Legendre's polynomials are orthogonal on the interval $[-1, 1]$, so the Theorem above says that

$$f(x) = c_0\, p_0(x) + c_1\, p_1(x) + c_2\, p_2(x),$$

where

$$c_i = \frac{f \cdot p_i}{p_i \cdot p_i}, \qquad i = 0, 1, 2.$$

We have seen in Example 6.2.1 that

$$p_0 \cdot p_0 = 2, \quad p_1 \cdot p_1 = \frac{2}{3}, \quad p_2 \cdot p_2 = \frac{2}{5}.$$

Therefore, we only need to compute the numerators in the expressions for the c_i. The result is,

$$f \cdot p_0 = \int_{-1}^{1} (1 + x + x^2)\, dx = \Big(x + \frac{x^2}{2} + \frac{x^3}{3}\Big)\Big|_{-1}^{1} = 2 + 0 + \frac{2}{3} \quad \Rightarrow \quad f \cdot p_0 = \frac{8}{3}.$$

This gives us

$$c_0 = \frac{f \cdot p_0}{p_0 \cdot p_0} = \frac{4}{3}.$$

For the next coefficient we have

$$f \cdot p_1 = \int_{-1}^{1} x\,(1 + x + x^2)\, dx = \Big(\frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4}\Big)\Big|_{-1}^{1} = 0 + \frac{2}{3} + 0 \quad \Rightarrow \quad f \cdot p_1 = \frac{2}{3}.$$

This gives us

$$c_1 = \frac{f \cdot p_1}{p_1 \cdot p_1} = 1.$$

For the last coefficient we compute

$$\begin{aligned}
f \cdot p_2 &= \int_{1}^{1} (1 + x + x^2)\frac{1}{2}(3x^2 - 1)\, dx \\
&= \frac{1}{2}\int_{-1}^{1} (3x^2 + 3x^3 + 3x^4 - 1 - x - x^2)\, dx \\
&= \frac{1}{2}\Big(x^3 + \frac{3}{4}x^4 + \frac{3}{5}x^5 - x - \frac{x^2}{2} - \frac{x^3}{3}\Big)\Big|_{-1}^{1} \\
&= \frac{1}{2}\Big(2 + 0 + \frac{6}{5} - 2 - 0 - \frac{2}{3}\Big) \\
&= \frac{4}{15}.
\end{aligned}$$

This gives us

$$c_2 = \frac{f \cdot p_2}{p_2 \cdot p_2} = \frac{2}{3}.$$

Therefore, we got the expansion

$$f(x) = \frac{4}{3}\, p_0(x) + p_1(x) + \frac{2}{3}\, p_2(x).$$

◁

Example 6.2.7. Find the eigenfunction expansion of the function

$$f(x) = \begin{cases} x & \text{for } x \in [0, 2], \\ 4 - x & \text{for } x \in [2, 4], \end{cases}$$

in terms of the unit eigenfunctions of the solution of the regular Sturm-Liouville system

$$-y'' = \lambda\, y, \quad x \in (0, 4),$$

and boundary conditions

$$y(0) = 0, \quad y(4) = 0.$$

Solution: We have solved the regular Sturm-Liouville system above in Example 6.2.2 and the eigenvalues and unit eigenfunctions are

$$\lambda_n = \Big(\frac{n\pi}{4}\Big)^2, \qquad y_n = \frac{1}{\sqrt{2}} \sin\Big(\frac{n\pi x}{4}\Big), \qquad n = 1, 2, 3, \cdots$$

Since function f is continuous on $[0, 4]$, then it can be expanded as the infinite series

$$f(x) = \sum_{n=0}^{\infty} c_n y_n(x) \quad \Rightarrow \quad f(x) = \sum_{n=0}^{\infty} c_n \frac{1}{\sqrt{2}} \sin\Big(\frac{n\pi x}{4}\Big),$$

where the coefficients c_n are given by

$$c_n = f \cdot y_n \quad \Rightarrow \quad c_n = \frac{1}{\sqrt{2}} \int_0^4 f(x) \sin\Big(\frac{n\pi x}{4}\Big)\, dx.$$

The rest of the problem is simply the calculation to compute the coefficient c_n. We start splitting the integral at $x = 2$,

$$c_n = \frac{1}{\sqrt{2}} \Big(\int_0^2 x \sin\Big(\frac{n\pi x}{4}\Big)\, dx + \int_2^4 (4 - x) \sin\Big(\frac{n\pi x}{4}\Big)\, dx \Big).$$

So, we need to do three integrals,

$$I_1 = \int_0^2 x \sin\Big(\frac{n\pi x}{4}\Big)\, dx, \qquad I_2 = 4 \int_2^4 \sin\Big(\frac{n\pi x}{4}\Big)\, dx, \qquad I_3 = -\int_2^4 x \sin\Big(\frac{n\pi x}{4}\Big)\, dx.$$

Recalling that

$$\int x \sin(ax)\, dx = -\frac{x}{a} \cos(ax) + \frac{1}{a^2} \sin(ax),$$

then we get

$$\begin{aligned} I_1 &= \Big(-\frac{4x}{n\pi} \Big) \cos\Big(\frac{n\pi x}{4}\Big) + \Big(\frac{4}{n\pi}\Big)^2 \sin\Big(\frac{n\pi x}{4}\Big) \Big) \Big|_0^2 \\ &= -\frac{8}{n\pi} \cos\Big(\frac{n\pi}{2}\Big) + \Big(\frac{4}{n\pi}\Big)^2 \sin\Big(\frac{n\pi}{2}\Big). \end{aligned}$$

$$\begin{aligned} I_2 &= -4\Big(\frac{4}{n\pi}\Big) \cos\Big(\frac{n\pi x}{4}\Big) \Big|_2^4 \\ &= -\frac{16}{n\pi} \cos(n\pi) + \Big(\frac{16}{n\pi}\Big) \cos\Big(\frac{n\pi}{2}\Big) \end{aligned}$$

$$\begin{aligned} I_3 &= \Big(\frac{4x}{n\pi}\Big) \cos\Big(\frac{n\pi x}{4}\Big) - \Big(\frac{4}{n\pi}\Big)^2 \sin\Big(\frac{n\pi x}{4}\Big) \Big) \Big|_2^4 \\ &= \frac{16}{n\pi} \cos(n\pi) - 0 - \Big(\frac{8}{n\pi}\Big) \cos\Big(\frac{n\pi}{2}\Big) + \Big(\frac{4}{n\pi}\Big)^2 \sin\Big(\frac{n\pi}{2}\Big). \end{aligned}$$

where in the last equation we used that $\sin(n\pi) = 0$. Them we can compute c_n,

$$\begin{aligned} c_n &= \frac{1}{\sqrt{2}} (I_1 + I_2 + I_3) \\ &= \frac{1}{\sqrt{2}} \Big[-\frac{8}{n\pi} \cos\Big(\frac{n\pi}{2}\Big) + \Big(\frac{4}{n\pi}\Big)^2 \sin\Big(\frac{n\pi}{2}\Big) - \frac{16}{n\pi} \cos(n\pi) + \Big(\frac{16}{n\pi}\Big) \cos\Big(\frac{n\pi}{2}\Big) \\ &\quad + \frac{16}{n\pi} \cos(n\pi) - \Big(\frac{8}{n\pi}\Big) \cos\Big(\frac{n\pi}{2}\Big) + \Big(\frac{4}{n\pi}\Big)^2 \sin\Big(\frac{n\pi}{2}\Big) \Big] \\ &= \frac{1}{\sqrt{2}} \Big(\frac{32}{n^2\pi^2}\Big) \sin\Big(\frac{n\pi}{2}\Big). \end{aligned}$$

Therefore, the function f can be written as

$$f(x) = \frac{32}{\sqrt{2}} \sum_{n=1}^{\infty} \Big(\frac{1}{n^2\pi^2}\Big) \sin\Big(\frac{n\pi}{2}\Big) \frac{1}{\sqrt{2}} \sin\Big(\frac{n\pi x}{4}\Big).$$

We can simplify the expression a little bit, because for $n = 2k$ or $n = 2k - 1$ we have

$$\sin\Big(\frac{2k\pi}{2}\Big) = \sin(k\pi) = 0, \quad \sin\Big(\frac{(2k-1)\pi}{2}\Big) = (-1)^{(k+1)}, \quad k = 1, 2, 3, \cdots .$$

Therefore, the expansion of function f can be written as

$$f(x) = 16 \sum_{k=1}^{\infty} \frac{(-1)^{(k+1)}}{(2k-1)^2\pi^2} \sin\Big(\frac{(2k-1)\pi x}{4}\Big).$$

◁

Example 6.2.8. Find the eigenfunction expansion of the function

$$f(x) = \begin{cases} x & \text{for } x \in [0, 5], \\ 5 & \text{for } x \in [5, 10], \end{cases}$$

in terms of the unit eigenfunctions of the solution of the regular Sturm-Liouville system

$$-y'' = \lambda\, y, \quad x \in (0, 10),$$

and boundary conditions

$$y(0) = 0, \quad y'(10) = 0.$$

Solution: We have solved the regular Sturm-Liouville system above in Example 6.2.4 and the eigenvalues and unit eigenfunctions are

$$\lambda_n = \Big(\frac{(2n-1)\pi}{20}\Big)^2, \qquad y_n = \frac{1}{\sqrt{5}} \sin\Big(\frac{(2n-1)\pi x}{20}\Big), \qquad n = 1, 2, 3, \cdots$$

Since function f is continuous on $[0, 10]$, then it can be expanded as the infinite series

$$f(x) = \sum_{n=0}^{\infty} c_n y_n(x) \quad \Rightarrow \quad f(x) = \sum_{n=0}^{\infty} c_n \frac{1}{\sqrt{5}} \sin\Big(\frac{(2n-1)\pi x}{20}\Big),$$

where the coefficients c_n are given by

$$c_n = f \cdot y_n \quad \Rightarrow \quad c_n = \frac{1}{\sqrt{5}} \int_0^{10} f(x) \sin\Big(\frac{(2n-1)\pi x}{20}\Big)\, dx.$$

The rest of the problem is simply the calculation to compute the coefficient c_n. We start splitting the integral at $x = 5$,

$$c_n = \frac{1}{\sqrt{5}} \Big(\int_0^5 x \sin\Big(\frac{(2n-1)\pi x}{20}\Big)\, dx + \int_5^{10} 5 \sin\Big(\frac{(2n-1)\pi x}{20}\Big)\, dx\Big).$$

So, we need to do two integrals,

$$I_1 = \int_0^5 x \sin\Big(\frac{(2n-1)\pi x}{20}\Big)\, dx, \qquad I_2 = 5 \int_5^{10} \sin\Big(\frac{(2n-1)\pi x}{20}\Big)\, dx.$$

Recalling that

$$\int x \sin(ax)\, dx = -\frac{x}{a} \cos(ax) + \frac{1}{a^2} \sin(ax),$$

then we get

$$\begin{aligned} I_1 &= \Big(-\frac{20x}{(2n-1)\pi}\Big)\cos\Big(\frac{(2n-1)\pi x}{20}\Big) + \Big(\frac{20}{(2n-1)\pi}\Big)^2 \sin\Big(\frac{(2n-1)\pi x}{20}\Big)\Big)\Big|_0^5 \\ &= -\frac{100}{(2n-1)\pi}\cos\Big(\frac{(2n-1)\pi}{4}\Big) + \Big(\frac{20}{(2n-1)\pi}\Big)^2 \sin\Big(\frac{(2n-1)\pi}{4}\Big). \end{aligned}$$

$$\begin{aligned} I_2 &= -5\Big(\frac{20}{(2n-1)\pi}\Big)\cos\Big(\frac{(2n-1)\pi x}{20}\Big)\Big|_5^{10} \\ &= -\frac{100}{(2n-1)\pi}\cos\Big(\frac{(2n-1)\pi}{2}\Big) + \frac{100}{(2n-1)\pi}\cos\Big(\frac{(2n-1)\pi}{4}\Big) \\ &= \frac{100}{(2n-1)\pi}\cos\Big(\frac{(2n-1)\pi}{4}\Big) \end{aligned}$$

where in the last equation we used that $\cos((2n-1)\pi/2) = 0$. Them we can compute c_n,

$$\begin{aligned} c_n &= \frac{1}{\sqrt{5}}\,(I_1 + I_2) \\ &= \frac{1}{\sqrt{5}}\Big[-\frac{100}{(2n-1)\pi}\cos\Big(\frac{(2n-1)\pi}{4}\Big) + \Big(\frac{20}{(2n-1)\pi}\Big)^2 \sin\Big(\frac{(2n-1)\pi}{4}\Big) \\ &\quad + \frac{100}{(2n-1)\pi}\cos\Big(\frac{(2n-1)\pi}{4}\Big)\Big] \\ &= \frac{1}{\sqrt{5}}\Big(\frac{20}{(2n-1)\pi}\Big)^2 \sin\Big(\frac{(2n-1)\pi}{4}\Big). \end{aligned}$$

Therefore, the function f can be written as

$$f(x) = \frac{(20)^2}{\sqrt{5}}\sum_{n=1}^{\infty}\Big(\frac{1}{(2n-1)^2\pi^2}\Big)\sin\Big(\frac{(2n-1)\pi}{4}\Big)\,\frac{1}{\sqrt{5}}\,\sin\Big(\frac{(2n-1)\pi x}{20}\Big),$$

which can be rewritten as

$$f(x) = 80\sum_{n=1}^{\infty}\Big(\frac{1}{(2n-1)^2\pi^2}\Big)\sin\Big(\frac{(2n-1)\pi}{4}\Big)\,\sin\Big(\frac{(2n-1)\pi x}{20}\Big),$$

◁

6.2.4. Solving a BVP. Eigenfunction expansions can be used to solve boundary value problems. The idea is to write the solution of the boundary value problem as an expansion in terms of a particular type of eigenfunctions. The coefficients in this expansion will be determined by the equation in our problem. The eigenfunctions used in the expansion are solutions of a Sturm-Liouville system associated to the original boundary value problem. We use this idea later on to solve more complicated equations, Partial Differential Equations, such as the heat equation.

The problem we want to solve is the following: Find a function $y(x)$ defined on an interval $[0, L]$, for some $L > 0$, solution of the boundary value problem

$$y''(x) + k^2\,y(x) = f(x), \qquad x \in (0, L), \tag{6.2.10}$$

$$y(0) = 0, \qquad y(L) = 0. \tag{6.2.11}$$

where $k > 0$ is a given constant and $f(x)$ is a given function on $[0, L]$.

Remarks:

(a) We can solve this problem using the Variation of Parameters method studied when we solved initial value problems, back in Section 2.3. Actually, this is left as an exercise.

(b) The Laplace transform method studied in Chapter 3 is not a good fit to solve boundary value problems, because the property of the Laplace transform,
$$\mathcal{L}[y''] = s^2\,\mathcal{L}[y] - s\,y(0) - y'(0),$$
involves both $y(0)$ and $y'(0)$, and at least one of these two values are not known in a boundary value problem.

The boundary problem in Eqs.(6.2.10), (6.2.11) determines a regular Sturm-Liouville system. Indeed, let's rewrite the boundary value problem using the operator $L(y) = -y''$. Then, the boundary value problem above is
$$L(y) = -k^2\,y - f, \qquad y(0) = 0, \qquad y(L) = 0.$$
The associated Sturm-Liouville system is
$$L(v) = \lambda\,v, \qquad v(0) = 0, \qquad v(L) = 0.$$
In other words, the associated Sturm-Liouville system is to find all numbers λ and nonzero functions $v(x)$ solutions of the eigenfunction problem
$$v''(x) + \lambda\,v(x) = 0, \qquad x \in (0, L), \tag{6.2.12}$$
$$v(0) = 0, \quad v(L) = 0. \tag{6.2.13}$$
We know from our work earlier in this section that a solution to the eigenfunction problem is
$$\lambda_n = \Big(\frac{\pi n}{L}\Big)^2, \qquad v_n(x) = \sin\Big(\frac{\pi n x}{L}\Big), \qquad n = 1, 2, 3, \cdots. \tag{6.2.14}$$
Use these eigenfunctions to write the function $f(x)$ in the boundary value problem as
$$f(x) = \sum_{n=1}^{\infty} \hat{f}_n\,v_n(x), \tag{6.2.15}$$
where the coefficients $\hat{f}_n$ are given by the usual formula
$$\hat{f}_n = \frac{f \cdot v_n}{v_n \cdot v_n}, \tag{6.2.16}$$
which in this case means
$$\hat{f}_n = \frac{2}{L}\int_0^L f(x)\,\sin\Big(\frac{\pi n x}{L}\Big)\,dx.$$
Now we can write our main result—formulas for the solutions of the boundary value problem in Eqs.(6.2.10), (6.2.11).

Theorem 6.2.9 (BVP). *Consider the boundary value problem*
$$y''(x) + k^2\,y(x) = f(x), \qquad x \in (0, L), \tag{6.2.17}$$
$$y(0) = 0, \qquad y(L) = 0, \qquad L > 0, \tag{6.2.18}$$
where $k > 0$ and f is a function that can be written as
$$f(x) = \sum_{n=1}^{\infty} \hat{f}_n\,v_n(x),$$
with the coefficients $\hat{f}_n$ defined in Eq. (6.2.16), and we denote λ_n and $v_n(x)$ as in Eq. (6.2.14).

(a) *If $k^2 \neq \lambda_n$ for all $n = 1, 2, 3, \cdots$, then the boundary value problem in Eqs.(6.2.17), (6.2.18) has a unique solution*

$$y(x) = \sum_{n=1}^{\infty} \frac{\hat{f}_n}{k^2 - \lambda_n}\, v_n(x). \tag{6.2.19}$$

(b) *If $k^2 = \lambda_{n_0}$ for some $n_0 \in \{1, 2, 3, \cdots\}$, then the boundary value problem in Eqs.(6.2.17), (6.2.18) has solution iff the function f is perpendicular to v_{n_0}, that is,*

$$f \cdot v_{n_0} = 0,$$

and in this case there are infinitely many solutions given by

$$y(x) = \hat{y}_{n_0}\, v_{n_0}(x) + \sum_{\substack{n=1\\ n\neq n_0}}^{\infty} \frac{\hat{f}_n}{k^2 - \lambda_n}\, v_n(x), \tag{6.2.20}$$

where $\hat{y}_{n_0}$ is an arbitrary constant.

Proof of Theorem 6.2.9: We want to solve the boundary value problem

$$y''(x) + k^2\, y(x) = f(x), \qquad x \in (0, L),$$
$$y(0) = 0, \qquad y(L) = 0, \qquad L > 0.$$

We assumed that f can be written in therms of the functions v_n,

$$f(x) = \sum_{n=1}^{\infty} \hat{f}_n\, v_n(x), \qquad \hat{f}_n = \frac{f \cdot v_n}{v_n \cdot v_n}.$$

We start writing the solution $y(x)$ in terms of the functions v_n,

$$y(x) = \sum_{n=1}^{\infty} \hat{y}_n\, v_n(x).$$

The boundary value problem fixes the coefficients $\hat{y}_n$, because using these expansions above in the differential equation we get

$$\sum_{n=1}^{\infty} \hat{y}_n \big(v_n'' + k^2\, v_n\big) = \sum_{n=1}^{\infty} \hat{f}_n\, v_n.$$

But the functions v_n satisfy

$$v_n'' = -\lambda_n\, v_n,$$

therefore, we obtain

$$\sum_{n=1}^{\infty} \big(\hat{y}_n \big(-\lambda_n + k^2\big) - \hat{f}_n\big)\, v_n = 0. \tag{6.2.21}$$

Now we use the orthogonality of the functions v_n. Multiply the equation above by a function v_m and distribute the product,

$$v_m \cdot \sum_{n=1}^{\infty} \big(\hat{y}_n \big(-\lambda_n + k^2\big) - \hat{f}_n\big)\, v_n = 0 \quad \Rightarrow \quad \sum_{n=1}^{\infty} \big(\hat{y}_n \big(-\lambda_n + k^2\big) - \hat{f}_n\big)\, (v_n \cdot v_m) = 0.$$

Since the functions v_n are mutually orthogonal,

$$v_n \cdot v_m = 0 \quad \text{for} \quad m \neq n,$$

which means that each term in the sum in Eq. (6.2.21) must vanish,

$$\hat{y}_n \big(-\lambda_n + k^2\big) - \hat{f}_n = 0.$$

Here we have two possibilities: If $k^2 \neq \lambda_n$ for all $n = 1, 2, 3, \cdots$, then we can compute $\hat{y}_n$ for all n, and we get

$$\hat{y}_n = \frac{\hat{f}_n}{k^2 - \lambda_n},$$

so the solution is

$$y(x) = \sum_{n=1}^{\infty} \frac{\hat{f}_n}{k^2 - \lambda_n} \, v_n(x).$$

This proves part (a). If there is an $n_0 \in 1, 2, 3, \cdots$ such that $k^2 = \lambda_{n_0}$, then we have two cases. In the case $n \neq n_0$ we can compute the corresponding coefficients $\hat{y}_n$ and we get

$$\hat{y}_n = \frac{\hat{f}_n}{k^2 - \lambda_n}, \qquad n \neq n_0.$$

In the case $n = n_0$ we have

$$\hat{y}_{n_0}\big(-\lambda_{n_0} + k^2\big) - \hat{f}_{n_0} = 0 \quad \Rightarrow \quad \hat{f}_{n_0} = 0 \quad \Rightarrow \quad f \cdot v_{n_0} = 0.$$

Notice two things. First, the equations above do not determine the coefficient $\hat{y}_{n_0}$, so it remains arbitrary. Second, This last equation means that f must be perpendicular to v_{n_0}. Only when f is perpendicular to v_{n_0} we have a solutions of the boundary value problem, and these solutions are given by

$$y(x) = \hat{y}_{n_0} \, v_{n_0}(x) + \sum_{\substack{n=1 \\ n \neq n_0}}^{\infty} \frac{\hat{f}_n}{k^2 - \lambda_n} \, v_n(x).$$

This proves part (b) in the theorem, and with that we establish the theorem. □

Example 6.2.9 (BVP)**.** Use an eigenfunction expansion to find the solution y of the boundary value problem

$$y'' + 4y = f(x), \quad x \in (0, 3),$$
$$y(0) = 0 \qquad y(3) = 0,$$

where $f(x) = u(x-1) - u(x-2)$ and $u(x)$ is the step function at $x = 0$.

Solution: We start writing the Sturm-Liouville problem associated to the boundary value problem above,

$$v'' + \lambda \, v = 0, \quad x \in (0, 3),$$
$$v(0) = 0 \qquad v(3) = 0.$$

We have seen earlier in this section that the solutions of the Sturm-Liouville problem above are

$$\lambda_n = \Big(\frac{\pi n}{3}\Big)^2, \qquad v_n(x) = \sin\Big(\frac{\pi n x}{3}\Big), \qquad n = 1, 2, 3, \cdots.$$

Now we use these eigenfunctions to write the right-hand side of the equation,

$$f(x) = \sum_{n=1} \hat{f}_n \, v_n(x),$$

where

$$\begin{aligned}\hat{f}_n &= \frac{2}{3}\int_0^3 f(x)\, v_n(x)\, dx \\ &= \frac{2}{3}\int_1^2 \sin\Big(\frac{\pi n x}{3}\Big)\, dx \\ &= \frac{1}{\pi n}\Big(-\cos\Big(\frac{2\pi n}{3}\Big) + \cos\Big(\frac{\pi n}{3}\Big)\Big) \\ &= \frac{1}{\pi n}\sin\Big(\frac{\pi n}{2}\Big)\, \sin\Big(\frac{\pi n}{6}\Big),\end{aligned}$$

whre in the last step we use the trigonometric identity

$$\cos(A) - \cos(B) = -2\sin\Big(\frac{(A+B)}{2}\Big)\, \sin\Big(\frac{(A-B)}{2}\Big).$$

Since $\pi n/3 \neq 2$ for all $n = 1, 2, 3, \cdots$, then the solution of the boundary value problem is given by Eq. (6.2.19),

$$y(x) = \sum_{n=1}^{\infty} \frac{1}{\pi n}\sin\Big(\frac{\pi n}{2}\Big)\, \sin\Big(\frac{\pi n}{6}\Big)\frac{1}{(4 - \frac{\pi^2 n^2}{3^2})}\sin\Big(\frac{\pi n x}{3}\Big).$$

This last expression can be rewritten as below, where we emphasize the constant factor multiplying the each eigenfunction,

$$y(x) = \sum_{n=1}^{\infty}\Big(\frac{9\sin\big(\frac{\pi n}{2}\big)\, \sin\big(\frac{\pi n}{6}\big)}{\pi n(36 - \pi^2 n^2)}\Big)\sin\Big(\frac{\pi n x}{3}\Big).$$

◁

6.2.5. Exercises.

6.2.1.- .

6.2.2.- .

6.3. Overview of Fourier Series

We have seen in a previous section that continuous functions on a closed interval can be expanded as an infinite sum of eigenfunctions solution of regular Sturm-Liouville system. In this section we focus on one of these infinite series expansion where the eigenfunctions are solutions of a *periodic* Sturm-Liouville system—the Fourier series expansion.

6.3.1. Periodic Sturm-Liouville System.

In the previous section we studied regular Sturm-Liouville systems. In this sections we are interested in the solutions of a similar problem, the periodic Sturm-Liouville system. The main difference between these two problems is in the boundary conditions at the extremes of the interval $[a, b]$ where the problem is defined. In the periodic Sturm-Liouville system we impose that the solution function and its derivative at the point a have the same values as at the point b of the interval $[a, b]$.

Definition 6.3.1. *A **periodic Sturm-Liouville system** is the eigenfunction problem*

$$-\big(p\,y'\big)' + q\,y = \lambda\,r\,y, \tag{6.3.1}$$

on $(a, b) \subset \mathbb{R}$, with $b > a$, satisfying the following conditions:

- ***(a)*** *The function p is differentiable, functions q and r are continuous.*
- ***(b)*** *Functions p and r are positive, meaning $p > 0$, $r > 0$ for all $x \in [a, b]$.*
- ***(c)*** *Each of the functions p, p', q, r satisfy the periodicity condition, which is that their values at $x = a$ is equal to their values at $x = b$.*
- ***(d)*** *The function y satisfies the periodic boundary conditions*

$$y(a) = y(b), \qquad y'(a) = y'(b). \tag{6.3.2}$$

All results mentioned in Theorem 6.2.7 hold for the periodic problem except part **(d)**. Recall that part **(d)** says that each eigenvalue, λ_n, of a regular Sturm-liouville problem has associated an eigenfunction, y_n, which is unique except for a multiplicative constant. This means that the eigenvalues of a regular Sturm-Liouville system have multiplicity one. In the periodic Sturm-Liouville system there are eigenvalues, λ_n, that have two linearly independent eigenfunctions, u_n and v_n, that is, these eigenfunctions are not proportional to each other. Therefore, in the periodic case there are eigenvalues with multiplicity two.

In this section we are interested in the solutions of one particular periodic Sturm-Liouville system, which we summarize the the following result.

Theorem 6.3.2 (Fourier Functions)**.** *The periodic Sturm-Liouville system*

$$-y'' = \lambda\,y,$$

on the interval (a, b), where $b > a$, with periodic boundary conditions

$$y(a) = y(b), \qquad y'(a) = y'(b),$$

has the solutions

$$\lambda_0 = 0, \quad u_0 = \frac{1}{2}, \tag{6.3.3}$$

and denoting $(b - a) = L > 0$, we have, for $n = 1, 2, 3, \cdots$,

$$\lambda_n = \Big(\frac{2\pi n}{L}\Big)^2, \quad u_n(x) = \cos\Big(\frac{2\pi n\,x}{L}\Big), \quad v_n(x) = \sin\Big(\frac{2\pi n\,x}{L}\Big). \tag{6.3.4}$$

Remark: A common particular case is when $[a, b] = [0, L]$, in which case the boundary conditions are given by

$$y(0) = y(L), \qquad y'(0) = y'(L).$$

Proof of Theorem 6.3.2: We need to solve the differential equation

$$y'' + \lambda\, y = 0.$$

We know that the solutions have the form $y(x) = e^{rx}$, which leads us to the characteristic equation for r given by

$$r^2 + \lambda = 0 \quad \Rightarrow \quad r^2 = -\lambda.$$

The solutions of this equation depend on whether $\lambda < 0$, or $\lambda = 0$, or $\lambda > 0$.

(a) If $\lambda < 0$, we write it as $\lambda = -\mu^2$ for $\mu > 0$, then we get $r_{\pm} = \pm\mu$, which gives the general solution

$$y(x) = c_1\, e^{\mu x} + c_2\, e^{-\mu x}$$

and its derivative is

$$y'(x) = c_1 \mu\, e^{\mu x} - c_2 \mu\, e^{-\mu x}.$$

The boundary condition $y(a) = y(b)$ implies

$$c_1\, e^{\mu a} + c_2\, e^{-\mu a} = c_1\, e^{\mu b} + c_2\, e^{-\mu b}. \tag{6.3.5}$$

The boundary condition $y'(a) = y'(b)$ implies

$$c_1\, \mu\, e^{\mu a} - c_2\, \mu\, e^{-\mu a} = c_1\, \mu\, e^{\mu b} - c_2\, \mu\, e^{-\mu b}. \tag{6.3.6}$$

Multiply Eq. (6.3.5) by μ and add it to Eq. (6.3.6), then we get

$$2c_1\, \mu\, e^{\mu a} = 2c_1\, \mu\, e^{\mu b}.$$

Since $\mu > 0$ we get the equation

$$c_1\, (e^{\mu b} - e^{\mu a}) = 0,$$

but $b \neq a$ therefore

$$c_1 = 0.$$

Now multiply Eq. (6.3.5) by μ but this time subtract from it to Eq. (6.3.6), then we get

$$2c_2\, \mu\, e^{-\mu a} = 2c_2\, \mu\, e^{-\mu b}.$$

Since $\mu > 0$ we get the equation

$$c_2\, (e^{-\mu b} - e^{-\mu a}) = 0,$$

but $b \neq a$ therefore

$$c_2 = 0.$$

Therefore, the only solution is $c_1 = c_2 = 0$. This means that λ cannot be negative.

(b) If $\lambda = 0$ the general solution of $y'' = 0$ is

$$y(x) = c_1 + c_2 x$$

and its derivative is

$$y'(x) = c_2.$$

The boundary condition $y(a) = y(b)$ implies

$$c_1 + c_2 a = c_1 + c_2 b \quad \Rightarrow \quad c_2(b - a) = 0 \quad \Rightarrow \quad c_2 = 0,$$

where we used again that $b \neq a$. Therefore, we get that $y(x) = c_1$, which always satisfies the other boundary condition given by $y'(a) = y'(b)$, since c_1 is a constant. We are free to choose any constant as a representative of the eigenfunction, so we choose

$$\lambda_0 = 0, \quad u_0 = \frac{1}{2}. \tag{6.3.7}$$

Later on we will comment on why we chose $1/2$ and not any other constant.

(c) If $\lambda > 0$, we write it as $\lambda = \mu^2$ for $\mu > 0$, then get

$$r^2 = -\mu^2 \quad \Rightarrow \quad r_{\pm} = \pm \mu\, i,$$

which gives us the general solution

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

and its derivative

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

The boundary condition $y(a) = y(b)$ implies

$$c_1 \cos(\mu a) + c_2 \sin(\mu a) = c_1 \cos(\mu b) + c_2 \sin(\mu b).$$

This equation can be rewritten as

$$-\, c_2\big(\sin(\mu b) - \sin(\mu a)\big) = c_1\big(\cos(\mu b) - \cos(\mu a)\big). \tag{6.3.8}$$

If we use the trigonometric identities

$$\sin(B) - \sin(A) = 2\sin\Big(\frac{(B-A)}{2}\Big)\, \cos\Big(\frac{(B+A)}{2}\Big)$$
$$\cos(B) - \cos(A) = -2\sin\Big(\frac{(B-A)}{2}\Big)\, \sin\Big(\frac{(B+A)}{2}\Big).$$

Using this identities in Eq. (6.3.8) we get

$$-\, 2c_2 \sin\Big(\frac{\mu\,(b-a)}{2}\Big)\, \cos\Big(\frac{\mu\,(b+a)}{2}\Big) = -2c_1 \sin\Big(\frac{\mu\,(b-a)}{2}\Big)\, \sin\Big(\frac{\mu\,(b+a)}{2}\Big). \tag{6.3.9}$$

Now we find a similar equation using the other boundary condition. The boundary condition $y'(a) = y'(b)$ implies

$$-\mu\, c_1 \sin(\mu a) + \mu\, c_2 \cos(\mu a) = -\mu\, c_1 \sin(\mu b) + \mu\, c_2 \cos(\mu b).$$

Since $\mu \neq 0$, this equation can be rewritten as

$$c_1\big(\sin(\mu b) - \sin(\mu a)\big) = c_2\big(\cos(\mu b) - \cos(\mu a)\big). \tag{6.3.10}$$

Using in Eq. (6.3.10) the same trigonometric identities mentioned above we get

$$2c_1 \sin\Big(\frac{\mu\,(b-a)}{2}\Big)\, \cos\Big(\frac{\mu\,(b+a)}{2}\Big) = -2c_2 \sin\Big(\frac{\mu\,(b-a)}{2}\Big)\, \sin\Big(\frac{\mu\,(b+a)}{2}\Big). \tag{6.3.11}$$

Let's focus on Eqs. (6.3.9), (6.3.11). Recalling $\mu \neq 0$, $b \neq a$, we have two possibilities:

(c1) Assume we have

$$\sin\big(\frac{\mu\,(b-a)}{2}\big) \neq 0,$$

then Eqs. (6.3.9) and (6.3.11) have the form

$$c_2 \cos\Big(\frac{\mu\,(b+a)}{2}\Big) = c_1 \sin\Big(\frac{\mu\,(b+a)}{2}\Big) \tag{6.3.12}$$
$$c_1 \cos\Big(\frac{\mu\,(b+a)}{2}\Big) = -c_2 \sin\Big(\frac{\mu\,(b+a)}{2}\Big) \tag{6.3.13}$$

If we multiply Eq. (6.3.12) by c_1 and Eq. (6.3.13) by c_2 we get

$$c_1^2 \sin\Big(\frac{\mu\,(b+a)}{2}\Big) = c_1 c_2 \cos\Big(\frac{\mu\,(b+a)}{2}\Big) = -c_2^2 \sin\Big(\frac{\mu\,(b+a)}{2}\Big),$$

which gives us

$$(c_1^2 + c_2^2)\sin\Big(\frac{\mu\,(b+a)}{2}\Big) = 0. \tag{6.3.14}$$

Now If we multiply Eq. (6.3.12) by c_2 and Eq. (6.3.13) by c_1 we get

$$c_2^2 \cos\Big(\frac{\mu\,(b+a)}{2}\Big) = c_2 c_1 \sin\Big(\frac{\mu\,(b+a)}{2}\Big) = -c_1^2 \cos\Big(\frac{\mu\,(b+a)}{2}\Big),$$

which gives us

$$(c_1^2 + c_2^2) \cos\Big(\frac{\mu\,(b+a)}{2}\Big) = 0. \tag{6.3.15}$$

If either $c_1 \neq 0$ or $c_2 \neq 0$ then we conclude that

$$\sin\Big(\frac{\mu\,(b+a)}{2}\Big) = 0, \qquad \cos\Big(\frac{\mu\,(b+a)}{2}\Big) = 0.$$

There is no value of μ that can satisfy both equations. Therefore both $c_1 = 0$ and $c_2 = 0$, and we have no solutions of the eigenfunction problem.

(c2) The other possibility is

$$\sin\big(\frac{\mu\,(b-a)}{2}\big) = 0.$$

In this case both Eq. (6.3.9) and (6.3.11) are satisfied, so both boundary conditions $y(a) = y(b)$ and $y'(a) = y'(b)$ are satisfied. This equation fixes the values of μ, since

$$\frac{\mu_n\,(b-a)}{2} = n\pi, \qquad n = 1, 2, 3, \cdots.$$

Therefore, using the notation $L = (b-a)$, which means $L > 0$, we got

$$\lambda_n = \Big(\frac{2\pi n}{L}\Big)^2 \quad y_{c_{1n}c_{2n}}(x) = c_{1n} \cos\Big(\frac{2\pi n x}{L}\Big) + c_{2n} \sin\Big(\frac{2\pi n x}{L}\Big), \tag{6.3.16}$$

where $n = 1, 2, 3, \cdots$ and the constants c_{1n} c_{2n} are arbitrary but not both zero.

We conclude that in the case $\lambda > 0$ we have the solutions given by Eq. (6.3.16). The eigenfunction $y_{c_{1n},c_{2n}}$ span a two dimensional space, since we have two arbitrary constants for each n. So, we choose the following representatives for the eigenfunctions,

$$\lambda_n = \Big(\frac{2\pi n}{L}\Big)^2, \qquad u_n(x) = \cos\Big(\frac{2\pi n x}{L}\Big) \qquad v_n(x) = \sin\Big(\frac{2\pi n x}{L}\Big), \tag{6.3.17}$$

for $n = 1, 2, 3, \cdots$.

This establishes the Theorem. □

Since the eigenfunctions u_n and v_n in Eqs. (6.3.3), (6.3.4) are solution of a periodic Sturm-Liouville problem, we already know that these functions are mutually orthogonal for different values of the index n. That is,

$$u_n \perp u_m, \qquad u_n \perp v_m, \qquad v_n \perp v_m, \quad \text{for all} \quad m \neq n,$$

where the symbol $\perp$ means "orthogonal to". This property can be verified by a direct calculation, which is left as an exercise. In our next result we show that for a fixed value of n we also have

$$u_n \perp v_n.$$

This last property is one reason why, from all possible eigenfunctions (6.3.16) for a fixed n, we have chosen the eigenfunctions given in Eq. (6.3.3) and Eq. (6.3.4).

Theorem 6.3.3 (Orthogonality of Fourier Functions)**.** *The functions u_0, u_n, v_n below, for $n = 1, 2, 3, \cdots$, and denoting $L = b - a$, are mutually orthogonal on the interval $[a, b]$,*

$$u_0 = \frac{1}{2}, \qquad u_n(x) = \cos\Big(\frac{2\pi n x}{L}\Big), \qquad v_n(x) = \sin\Big(\frac{2\pi n x}{L}\Big). \tag{6.3.18}$$

Furthermore, the length squared of these functions is

$$\|u_0\|^2 = \frac{L}{4}, \qquad \|u_n\|^2 = \|v_n\|^2 = \frac{L}{2}.$$

Proof of Theorem 6.3.3: As we mentioned above, since the functions u_0, u_n, v_n in Eq.(6.3.18) are solutions of a periodic Sturm-Liouville problem, then they are mutually orthogonal for different values of the index n. We only need to show that for a fixed value of $n > 0$, then functions u_n are orthogonal to the functions v_n. So, we compute the ineer product

$$u_n \cdot v_n = \int_a^b \cos\Big(\frac{2\pi nx}{L}\Big) \sin\Big(\frac{2\pi nx}{L}\Big)\, dx$$

for $n = 1, 2, 3, \cdots$. We use the trigonometric identity

$$\cos(A)\sin(A) = \frac{1}{2}\sin(2A).$$

This last identity in our integral gives

$$\begin{aligned} u_n \cdot v_n &= \frac{1}{2}\int_a^b \sin\Big(\frac{4\pi nx}{L}\Big)\, dx \\ &= -\frac{L}{8\pi n}\Big(\cos\Big(\frac{4\pi nb}{L}\Big) - \cos\Big(\frac{4\pi na}{L}\Big)\Big). \end{aligned}$$

Now the trigonometric identity

$$\cos(B) - \cos(A) = -2\sin\Big(\frac{(B+A)}{2}\Big)\sin\Big(\frac{(B-A)}{2}\Big)$$

together with $b - a = L$, imply

$$u_n \cdot v_n = \frac{L}{4\pi n}\sin\Big(\frac{2\pi n(b+a)}{L}\Big)\sin(2\pi n)$$

But $\sin(2\pi n) = 0$ for $n = 1, 2, 3, \cdots$, so we conclude

$$u_n \cdot v_n = 0 \quad \Rightarrow \quad u_n \perp v_n, \qquad n = 1, 2, 3, \cdots.$$

The furthermore part is proven by a direct calculation. Recalling $b - a = L$, we get

$$\|u_0\|^2 = u_0 \cdot u_0 = \int_a^b \frac{1}{4}\, dx \quad \Rightarrow \quad \|u_0\|^2 = \frac{L}{4}.$$

Now, for $n = 1, 2, 3, \cdots$ we have

$$\|u_n\|^2 = \int_a^b \cos^2\Big(\frac{2\pi nx}{L}\Big)\, dx.$$

The trigonometric identity

$$\cos^2(A) = \frac{1}{2}\big(1 + \cos(2A)\big)$$

allows us to compute the integral,

$$\begin{aligned} \|u_n\|^2 &= \frac{1}{2}\int_a^b \Big(1 + \cos\Big(\frac{4\pi nx}{L}\Big)\Big)\, dx \\ &= \frac{L}{2} + \frac{L}{8\pi n}\Big(\sin\Big(\frac{4\pi nb}{L}\Big) - \sin\Big(\frac{4\pi na}{L}\Big)\Big). \end{aligned}$$

One more trigonometric identity,

$$\sin(B) - \sin(A) = 2\sin\Big(\frac{(B-A)}{2}\Big)\cos\Big(\frac{(B+A)}{2}\Big), \tag{6.3.19}$$

gives us the result

$$\|u_n\|^2 = \frac{L}{2} + \frac{L}{8\pi n}\sin(4\pi n)\,\cos\Big(\frac{4\pi n(b+a)}{L}\Big).$$

Since $\sin(4\pi n) = 0$ for $n = 1, 2, 3, \cdots$, we obtain

$$\|u_n\|^2 = \frac{L}{2}.$$

Finally, we need to compute

$$\|v_n\|^2 = \int_a^b \sin^2\Big(\frac{2\pi n x}{L}\Big)\,dx.$$

The trigonometric identity

$$\sin^2(A) = \frac{1}{2}\big(1 - \cos(2A)\big)$$

allows us to compute the integral,

$$\begin{aligned}\|v_n\|^2 &= \frac{1}{2}\int_a^b \Big(1 - \cos\Big(\frac{4\pi n x}{L}\Big)\Big)\,dx\\ &= \frac{L}{2} - \frac{L}{8\pi n}\Big(\sin\Big(\frac{4\pi n b}{L}\Big) - \sin\Big(\frac{4\pi n a}{L}\Big)\Big).\end{aligned}$$

If we use one more time the trigonometric identity in Eq. (6.3.19) we get

$$\|v_n\|^2 = \frac{L}{2} - \frac{L}{8\pi n}\sin(4\pi n)\,\cos\Big(\frac{4\pi n(b+a)}{L}\Big).$$

Since $\sin(4\pi n) = 0$ for $n = 1, 2, 3, \cdots$, we obtain

$$\|v_n\|^2 = \frac{L}{2}.$$

This establishes the Theorem. □

Remark: We can use the length of the eigenfunctions to normalize these eigenfunctions. The result is the following set of *orthonormal* vectors, which is sometimes used in the literature on Fourier series, where $L = (b - a)$ and $n = 1, 2, 3, \cdots$,

$$\tilde{u}_0 = \frac{1}{\sqrt{L}}, \quad \tilde{u}_n(x) = \sqrt{\frac{2}{L}}\cos\Big(\frac{2\pi n x}{L}\Big), \quad \tilde{v}_n(x) = \sqrt{\frac{2}{L}}\sin\Big(\frac{2\pi n x}{L}\Big).$$

We mentioned that a common particular case of Theorem 6.3.2 is when the interval is $[a, b] = [0, L]$. Another common particular case of this theorem is when the interval is $[a, b] = [-\ell, \ell]$ for some $\ell > 0$. In this case $L = (b - a)$ has the form $L = 2\ell$. The formulas for the solutions of the periodic Sturm-Liouville problem on $[-\ell, \ell]$ are summarized below.

Corollary 6.3.4 (Particular Case)**.** *The periodic Sturm-Liouville system on the interval* $[-\ell, \ell]$*, which is given by*

$$-y'' = \lambda\, y, \qquad y(-\ell) = y(\ell), \qquad y'(-\ell) = y'(\ell), \qquad \ell > 0,$$

has the solutions

$$\lambda_0 = 0, \quad u_0 = \frac{1}{2}, \tag{6.3.20}$$

and, for $n = 1, 2, 3, \cdots$ *we also have*

$$\lambda_n = \Big(\frac{\pi n}{\ell}\Big)^2, \quad u_n(x) = \cos\Big(\frac{\pi n x}{\ell}\Big), \quad v_n(x) = \sin\Big(\frac{\pi n x}{\ell}\Big). \tag{6.3.21}$$

6.3.2. Fourier Series Expansion. The eigenfunctions in Eqs. (6.3.3), (6.3.4) can be used to span a particular type of continuous (and also piecewise continuous) functions on the interval $[a,b]$.

Theorem 6.3.5 (Fourier Series Expansion)**.** *If a function f is continuous with piecewise continuous derivative on an interval $[a,b]$, with $b>a$, then*

$$f(x)=F(x)$$

for all $x\in[a,b]$, where F is given by

$$F(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\Big(a_n\cos\Big(\frac{2\pi nx}{L}\Big)+b_n\sin\Big(\frac{2\pi nx}{L}\Big)\Big),$$

with $L=(b-a)$, and the coefficients above are given by the formulas

$$a_0=\frac{2}{L}\int_a^b f(x)\,dx,\quad a_n=\frac{2}{L}\int_a^b f(x)\cos\Big(\frac{2\pi nx}{L}\Big)\,dx,\quad b_n=\frac{2}{L}\int_a^b f(x)\sin\Big(\frac{2\pi nx}{L}\Big)\,dx,$$

with $n=1,2,3,\cdots$. Furthermore, if a function f and its derivative f' are piecewise continuous on $[a,b]$, then the function F above satisfies that

$$F(x)=f(x)$$

for all $x\in[a,b]$ where f is continuous, while

$$F(x_0)=\frac{1}{2}\Big(\lim_{x\to x_0^+}f(x)+\lim_{x\to x_0^-}f(x)\Big)$$

for all x_0 where f is discontinuous.

Remarks:

(a) We have chosen the basis function $u_0=1/2$ instead of just $u_0=1$, because for this constant $1/2$ the formula for the coefficient a_0 has exactly the same factor in front, $2/L$, as the formulas for a_n and b_n for $n\geqslant 1$.
(b) The proof of that these orthogonal functions span the set of continuous functions on $[a,b]=[-\ell,\ell]$ in chapter 11 in Birkhoff [3]. This proof can be generalized to an arbitrary interval $[a,b]$, with finite constants $b>a$.

Proof of the Coefficient Formulas in Theorem 6.3.5: We know that the eigenfunctions of a periodic Sturm-Liouville system span the set of continuous functions on $[a,b]$. So, every continuous function f can be written as $f(x)=F(x)$, where

$$F(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\Big(a_n\cos\Big(\frac{2\pi nx}{L}\Big)+b_n\sin\Big(\frac{2\pi nx}{L}\Big)\Big).$$

Since the eigenfunctions

$$u_0=\frac{1}{2},\qquad u_n(x)=\cos\Big(\frac{2\pi nx}{L}\Big),\qquad v_n(x)=\Big(\frac{2\pi nx}{L}\Big),\qquad n=1,2,\cdots$$

are mutually orthogonal, then Theorem 6.2.8 implies that the coefficients in the expansion above are given by

$$a_0=\frac{f\cdot u_0}{u_0\cdot u_0},\qquad a_n=\frac{f\cdot u_n}{u_n\cdot u_n},\qquad b_n=\frac{f\cdot v_n}{v_n\cdot v_n},$$

where we used the inner product on the interval $[a,b]$ given by

$$g \cdot h = \int_a^b g(x)\,h(x)\,dx.$$

It is not difficult to see that

$$u_0 \cdot u_0 = \int_a^b \frac{1}{4}\,dx \quad \Rightarrow \quad u_0 \cdot u_0 = \frac{L}{4},$$

where $L = b - a$. In Theorem 6.3.3 we found that

$$u_n \cdot u_n = \int_a^b \cos^2\Big(\frac{n\pi x}{L}\Big)\,dx = \frac{L}{2}, \qquad v_n \cdot v_n = \int_a^b \sin^2\Big(\frac{n\pi x}{L}\Big)\,dx = \frac{L}{2},$$

for $n = 1, 2, 3, \cdots$. Then, the formula for the coefficient a_0 is

$$a_0 = \frac{f \cdot u_0}{u_0 \cdot u_0} = \frac{4}{L}\int_a^b f(x)\,\frac{1}{2}\,dx \quad \Rightarrow \quad a_0 = \frac{2}{L}\int_a^b f(x)\,dx.$$

Next we focus on the formula for the a_n, for $n = 1, 2, 3, \cdots$,

$$a_n = \frac{f \cdot u_n}{u_n \cdot u_n} \quad \Rightarrow \quad a_n = \frac{2}{L}\int_a^b f(x)\cos\Big(\frac{n\pi x}{L}\Big)\,dx.$$

Lastly, we have the formula for the coefficients b_n, for $n = 1, 2, 3, \cdots$

$$b_n = \frac{f \cdot v_n}{v_n \cdot v_n} \quad \Rightarrow \quad b_n = \frac{2}{L}\int_a^b f(x)\sin\Big(\frac{n\pi x}{L}\Big)\,dx.$$

This establishes the Theorem. □

Now we write the particular case of Theorem 6.3.5 in the interval $[-\ell, \ell]$.

Corollary 6.3.6 (Particular Case)**.** *If a function f is continuous with piecewise continuous derivative on an interval $[-\ell, \ell]$, with $\ell > 0$, then*

$$f(x) = F(x)$$

for all $x \in [-\ell, \ell]$, where F is given by

$$F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\Big(a_n \cos\Big(\frac{\pi n x}{\ell}\Big) + b_n \sin\Big(\frac{\pi n x}{\ell}\Big)\Big),$$

and the coefficients above are given by the formulas

$$a_0 = \frac{1}{\ell}\int_{-\ell}^{\ell} f(x)\,dx, \quad a_n = \frac{1}{\ell}\int_{-\ell}^{\ell} f(x)\cos\Big(\frac{\pi n x}{\ell}\Big)\,dx, \quad b_n = \frac{1}{\ell}\int_{-\ell}^{\ell} f(x)\sin\Big(\frac{\pi n x}{\ell}\Big)\,dx,$$

with $n = 1, 2, 3, \cdots$. The furthermore part of Theorem 6.3.5 is exactly the same.

We now use the formulas in the Theorem and the Corollary above to compute the Fourier series expansion of a few continuous functions.

Example 6.3.1. Find the Fourier expansion, F, of $f(x) = \begin{cases} \dfrac{x}{3}, & \text{for} \quad x \in [0,3] \\ 0, & \text{for} \quad x \in [-3,0). \end{cases}$

Solution: The Fourier expansion of f is

$$F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\Big(\frac{\pi n x}{\ell}\Big) + b_n \sin\Big(\frac{\pi n x}{\ell}\Big)$$

In our case $\ell = 3$. We start computing b_n for $n \geqslant 1$,

$$\begin{aligned}
b_n &= \frac{1}{3}\int_{-3}^{3} f(x)\sin\Big(\frac{n\pi x}{3}\Big)\,dx\\
&= \frac{1}{3}\int_{0}^{3} \frac{x}{3}\sin\Big(\frac{n\pi x}{3}\Big)\,dx\\
&= \frac{1}{9}\Big(-\frac{3x}{n\pi}\cos\Big(\frac{n\pi x}{3}\Big)+\frac{9}{n^2\pi^2}\sin\Big(\frac{n\pi x}{3}\Big)\Big)\Big|_0^3\\
&= \frac{1}{9}\Big(-\frac{9}{n\pi}\cos(n\pi)+0+0-0\Big),
\end{aligned}$$

therefore we get

$$b_n = \frac{(-1)^{(n+1)}}{n\pi}.$$

A similar calculation gives us $a_n = 0$ for $n \geqslant 1$,

$$\begin{aligned}
a_n &= \frac{1}{3}\int_{-3}^{3} f(x)\cos\Big(\frac{n\pi x}{3}\Big)\,dx\\
&= \frac{1}{3}\int_{0}^{3} \frac{x}{3}\cos\Big(\frac{n\pi x}{3}\Big)\,dx\\
&= \frac{1}{9}\Big(\frac{3x}{n\pi}\sin\Big(\frac{n\pi x}{3}\Big)+\frac{9}{n^2\pi^2}\cos\Big(\frac{n\pi x}{3}\Big)\Big)\Big|_0^3\\
&= \frac{1}{9}\Big(0+\frac{9}{n^2\pi^2}\cos(n\pi)-0-\frac{9}{n^2\pi^2}\Big),
\end{aligned}$$

therefore we get

$$a_n = \frac{((-1)^n-1)}{n^2\pi^2}.$$

Finally, we compute a_0,

$$a_0 = \frac{1}{3}\int_0^3 \frac{x}{3}\,dx = \frac{1}{9}\,\frac{x^2}{2}\Big|_0^3 = \frac{1}{2}, \quad \Rightarrow \quad a_0 = \frac{1}{2}.$$

Therefore, we get

$$F(x) = \frac{1}{4} + \sum_{n=1}^{\infty}\Big(\frac{((-1)^n-1)}{n^2\pi^2}\cos\Big(\frac{n\pi x}{3}\Big) + \frac{(-1)^{(n+1)}}{n\pi}\sin\Big(\frac{n\pi x}{3}\Big)\Big).$$

◁

6.3.3. Even or Odd Functions. The Fourier series expansion of a function given in an interval of the form $[-\ell,\ell]$ takes a simpler form in case the function is either even or odd. Since the cosine functions are even and the sine functions are odd, it will not be a big surprise when we show that even functions have Fourier series expansion with only the cosine terms and the constant term, while odd functions have Fourier series expansions with only the sine terms.

Definition 6.3.7. *A function f on $[-\ell,\ell]$ is:*

- ***even*** *iff $f(-x) = f(x)$ for all $x \in [-\ell,\ell]$;*
- ***odd*** *iff $f(-x) = -f(x)$ for all $x \in [-\ell,\ell]$.*

The graph of an even function is symmetrical about the vertical axis, while the graph of an odd function is symmetrical about the origin.

Example 6.3.2. Examples of even functions and odd functions.

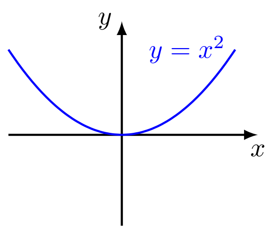


FIGURE 1. $y = x^2$ is even.

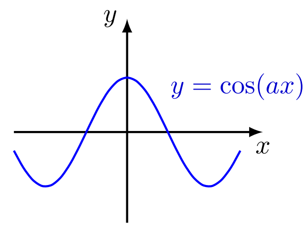


FIGURE 2. $y = \cos(ax)$ is even.

◁

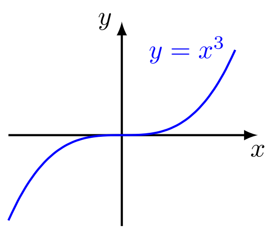


FIGURE 3. $y = x^3$ is odd.

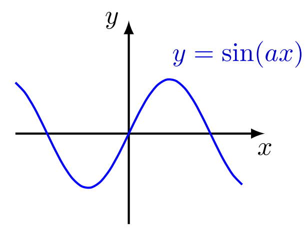


FIGURE 4. $y = \sin(ax)$ is odd.

◁

Remark: Most functions are neither even nor odd. The function $y = e^x$ is neither even nor odd. And in the case that a function is even, such as $y(x) = \cos(ax)$, or odd, such as $y(x) = \sin(ax)$, it is very simple to break that symmetry, for example by shifting the function horizontally, $y(x) = \cos(x - \pi/4)$, or in the case of odd functions by shifting them vertically, $y(x) = 1 + \sin(x)$.

We now summarize a few properties of even functions and odd functions.

Theorem 6.3.8. *If f_e, g_e are even and p_o, q_o are odd functions, then:*

(1) $a\, f_e + b\, g_e$ *is even for all* $a,\, b \in \mathbb{R}$.
(2) $a\, p_o + b\, q_o$ *is odd for all* $a,\, b \in \mathbb{R}$.
(3) $f_e\, g_e$ *is even.*
(4) $p_o\, q_o$ *is even.*
(5) $f_e\, p_o$ *is odd.*
(6) $\displaystyle\int_{-\ell}^{\ell} f_e\, dx = 2\int_0^{\ell} f_e\, dx$.
(7) $\displaystyle\int_{-\ell}^{\ell} p_o\, dx = 0$.

Remark: We leave proof as an exercise. Notice that the last two equations above are simple to understand, just by looking at the figures below.

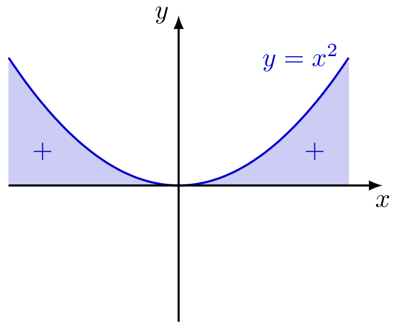


FIGURE 5. Integral of an even function.

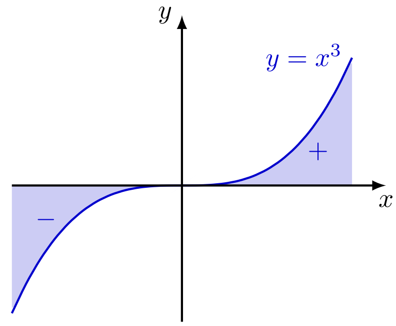


FIGURE 6. Integral of an odd function.

In the case that a function is either even or odd, half of its Fourier series expansion coefficients vanish. In this case the Fourier series is called either a sine or a cosine series.

Theorem 6.3.9 (Cosine and Sine Series)**.** *Let f be a function on $[-\ell, \ell]$ with a Fourier series expansion*

$$F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \Big(a_n \cos\Big(\frac{\pi n x}{\ell}\Big) + b_n \sin\Big(\frac{\pi n x}{\ell}\Big) \Big).$$

(a) If f is even, then $b_n = 0$, and the Fourier series is called a cosine series,

$$F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\Big(\frac{\pi n x}{\ell}\Big).$$

(b) If f is odd, then $a_n = 0$ and Fourier series is called a sine series,

$$F(x) = \sum_{n=1}^{\infty} b_n \sin\Big(\frac{\pi n x}{\ell}\Big).$$

Proof of Theorem 6.3.9:
Part (a): Suppose that f is even, then for $n \geqslant 1$ we get

$$b_n = \frac{1}{\ell} \int_{-\ell}^{\ell} f(x) \sin\Big(\frac{\pi n x}{\ell}\Big)\, dx,$$

but f is even and the sine is odd, so the integrand is odd. Therefore $b_n = 0$.
Part (b): Suppose that f is odd, then for $n \geqslant 1$ we get

$$a_n = \frac{1}{\ell} \int_{-\ell}^{\ell} f(x) \cos\Big(\frac{\pi n x}{\ell}\Big)\, dx,$$

but f is odd and the cosine is even, so the integrand is odd. Therefore $a_n = 0$. Finally

$$a_0 = \frac{1}{\ell} \int_{-\ell}^{\ell} f(x)\, dx,$$

but f is odd, hence $a_0 = 0$. This establishes the Theorem. □

Example 6.3.3. Find the Fourier expansion, F, of $f(x) = \begin{cases} 1, & \text{for} \quad x \in [0, 3] \\ -1, & \text{for} \quad x \in [-3, 0). \end{cases}$

Solution: The function f is odd, so its Fourier series expansion

$$F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\Big(\frac{\pi n x}{\ell}\Big) + b_n \sin\Big(\frac{\pi n x}{\ell}\Big)$$

is actually a sine series. Therefore, all the coefficients $a_n = 0$ for $n \geqslant 0$. So we only need to compute the coefficients b_n. Since in our case $\ell = 3$, we have

$$\begin{aligned} b_n &= \frac{1}{3}\int_{-3}^{3} f(x) \sin\Big(\frac{n\pi x}{3}\Big)\,dx \\ &= \frac{1}{3}\Big(\int_{-3}^{0} (-1) \sin\Big(\frac{n\pi x}{3}\Big)\,dx + \int_{0}^{3} \sin\Big(\frac{n\pi x}{3}\Big)\,dx\Big) \\ &= \frac{2}{3}\int_{0}^{3} \sin\Big(\frac{n\pi x}{3}\Big)\,dx \\ &= \frac{2}{3}\,\frac{3}{n\pi}(-1)\cos\Big(\frac{n\pi x}{3}\Big)\Big|_0^3 \\ &= \frac{2}{n\pi}\big(-(-1)^n + 1\big) \quad \Rightarrow \quad b_n = \frac{2}{n\pi}\,((-1)^{(n+1)} + 1). \end{aligned}$$

Therefore, we get

$$F(x) = \sum_{n=1}^{\infty} \frac{2}{n\pi}\,((-1)^{(n+1)} + 1)\, \sin\Big(\frac{n\pi x}{L}\Big).$$

◁

Example 6.3.4. Find the Fourier series expansion, F, of the function

$$f(x) = \begin{cases} x & x \in [0,1], \\ -x & x \in [-1,0). \end{cases}$$

Solution: Since f is even, then $b_n = 0$. And since $\ell = 1$, we get

$$F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(\pi n x),$$

We start with a_0. Since f is even, a_0 is given by

$$a_0 = 2\int_0^1 f(x)\,dx = 2\int_0^1 x\,dx = 2\,\frac{x^2}{2}\Big|_0^1 \quad \Rightarrow \quad a_0 = 1.$$

Now we compute the a_n for $n \geqslant 1$. Since f and the cosines are even, so is their product,

$$\begin{aligned} a_n &= 2\int_0^1 x\cos(n\pi x)\,dx \\ &= 2\Big(\frac{x}{n\pi}\sin(n\pi x) + \frac{1}{n^2\pi^2}\cos(n\pi x)\Big)\Big|_0^1 \\ &= \frac{2}{n^2\pi^2}\big(\cos(n\pi) - 1\big) \quad \Rightarrow \quad a_n = \frac{2}{n^2\pi^2}\big((-1)^n - 1\big). \end{aligned}$$

Therefore, the Fourier expansion of the function f above is

$$F(x) = \frac{1}{2} + \sum_{n=1}^{\infty} \frac{2}{n^2\pi^2}\big((-1)^n - 1\big)\,\cos(n\pi x).$$

◁

Example 6.3.5. Find the Fourier series expansion, F, of the function

$$f(x) = \begin{cases} 1 - x & x \in [0, 1] \\ 1 + x & x \in [-1, 0). \end{cases}$$

Solution: Since f is even, then $b_n = 0$. And since $L\ell = 1$, we get

$$F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(\pi n x),$$

We start computing a_0,

$$\begin{aligned} a_0 &= \int_{-1}^{1} f(x)\,dx \\ &= \int_{-1}^{0} (1 + x)\,dx + \int_{0}^{1} (1 - x)\,dx \\ &= \Big(x + \frac{x^2}{2}\Big)\Big|_{-1}^{0} + \Big(x - \frac{x^2}{2}\Big)\Big|_{0}^{1} \\ &= \Big(1 - \frac{1}{2}\Big) + \Big(1 - \frac{1}{2}\Big) \quad \Rightarrow \quad a_0 = 1. \end{aligned}$$

Similarly,

$$\begin{aligned} a_n &= \int_{-1}^{1} f(x)\, \cos(n\pi x)\,dx \\ &= \int_{-1}^{0} (1 + x)\, \cos(n\pi x)\,dx + \int_{0}^{1} (1 - x)\, \cos(n\pi x)\,dx. \end{aligned}$$

Recalling the integrals

$$\begin{aligned} \int \cos(n\pi x)\,dx &= \frac{1}{n\pi}\, \sin(n\pi x), \\ \int x\, \cos(n\pi x)\,dx &= \frac{x}{n\pi}\, \sin(n\pi x) + \frac{1}{n^2\pi^2}\, \cos(n\pi x), \end{aligned}$$

it is not difficult to see that

$$\begin{aligned} a_n &= \frac{1}{n\pi}\, \sin(n\pi x)\Big|_{-1}^{0} + \Big[\frac{x}{n\pi}\, \sin(n\pi x) + \frac{1}{n^2\pi^2}\, \cos(n\pi x)\Big]\Big|_{-1}^{0} \\ &\quad + \frac{1}{n\pi}\, \sin(n\pi x)\Big|_{0}^{1} - \Big[\frac{x}{n\pi}\, \sin(n\pi x) + \frac{1}{n^2\pi^2}\, \cos(n\pi x)\Big]\Big|_{0}^{1} \\ &= \Big[\frac{1}{n^2\pi^2} - \frac{1}{n^2\pi^2}\cos(-n\pi)\Big] - \Big[\frac{1}{n^2\pi^2}\cos(-n\pi) - \frac{1}{n^2\pi^2}\Big], \end{aligned}$$

we then conclude that

$$a_n = \frac{2}{n^2\pi^2}\big[1 - \cos(-n\pi)\big] = \frac{2}{n^2\pi^2}\big(1 - (-1)^n\big).$$

Therefore, the Fourier expansion of the function f above is

$$F(x) = \frac{1}{2} + \sum_{n=1}^{\infty} \frac{2}{n^2\pi^2}\big(1 - (-1)^n\big)\, \cos(n\pi x).$$

◁

6.3.4. Solving an IVP. We have seen in Section 6.2 that eigenfunction expansions are useful to solve boundary value problems. Fourier series are a particular type of eigenfunction expansions, so they are useful to solve particular types of boundary value problems. The calculation is pretty similar to the calculations done in Theorem 6.2.9. However, in this section we apply Fourier series to a different type of problem, to an initial value problem.

Let us denote by $y(t)$ the vertical displacement of the spring-mass from its equilibrium position, positive downwards, as function of time $t \in \mathbb{R}$. Let $k > 0$ be the spring constant and $m > 0$ the mass of the object attached to the spring. Denote by $f(t)$ the external force on the spring as function of time. The motion of this system is described by Newton's equation

$$m\, y''(t) + k\, y(t) = f(t), \tag{6.3.22}$$

$$y(0) = y_0, \qquad y'(0) = v_0, \tag{6.3.23}$$

where y_0, y_1 are arbitrary constants interpreted as the initial position and initial velocity of the spring-mass system. Since we want to solve this problem using Fourier series, we assume that the external force f is periodic in time, with period T, which means

$$f(t) = f(t+T)$$

for all $t \in \mathbb{R}$. The periodicity of the source force f means we do not need to study this problem for all $t \in \mathbb{R}$, instead we can restrict the problem to the time interval $[-\tau, \tau]$, where $\tau = T/2$. Then, the force and the solution outside this interval can be obtained form their values inside this interval. Finally, we assume one more property on the source force, we also assume that this force f is odd, that is,

$$f(-t) = -f(t) \quad \text{for all} \quad t \in [-\tau, \tau].$$

Remarks:

(a) As we said above, we assumed the function f is periodic because then we can use Fourier series on the interval in time $[-\tau, \tau]$,

$$f(t) = \frac{\tilde{f}_0}{2} + \sum_{n=1}^{\infty} \Big(\tilde{f}_n \cos\Big(\frac{n\pi t}{\tau}\Big) + \hat{f}_n \sin\Big(\frac{n\pi t}{\tau}\Big)\Big).$$

(b) The initial value problem above can be solved for a general function f on $[-\tau, \tau]$, functions that is neither even nor odd. The only reason we restrict the function f to be odd is to simplify the calculation, because in this case its Fourier expansion is a sine series,

$$f(t) = \sum_{n=1}^{\infty} \hat{f}_n \sin\Big(\frac{n\pi t}{\tau}\Big). \tag{6.3.24}$$

(c) We could solve this problem using the Variation of Parameters method studied when we solved initial value problems, back in Section 2.3. In the Variation of Parameters we find a particular solution of the nonhomogeneous equation, $y_p(t)$, in terms of the fundamental solutions of the homogeneous equation, $y_1(t)$, $y_2(t)$, by a formula

$$y_p(t) = u_1(t)\, y_1(t) + u_2(t)\, y_2(t),$$

where $u_1(t)$, $u_2(t)$ are given by integrals of the force function $f(t)$, the fundamental solutions $y_1(t)$, $y_2(t)$, and their Wronskian. In the case when the force $f(t)$ is a very complicated function we may not be able to integrate to find $u_1(t)$, $u_2(t)$. In such case the Variation of Parameters is not useful to solve the equation. However, the Fourier series method described in this section will still work and it will give us an approximation of the solution, the better approximation the more terms we add in the solution.

(d) This problem can also be solved with the Laplace transform method studied in Chapter 3. Similarly as it happens in the Variation of Parameters method, if the force function $f(t)$ is very complicated we may not be able to invert the Laplace transform of the solution $y(t)$. However, as we said above, the Fourier series method described in this section will give us an approximation of the solution.

Theorem 6.3.10 (IVP)**.** *Consider the initial value problem*

$$m\,y''(t) + k\,y(t) = f(t), \qquad t \in (-\tau, \tau), \qquad \tau > 0, \tag{6.3.25}$$

$$y(0) = y_0, \qquad y'(0) = v_0, \tag{6.3.26}$$

where object mass $m > 0$, spring constant $k > 0$, the initial position y_0 and initial velocity y_1 are arbitrary constants, and f is an odd function that can be written as a sine series,

$$f(t) = \sum_{n=1}^{\infty} \hat{f}_n \sin\Big(\frac{n\pi t}{\tau}\Big).$$

If the angular frequency $\omega = \sqrt{k/m}$ satisfies that $\omega \neq \dfrac{n\pi}{\tau}$ for all $n = 1, 2, 3, \cdots$, then the initial value problem in Eqs.(6.3.25), (6.3.26) has the unique solution

$$y(t) = y_0 \cos(\omega t) + \frac{v_0}{\omega} \sin(\omega t) + \sum_{n=1}^{\infty} \frac{\hat{f}_n}{m\omega\Delta_n} \Big(\omega \sin\Big(\frac{n\pi t}{\tau}\Big) - \Big(\frac{n\pi}{\tau}\Big) \sin(\omega t)\Big), \tag{6.3.27}$$

where we introduced the notation

$$\Delta_n = \Big(\omega^2 - \Big(\frac{n\pi}{\tau}\Big)^2\Big).$$

If there exists an $n_0 \in \{1, 2, 3, \cdots\}$ such that $\omega = \dfrac{n_0\pi}{\tau}$, then the initial value problem in Eqs.(6.3.25), (6.3.26) has the unique solution

$$\begin{aligned} y(t) =&\; y_0 \cos(\omega t) + \frac{v_0}{\omega} \sin(\omega t) + \frac{\hat{f}_{n_0}}{2m\omega^2} \big(\sin(\omega t) - \omega t \cos(\omega t)\big) \\ &+ \sum_{\substack{n=1 \\ n \neq n_0}}^{\infty} \frac{\hat{f}_n}{m\omega\Delta_n} \Big(\omega \sin\Big(\frac{n\pi t}{\tau}\Big) - \Big(\frac{n\pi}{\tau}\Big) \sin(\omega t)\Big). \end{aligned} \tag{6.3.28}$$

Remarks:

(a) Notice that we can decompose the solution $y(t)$ of the initial value problem in Eqs.(6.3.25), (6.3.26) as

$$y(t) = y_h(t) + y_p(t)$$

where $y_h(t)$, is the unique solution of the *homogeneous* differential equation with non-homogeneous initial conditions,

$$m\,y_h''(t) + k\,y_h(t) = 0,$$

$$y_h(0) = y_0, \qquad y_h'(0) = v_0,$$

and $y_p(t)$, is the unique solution of the *nonhomogeneous* differential equation with homogeneous initial conditions

$$m\, y_p''(t) + k\, y_p(t) = f(t),$$
$$y_p(0) = 0, \qquad y_p'(0) = 0.$$

(b) The first initial value problem for y_h is simple to solve and we will show that

$$y_h(t) = y_0 \cos(\omega t) + \frac{v_0}{\omega}\sin(\omega t).$$

The second initial value problem for y_p is more complicated to solve. Here is where we use the Fourier series expansion of f given in Eq.(6.3.24). We will show that for $\omega \neq \dfrac{n\pi}{\tau}$ for all $n = 1, 2, 3, \cdots$ we get

$$y_p(t) = \sum_{n=1}^{\infty} \frac{\hat{f}_n}{m\omega\Delta_n}\Big(\omega \sin\Big(\frac{n\pi t}{\tau}\Big) - \Big(\frac{n\pi}{\tau}\Big)\sin(\omega t)\Big),$$

but in the case that there is an n_0 such that $\omega = \dfrac{n_0\pi}{\tau}$ the function $y_p(t)$ is

$$y_p(t) = \frac{\hat{f}_{n_0}}{2m\omega^2}\big(\sin(\omega t) - \omega t\cos(\omega t)\big) + \sum_{\substack{n=1\\ n\neq n_0}}^{\infty} \frac{\hat{f}_n}{m\omega\Delta_n}\Big(\omega\sin\Big(\frac{n\pi t}{\tau}\Big) - \Big(\frac{n\pi}{\tau}\Big)\sin(\omega t)\Big).$$

Proof of Theorem 6.3.10: We split the solution $y(t)$ of the initial value problem in Eqs.(6.3.25), (6.3.26) in two functions,

$$y(t) = y_h(t) + y_p(t)$$

where $y_h(t)$, is the unique solution of the homogeneous differential equation with nonhomogeneous initial conditions,

$$y_h''(t) + \omega^2\, y_h(t) = 0, \tag{6.3.29}$$
$$y_h(0) = y_0, \qquad y_h'(0) = v_0, \tag{6.3.30}$$

where we used that $\omega^2 = k/m$, and $y_p(t)$, is the unique solution of the nonhomogeneous differential equation with homogeneous initial conditions

$$y_p''(t) + \omega^2\, y_p(t) = \frac{f(t)}{m}, \tag{6.3.31}$$
$$y_p(0) = 0, \qquad y_p'(0) = 0. \tag{6.3.32}$$

We have seen in Section 2.2 that the general solution of the homogeneous differential equation in (6.3.29) is

$$y_h(t) = c\cos(\omega t) + d\sin(\omega t).$$

The nonhomogeneous initial conditions fix the constants c and d, because

$$y_0 = y(0) = c, \qquad v_0 = y'(0) = \omega d,$$

therefore we get

$$y_h(t) = y_0\cos(\omega t) + \frac{v_0}{\omega}\sin(\omega t).$$

Now we focus on finding the solution $y_p(t)$ of the nonhomogeneous initial value problem in (6.3.31), (6.3.32), here we use that the source force, $f(t)$ can be written as a sine series,

$$f(t) = \sum_{n=1}^{\infty} \hat{f}_n \sin\Big(\frac{n\pi t}{\tau}\Big).$$

If we denote by $y_n(t)$ the solution of the initial value problem

$$y_n''(t) + \omega^2 \, y_n(t) = \frac{\hat{f}_n}{m} \sin\Big(\frac{n\pi t}{\tau}\Big), \tag{6.3.33}$$

$$y_n(0) = 0, \qquad y_n'(0) = 0. \tag{6.3.34}$$

then the solution $y_p(t)$ of (6.3.31), (6.3.31) is given by

$$y_p(t) = \sum_{n=1}^{\infty} y_n(t), \tag{6.3.35}$$

since each $y_n(t)$ satisfies homogeneous boundary conditions. Finally, the functions $y_n(t)$ are not difficult to find. We can use the Undetermined Coefficients method studied in Section 2.3. A guess for a particular solution of the differential equation in (6.3.33) is

$$y_{p_n}(t) = c_n \cos\Big(\frac{n\pi t}{\tau}\Big) + d_n \sin\Big(\frac{n\pi t}{\tau}\Big). \tag{6.3.36}$$

This is the correct guess to the differential equation because

$$\omega \neq \frac{n\pi}{\tau} \quad \text{for all} \quad n = 1, 2, 3, \cdots .$$

Indeed, this condition above guaranties the all the y_{p_n} are not solutions of the homogeneous differential equation. Then we compute

$$y_{p_n}''(t) = -\Big(\frac{n\pi}{\tau}\Big)^2 c_n \cos\Big(\frac{n\pi t}{\tau}\Big) - \Big(\frac{n\pi}{\tau}\Big)^2 d_n \sin\Big(\frac{n\pi t}{\tau}\Big)$$

and we put y_{p_n} and y_{p_n}'' in the differential equation (6.3.33) and we get

$$c_n\Big(\omega^2 - \Big(\frac{n\pi}{\tau}\Big)^2\Big) \cos\Big(\frac{n\pi t}{\tau}\Big) + \Big(d_n\big(\omega^2 - \Big(\frac{n\pi}{\tau}\Big)^2\big) - \frac{\hat{f}_n}{m}\Big) \sin\Big(\frac{n\pi t}{\tau}\Big) = 0.$$

This last equation must hold for all $t \in [-\tau, \tau]$, therefore we get

$$c_n\Big(\omega^2 - \Big(\frac{n\pi}{\tau}\Big)^2\Big) = 0, \qquad d_n\big(\omega^2 - \Big(\frac{n\pi}{\tau}\Big)^2\big) - \frac{\hat{f}_n}{m} = 0.$$

Since $\omega \neq n\pi/\tau$ for all n, then equation on the left says $c_n = 0$. The equation on the right gives us the d_n,

$$d_n = \frac{\hat{f}_n}{m(\omega^2 - (\frac{n\pi}{\tau})^2)}.$$

If we introduce the notation

$$\Delta_n = \big(\omega^2 - \Big(\frac{n\pi}{\tau}\Big)^2\big),$$

then the coefficient d_n has the form

$$d_n = \frac{\hat{f}_n}{m\Delta_n}.$$

So, we got our particular solution

$$y_{p_n}(t) = \frac{\hat{f}_n}{m\Delta_n} \sin\Big(\frac{n\pi t}{\tau}\Big).$$

Then, the general solution of (6.3.33) is

$$y_n(t) = \tilde{c}_n \cos(\omega t) + \tilde{d}_n \sin(\omega t) + \frac{\hat{f}_n}{m\Delta_n} \sin\Big(\frac{n\pi t}{\tau}\Big).$$

Now we impose the homogeneous initial condition in (6.3.34). The first condition is

$$0 = y_n(0) = \tilde{c}_n.$$

Therefore, after the first initial condition we got

$$y_n(t) = \tilde{d}_n \sin(\omega t) + \frac{\hat{f}_n}{m\Delta_n} \sin\Big(\frac{n\pi t}{\tau}\Big).$$

The second initial condition in (6.3.34) involves $y_n'(t)$, that is,

$$y_n'(t) = \omega\tilde{d}_n \cos(\omega t) + \Big(\frac{n\pi}{\tau}\Big)\frac{\hat{f}_n}{m\Delta_n} \cos\Big(\frac{n\pi t}{\tau}\Big).$$

The boundary condition says

$$0 = y_n'(0) = \omega\tilde{d}_n + \Big(\frac{n\pi}{\tau}\Big)\frac{\hat{f}_n}{m\Delta_n} \quad \Rightarrow \quad \tilde{d}_n = -\Big(\frac{n\pi}{\tau}\Big)\frac{\hat{f}_n}{m\omega\Delta_n}.$$

We conclude that the functions $y_n(t)$ solutions of the initial value problem in (6.3.33), (6.3.34) are given by

$$y_n(t) = \frac{\hat{f}_n}{m\omega\Delta_n}\Big(\omega \sin\Big(\frac{n\pi t}{\tau}\Big) - \Big(\frac{n\pi}{\tau}\Big)\sin(\omega t)\Big).$$

According to Eq. (6.3.35) we now add all these solutions $y_n(t)$ and we get the particular solution $y_p(t)$ of the nonhomogeneous initial value problem (6.3.31), (6.3.32),

$$y_p(t) = \sum_{n=1}^{\infty} \frac{\hat{f}_n}{m\omega\Delta_n}\Big(\omega \sin\Big(\frac{n\pi t}{\tau}\Big) - \Big(\frac{n\pi}{\tau}\Big)\sin(\omega t)\Big).$$

The solution of our original initial value problem is

$$y(t) = y_h(t) + y_p(t).$$

This establishes the first part of the Theorem. For the second part, when there is an n_0 such that

$$\omega = \frac{n_0\pi}{\tau},$$

all the calculations of the previous part are the same for $n \neq n_0$. Now we only need to compute the part of the solution when $n = n_0$. In this case the guess for the particular solution y_{p_n} given in Eq. (6.3.36) is incorrect, because that guess for $n = n_0$ can be written in terms of ω as

$$y_{p_{n_0}}(t) = c_{n_0} \cos\Big(\frac{n_0\pi t}{\tau}\Big) + d_{n_0} \sin\Big(\frac{n_0\pi t}{\tau}\Big) \quad \Rightarrow \quad y_{p_{n_0}}(t) = c_{n_0}\cos(\omega t) + d_{n_0}\sin(\omega t).$$

We see that this guess is incorrect because it is solution of the homogeneous equation

$$y'' + \omega^2 y = 0.$$

Therefore, the correct guess is

$$y_{p_{n_0}}(t) = c_{n_0}\, t\cos(\omega t) + d_{n_0}\, t\sin(\omega t).$$

Then, for this particular n_0 we compute $y_{p_{n_0}}''$, which gives us

$$y_{p_{n_0}}''(t) = -c_{n_0}\omega^2 t\cos(\omega t) - d_{n_0}\omega^2 t\sin(\omega t) - 2c_{n_0}\omega\sin(\omega t) + 2d_{n_0}\omega\cos(\omega t).$$

We now put $y_{p_{n_0}}$ and $y_{p_{n_0}}''$ in the differential equation (6.3.33), a few terms simplify, and we get

$$-2c_{n_0}\omega\sin(\omega t) + 2d_{n_0}\omega\cos(\omega t) = \frac{\hat{f}_{n_0}}{m}\sin(\omega t).$$

We conclude that

$$d_{n_0} = 0, \qquad c_{n_0} = -\frac{\hat{f}_{n_0}}{2m\omega}.$$

Therefore we get

$$y_{p_{n_0}}(t) = -\frac{\hat{f}_{n_0}}{2m\omega} t\cos(\omega t).$$

Then, the general solution of (6.3.33) is

$$y_{n_0}(t) = \tilde{c}_{n_0} cos(\omega t) + \tilde{d}_{n_0}\sin(\omega t) - \frac{\hat{f}_{n_0}}{2m\omega} t\cos(\omega t).$$

Now we need to find the solution that satisfies the initial conditions in (6.3.34). The first initial condition says

$$0 = y_{n_0}(0) = \tilde{c}_{n_0},$$

so the solution reduces to

$$y_{n_0}(t) = \tilde{d}_{n_0}\sin(\omega t) - \frac{\hat{f}_{n_0}}{2m\omega} t\cos(\omega t).$$

Its derivative is

$$y_{n_0}(t) = \tilde{d}_{n_0}\omega\cos(\omega t) - \frac{\hat{f}_{n_0}}{2m\omega}\cos(\omega t) + \frac{\hat{f}_{n_0}}{2m} t\sin(\omega t),$$

then the other boundary condition is

$$0 = y'_{n_0}(0) = \tilde{d}_{n_0}\omega - \frac{\hat{f}_{n_0}}{2m\omega} \quad \Rightarrow \quad \tilde{d}_{n_0} = \frac{\hat{f}_{n_0}}{2m\omega^2}.$$

Therefore we conclude that

$$y_{n_0}(t) = \frac{\hat{f}_{n_0}}{2m\omega^2}\big(\sin(\omega t) - \omega t\cos(\omega t)\big).$$

This is the term $n = n_0$ in Eq. (6.3.28), which was the only term we needed to compute in that expression. This establishes the Theorem. □

Example 6.3.6. Use Fourier series to find the solution $y(t)$ of the initial value problem

$$y''(t) + 4\,y(t) = f(t), \quad t \in (-3, 3),$$
$$y(0) = 1 \qquad y'(0) = 2,$$

where function $f(t)$ is given by

$$f(t) = \begin{cases} -1 & t \in [-3, 0) \\ 0 & t = 0, \\ 1 & t \in (0, 3]. \end{cases}$$

Solution: Instead of using the solution formula in (6.3.27) we repeat the steps in the proof of Theorem 6.3.10. In this problem we have $m = 1$, $k = 4$, so $\omega = 2$, and $\tau = 3$. Since $f(t)$ is odd in $[-3, 3]$, we write it as a sine series

$$f(t) = \sum_{n=1} \hat{f}_n \sin\Big(\frac{\pi n t}{3}\Big),$$

where the coefficients $\hat{f}_n$ are given by

$$\begin{aligned}\hat{f}_n &= \frac{2}{3}\int_0^3 \sin\Big(\frac{\pi n t}{3}\Big)\, dt\\ &= -\frac{2}{\pi n}\cos\Big(\frac{\pi n t}{3}\Big)\Big|_0^3\\ &= -\frac{2}{\pi n}\big(\cos(\pi n) - 1\big)\\ &= \frac{2}{\pi n}(1 - (-1)^n),\end{aligned}$$

where in the last step we used that $\cos(\pi n) = (-1)^n$. Notice that $2 \neq \pi n/3$ for all $n = 1, 2, 3, \cdots$, so the proof of Theorem 6.3.10 works in this case. In particular we write solution $y(t)$ of the initial value problem in this example as

$$y(t) = y_h(t) + y_p(t)$$

where $y_h(t)$ is the solution of

$$y_h'' + 4\, y_h = 0, \qquad y_h(0) = 1 \qquad y_h'(0) = 2,$$

and $y_p(t)$ is solution of

$$y_p'' + 4\, y_p = f, \qquad y_p(0) = 0 \qquad y_p'(0) = 0.$$

It is then clear that $y(t)$ solves the original initial value problem in this example. Now, in Section 2.2 we saw how to find the solution y_h. The characteristic polynomial of this equation is

$$r^2 + 4 = 0 \quad \Rightarrow \quad r = \pm 2i$$

which means that the real general solution is

$$y_h(t) = c\cos(2t) + d\sin(2t).$$

The initial conditions determine the constants c, d, because

$$1 = y(0) = c, \qquad 2 = y'(0) = 2d \quad \Rightarrow \quad c = 1, \qquad d = 1,$$

therefore,

$$y_h(t) = \cos(2t) + \sin(2t). \tag{6.3.37}$$

Now we compute $y_p(t)$ solution of

$$y_p'' + 4\, y_p = f, \qquad y_p(0) = 0 \qquad y_p'(0) = 0.$$

Here is where we use the Fourier series decomposition of $f(t)$,

$$f(t) = \sum_{n=1}^{\infty} \hat{f}_n \, \sin\Big(\frac{n\pi t}{3}\Big), \qquad \hat{f}_n = \frac{2}{n\pi}(1 - (-1)^n).$$

So, if we find the functions $y_n(t)$ solutions for each term of the sum above,

$$y_n'' + 4\, y_n = \hat{f}_n \, \sin\Big(\frac{n\pi t}{3}\Big), \qquad y_n(0) = 0 \qquad y_n'(0) = 0,$$

then the $y_p(t)$ must be the sum of all these $y_n(t)$,

$$y_p(t) = \sum_{n=1}^{\infty} y_n(t),$$

because each of these $y_n(t)$ satisfy the homogeneous initial conditions

$$y_n(0) = 0 \qquad y_n'(0) = 0.$$

So now we focus on finding the functions $y_n(t)$. We first guess particular solutions $y_{p_n}(t)$ for each n. Our first guess is

$$y_{p_n}(t) = c_n \cos\Big(\frac{n\pi t}{3}\Big) + d_n \sin\Big(\frac{n\pi t}{3}\Big).$$

This is the correct guess because the solutions of the homogeneous differential equation are

$$y_1(t) = \cos(2t), \qquad y_2(t) = \sin(2t),$$

and $2 \neq n\pi/3$ for all $n = 1, 2, 3, \cdots$. Now we can compute y''_{p_n}, which is,

$$y''_{p_n}(t) = -\Big(\frac{n\pi}{3}\Big)^2 c_n \cos\Big(\frac{n\pi t}{3}\Big) - \Big(\frac{n\pi}{3}\Big)^2 d_n \sin\Big(\frac{n\pi t}{3}\Big).$$

and then we can put y''_{p_n} and y_{p_n} in the nonhomogeneous equation

$$y''_{p_n} + 4\, y_{p_n} = \hat{f}_n \, \sin\Big(\frac{n\pi t}{3}\Big).$$

We get

$$\big(-\Big(\frac{n\pi}{3}\Big)^2 + 2\big) c_n \cos\Big(\frac{n\pi t}{3}\Big) + \big(-\Big(\frac{n\pi}{3}\Big)^2 + 2\big) d_n \sin\Big(\frac{n\pi t}{3}\Big) = \hat{f}_n \, \sin\Big(\frac{n\pi t}{3}\Big),$$

which means

$$\big(-\Big(\frac{n\pi}{3}\Big)^2 + 2\big) c_n = 0 \qquad \big(-\Big(\frac{n\pi}{3}\Big)^2 + 2\big) d_n = \hat{f}_n.$$

The equation on the left says $c_n = 0$, the equation on the right says

$$d_n = \frac{\hat{f}_n}{(2 - (\frac{n\pi}{3})^2)}.$$

If we denote

$$\Delta_n = \big(2 - \Big(\frac{n\pi}{3}\Big)^2\big),$$

then the coefficient d_n above is

$$d_n = \frac{\hat{f}_n}{\Delta_n}.$$

So we got our particular solution for each n, which is

$$y_{p_n}(t) = \frac{\hat{f}_n}{\Delta_n} \sin\Big(\frac{n\pi t}{3}\Big).$$

Now we need to find the particular solution that satisfies the initial condition. For that we write the general solution

$$y_n(t) = \tilde{c}_n \cos(2t) + \tilde{d}_n \sin(2t) + \frac{\hat{f}_n}{\Delta_n} \sin\Big(\frac{n\pi t}{3}\Big)$$

and we find the constants $\tilde{c}_n$, $\tilde{d}_n$ using the initial conditions $y_n(0) = 0$ and $y'_n(0) = 0$. The first condition says

$$0 = y_n(0) = \tilde{c}_n,$$

therefore, the solution after this first condition is

$$y_n(t) = \tilde{d}_n \sin(2t) + \frac{\hat{f}_n}{\Delta_n} \sin\Big(\frac{n\pi t}{3}\Big)$$

The second initial condition says

$$0 = y'_n(0) = 2\tilde{d}_n + \Big(\frac{n\pi}{3}\Big) \frac{\hat{f}_n}{\Delta_n},$$

which means

$$\tilde{d}_n = -\Big(\frac{n\pi}{6}\Big)\frac{\hat{f}_n}{\Delta_n}.$$

We arrived at a formula for the solutions $y_n(t)$,

$$y_n(t) = \frac{\hat{f}_n}{2\Delta_n}\Big(2\sin\Big(\frac{n\pi t}{3}\Big) - \Big(\frac{n\pi}{3}\Big)\sin(2t)\Big).$$

If we add the $y_n(t)$ for all n we get our function $y_p(t)$,

$$y_p(t) = \sum_{n=1}^{\infty} \frac{\hat{f}_n}{2\Delta_n}\Big(2\sin\Big(\frac{n\pi t}{3}\Big) - \Big(\frac{n\pi}{3}\Big)\sin(2t)\Big). \tag{6.3.38}$$

Then we are done, because the solution of the initial value problem in this example is

$$y(t) = y_h(t) + y_p(t),$$

and equations (6.3.37), (6.3.38) imply

$$y(t) = \cos(2t) + \sin(2t) + \sum_{n=1}^{\infty} \frac{\hat{f}_n}{2\Delta_n}\Big(2\sin\Big(\frac{n\pi t}{3}\Big) - \Big(\frac{n\pi}{3}\Big)\sin(2t)\Big),$$

where we denoted

$$\hat{f}_n = \frac{2}{n\pi}(1-(-1)^n), \qquad \Delta_n = \Big(2 - \Big(\frac{n\pi}{3}\Big)^2\Big).$$

$\lhd$

6.3.5. Fourier Series of Extensions. We have seen in Theorem 6.3.5 that any piecewise continuous function having a piecewise continuous derivative on an interval $[a, b]$ can be expanded in a Fourier series. Then we saw that a particular case of this theorem, Corollary 6.3.6, when the interval is $[a, b] = [-\ell, \ell]$. Usually in the literature people prove our Corollary but not our Theorem, because the former has a simpler proof than the latter. Then, they usually run into a problem because somewhere along the way they need to compute a Fourier expansion of a function defined only on an interval of the form $[0, L]$.

One solution to this problem is to recompute the Fourier Theorem in an interval of the form $[0, L]$, which defeats the choice of proving this theorem on a simpler interval $[\ell, \ell]$. Another solution is to extend the function originally defined in $[0, L]$ to an interval $[-L, L]$, in this way we can use the Fourier formulas obtained on intervals symmetric around zero. extension of the function

Given a function f defined on an interval $(0, L]$, people extend it to an interval $[-L, L]$ by defining an extension as follows,

$$f_{\rm ext}(x) = f(x), \qquad x \in (0, L],$$
$$f_{\rm ext}(x) = \text{Whatever we want}, \qquad x \in [-L, 0].$$

Then, we compute the Fourier series expansion of the extension, where they can use the Fourier formulas on the interval $[-L, L]$ (our Corollary 6.3.6) on this extension. Since the extension $f_{\rm ext}$ is *not unique*, then they have infinitely many Fourier expansions of the original function f on $(0, L]$, one expansion for each extension.

In this section we explain the most common extensions done in the literature and we compute the formulas for their Fourier expansions.

We can choose the extension $f_{\rm ext}$ as we please, for example the simplest extension is

$$f_{\rm ext}(x) = 0 \qquad x \in [-L, 0].$$

Another possibilities, that actually give simpler Fourier expansions, is to extend the function f into $[-L, 0]$ so that the f_{ext} is either even or odd. In the following result we summarize four extensions of f so that their corresponding Fourier series expansions are simple to compute.

Theorem 6.3.11 (Fourier Series of Extensions)**.** *Any piecewise continuous function f defined on an interval $(0, L]$, for $L > 0$ have the following Fourier series expansions:*

(a) $$f(x) = \sum_{n=1}^{\infty} c_n \sin\Big(\frac{n\pi x}{L}\Big), \quad c_n = \frac{2}{L}\int_0^L f(x)\,\sin\Big(\frac{n\pi x}{L}\Big)\,dx, \quad n = 1, 2, \cdots .$$

(b) $$f(x) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n \cos\Big(\frac{n\pi x}{L}\Big), \quad c_n = \frac{2}{L}\int_0^L f(x)\,\cos\Big(\frac{n\pi x}{L}\Big)\,dx, \quad n = 0, 1, 2, \cdots .$$

(c) $$f(x) = \sum_{n=1}^{\infty} c_n \sin\Big(\frac{(2n-1)\pi x}{2L}\Big),\ c_n = \frac{2}{L}\int_0^L f(x)\,\sin\Big(\frac{(2n-1)\pi x}{2L}\Big)\,dx,\ n = 1, 2, \cdots .$$

(d) $$f(x) = \sum_{n=1}^{\infty} c_n \cos\Big(\frac{(2n-1)\pi x}{2L}\Big),\ c_n = \frac{2}{L}\int_0^L f(x)\,\cos\Big(\frac{(2n-1)\pi x}{2L}\Big)\,dx,\ n = 1, 2, \cdots .$$

Remark: The different Fourier expansions of the function f defined in $(0, L]$ come from the different extensions f_{ext} defined either on $[-L, L]$ or on $[-2L, 2L]$. More specifically, the formulas above come from the following extensions:

(a) The f_{ext} is defined on $[-L, L]$ as the odd extension of f.
(b) The f_{ext} is defined on $[-L, L]$ as the even extension of f.
(c) The f_{ext} is defined on $[-2L, 2L]$ as the symmetric odd extension of f.
(d) The f_{ext} is defined on $[-2L, 2L]$ as the negative-symmetric even extension of f.

Proof of Theorem 6.3.11:
Part **(a):** We extend f defined $(0, L]$ into the domain $[-L, L]$ as an odd function,

$$f_{\text{odd}}(x) = f(x), \quad x \in (0, L], \quad f_{\text{odd}}(0) = 0, \quad f_{\text{odd}}(x) = -f(-x), \qquad x \in [-L, 0).$$

Since f_{odd} is piecewise continuous on $[-L, L]$, then it has a Fourier series expansion. Since f_{odd} is odd, the Fourier series is a sine series

$$f_{\text{odd}}(x) = \sum_{n=1}^{\infty} c_n \sin\Big(\frac{n\pi x}{L}\Big) \tag{6.3.39}$$

and the coefficients are given by the formula

$$c_n = \frac{1}{L}\int_{-L}^{L} f_{\text{odd}}(x)\sin\Big(\frac{n\pi x}{L}\Big)\,dx = \frac{2}{L}\int_0^L f(x)\sin\Big(\frac{n\pi x}{L}\Big)\,dx$$

where we used $f_{\text{odd}}(x) = f(x)$ for $x \in (0, L]$. Restricting Eq. (6.3.39) to $x \in (0, L]$ we get,

$$f(x) = \sum_{n=1}^{\infty} c_n \sin\Big(\frac{n\pi x}{L}\Big).$$

This establishes Part (**a**) in the Theorem.

Part **(b):** We extend f defined $(0, L]$ into the domain $[-L, L]$ as an even function,

$$f_{\text{even}}(x) = f(x), \quad x \in (0, L], \ \ f_{\text{even}}(0) = \lim_{x\to 0^+} f(x), \ \ f_{\text{even}}(x) = f(-x), \ \ x \in [-L, 0).$$

Since $f_{\rm even}$ is piecewise continuous on $[-L, L]$, then it has a Fourier series expansion. Since $f_{\rm even}$ is even, the Fourier series is a cosine series

$$f_{\rm odd}(x) = \sum_{n=1}^{\infty} c_n \cos\Big(\frac{n\pi x}{L}\Big) \tag{6.3.40}$$

and the coefficients are given by the formula

$$c_n = \frac{1}{L}\int_{-L}^{L} f_{\rm even}(x)\cos\Big(\frac{n\pi x}{L}\Big)\,dx = \frac{2}{L}\int_0^L f(x)\cos\Big(\frac{n\pi x}{L}\Big)\,dx$$

where we used $f_{\rm even}(x) = f(x)$ for $x \in (0, L]$. Restricting Eq. (6.3.40) to $x \in (0, L]$ we get,

$$f(x) = \sum_{n=1}^{\infty} c_n \cos\Big(\frac{n\pi x}{L}\Big).$$

This establishes Part (**b**) in the Theorem.

Part **(c):** We make two extensions of f, which is defined $(0, L]$. First we make a symmetric extension of f to $(0, 2L]$, and then an odd extension to $[-2L, 2L]$. A symmetric extension, $f_{\rm s}$, of function f to $(0, 2L]$ is given by

$$f_{\rm s}(x) = \begin{cases} f(x) & \text{for} \quad x \in (0, L], \\ f(2L - x) & \text{for} \quad x \in (L, 2L], \end{cases}$$

followed by an odd extension, $f_{\rm s,odd}$, of the function $f_{\rm s}$ to $[-2L, 2L]$, which is

$$f_{\rm s,odd}(x) = \begin{cases} -f_{\rm s}(-x) & \text{for} \quad x \in [-2L, 0), \\ 0 & \text{for} \quad x = 0, \\ f_{\rm s}(x) & \text{for} \quad x \in (0, 2L]. \end{cases}$$

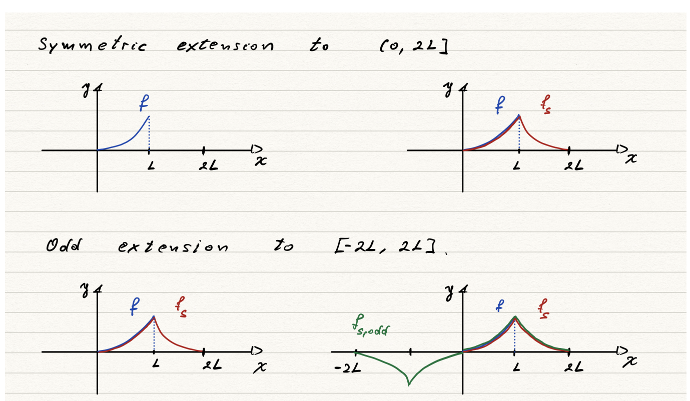


FIGURE 7. Symmetric extension, $f_{\rm s}$, followed by an odd extension, $f_{\rm s,odd}$, of a function f.

An example of these extensions is given in Figure 7. This extension $f_{\rm s,odd}$ has a Fourier series expansion on the interval $[-2L, 2L]$, but since the extension is odd, the Fourier series contains only sine terms,

$$f_{\rm s,odd}(x) = \sum_{n=1}^{\infty} \tilde{c}_n \sin\Big(\frac{n\pi x}{2L}\Big), \tag{6.3.41}$$

with the coefficients $\tilde{c}_n$ given by the usual Fourier formula, this time in the interval $[-2L, 2L]$,

$$\tilde{c}_n = \frac{1}{2L} \int_{-2L}^{2L} f_{\rm s,odd}(x) \sin\Big(\frac{n\pi x}{2L}\Big)\, dx.$$

Since the function $f_{\rm s,odd}$ is odd, the integrand above is even, then

$$\tilde{c}_n = \frac{2}{2L} \int_{0}^{2L} f_{\rm s}(x) \sin\Big(\frac{n\pi x}{2L}\Big)\, dx.$$

We now split the integral in two integrals and use the definition of the symmetric extension,

$$\tilde{c}_n = \frac{1}{L} \Big(\int_0^L f(x) \sin\Big(\frac{n\pi x}{2L}\Big)\, dx + \int_L^{2L} f(2L-x) \sin\Big(\frac{n\pi x}{2L}\Big)\, dx \Big).$$

The change of variables $y = 2L - x$ in the second term above implies

$$\begin{aligned}
\int_L^{2L} f(2L-x) \sin\Big(\frac{n\pi x}{2L}\Big)\, dx &= \int_L^0 f(y) \sin\Big(\frac{n\pi(2L-y)}{2L}\Big)\, (-dy)\\
&= \int_0^L f(y) \sin\Big(n\pi - \frac{n\pi y}{2L}\Big)\, dy\\
&= \int_0^L f(x) \sin\Big(n\pi - \frac{n\pi x}{2L}\Big)\, dx.
\end{aligned}$$

The trigonometric identity $\sin(\theta+\phi) = \sin(\theta)\cos(\phi) + \sin(\phi)\cos(\theta)$ implies

$$\sin\Big(n\pi - \frac{n\pi x}{2L}\Big) = \sin(n\pi)\cos\Big(\frac{n\pi x}{2L}\Big) - \sin\Big(\frac{n\pi x}{2L}\Big)\cos(n\pi) = -(-1)^n \sin\Big(\frac{n\pi x}{2L}\Big),$$

and this equation in the integral above gives

$$\int_L^{2L} f(2L-x) \sin\Big(\frac{n\pi x}{2L}\Big)\, dx = -(-1)^n \int_0^L f(x) \sin\Big(\frac{n\pi x}{2L}\Big)\, dx.$$

Therefore, the coefficient $\tilde{c}_n$ is

$$\tilde{c}_n = \big(1 - (-1)^n\big) \frac{1}{L} \int_0^L f(x) \sin\Big(\frac{n\pi x}{2L}\Big)\, dx.$$

Half of these coefficients are zero. Indeed, if $n = 2k$ we have

$$1 - (-1)^{2k} = 1 - 1^k = 1 - 1 = 0.$$

When $n = 2k - 1$ we have

$$1 - (-1)^{2k-1} = 1 - (-1)(-1)^{2k} = 1 + 1 = 2.$$

Therefore, we only get the odd values of the index n,

$$\tilde{c}_{2k} = 0, \quad \tilde{c}_{2k-1} = \frac{2}{L}\frac{1}{L} \int_0^L f(x) \sin\Big(\frac{(2k-1)\pi x}{2L}\Big)\, dx.$$

Finally, relabel k into n,

$$\tilde{c}_{2n} = 0, \quad \tilde{c}_{2n-1} = \frac{2}{L} \int_0^L f(x) \sin\Big(\frac{(2n-1)\pi x}{2L}\Big)\, dx.$$

Therefore, half the terms in Eq. (6.3.41) vanish, and denoting $c_n = \tilde{c}_{2n-1}$, we obtain that the function f on $(0, L]$ can be expanded as

$$f(x) = \sum_{n=1}^{\infty} c_n \sin\Big(\frac{(2n-1)\pi x}{2L}\Big),$$

with

$$c_n = \frac{2}{L}\int_0^L f(x)\sin\Big(\frac{(2n-1)\pi x}{2L}\Big)\,dx.$$

This establishes the Theorem for Part **(c)**.

Part **(d):** We make two extensions of f, which is defined $(0, L]$. First we make a negative-symmetric extension of f to $(0, 2L]$, and then an even extension to $[-2L, 2L]$. A negative-symmetric extension, $f_{\rm ns}$, of function f to $(0, 2L]$ is given by

$$f_{\rm ns}(x) = \begin{cases} f(x) & \text{for} \quad x \in (0, L], \\ -\,f(2L - x) & \text{for} \quad x \in (L, 2L], \end{cases}$$

followed by an even extension, $f_{\rm ns,even}$, of the function $f_{\rm ns}$ to $[-2L, 2L]$, which is is

$$f_{\rm ns,odd}(x) = \begin{cases} f_{\rm ns}(-x) & \text{for} \quad x \in [-2L, 0), \\ \lim\limits_{x\to 0^+} f(x) & \text{for} \quad x = 0, \\ f_{\rm ns}(x) & \text{for} \quad x \in (0, 2L]. \end{cases}$$

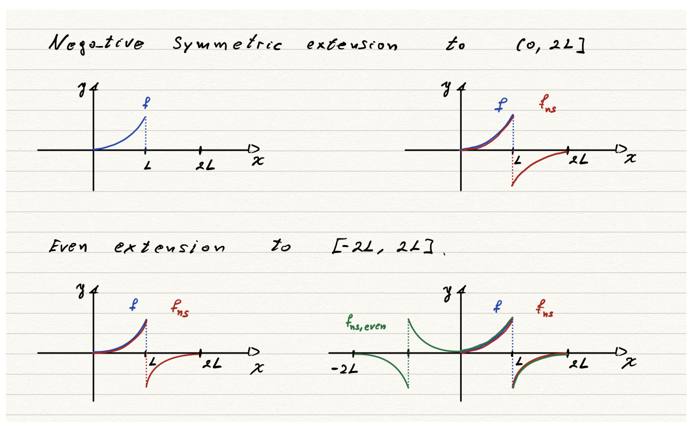


FIGURE 8. Negative symmetric extension, $f_{\rm ns}$, followed by an even extension, $f_{\rm ns,even}$, of a function f.

An example of these extensions is given in Figure 8. This extension $f_{\rm ns,even}$ has a Fourier series expansion on the interval $[-2L, 2L]$, but since the extension is even, the Fourier series contains only cosine terms,

$$f_{\rm ns,even}(x) = \frac{\tilde{c}_0}{2} + \sum_{n=1}^{\infty} \tilde{c}_n \cos\Big(\frac{n\pi x}{2L}\Big), \tag{6.3.42}$$

with the coefficients $\tilde{c}_n$ given by the usual Fourier formula, this time in the interval $[-2L, 2L]$,

$$\tilde{c}_n = \frac{1}{2L}\int_{-2L}^{2L} f_{\rm ns,even}(x)\cos\Big(\frac{n\pi x}{2L}\Big)\,dx, \qquad n = 0, 1, 2, \cdots,$$

Since the function $f_{\rm ns,even}$ is even, the integrand above is even, then

$$\tilde{c}_n = \frac{2}{2L}\int_0^{2L} f_{\rm ns}(x)\cos\Big(\frac{n\pi x}{2L}\Big)\,dx.$$

Now we split the integral in two integrals and use the definition of the negative symmetric extension,

$$\tilde{c}_n = \frac{1}{L}\Big(\int_0^L f(x)\cos\Big(\frac{n\pi x}{2L}\Big)\,dx - \int_L^{2L} f(2L-x)\cos\Big(\frac{n\pi x}{2L}\Big)\,dx\Big).$$

The change of variables $y = 2L - x$ in the second term above implies

$$\begin{aligned}\int_L^{2L} f(2L-x)\cos\Big(\frac{n\pi x}{2L}\Big)\,dx &= \int_L^0 f(y)\cos\Big(\frac{n\pi(2L-y)}{2L}\Big)\,(-dy)\\ &= \int_0^L f(y)\cos\Big(n\pi - \frac{n\pi y}{2L}\Big)\,dy\\ &= \int_0^L f(x)\cos\Big(n\pi - \frac{n\pi x}{2L}\Big)\,dx.\end{aligned}$$

The trigonometric identity $\cos(\theta+\phi) = \cos(\theta)\cos(\phi) - \sin(\phi)\sin(\theta)$ implies

$$\cos\Big(n\pi - \frac{n\pi x}{2L}\Big) = \cos(n\pi)\cos\Big(\frac{n\pi x}{2L}\Big) - \sin\Big(\frac{n\pi x}{2L}\Big)\sin(n\pi) = (-1)^n\cos\Big(\frac{n\pi x}{2L}\Big),$$

and this equation in the integral above gives

$$\int_L^{2L} f(2L-x)\cos\Big(\frac{n\pi x}{2L}\Big)\,dx = (-1)^n\int_0^L f(x)\cos\Big(\frac{n\pi x}{2L}\Big)\,dx.$$

Therefore, the coefficient $\tilde{c}_n$ is

$$\tilde{c}_n = (1-(-1)^n)\frac{1}{L}\int_0^L f(x)\cos\Big(\frac{n\pi x}{2L}\Big)\,dx.$$

Half of these coefficients are zero. Indeed, if $n = 2k$ we have

$$1-(-1)^{2k} = 1 - 1^k = 1 - 1 = 0.$$

When $n = 2k-1$ we have

$$1-(-1)^{2k-1} = 1-(-1)(-1)^{2k} = 1+1 = 2.$$

Therefore, we only get the odd values of the index n,

$$\tilde{c}_{2k} = 0, \quad \tilde{c}_{2k-1} = \frac{2}{L}\frac{1}{L}\int_0^L f(x)\sin\Big(\frac{(2k-1)\pi x}{2L}\Big)\,dx.$$

Finally, relabel k into n,

$$\tilde{c}_{2n} = 0, \quad \tilde{c}_{2n-1} = \frac{2}{L}\int_0^L f(x)\sin\Big(\frac{(2n-1)\pi x}{2L}\Big)\,dx.$$

Therefore, half the terms in Eq. (6.3.42) vanish, including $\tilde{c}_0$, and denoting $c_n = \tilde{c}_{2n-1}$, we obtain that the function f on $(0, L]$ can be expanded as

$$f(x) = \sum_{n=1}^{\infty} c_n\cos\Big(\frac{(2n-1)\pi x}{2L}\Big),$$

with

$$c_n = \frac{2}{L}\int_0^L f(x)\cos\Big(\frac{(2n-1)\pi x}{2L}\Big)\,dx.$$

This establishes the Theorem for Part **(d)** of the theorem, then the whole Theorem. □

Example 6.3.7. Find the Fourier expansion, F_o, of the odd extension of the function

$$f(x) = 1 - x \quad \text{for} \quad x \in (0, 1].$$

Solution: Since the function is $f(x) = 1 - x$ for $x \in (0, 1]$, its odd extension is

$$f_o(x) = \begin{cases} 1 - x, & x \in (0, 1] \\ 0, & x = 0 \\ -1 - x, & x \in [-1, 0). \end{cases}$$

Since f_o is a piecewise function defined in $[-1, 1]$, it has a Fourier expansion F_o. Since f_o is odd, its Fourier expansion is a sine series,

$$F_o(x) = \sum_{n=1}^{\infty} b_n \sin(n\pi x),$$

where we used that $L = 1$. And the coefficients b_n can be obtained from the formula

$$b_n = \int_{-1}^{1} f_o(x) \sin(n\pi x)\, dx = 2\int_0^1 f(x) \sin(n\pi x)\, dx = 2\int_0^1 (1 - x) \sin(n\pi x)\, dx.$$

Recalling that $\int x \sin(ax)\, dx = -(x/a)\cos(ax) + (1/a^2)\sin(ax)$, then we get

$$\begin{aligned} b_n &= 2\Big(\frac{(-1)}{n\pi}\cos(n\pi x) + \frac{x}{n\pi}\cos(n\pi x) - \frac{1}{n^2\pi^2}\sin(n\pi x)\Big)\Big|_0^1 \\ &= 2\Big(\big(\frac{(-1)}{n\pi}\cos(n\pi) + \frac{1}{n\pi}\cos(n\pi) - 0\big) - \big(\frac{(-1)}{n\pi} + 0 - 0\big)\Big) \\ &= \frac{2}{n\pi}. \end{aligned}$$

Therefore, the Fourier expansion of the odd extension of the function f above is

$$F_o(x) = \sum_{n=1}^{\infty} \frac{2}{n\pi} \sin(n\pi x).$$

◁

Example 6.3.8. Find the Fourier expansion, F_e, of the even extension of the function

$$f(x) = 1 - x \quad \text{for} \quad x \in (0, 1].$$

Solution: Since the function is $f(x) = 1 - x$ for $x \in (0, 1]$, its even extension is

$$f_e(x) = \begin{cases} 1 - x, & x \in (0, 1] \\ 1, & x = 0 \\ 1 + x, & x \in [-1, 0). \end{cases}$$

Since f_e is a continuous function defined in $[-1, 1]$, it has a Fourier expansion F_e. Since f_e is oeven, its Fourier expansion is a cosine series,

$$F_e(x) = \sum_{n=1}^{\infty} a_n \cos(n\pi x),$$

where we used that $L = 1$. And the coefficients a_n can be obtained from the formula

$$b_n = \int_{-1}^{1} f_e(x) \cos(n\pi x)\, dx = 2\int_0^1 f(x) \cos(n\pi x)\, dx = 2\int_0^1 (1 - x) \cos(n\pi x)\, dx.$$

Recalling that $\int x\cos(ax)\,dx = (x/a)\sin(ax) + (1/a^2)\cos(ax)$, then we get

$$\begin{aligned} a_n &= 2\Big(\frac{1}{n\pi}\sin(n\pi x) - \frac{x}{n\pi}\sin(n\pi x) - \frac{1}{n^2\pi^2}\cos(n\pi x)\Big)\Big|_0^1 \\ &= 2\Big(\big(0 - 0 - \frac{1}{n^2\pi^2}\cos(n\pi)\big) - (0 - 0 - \frac{1}{n^2\pi^2})\Big) \\ &= \frac{2}{n^2\pi^2}(1 - (-1)^n). \end{aligned}$$

Therefore, the Fourier expansion of the even extension of the function f above is

$$F_e(x) = \sum_{n=1}^{\infty} \frac{2}{n^2\pi^2}(1 - (-1)^n)\cos(n\pi x).$$

$\lhd$

6.3.6. Exercises.

6.3.1.- .

6.3.2.- .

6.4. The Heat Equation

We now solve our first *partial* differential equation—the heat equation—which describes the temperature of a solid material as function of time and space. This is a partial differential equation because it contains partial derivatives of both time and space variables.

Partial differential equations have infinitely many solutions, but one can find appropriate boundary conditions and initial conditions so that the solution is unique. In this section we first solve the equation with general boundary conditions, called Robin boundary conditions. The we explicitly compute the solutions in particular cases, called Dirichlet, Neumann, and Mixed boundary conditions.

The Dirichlet condition keeps the temperature constant on the sides of the material, the Neumann condition prevents heat from entering or leaving the material, and Mixed conditions are a mix of Dirichlet on one side and Neumann on the other side. For either type of boundary conditions, the initial condition is the same—the initial temperature of the material.

We use both the Sturm-Liouville theory and the separation of variables method to solve the heat equation. In the latter method we transform the problem with the partial differential equation into two problems for ordinary differential equations. One of these problems is an initial value problem for a first order equation and the other problem is an eigenfunction problem for a second order equation.

6.4.1. Overview of the Heat Equation. We want to describe how the temperature changes inside a solid material. Our main variable is the function $T(t, x, y, z)$ representing the temperature of a solid material at time t in a position (x, y, z). To fix ideas consider that the material is a rectangular box and our coordinate system is set as in the left side of Fig. 9. The temperature of this material is described by the three-space dimensional heat equation. But in this section, however, we study the one-space dimensional heat equation, because it is simper than the original three-dimensional problem.

Notice that the one-space dimensional heat equation has applications in our three-dimensional world. We can transform our three-dimensional problem into a one-dimensional problem by doing three things. First we add thermal insulation on the top, bottom, front and back sides of the rectangular box. Second, we choose an initial temperature of the bar to be function of only the x-coordinate. The latter means that all points in the red square in the picture on the right side in Fig. 9 have the same temperature initially. Third, any other additional conditions on the temperature must be a function of x alone. Then, one can show that the three-space dimensional heat equation implies that the temperature of this system will depend only one space variable, x, and on time, t, of course.

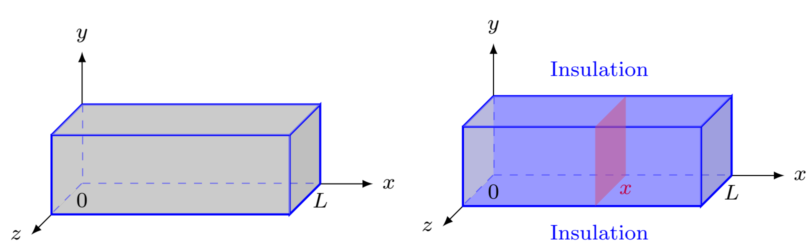


FIGURE 9. On the left we have a rectangular-shaped solid material and our coordinate system. On the right we prepare the system with thermal insulation on four of its six sides and an initial Temperature that depends only on the x variable.

The heat equation describes the temperature changes inside a solid material. Fluid materials are not described by this equation, because warmer parts of a fluid have different densities that colder parts, which originates movement inside the fluid. These movements, called convection currents, are not described by the heat equation above.

Now, let us introduce the one-space dimensional heat equation

Definition 6.4.1. *The **heat equation** in one-space dimension, for a function $T(t,x)$ is*

$$\partial_t T(t,x) = \alpha\, \partial_x^2 T(t,x), \quad \text{for} \quad t \in (0,\infty), \quad x \in (a,b),$$

with $\alpha > 0$, $b > a$ constants and ∂_t, ∂_x partial derivatives with respect to t and x.

The function T represents the temperature of a solid material, such as the rectangular box in Fig. 9, the variable t is a time coordinate, and the variable x is a space coordinate. We usually work on the interval $[a,b] = [0,L]$. The constant $\alpha > 0$ is the thermal diffusivity, which has units of $[\alpha] = [x]^2/[t]$. This constant characterizes how fast the heat travels along the material. The higher the value of α the faster the heats moves inside the material. Metals have higher values of α than thermal insulators such as plastic or wood. A constant thermal diffusivity is a good approximation for homogeneous in space materials, but more complex materials are described by a nonconstant function α that may depend on (x,y,t) and even on time t. In this section we assume that α is constant.

Before we start any detailed calculation to solve the heat equation we show the qualitative behavior of its solutions. The meaning of the left-hand side and the right-hand side of the heat equation above are the following

$$\left.\begin{array}{l}\text{How fast the temperature}\\ \text{increases of decreases.}\end{array}\right\} \quad = \alpha \ \ (>0) \quad \left\{\begin{array}{l}\text{The concavity of the graph of } T\\ \text{in the variable } x \text{ at a given time.}\end{array}\right.$$

Suppose that at a fixed time $t \geqslant 0$ the graph of the temperature T as function of x is given by Fig. 10. We assume that the boundary conditions are $T(t,0) = T_0 = 0$ and $T(t,L) = T_L > 0$. Since the heat equation relates the time variation of the temperature, $\partial_t T$, to the curvature of the function T in the x variable, $\partial_x^2 T$, then we have the following:

(a) In the regions where the function T is concave up, hence $\partial_x^2 T > 0$, the heat equation says that the temperature must increase $\partial_t T > 0$.
(b) In the regions where the function T is concave down, hence $\partial_x^2 T < 0$, the heat equation says that the temperature must decrease $\partial_t T < 0$.

Therefore, the temperature will evolve in time following the red arrows in Fig. 10.

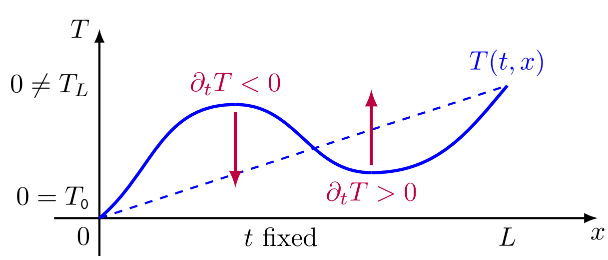


FIGURE 10. Qualitative behavior of a solution to the heat equation.

We conclude that the heat equation tries to make the temperature along the material to vary the least possible in a way that is consistent with the boundary conditions. In the case of the figure below, the temperature will try to get to the dashed line as $t \to \infty$.

We end this overview of the heat equation mentioning a few generalizations and a couple of similar equations.

(a) The heat equation in three space dimensions is

$$\partial_t T = \nabla \cdot \big(\alpha(\boldsymbol{x})\, \nabla T\big),$$

where the thermal diffusivity is a positive, differentiable function of $\boldsymbol{x} = (x, y, z)$, and we denoted by $\nabla T = \langle \partial_x T, \partial_y T, \partial_z T\rangle$ the gradient of the temperature. In the case that the thermal diffusivity k is constant, we get

$$\partial_t T = \alpha\, \nabla^2 T,$$

where $\nabla^2 T = \partial_x^2 T + \partial_y^2 T + \partial_z^2 T$. In the case the Temperature depends only on one-space dimension, say x, we reobtain the equation in Def. 6.4.1,

$$\partial_t T = \alpha\, \partial_x^2 T.$$

The method we use in this section to solve the one-space dimensional equation can be generalized to solve the three-space dimensional equation.

(b) The wave equation in three-space dimensions is

$$\partial_t^2 u = v^2\, \nabla^2 u.$$

This equation describes how waves propagate in a medium. The constant v has units of velocity, and it is the wave's speed. The function u describe a property of the medium, such as pressure in a liquid. In this case the equation would describe how sound propagates in the liquid or pressure in a gas, such as air.

(c) The Schrödinger equation of Quantum Mechanics is

$$i\hbar\, \partial_t \psi = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(t, \boldsymbol{x})\, \psi,$$

where m is the mass of a particle and $\hbar$ is the Planck constant divided by 2π, while $i^2 = -1$, and $\psi(t, \boldsymbol{x})$ is the probability density of finding the particle at the position $\boldsymbol{x}$ in space and at the time t. The presence of the complex i in the equation has an important effect on the solutions—the solutions of the Schrödinger equation behave more like the solutions of the wave equation than the solutions of the heat equation.

6.4.2. Boundary Conditions. The heat equation contains partial derivatives with respect to time and space. Solving the equation means to integrate in space and time variables. When integrate in the space variable we introduce an arbitrary function of time, and when we integrate in time we introduce an arbitrary function of the space variable. This means we have infinitely many solutions to the heat equation, one for each choice of these integration functions.

The infinitely many solutions to the heat equation correspond to the infinitely many different ways we can set up our solid bar. Different solutions for the temperature could correspond to different initial temperature in the bar. We can also set up the bar so that heat is allowed to flow inside or outside the bar from one side and not from another side.

Our physical experience tells us how to set up the bar so that we have a unique solution to the heat equation. First, we need to set an initial temperature for the bar. This is called an initial condition for the heat equation and it means to fix a function $\tau(x)$ so that

$$T(t = 0, x) = \tau(x).$$

The initial temperature is not enough to determine the behavior of the temperature inside the material. Second, we also need to control the heat coming in or going out of the material through its surface. The conditions we fix on this surface are called boundary conditions.

In Figures 11-14 we list the four most common boundary conditions on the interval $[a,b]=[0,L]$, called Dirichlet, Neumann, and mixed boundary conditions.

$$\text{Dirichlet BC: } \begin{cases} T(t,0)=0,\\ T(t,L)=0. \end{cases} \qquad \text{Neumann BC: } \begin{cases} \partial_x T(t,0)=0,\\ \partial_x T(t,L)=0. \end{cases}$$

$$\text{Mixed BC: } \begin{cases} T(t,0)=0,\\ \partial_x T(t,L)=0. \end{cases} \qquad \text{Mixed BC: } \begin{cases} \partial_x T(t,0)=0,\\ T(t,L)=0. \end{cases}$$

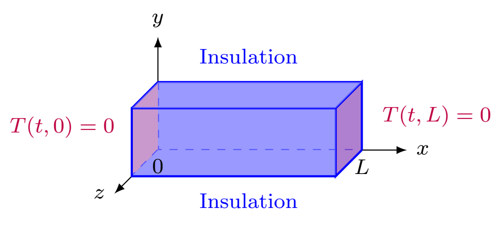


FIGURE 11. Dirichlet BC.

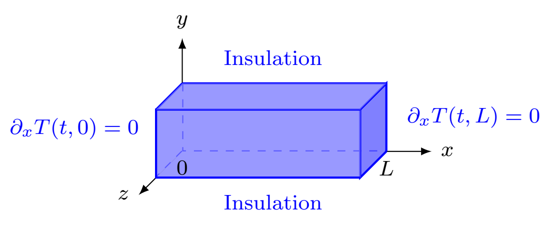


FIGURE 12. Neumann BC.

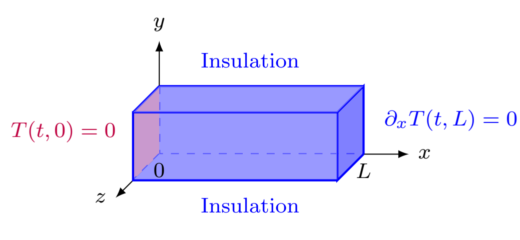


FIGURE 13. Mixed BC.

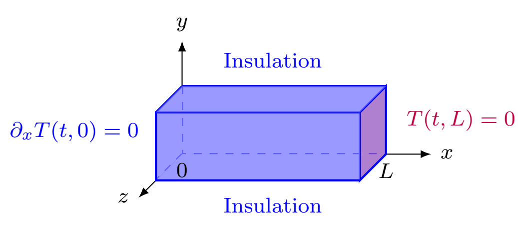


FIGURE 14. Mixed BC.

Dirichlet boundary conditions are named after Johann Peter Gustav Lejeune Dirichlet (1805-1859) who worked this problem around 1850. These conditions fix the temperature on the sides $x=0$ and $x=L$ of our solid bar. If the temperature inside the bar is larger than the temperatures at the boundary conditions, then heat will go out of the bar. If the temperature inside the bar is lower than the temperatures at the boundary conditions, then heat will go into the bar.

Neumann boundary conditions are named after Cal Gottfried Neumann (1832-1925) who worked in this problem in the 1860s. To understand the physical meaning of these conditions we need to know that the heat flux in a solid material is proportional to the negative gradient of the temperature, that is

$$\boldsymbol{q} = -k\nabla T,$$

where $\boldsymbol{q}$ is a the heat flux, which has units of $[\text{Energy}]/([x]^2\,[t])$, $k>0$ is the thermal conductivity, with units $[k]=[\text{Energy}]/([x]\,[t]\,[T])$, and ∇T is the gradient of the temperature. This is the case because the temperature gradient points in the direction where the temperature increases. Since the heat propagates from hot places to cold places, then the heat flux must point in the opposite direction of the temperature gradient. Neumann boundary condition says that the component of the heat flux normal to the surfaces $x = 0$ and $x = L$ is zero. This means that no heat is coming in or going out the bar through these boundaries. In other words, we have thermal insulation on these boundaries. Mixed boundaries conditions are just a combination of Dirichelt conditions on one boundary and Neumann conditions on the other boundary.

Before we write our main result, we show one more picture. Fig. 15 we shows the domain in the tx-space where the solution of the heat equation is defined. We highlight the part of the domain where we prescribe the initial data (green) and the boundary conditions (red). The values of the solution in the blue region are obtained solving the differential equation.

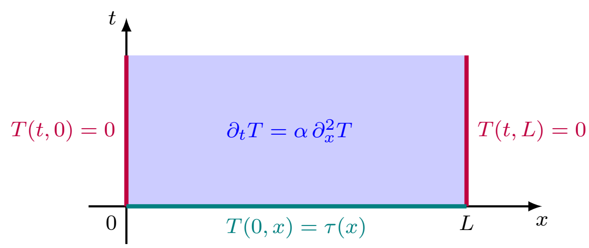


FIGURE 15. Sketch of the initial-boundary value problem on the tx-plane for Dirichlet boundary conditions.

6.4.3. Initial Boundary Value Problem. The problem we are going to solve is called an initial boundary value problem (IBVP) for the heat equation. We are going to find a function temperature solution of the heat equation that satisfy a prescribed initial temperature and prescribed boundary conditions at $x = a$ and $x = b$. We now summarize our main result and we prove it using the Sturm-Liouville theory. After that we show how to find solutions of this problem using the separation of variables method.

Theorem 6.4.2 (General IBVP)**.** *Consider the initial boundary value problem for the one-space dimensional heat equation,*

$$\partial_t T = \alpha\, \partial_x^2 T, \qquad t \in (0, \infty), \qquad x \in (a, b),$$

where $\alpha > 0$, $b > a$ are constants, with homogeneous boundary conditions

$$a_1\, T(t, a) + a_2\, \partial_x T(t, a) = 0, \tag{6.4.1}$$

$$b_1\, T(t, b) + b_2\, \partial_x T(t, b) = 0, \tag{6.4.2}$$

where a_1, a_2, b_1, b_2 are given constants, and with initial condition

$$T(0, x) = \tau(x), \qquad x \in [a, b].$$

The initial boundary value problem above has a unique solution, $T(t, x)$, given by

$$T(t, x) = \sum_{n=0}^{\infty} c_n\, e^{-\alpha \lambda_n t}\, w_n(x),$$

where the constants λ_n and the non-zero functions $w_n(x)$, for $n = 0, 1, 2, \cdots$, are eigenvalues and unit eigenfunctions solutions of the Regular Sturm-Liouville System

$$-w'' = \lambda\, w, \qquad \begin{aligned} a_1\, w(a) + a_2\, w'(a) &= 0, \\ b_1\, w(b) + b_2\, w'(b) &= 0, \end{aligned}$$

and the coefficients c_n *are determined from the initial temperature,*

$$c_n = \int_a^b \tau(x)\, w_n(x)\, dx.$$

Remarks: We see in the result above that we do not write the functions $w_n(x)$ explicitly. The explicit form of these solutions depends on the values of a_1, a_2, b_1, b_2 in boundary conditions. The boundary conditions in the theorem above include the following particular cases.

(a) *Dirichlet Boundary Conditions:*

$$T(t,a) = 0, \qquad T(t,b) = 0,$$

which corresponds to $a_1 = 1$, $a_2 = 0$, $b_1 = 1$, $b_2 = 0$. These conditions fix the temperature at the ends of the object. While the temperature is constant at the border of the object, heat could be flowing through the boundary, since the heat flux is controlled by $\partial_x T$. As we mentioned earlier, these conditions are named after Peter Dirichlet, who worked in this problem in the 1850s.

(b) *Neumann Boundary Conditions:*

$$\partial_x T(t,a) = 0, \qquad \partial_x T(t,b) = 0,$$

which correspond to the case $a_1 = 0$, $a_2 = 1$, $b_1 = 0$, $b_2 = 1$. These conditions fix the heat flux at the boundary of the object, they say that no heat is going through the boundary of the object. In other words, the boundary is thermally insulated. In this case the temperature of the object could be changing at the boundary, since there is no condition on the temperature values. As we mentioned earlier, these conditions are named after Cal Neumann, who worked in this problem in the 1860s.

(c) *Mixed Boundary Conditions:*

$$T(t,a) = 0, \quad \partial_x T(t,b) = 0, \qquad \text{or} \qquad \partial_x T(t,a) = 0, \quad T(t,b) = 0,$$

which correspond to either $a_1 = 1$, $a_2 = 0$, $b_1 = 0$, $b_2 = 1$, or $a_1 = 0$, $a_2 = 1$, $b_1 = 1$, $b_2 = 0$, respectively. These conditions combine Dirichlet conditions on one side and Neumann conditions on the other side.

(d) *Steklov Boundary Conditions:*

$$a_1\, T(t,a) + a_2\, \partial_x T(t,a) = 0,$$
$$b_1\, T(t,b) + b_2\, \partial_x T(t,b) = 0,$$

which correspond to the case $a_1 \neq 0$, $a_2 \neq 0$, $b_1 \neq 0$, $b_2 \neq 0$. These conditions describe more complicated situations involving semipermeable boundaries for the heat equation and elastic conditions for the wave equation. These boundary condition are usually named Robin boundary conditions, after Victor Gustave Robin (1855-1897). However, Robin never wrote these conditions nor solved a partial differential equation with these boundary conditions. It seems that Vladimir Andreevich Steklov (1864-1926) first worked in this problem in 1900.

To understand the idea of the proof of this theorem we need to recall Theorem 6.2.8 in § 6.2. This result says that any continuous function $f(x)$ can be expanded in terms of the solutions of a Regular Sturm-Liouville System $y_n(x)$, for $n = 0, 1, 2, \cdots$, as follows,

$$f(x) = \sum_{n=0}^{\infty} c_n\, y_n(x), \qquad c_n = \frac{f \cdot y_n}{y_n \cdot y_n},$$

where we used the inner product

$$g \cdot h = \int_a^b g(x)\, h(x), dx.$$

If the function f is function of two variables, t, and x, we can still expand f in using the eigenfunctions $y_n(x)$, only in this case the coefficients c_n depend on t, that is,

$$f(t,x) = \sum_{n=0}^{\infty} c_n(t)\, y_n(x), \qquad c_n(t) = \frac{f \cdot y_n}{y_n \cdot y_n}.$$

The idea of the proof of Theorem 6.4.2 is to decompose the temperature $T(t,x)$ in terms of an orthogonal set of functions, $w_n(x)$, as follows

$$T(t,x) = \sum_{n=0}^{\infty} v_n(t), w_n(x),$$

where the coefficients in this expansion, v_n, depend on the variable t. We put this expression in the heat equation and choose the functions w_n as solutions of a Sturm-Liouville problem

$$w_n'' + \lambda\, w_n = 0,$$

satisfying the boundary conditions of the heat equation. The orthogonality of the eigenfunctions w_n will simplify the equation for the coefficients $v_n(t)$ because they decouple the equations for each value of the index n. One more comment, in the proof of Theorem 6.4.2 we use unit eigenfunctions, so $w_n \cdot w_n = 1$.

Proof of Theorem 6.4.2: Let $w_n(x)$ be unit eigenfunctions of a Regular Sturm-Liouville System, which we will specify later. Then, no matter what that Regular Sturm-Liouville System is, we can always expand the temperature function $T(t,x)$ in terms of eigenfunctions $w_n(x)$,

$$T(t,x) = \sum_{n=0}^{\infty} v_n(t)\, w_n(x), \qquad v_n(t) = \frac{T \cdot w_n}{w_n \cdot w_n}. \tag{6.4.3}$$

Just to be clear, we have named the coefficients in the expansion as $v_n(t)$ and the unit eigenfunctions solutions of a Regular Sturm-Liouville System—to be chosen later—as $w_n(x)$. Now we want to put that expression for $T(t,x)$ into the heat equation. In order to do that we need to compute the time derivative,

$$\partial_t T(t,x) = \sum_{n=0}^{\infty} \Big(\frac{d}{dt} v_n(t)\Big)\, w_n(x) = \sum_{n=0}^{\infty} \dot{v}_n(t)\, w_n(x),$$

where we used the notation $\frac{d}{dt} v_n = \dot{v}_n$. Now we compute the space derivatives,

$$\partial_x^2 T(t,x) = \sum_{n=0}^{\infty} v_n(t) \Big(\frac{d^2}{dx^2} w_n(x)\Big) = \sum_{n=0}^{\infty} v_n(t)\, w_n''(x),$$

where we used the notation $\frac{d}{dx} w_n = w_n'$. Then, the heat equation, $\partial_t T = k\, \partial_x^2 T$ implies,

$$\sum_{n=0}^{\infty} \dot{v}_n(t)\, w_n(x) = \alpha \sum_{n=0}^{\infty} v_n(t)\, w_n''(x). \tag{6.4.4}$$

Now, we turn to the boundary conditions. If we introduce the expansion in Eq. 6.4.3 into the boundary conditions we get,

$$\sum_{n=0}^{\infty} v_n(t)\,\big(a_1\,w_n(a) + a_2\,w_n'(a)\big) = 0, \tag{6.4.5}$$

$$\sum_{n=0}^{\infty} v_n(t)\,\big(b_1\,w_n(b) + b_2\,w_n'(b)\big) = 0. \tag{6.4.6}$$

Now is when we choose the functions $w_n(x)$. We choose these functions to be the unit eigenfunctions solution of the Regular Sturm-Liouville System,

$$-w_n'' = \lambda_n\,w_n, \qquad \begin{aligned} a_1\,w_n(a) + a_2\,w_n'(a) = 0,\\ b_1\,w_n(b) + b_2\,w_n'(b) = 0. \end{aligned}$$

The boundary conditions above for w_n imply that every term in the sums in Eqs. (6.4.5), (6.4.6) vanish. Therefore, the temperature function $T(t,x)$ given in Eq. (6.4.3) satisfies the boundary conditions in our problem, given in Eqs. (6.4.1), (6.4.2). The differential equation satified by w_n implies that the heat equation Eq. (6.4.4) is now given by

$$\sum_{n=0}^{\infty} \dot v_n(t)\,w_n(x) = \alpha \sum_{n=0}^{\infty} v_n(t)\,(-\lambda_n)\,w_n(x),$$

which gives us the equation

$$\sum_{n=0}^{\infty} \big(\dot v_n(t) + \alpha\lambda_n\,v_n(t)\big)\,w_n(x) = 0.$$

Since the solutions, w_n, of the Regular Sturm-Liouville System are mutually orthogonal, the equation above implies that the equation above is satisfied term by term,

$$\dot v_n(t) = -\alpha\lambda_n\,v_n(t), \qquad n = 0, 1, 2, \cdots.$$

We know how to solve this equation for v_n, since it is the radioactive decay equation,

$$v_n(t) = c_n\,e^{-\alpha\lambda_n t}.$$

Therefore, we obtain that $T(t,x)$ satisfies the boundary conditions and it is given by

$$T(t,x) = \sum_{n=0}^{\infty} c_n\,e^{-\alpha\lambda_n t}\,w_n(x).$$

The coefficients c_n are determined by the initial temperature $T(0,x) = \tau(x)$. Indeed,

$$\tau(x) = T(0,x) = \sum_{n=0}^{\infty} c_n\,w_n(x)$$

implies the coefficients c_n are given by

$$c_n = \frac{\tau \cdot w_n}{w_n \cdot w_n},$$

but since the eigenfunctions are unit, we get

$$c_n = \tau \cdot w_n \quad \Rightarrow \quad c_n = \int_a^b \tau(x)\,w_n(x)\,dx.$$

This establishes the Theorem. □

Remark: It is common in the literature to solve the heat equation using the *separation of variables method*, which can be summarized as follows.

(a) We start looking for simple solutions of the boundary value problem.
(b) The superposition property says that additions of simple solutions is also a solution.
(c) We determine the free constants in the superposition with the initial condition.

Now we show the separation of variables method and then we compare both calculations.

Separation of Variables and Theorem 6.4.2: The separation of variables method consists in looking for simple solutions of the heat equation having the particular form

$$u(t,x) = v(t)\, w(x).$$

So we look for solutions having the variables separated into two functions. Introduce this particular function in the heat equation,

$$\dot{v}(t)\, w(x) = \alpha\, v(t)\, w''(x) \quad \Rightarrow \quad \frac{1}{\alpha}\,\frac{\dot{v}(t)}{v(t)} = \frac{w''(x)}{w(x)},$$

where we used the notation $\frac{d}{dt}v = \dot{v}$ and $\frac{d}{dx}w = w'$. The separation of variables in the function u implies a separation of variables in the heat equation. The left-hand side in the last equation above depends only on t and the right-hand side depends only on x. The only possible solution is that both sides are equal the same constant, call it $-\lambda$. So we end up with two equations

$$\frac{1}{\alpha}\,\frac{\dot{v}(t)}{v(t)} = -\lambda, \qquad \text{and} \qquad \frac{w''(x)}{w(x)} = -\lambda.$$

The equation on the left is first order and simple to solve. The solution depends on λ,

$$v_\lambda(t) = c_\lambda\, e^{-\alpha\lambda t}, \qquad c_\lambda = v_\lambda(0).$$

The second equation leads to an eigenfunction problem for w once boundary conditions are provided. These boundary conditions come from the boundary conditions for the heat equation,

$$\left.\begin{aligned} a_1 u(t,a) + a_2 \partial_x u(t,a) = 0 \quad \text{for all } t \geqslant 0,\\ b_1 u(t,b) + b_2 \partial_x u(t,b) = 0 \quad \text{for all } t \geqslant 0, \end{aligned}\right\} \quad \Rightarrow \quad \begin{cases} v(t)\,\big(a_1 w(a) + a_2 w'(a)\big) = 0,\\ v(t)\,\big(b_1 w(b) + b_2 w'(b)\big) = 0. \end{cases}$$

Since we have found that the functions $v_\lambda(t)$ are nonzero for all t, we get conditions on the functions $w(x)$,

$$\begin{aligned} a_1 w(a) + a_2 w'(a) = 0,\\ b_1 w(b) + b_2 w'(b) = 0. \end{aligned}$$

So we need to solve the following eigenfunction problem for $w(x)$;

$$w'' + \lambda\, w = 0, \qquad \begin{cases} a_1 w(a) + a_2 w'(a) = 0,\\ b_1 w(b) + b_2 w'(b) = 0. \end{cases}$$

This is a regular Sturm-Liouville system and we denote its solutions to be, respectively, the eigenvalues and unit eigenfunctions

$$\lambda_n, \quad w_n(x), \quad n = 0, 1, 2, \cdots .$$

Since we now know the values of λ_n, we introduce them in $v_n(t) = v_{\lambda_n}(t)$,

$$v_n(t) = c_n\, e^{-\alpha\lambda_n t}.$$

Therefore, we got a simple solution of the heat equation that solves the boundary conditions in the problem,

$$u_n(t,x) = c_n\, e^{-\alpha\lambda_n t}\, w_n(x),$$

where $n = 1, 2, \cdots$. Since the boundary conditions for u_n are homogeneous, then any linear combination of the solutions u_n is also a solution of the heat equation with homogenous boundary conditions. Hence the function

$$T(t,x) = \sum_{n=0}^{\infty} c_n \, e^{-\alpha\lambda_n t} \, w_n(x)$$

is solution of the heat equation with the boundary conditions in (6.4.1), (6.4.2). Here the c_n are arbitrary constants. Notice that at $t = 0$ we have

$$\tau(x) = T(0,x) = \sum_{n=0}^{\infty} c_n \, w_n(x).$$

This is the orthogonal decomposition of the initial temperature $\tau(x)$ in terms of the unit eigenfunctions $w_n(x)$. We know that any continuous function with piecewise continuous derivative on $[a,b]$ can be written in that form, and we also know that the coefficients, c_n, for $n = 0, 1, 2, \cdots$, are given by

$$c_n = \tau \cdot w_n = \int_a^b \tau(x) \, w_n(x) \, dx.$$

This establishes the Theorem. □

The main difference between the two calculations above to find solutions of the heat equation is that in the second calculation, the separation of variables method, we do not show that the solution found is the only solution of the initial boundary value problem. As far as that the separation of variables method goes, there could be another solution that we do not know about, which is not possible to write as a sum of simple solutions. However, the first calculation using the Sturm-Liouville theory says that this is not the case. The uniqueness of the solution to the initial boundary value problem for the heat equation is proven only when we use the Sturm-Liouville theory.

6.4.4. Dirichlet, Neumann, and Mixed Problems. We have seen how to solve the heat equation for general boundary conditions. We have also seen that these conditions include the Dirichlet, Neumann, and two Mixed boundary conditions. In these cases we can solve explicitly the Sturm-Liouville problem for the eigenvalues λ_n and unit eigenfunctions $w_n(x)$. Now we summarize the solution of the heat equation for these four boundary conditions in the interval $[0, L]$.

Corollary 6.4.3 (Dirichlet, Neumann, Mixed IBVP)**.** *Consider the initial boundary value problems for the one-space dimensional heat equation,*

$$\partial_t T = \alpha \, \partial_x^2 T, \qquad t \in (0,\infty), \qquad x \in (0,L),$$

where $\alpha > 0$, $L > 0$ are constants, with boundary conditions of any of the following types:

Dirichlet:	$T(t,0) = 0,$	$T(t,L) = 0,$	(6.4.7)
Neumann:	$\partial_x T(t,0) = 0,$	$\partial_x T(t,L) = 0,$	(6.4.8)
Mixed 1:	$T(t,0) = 0,$	$\partial_x T(t,L) = 0,$	(6.4.9)
Mixed 2:	$\partial_x T(t,0) = 0,$	$T(t,L) = 0,$	(6.4.10)

and with initial condition

$$T(0,x) = \tau(x), \qquad x \in [0,L].$$

Each of these initial boundary value problems in (1)-(4) has a unique solution, $T(t,x)$,

$$T(t,x) = \sum_{n=0 \text{ or } n=1}^{\infty} c_n \, e^{-\alpha\lambda_n t} \, w_n(x),$$

where λ_n and $w_n(x)$ are the eigenvalues and eigenfunctions solutions of the regular Sturm-Liouville systems given by the differential equation

$$w''(x) + \lambda \, w(x) = 0,$$

together with the respective boundary conditions

$$\begin{aligned}
&\textit{Dirichlet:} \quad w(0)=0, \quad w(L)=0, \quad \Rightarrow \quad \lambda_n = \Big(\frac{n\pi}{L}\Big)^2, \quad w_n(x) = \sin\Big(\frac{n\pi x}{L}\Big),\\
&\textit{Neumann:} \quad w'(0)=0, \quad w'(L)=0, \quad \Rightarrow \quad \begin{cases} \lambda_0 = 0, & w_0(x) = \dfrac{1}{2},\\ \lambda_n = \Big(\dfrac{n\pi}{L}\Big)^2, & w_n(x) = \cos\Big(\dfrac{n\pi x}{L}\Big),\end{cases}\\
&\textit{Mixed 1:} \quad w(0)=0, \quad w'(L)=0, \quad \Rightarrow \quad \lambda_n = \Big(\frac{(2n-1)\pi}{2L}\Big)^2, \quad w_n(x) = \sin\Big(\frac{(2n-1)\pi x}{2L}\Big),\\
&\textit{Mixed 2:} \quad w'(0)=0, \quad w(L)=0, \quad \Rightarrow \quad \lambda_n = \Big(\frac{(2n-1)\pi}{2L}\Big)^2, \quad w_n(x) = \cos\Big(\frac{(2n-1)\pi x}{2L}\Big),
\end{aligned}$$

where $n = 1, 2, 3, \cdots$, and the coefficients c_n are determined from the initial temperature,

$$c_0 = \frac{2}{L}\int_0^L \tau(x)\, dx, \qquad c_n = \frac{2}{L}\int_0^L \tau(x)\, w_n(x)\, dx.$$

Remarks:

(a) This corollary follows from Theorem 6.4.2 and Examples 6.2.2-6.2.5.
(b) The eigenfunctions $w_n(x)$ we use in this corollary are not unit functions.

Example 6.4.1 (Dirichlet)**.** Use the Sturm-Liouville theory to find the solution to the initial-boundary value problem

$$4\,\partial_t T = \partial_x^2 T, \quad t > 0, \quad x \in (0, 2),$$

with initial conditions $\tau(x) = T(0,x)$ and boundary conditions given by

$$\text{IC:} \quad \tau(x) = \begin{cases} 0 & x \in \big[0, \frac{2}{3}\big),\\ 5 & x \in \big[\frac{2}{3}, \frac{4}{3}\big],\\ 0 & x \in \big(\frac{4}{3}, 2\big], \end{cases} \qquad \text{BC:} \quad \begin{cases} T(t,0) = 0,\\ T(t,2) = 0. \end{cases}$$

Solution: In this problem we have $\alpha = 1/4$ and $L = 2$. We start writing the temperature as an expansion in orthogonal functions $w_n(x)$ that will be chosen later on. So we get

$$T(t,x) = \sum_{n=1}^{\infty} v_n(t)\, w_n(x).$$

We now compute its time and space derivatives,

$$\partial_t T(t,x) = \sum_{n=1}^{\infty} \dot v_n(t)\, w_n(x), \qquad \partial_x^2 T(t,x) = \sum_{n=1}^{\infty} v_n(t)\, w_n''(x),$$

which give us the heat equation

$$\sum_{n=1}^{\infty} \dot v_n(t)\, w_n(x) = \frac{1}{4} \sum_{n=1}^{\infty} v_n(t)\, w_n''(x), \tag{6.4.11}$$

and the Dirichlet boundary conditions,

$$T(t,0) = \sum_{n=0}^{\infty} v_n(t)\, w_n(0) = 0, \qquad T(t,0) = \sum_{n=0}^{\infty} v_n(t)\, w_n(2) = 0. \tag{6.4.12}$$

Now it is time to choose the orthogonal functions $w_n(x)$ as solution of the Regular Sturm-Liouville System

$$-w''(x) = \lambda\, w(x), \qquad w(0) = 0, \qquad w(L) = 0.$$

We have seen in Example 6.2.2 that the solutions of this eigenfunction problem are

$$\lambda_n = \Big(\frac{n\pi}{2}\Big)^2 \qquad w_n(x) = \sin\Big(\frac{n\pi x}{2}\Big), \qquad n = 1, 2, 3, \cdots .$$

Since these functions $w_n(x)$ satisfy the boundary conditions $w_n(0) = 0$ and $w_n(2) = 0$, then the boundary conditions in Eq. (6.4.12) are satisfied term by term. And since the function w_n satisfies the differential equation $w_n'' = -\lambda_n\, w_n$, then the heat equation in (6.4.11) has the form

$$\sum_{n=1}^{\infty} \dot v_n(t)\, w_n(x) = \sum_{n=1}^{\infty} \Big(-\frac{1}{4}\,\lambda_n\Big)\, v_n(t)\, w_n(x).$$

The orthogonality of the $w_n(x)$ implies the equation above is satisfied term by term,

$$\dot v_n(t) = -\frac{1}{4}\,\lambda_n\, v_n(t).$$

The solution of this equation, which is separable, and it is called the radioactive decay equation, is

$$v_n(t) = c_n\, e^{-\frac{1}{4}(\frac{n\pi}{2})^2 t}.$$

Therefore, the solution of the heat equation is

$$T(t,x) = \sum_{n=1}^{\infty} c_n\, e^{-\frac{1}{4}(\frac{n\pi}{2})^2 t}\, \sin\Big(\frac{n\pi x}{2}\Big).$$

The coefficients c_n are determined from the initial condition,

$$\tau(x) = T(0,x) = \sum_{n=1}^{\infty} c_n\, \sin\Big(\frac{n\pi x}{2}\Big).$$

Since the functions $w_n(x)$ are orthogonal, we get that

$$c_n = \frac{\tau \cdot w_n}{w_n \cdot w_n}.$$

In Example 6.2.2 we computed

$$w_n \cdot w_n = \frac{L}{2}, \qquad L = 2 \quad \Rightarrow \quad w_n \cdot w_n = 1.$$

Then, the formula for the coefficients c_n is

$$c_n = \int_0^2 \tau(x)\, \sin\Big(\frac{n\pi x}{2}\Big)\, dx.$$

If we use the particular form of the initial temperature in this problem, we get

$$\begin{aligned} c_n &= 5\int_{2/3}^{4/3} \sin\Big(\frac{n\pi x}{2}\Big)\,dx \\ &= -5\,\frac{2}{n\pi}\cos\Big(\frac{n\pi x}{2}\Big)\Big|_{2/3}^{4/3} \\ &= -\frac{10}{n\pi}\Big(\cos\Big(\frac{2n\pi}{3}\Big) - \cos\Big(\frac{n\pi x}{3}\Big)\Big). \end{aligned}$$

Therefore, the solution of the heat equation in this problem is

$$T(t,x) = \sum_{n=1}^{\infty} \frac{10}{n\pi}\,\Big(\cos\Big(\frac{n\pi x}{3}\Big) - \cos\Big(\frac{2n\pi}{3}\Big)\Big)\, e^{-\frac{1}{4}(\frac{n\pi}{2})^2 t}\,\sin\Big(\frac{n\pi x}{2}\Big).$$

◁

Example 6.4.2 (Dirichlet)**.** Use the separation of variables method to find the solution to the initial-boundary value problem (same as the previous example)

$$4\,\partial_t T = \partial_x^2 T, \quad t > 0, \quad x \in (0,2),$$

with initial conditions $\tau(x) = T(0,x)$ and boundary conditions given by

$$\text{IC:}\quad \tau(x) = \begin{cases} 0 & x \in \big[0, \frac{2}{3}\big), \\ 5 & x \in \big[\frac{2}{3}, \frac{4}{3}\big], \\ 0 & x \in \big(\frac{4}{3}, 2\big], \end{cases} \qquad \text{BC:}\quad \begin{cases} T(t,0) = 0, \\ T(t,2) = 0. \end{cases}$$

Solution: We use the separation of variable method and we look for simple solutions of the form

$$u(t,x) = v(t)\,w(x).$$

We put this function into the heat equation and we get

$$4w(x)\,\frac{dv}{dt}(t) = v(t)\,\frac{d^2w}{dx^2}(x) \quad \Rightarrow \quad \frac{4\dot v(t)}{v(t)} = \frac{w''(x)}{w(x)} = -\lambda.$$

So, the equations for v and w are

$$\dot v(t) = -\frac{\lambda}{4}\,v(t), \qquad w''(x) + \lambda\,w(x) = 0.$$

The solution for v depends on λ, and is given by

$$v_\lambda(t) = c_\lambda\, e^{-\frac{\lambda}{4}t}, \qquad c_\lambda = v_\lambda(0).$$

Next we turn to the the equation for w, and we solve the boundary value problem

$$w''(x) + \lambda\,w(x) = 0, \quad \text{with BC} \quad w(0) = w(2) = 0.$$

This is an eigenfunction problem for w and λ. This problem has solution only for $\lambda > 0$, since only in that case the characteristic polynomial has complex roots. Let $\lambda = \mu^2$, then

$$p(r) = r^2 + \mu^2 = 0 \quad \Rightarrow \quad r_\pm = \pm\mu\, i.$$

The general solution of the differential equation is

$$w_n(x) = c_1\cos(\mu x) + c_2\sin(\mu x).$$

The first boundary conditions on w implies

$$0 = w(0) = c_1, \quad \Rightarrow \quad w(x) = c_2 \sin(\mu x).$$

The second boundary condition on w implies

$$0 = w(2) = c_2 \sin(\mu 2), \quad c_2 \neq 0, \quad \Rightarrow \quad \sin(\mu 2) = 0.$$

Then, $\mu_n 2 = n\pi$, that is, $\mu_n = \dfrac{n\pi}{2}$. Choosing $c_2 = 1$, we conclude,

$$\lambda_n = \Big(\frac{n\pi}{2}\Big)^2, \qquad w_n(x) = \sin\Big(\frac{n\pi x}{2}\Big), \quad n = 1, 2, \cdots .$$

Using the values of λ_n found above in the formula for v_λ we get

$$v_n(t) = c_n\, e^{-\frac{1}{4}(\frac{n\pi}{4})^2 t}, \qquad c_n = v_n(0).$$

Therefore, we get

$$T(t,x) = \sum_{n=1}^{\infty} c_n\, e^{-(\frac{n\pi}{4})^2 t} \sin\Big(\frac{n\pi x}{2}\Big).$$

The initial condition is

$$\tau(x) = T(0,x) = \begin{cases} 0 & x \in \big[0, \frac{2}{3}\big), \\ 5 & x \in \big[\frac{2}{3}, \frac{4}{3}\big], \\ 0 & x \in \big(\frac{4}{3}, 2\big]. \end{cases}$$

Therefore, the coefficients b_n are given by

$$\begin{aligned} c_n &= \frac{2}{2} \int_0^2 f(x) \sin\Big(\frac{n\pi x}{2}\Big)\, dx \\ &= \int_{2/3}^{4/3} 5 \sin\Big(\frac{n\pi x}{2}\Big)\, dx \\ &= -\frac{10}{n\pi} \cos\Big(\frac{n\pi x}{2}\Big)\Big|_{2/3}^{4/3}, \end{aligned}$$

then, we get

$$c_n = -\frac{10}{n\pi}\Big(\cos\Big(\frac{2n\pi}{3}\Big) - \cos\Big(\frac{n\pi}{3}\Big)\Big).$$

We conclude that the solution of the initial boundary value problem for the heat equation contains is

$$T(t,x) = \sum_{n=1}^{\infty} \frac{10}{n\pi}\Big(\cos\Big(\frac{n\pi}{3}\Big) - \cos\Big(\frac{2n\pi}{3}\Big)\Big)\, e^{-\frac{1}{4}(\frac{n\pi}{2})^2 t} \sin\Big(\frac{n\pi x}{2}\Big).$$

◁

Example 6.4.3 (Dirichlet)**.** Find the solution to the initial-boundary value problem

$$\partial_t T = 4\, \partial_x^2 T, \quad t > 0, \quad x \in [0, 2],$$

with initial condition $\tau(x) = T(0,x)$ and boundary conditions given by

$$\tau(x) = 3\sin(\pi x/2), \quad T(t,0) = 0, \quad T(t,2) = 0.$$

Solution: We use the separation of variables method. We look for simple solutions

$$u(t,x) = v(t)\, w(x).$$

We put this function into the heat equation and we get

$$w(x)\,\dot v(t) = 4\,v(t)\,w''(x) \quad\Rightarrow\quad \frac{\dot v(t)}{4\,v(t)} = \frac{w''(x)}{w(x)} = -\lambda_n.$$

The equations for v and w are

$$\dot v(t) = -4\lambda\,v(t), \qquad w''(x) + \lambda\,w(x) = 0.$$

We solve for v, which depends on the constant λ, and we get

$$v_\lambda(t) = c_\lambda\,e^{-4\lambda t},$$

where $c_\lambda = v_\lambda(0)$. Next we turn to the boundary value problem for w. We need to find the solution of

$$w''(x) + \lambda\,w(x) = 0, \quad \text{with} \quad w(0) = w(2) = 0.$$

This is an eigenfunction problem for w and λ. From Example 6.2.2 we know that this problem has solutions only for $\lambda > 0$, which is when the characteristic polynomial of the equation for w has complex roots. So we write $\lambda = \mu^2$ for $\mu > 0$. The characteristic polynomial of the differential equation for w is

$$p(r) = r^2 + \mu^2 = 0 \quad\Rightarrow\quad r_\pm = \pm\mu i.$$

The general solution of the differential equation is

$$w(x) = \tilde c_1\,\cos(\mu x) + \tilde c_2\,\sin(\mu x).$$

The first boundary conditions on w implies

$$0 = w(0) = \tilde c_1, \quad\Rightarrow\quad w(x) = \tilde c_2\,\sin(\mu x).$$

The second boundary condition on w implies

$$0 = w(2) = \tilde c_2\,\sin(\mu 2), \quad \tilde c_2 \neq 0, \quad\Rightarrow\quad \sin(\mu 2) = 0.$$

Then, $\mu_n 2 = n\pi$, that is, $\mu_n = n\pi/2$, for $n \geqslant 1$. Choosing $\tilde c_2 = 1$, we conclude,

$$\lambda_n = \Big(\frac{n\pi}{2}\Big)^2, \qquad w_n(x) = \sin\Big(\frac{n\pi x}{2}\Big), \quad n = 1, 2, \cdots.$$

Using these λ_n in the expression for v_λ we get

$$v_n(t) = c_n\,e^{-4(\frac{n\pi}{2})^2 t} \quad\Rightarrow\quad v_n(t) = c_n\,e^{-(n\pi)^2 t}.$$

The expressions for v_n and w_n imply that the simple solution solution u_n has the form

$$u_n(t,x) = c_n\,e^{-(n\pi)^2 t}\,\sin\Big(\frac{n\pi x}{2}\Big).$$

Since any linear combination of the function above is also a solution, we get that the temperature function is given by

$$T(t,x) = \sum_{n=1}^{\infty} c_n\,e^{-(n\pi)^2 t}\,\sin\Big(\frac{n\pi x}{2}\Big).$$

The initial condition is

$$3\,\sin\Big(\frac{\pi x}{2}\Big) = \sum_{n=1}^{\infty} c_n\,\sin\Big(\frac{n\pi x}{2}\Big).$$

Since the functions w_n are mutually orthogonal, we conclude that

$$c_1 = 3, \quad \text{and} \quad c_m = 0 \quad \text{for} \quad m \neq 1.$$

Therefore, the solution of the initial boundary value problem for the heat equation is

$$T(t,x) = 3\,e^{-4\pi^2 t}\,\sin\Big(\frac{\pi x}{2}\Big).$$

◁

Example 6.4.4 (Neumann)**.** Find the solution to the initial-boundary value problem

$$\partial_t T = \partial_x^2 T, \quad t > 0, \quad x \in (0, 3),$$

with initial and boundary conditions given by

$$\text{IC:} \quad \tau(x) = T(0, x) = \begin{cases} 7, & x \in \big[\frac{3}{2}, 3\big], \\ 0, & x \in \big[0, \frac{3}{2}\big), \end{cases} \qquad \text{NBC:} \quad \begin{cases} \partial_x u(t, 0) = 0, \\ \partial_x u(t, 3) = 0. \end{cases}$$

Solution: We use the separation of variable method and we look for simple solutions of the form

$$u(t, x) = v(t)\, w(x).$$

We put this simple function into the heat equation and we get

$$w(x)\, \frac{dv}{dt}(t) = v(t)\, \frac{d^2 w}{dx^2}(x) \quad \Rightarrow \quad \frac{\dot{v}(t)}{v(t)} = \frac{w''(x)}{w(x)} = -\lambda.$$

So, the equations for v and w are

$$\dot{v}(t) = -\lambda\, v(t), \qquad w''(x) + \lambda\, w(x) = 0.$$

The solution for v depends on λ, and is given by

$$v_\lambda(t) = c_\lambda\, e^{-\lambda t}, \qquad c_\lambda = v_\lambda(0).$$

Next we turn to the the equation for w, and we solve the Regular Sturm-Liouville System

$$w''(x) + \lambda\, w(x) = 0, \qquad w'(0) = 0, \qquad w'(3) = 0.$$

This is an eigenfunction problem for w and λ. This problem has solution for $\lambda = 0$ and $\lambda > 0$. For $\lambda = 0$ we have

$$w''(x) = 0 \quad \Rightarrow \quad w(x) = c_1 + c_2 x.$$

The boundary conditions are for $w'(x) = c_2$, and they imply $c_2 = 0$. So we got that $w_0(x) = c_1$ is an eigenfunction with eigenvalue $\lambda_0 = 0$. Let's choose $c_0 = \frac{1}{2}$, then we have

$$\lambda_0 = 0, \quad w_0(x) = \frac{1}{2}.$$

In the case $\lambda > 0$ the characteristic polynomial of $w'' + \lambda w = 0$ has complex roots,

$$p(r) = r^2 + \lambda = 0 \quad \Rightarrow \quad r_\pm = \pm\sqrt{\lambda}\, i.$$

The general solution of the differential equation is

$$w_n(x) = c_1 \cos(\sqrt{\lambda}\, x) + c_2 \sin(\sqrt{\lambda}\, x).$$

Its derivative is

$$w'(x) = -\sqrt{\lambda}\, c_1 \sin(\sqrt{\lambda}\, x) + \sqrt{\lambda}\, c_2 \cos(\sqrt{\lambda}\, x).$$

The first boundary condition on w implies

$$0 = w'(0) = \sqrt{\lambda}\, c_2, \quad \Rightarrow \quad c_2 = 0 \quad \Rightarrow \quad w(x) = c_1\, \cos(\sqrt{\lambda}\, x).$$

The second boundary condition on w implies

$$0 = w'(3) = -\sqrt{\lambda}\, c_1\, \sin(\sqrt{\lambda}\, 3), \quad c_1 \neq 0, \quad \Rightarrow \quad \sin(\sqrt{\lambda}\, 3) = 0.$$

Then, $3\sqrt{\lambda_n} = n\pi$. If we choose $c_1 = 1$, we conclude,

$$\lambda_n = \Big(\frac{n\pi}{3}\Big)^2, \qquad w_n(x) = \cos\Big(\frac{n\pi x}{3}\Big), \quad n = 1, 2, 3, \cdots .$$

Recall we already have $\lambda_0 = 0$ with $w_0(x) = \frac{1}{2}$. Using the values of λ_n found above in the formula for v_λ we get

$$v_0(t) = c_0, \quad v_n(t) = c_n\, e^{-(\frac{n\pi}{3})^2 t}, \qquad c_n = v_n(0), \quad n = 1, 2, 3, \cdots .$$

Therefore, the superposition property of the solutions of the heat equation implies that any linear combination of simple solutions is a solution, so we get

$$T(t, x) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n\, e^{-(\frac{n\pi}{3})^2 t} \cos\Big(\frac{n\pi x}{3}\Big).$$

The initial condition is

$$\tau(x) = T(0, x) = \begin{cases} 7 & x \in \big[\frac{3}{2}, 3\big], \\ 0 & x \in \big[0, \frac{3}{2}\big). \end{cases}$$

Therefore, we get the orthogonal expansion for $\tau(x)$,

$$\tau(x) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n \cos\Big(\frac{n\pi x}{3}\Big).$$

Since this is an orthogonal expansion, we know how to compute the coefficients. The coefficient c_0 is given by

$$c_0 = \frac{2}{3} \int_0^3 \tau(x)\, dx = \frac{2}{3} \int_{3/2}^3 7\, dx = \frac{2}{3}\, 7\, \frac{3}{2} \quad \Rightarrow \quad c_0 = 7.$$

Now the coefficients c_n for $n \geqslant 1$ are given by

$$c_n = \frac{2}{3} \int_0^3 \tau(x) \cos\Big(\frac{n\pi x}{3}\Big)\, dx = \frac{2}{3} \int_{3/2}^3 7 \cos\Big(\frac{n\pi x}{3}\Big)\, dx,$$

that is,

$$c_n = \frac{14}{3}\, \frac{3}{n\pi}\, \sin\Big(\frac{n\pi x}{3}\Big)\Big|_{3/2}^3 = \frac{14}{n\pi} \Big(0 - \sin\Big(\frac{n\pi}{2}\Big)\Big) = -\frac{14}{n\pi} \sin\Big(\frac{n\pi}{2}\Big).$$

But for $n = 2k$ we have that

$$\sin\Big(\frac{2k\pi}{2}\Big) = \sin(k\pi) = 0,$$

while for $n = 2k - 1$ we have that

$$\sin\Big(\frac{(2k-1)\pi}{2}\Big) = (-1)^{k-1}.$$

Therefore, we obtain

$$c_{2k} = 0, \qquad c_{2k-1} = \frac{14(-1)^k}{(2k-1)\pi}, \quad k = 1, 2, \cdots .$$

So the solution of the initial-boundary value problem for the heat equation is

$$T(t, x) = \frac{7}{2} + \sum_{k=1}^{\infty} \frac{14(-1)^k}{(2k-1)\pi}\, e^{-(\frac{(2k-1)\pi}{3})^2 t} \cos\Big(\frac{(2k-1)\pi x}{3}\Big).$$

◁

Example 6.4.5 (Mixed)**.** Find the solution to the initial-boundary value problem

$$\partial_t T = \partial_x^2 T, \quad t > 0, \quad x \in [0, 3],$$

with initial and boundary conditions given by

$$\text{IC:} \quad T(0, x) = \begin{cases} 7 & x \in \big[\frac{3}{2}, 3\big], \\ 0 & x \in \big[0, \frac{3}{2}\big), \end{cases} \qquad \text{BC:} \quad \begin{cases} T(t, 0) = 0, \\ \partial_x T(t, 3) = 0. \end{cases}$$

Solution: This is a heat equation with mixed boundary conditions and we solve it with the separation of variables method. As usual, we look for simple solutions of the form

$$u(t, x) = v(t)\, w(x).$$

We put this function into the heat equation and we get

$$w(x)\, \frac{dv}{dt}(t) = v(t)\, \frac{d^2 w}{dx^2}(x) \quad \Rightarrow \quad \frac{\dot{v}(t)}{v(t)} = \frac{w''(x)}{w(x)} = -\lambda.$$

So, the equations for v and w are

$$\dot{v}(t) = -\lambda\, v(t), \qquad w''(x) + \lambda\, w(x) = 0.$$

The solution for v depends on λ, and is given by

$$v_\lambda(t) = c_\lambda\, e^{-\lambda t}, \qquad c_\lambda = v_\lambda(0).$$

Next we turn to the the equation for w, and we solve the BVP

$$w''(x) + \lambda\, w(x) = 0, \quad \text{with BC} \quad w(0) = 0, \quad w'(3) = 0.$$

This is an eigenfunction problem for w and λ. This problem has nonzero solutions only for $\lambda > 0$. In this case the characteristic polynomial of $w'' + \lambda w = 0$ has complex roots. Let $\lambda = \mu^2$, then

$$p(r) = r^2 + \mu^2 = 0 \quad \Rightarrow \quad r_\pm = \pm \mu\, i.$$

The general solution of the differential equation is

$$w_n(x) = c_1 \cos(\mu x) + c_2 \sin(\mu x).$$

Its derivative is

$$w'(x) = -\mu\, c_1 \sin(\mu x) + \mu\, c_2 \cos(\mu x).$$

The first boundary conditions on w implies

$$0 = w(0) = c_1 + c_2\, 0, \quad \Rightarrow \quad c_1 = 0 \quad \Rightarrow \quad w(x) = c_2\, \sin(\mu x).$$

The second boundary condition on w implies

$$0 = w'(3) = -\mu\, c_2\, \cos(\mu 3), \quad c_1 \neq 0, \quad \mu > 0 \quad \Rightarrow \quad \cos(\mu 3) = 0.$$

Then, $3\mu_n = (2n-1)\pi/2$, that is, $\mu_n = \dfrac{(2n-1)\pi}{6}$. Choosing $c_2 = 1$, we conclude,

$$\lambda_n = \Big(\frac{(2n-1)\pi}{6}\Big)^2, \qquad w_n(x) = \sin\Big(\frac{(2n-1)\pi x}{6}\Big), \quad n = 1, 2, 3, \cdots.$$

Using the values of λ_n found above in the formula for v_λ we get

$$v_n(t) = c_n\, e^{-(\frac{(2n-1)\pi}{6})^2 t}, \qquad c_n = v_n(0).$$

Therefore, we get

$$T(t, x) = \sum_{n=1}^{\infty} c_n\, e^{-(\frac{(2n-1)\pi}{6})^2 t}\, \sin\Big(\frac{(2n-1)\pi x}{6}\Big),$$

where, as we showed in Theorem **??**, the coefficients c_n above are given by

$$c_n = \frac{2}{L} \int_0^L \tau(x) \sin\Big(\frac{(2n-1)\pi x}{2L}\Big)\, dx,$$

awith f the initial temperature. In our case, the initial condition for the temperature is

$$\tau(x) = u(0,x) = \begin{cases} 7 & x \in [\frac{3}{2}, 3], \\ 0 & x \in [0, \frac{3}{2}), \end{cases}$$

Therefore, the coefficients c_n are given by

$$\begin{aligned} c_n &= \frac{2}{3} \int_0^3 \tau(x) \sin\Big(\frac{(2n-1)\pi x}{6}\Big)\, dx \\ &= \frac{2}{3} \int_{3/2}^3 7 \sin\Big(\frac{(2n-1)\pi x}{6}\Big)\, dx \\ &= \frac{14}{3}\,(-1)\,\frac{6}{(2n-1)\pi}\,\cos\Big(\frac{(2n-1)\pi x}{6}\Big)\Big|_{3/2}^{3} \\ &= -\frac{14}{(2n-1)\pi}\,\Big(\cos\Big(\frac{(2n-1)\pi}{2}\Big) - \cos\Big(\frac{(2n-1)\pi}{4}\Big)\Big) \\ &= \frac{14}{(2n-1)\pi}\,\cos\Big(\frac{(2n-1)\pi}{4}\Big). \end{aligned}$$

Therefore, we got

$$T(t,x) = \sum_{n=1}^{\infty} \frac{14}{(2n-1)\pi}\, e^{-(\frac{(2n-1)\pi}{6})^2 t}\, \sin\Big(\frac{(2n-1)\pi x}{6}\Big).$$

◁

6.4.5. Non-Homogeneous Conditions. In this subsection we find solutions to the heat equation with Dirichlet boundary conditions, but this time the boundary conditions are non-homogenous. This means that instead of homogeneous conditions

$$T(t,0) = 0, \qquad T(t,L) = 0,$$

we have boundary conditions of the form

$$T(t,0) = T_0, \qquad T(t,L) = T_L,$$

with T_0 and/or T_L nonzero.

Definition 6.4.4. *The Dirichlet initial-boundary value problem for the one-space dimensional heat equation is to find solutions $T(t,x)$ of the equation*

$$\partial_t T = \alpha\, \partial_x^2 T, \qquad t \in (0,\infty), \qquad x \in (0,L),$$

with α, L positive constants, and with boundary conditions

$$T(t,0) = T_0, \qquad T(t,L) = T_L,$$

where T_0 and T_L are arbitrary constants. The problem is called ***non-homogeneous*** *when at least one of the constant T_0 and T_L is non-zero; the problem is called* ***homogeneous*** *when both constants are zero.*

Unfortunately, the Sturm-Liouville theory and the separation of variables method we used to solve the homogeneous Dirichlet boundary value problem in 6.4.4 do not work for non-homogeneous boundary conditions. For example, if we look for simple solutions of the form

$$u(t,x) = v(t)\,w(x),$$

then the function w must be solution of

$$-w''(x) = \lambda\, w(x), \qquad w(0) = T_0, \quad w(L) = T_L.$$

When T_0 and/or T_L are nonzero, this problem does not determine the constant λ. The reason is that this differential operator

$$L(w) = -w'', \qquad w(0) = T_0, \quad w(L) = T_L,$$

is a not a linear operator when T_0 and/or T_L are nonzero. One way to solve this problem is to shift the non-homogeneous boundary conditions for T into an homogeneous boundary condition for another function. The shift is done subtracting the equilibrium solution of the heat equation with the appropriate non-homogeneous boundary condition.

Definition 6.4.5. *A function T_e is an **equilibrium solution** of the Dirichlet problem in Def. 6.4.4 iff the function T_e is time-independent and solution of the boundary value problem*

$$\partial_x^2 T_e = 0, \qquad T_e(0) = T_0, \qquad T_e(L) = T_L. \tag{6.4.13}$$

The equilibrium solutions are time-independent solutions of the Dirichlet initial-boundary value problem. These equilibrium solutions are simple to find.

Theorem 6.4.6 (Equilibrium)**.** *There is a unique solution of Eqs. (6.4.13) given by*

$$T_e(x) = T_0 + (T_L - T_0)\,\frac{x}{L}.$$

Proof of Theorem 6.4.6. We look for time independent solutions, that is, $T_e = T_e(\not{t}, x)$, so the heat equation is

$$0 = \partial_t T_e(x) = \alpha\, \partial_x^2 T_e(x) = \alpha\, T_e''(x) \quad \Rightarrow \quad T_e''(x) = 0.$$

All the solutions of the equation above are

$$T_e(x) = c_1 + c_2\, x,$$

where c_1 and c_2 are arbitrary constants. The boundary conditions fix these constants. The first condition implies

$$T_0 = T_e(0) = c_1 \quad \Rightarrow \quad T_e(x) = T_0 + c_2\, x.$$

The second condition implies

$$T_L = T_e(L) = T_0 + c_2\, L \quad \Rightarrow \quad c_2 = \frac{(T_L - T_0)}{L}.$$

so we get that

$$T_e(x) = T_0 + (T_L - T_0)\,\frac{x}{L}.$$

This establishes the Theorem. □

Since the functions T and T_e satisfy the same non-homogeneous boundary conditions, we use the equilibrium solution found above to shift the non-homogeneous boundary value problem for T into an homogeneous boundary value problem for $T_h = T - T_e$.

Theorem 6.4.7 (Dirichlet Nonhomogeneous)**.** *Consider the initial boundary value problem for the one-space dimensional heat equation,*

$$\partial_t T = \alpha\, \partial_x^2 T, \qquad t \in (0, \infty), \quad x \in (0, L),$$

with α and L positive constants and non-homogeneous Dirichlet boundary conditions

$$T(t, 0) = T_0, \qquad T(t, L) = T_L,$$

with T_0, T_L arbitrary constants, and intial condition

$$T(0, x) = \tau(x), \qquad x \in [0, L].$$

The initial boundary value problem above has a unique solution, $T(t, x)$, given by

$$T(t, x) = T_e(x) + T_h(t, x)$$

where T_e is the equilibrium solution

$$T_e(x) = T_0 + (T_L - T_0)\, \frac{x}{L}.$$

and T_h is the unique solution of the heat equation

$$\partial_t T_h = \alpha\, \partial_x^2 T_h, \qquad t \in (0, \infty), \quad x \in (0, L),$$

with homogeneous Dirichlet boundary conditions

$$T_h(t, 0) = 0, \quad T_h(t, L) = 0,$$

and initial condition

$$T_h(0, x) = \tau(x) - T_e(x),$$

which is given by the formula

$$T_h(t, x) = \sum_{n=1}^{\infty} c_n\, e^{-\alpha(\frac{n\pi}{L})^2 t} \sin\Big(\frac{n\pi x}{L}\Big),$$

and the coefficients c_n above are determined by the initial temperature, $T_h(0, x)$,

$$c_n = \frac{2}{L} \int_0^L \big(\tau(x) - T_e(x)\big) \sin\Big(\frac{n\pi x}{L}\Big)\, dx.$$

Remark: The main ideas to prove the theorem above are as follows:

(a) We write $T(t, x) = T_e(x) + T_h(t, x)$, where T_e is the equilibrium solution of the heat equation with the same boundary conditions as T.

(b) We show that the reminder function, T_h, satisfies the heat equation with homogeneous boundary conditions.

(c) We use the separation of variables method to find the function T_h.

(d) We determine the free constants in T_h with the initial condition for T_h, which is determined by the initial condition on T.

Proof of the Theorem 6.4.7: We write function T as follows,

$$T(t, x) = T_e(x) + T_h(t, x),$$

where T_e is the equilibrium (time independent) solution of the heat equation with non-homogeneous boundary condition,

$$T_e''(x) = 0, \qquad T_e(0) = T_0, \quad T_e(L) = T_L.$$

Theorem 6.4.6 says that the solution is

$$T_e(x) = T_0 + (T_L - T_0)\, \frac{x}{L}.$$

Now we look for the equation and boundary conditions satisfied by T_h. Since

$$\partial_t T(t,x) = \partial_t T_e(x) + \partial_t T_h(t,x) = \partial_t T_h(t,x),$$
$$\partial_x^2 T(t,x) = \partial_x^2 T_e(x) + \partial_x^2 T_h(t,x) = \partial_x^2 T_h(t,x),$$

then T is solution of the heat equation if and only if T_h is,

$$\partial_t T(t,x) = \alpha\, \partial_x^2 T(t,x) \quad \Leftrightarrow \quad \partial_t T_h(t,x) = \alpha\, \partial_x^2 T_h(t,x).$$

The function T_h satisfies homogeneous Dirichlet boundary conditions, since

$$T_0 = T(t,0) = T_e(0) + T_h(t,0) = T_0 + T_h(t,0) \quad \Rightarrow \quad T_h(t,0) = 0,$$
$$T_L = T(t,L) = T_e(L) + T_h(t,L) = T_L + T_h(t,L) \quad \Rightarrow \quad T_h(t,L) = 0.$$

Then, Theorem 6.4.2 says that all the solutions T_h for that problem are given by

$$T_h(t,x) = \sum_{n=1}^{\infty} c_n\, e^{-\alpha(\frac{n\pi}{L})^2 t} \sin\Big(\frac{n\pi x}{L}\Big).$$

The coefficients c_n are fixed by the initial condition on T_h, which is given by the initial condition on $T(0,x) = \tau(x)$, since $T_h = T - T_e$, then

$$T_h(0,x) = \tau(x) - T_e(x).$$

Therefore, Theorem 6.4.2 says that the coefficients c_n are given by the formula

$$c_n = \frac{2}{L} \int_0^L \big(\tau(x) - T_e(x)\big) \sin\Big(\frac{n\pi x}{L}\Big)\, dx.$$

This establishes the Theorem. □

Example 6.4.6 (Non-Homogeneous Dirichlet)**.** Find the solution of the heat equation

$$\partial_t T = \alpha\, \partial_x^2 T, \qquad t \in (0,\infty), \quad x \in (0,L),$$

with boundary and initial conditions given by

$$\text{NHDBC:} \quad \begin{cases} T(t,0) = T_0, \\ T(t,5) = T_L. \end{cases} \qquad \text{IC:} \quad \tau(x) = T(0,x) = \begin{cases} T_0, & x \in \big[0, \frac{L}{2}\big], \\ T_L, & x \in \big(\frac{L}{2}, L\big], \end{cases}$$

Solution: We first find the equilibrium solution $T_e(x)$, which is solution of the problem

$$T_e''(x) = 0, \qquad T_e(0) = T_0, \quad T_e(L) = T_L.$$

The general solution of the differential equation is

$$T_e(x) = c_1 + c_2\, x,$$

for arbitrary constants c_1, c_2. The boundary conditions fix these constants. The first condition is

$$T_0 = T_e(0) = c_1 \quad \Rightarrow \quad T_e(x) = T_0 + c_2\, x.$$

The second condition is

$$T_L = T_e(L) = T_L + c_2\, L \quad \Rightarrow \quad c_2 = \frac{(T_L - T_0)}{L}.$$

Therefore, the equilibrium solution for this problem is

$$T_e(x) = T_0 + (T_L - T_0)\, \frac{x}{L}.$$

Then, the solution of the heat equation can be written as

$$T(t,x) = T_0 + (T_L - T_0)\,\frac{x}{L} + T_h(t,x).$$

Since functions T and T_h differ on a linear function of x,

$$\partial_t T = \partial_t T_h \quad \text{and} \quad \partial_x^2 T = \partial_x^2 T_h,$$

then T is solution of the heat equation if and only if T_h is, that is,

$$\partial_t T = \alpha\,\partial_x^2 T \quad \Leftrightarrow \quad \partial_t T_h = \alpha\,\partial_x^2 T_h.$$

The boundary conditions on T_h are homogeneous, since T and T_e satisfy the same conditions,

$$T_0 = T(t,0) = T_0 + T_h(t,0) \quad \Rightarrow \quad T_h(t,0) = 0,$$
$$T_L = T(t,L) = T_0 + (T_L - T_0) + T_h(t,L) \quad \Rightarrow \quad T_h(t,L) = 0.$$

Therefore, Theorem 6.4.2 says that all the solutions T_h for that problem are given by

$$T_h(t,x) = \sum_{n=1}^{\infty} c_n\, e^{-\alpha(\frac{n\pi}{L})^2 t} \sin\Big(\frac{n\pi x}{L}\Big).$$

The coefficients c_n are fixed by the initial condition on T_h, which is given by the initial condition on $T(0,x) = \tau(x)$, since $T_h = T - T_e$, then

$$T_h(0,x) = \tau(x) - T_e(x).$$

Therefore, Theorem 6.4.2 says that the coefficients c_n are given by the formula

$$c_n = \frac{2}{L}\int_0^L \Big(\tau(x) - T_0 - (T_L - T_0)\,\frac{x}{L}\Big) \sin\Big(\frac{n\pi x}{L}\Big)\,dx.$$

Now, we split the integral on $[0,L]$ into an integral in $[0,L/2]$ and in $[L/2,L]$, and we use the definition of $\tau(x)$ in these intervals,

$$\begin{aligned}
c_n &= \frac{2}{L}\Big[\int_0^{\frac{L}{2}} \Big(T_0 - T_0 - (T_L - T_0)\,\frac{x}{L}\Big) \sin\Big(\frac{n\pi x}{L}\Big)\,dx \\
&\quad + \int_{\frac{L}{2}}^{L} \Big(T_L - T_0 - (T_L - T_0)\,\frac{x}{L}\Big) \sin\Big(\frac{n\pi x}{L}\Big)\,dx\Big] \\
&= -\frac{2}{L}\,\frac{(T_L - T_0)}{L}\int_0^{\frac{L}{2}} x\sin\Big(\frac{n\pi x}{5}\Big)\,dx + \frac{2}{L}(T_L - T_0)\int_{\frac{L}{2}}^{L} \sin\Big(\frac{n\pi x}{L}\Big)\,dx \\
&\quad - \frac{2}{L}\,\frac{(T_L - T_0)}{L}\int_{\frac{L}{2}}^{L} x\sin\Big(\frac{n\pi x}{L}\Big)\,dx \\
&= -\frac{2(T_L - T_0)}{L^2}\int_0^{L} x\sin\Big(\frac{n\pi x}{L}\Big)\,dx + \frac{2(T_L - T_0)}{L}\int_{\frac{L}{2}}^{L} \sin\Big(\frac{n\pi x}{L}\Big)\,dx.
\end{aligned}$$

If we recall the integral

$$\int x\sin(ax)\,dx = -\frac{x}{a}\cos(ax) + \frac{1}{a^2}\sin(ax),$$

and set $a = \frac{n\pi}{L}$, we get

$$\begin{aligned} c_n &= -\frac{2(T_L - T_0)}{L^2}\Big[\big(-\frac{Lx}{n\pi}\big)\cos\Big(\frac{n\pi x}{L}\Big) + \frac{L^2}{n^2\pi^2}\sin\Big(\frac{n\pi x}{L}\Big)\Big]\Big|_0^L \\ &\quad + \frac{2(T_L - T_0)}{L}\Big(-\frac{L}{n\pi}\Big)\cos\Big(\frac{n\pi x}{L}\Big)\Big|_{\frac{L}{2}}^L \\ &= -\frac{2(T_L - T_0)}{L^2}\Big[\Big(-\frac{L^2}{n\pi}\cos(n\pi)\Big) + 0 - (0+0)\Big] - \frac{2(T_L - T_0)}{n\pi}\Big[\cos(n\pi) - \cos\Big(\frac{n\pi}{2}\Big)\Big] \\ &= \frac{2(T_L - T_0)}{n\pi}\cos(n\pi) - \frac{2(T_L - T_0)}{n\pi}\cos(n\pi) + \frac{2(T_L - T_0)}{n\pi}\cos\Big(\frac{n\pi}{2}\Big), \end{aligned}$$

so we conclude that

$$c_n = \frac{2(T_L - T_0)}{n\pi}\cos\Big(\frac{n\pi}{2}\Big).$$

Therefore, the solution of the heat equation is

$$u(t,x) = T_0 + (T_L - T_0)\frac{x}{L} + \sum_{n=1}^{\infty}\frac{2(T_L - T_0)}{n\pi}\cos\Big(\frac{n\pi}{2}\Big)\,e^{-k(\frac{n\pi}{L})^2 t}\sin\Big(\frac{n\pi x}{L}\Big).$$

There is one more simplification that can be made, since half of the coefficients c_n are zero. If we recall that

$$\cos\Big(\frac{(2k-1)\pi}{2}\Big) = 0, \quad \text{and} \quad \cos\Big(\frac{2k\pi}{2}\Big) = \cos(k\pi) = (-1)^k,$$

then, for all $k = 1, 2, \cdots$ we get that

$$c_{2k-1} = 0 \quad \text{and} \quad c_{2k} = \frac{(T_L - T_0)}{k\pi}(-1)^k.$$

Using these formulas for the coefficients, the solution u can be written as

$$T(t,x) = T_0 + (T_L - T_0)\,\frac{x}{L} + \sum_{k=1}^{\infty}\frac{(T_L - T_0)}{k\pi}(-1)^k\,e^{-\alpha(\frac{2k\pi}{L})^2 t}\sin\Big(\frac{2k\pi x}{L}\Big).$$

Finally, renaming the summation index back to n we get

$$u(t,x) = T_0 + (T_L - T_0)\frac{x}{L} + \sum_{n=1}^{\infty}\frac{(T_L - T_0)}{n\pi}(-1)^n\,e^{-\alpha(\frac{2n\pi}{L})^2 t}\sin\Big(\frac{2n\pi x}{L}\Big).$$

◁

6.4.6. Exercises.

6.4.1.- .

6.4.2.- .

CHAPTER 7

Appendices

A. Overview of Complex Numbers

The first public appearance of complex numbers was in 1545 Gerolamo Cardano's Ars Magna, when he published a way to find solutions of a cubic equation $ax^3 + bx + c = 0$. The solution formula was not his own but given to him sometime earlier by Scipione del Ferro. In order to get such formula there was a step in the calculation involving a $\sqrt{-1}$, which was a mystery for the mathematicians of that time. There is no real number so that its square is -1, so what does this symbol, $\sqrt{-1}$, even mean? More intriguing, a few steps later during the calculation, this $\sqrt{-1}$ cancels out, and it does not appear in the final formula for the roots of the cubic equation. It was like a ghost entered your calculation and walked out of it without leaving a trace. Maybe we should call them ghost numbers, or magic numbers.

Everything in nature is magic until we understand how it works, then knowledge advances and magic retreats, one step at a time. It took a while, until the beginning of the 19th century with the—independent but almost simultaneous—works of Karl Gauss and William Hamilton, but our magic numbers were finally understood and they became the complex numbers.

In spite of their name, there is nothing complex about complex numbers. *Planar numbers* is a better fit to what they are—the set of all ordered pairs of real numbers together with specific addition and multiplication rules. Complex numbers can be identified with points on a plane, in the same way that real numbers can be identified with points on a line.

Definition A.1. ***Complex numbers*** *are numbers of the form*

$$(a, b),$$

where a and b are real numbers, together with the operations of addition,

$$(a, b) + (c, d) = (a + c, b + d), \tag{A.1}$$

and multiplication,

$$(a, b)(c, d) = (ac - bd, ad + bc). \tag{A.2}$$

The operation of addition is simple to understand because it is exactly how we add vectors on a plane,

$$\langle a, b\rangle + \langle c, d\rangle = \langle (a + c), (b + d)\rangle.$$

It is the multiplication what distinguishes complex numbers from vectors on the plane. To understand these operations it is useful to start with the following properties.

Theorem A.2. *The addition and multiplication of complex number are commutative, associative, and distributive. That is, given arbitrary complex numbers x, y, and z holds*

(a) Commutativity: $x + y = y + x$ and $x\,y = y\,x$.
(b) Associativity: $x + (y + z) = (x + y) + z$ and $x\,(y\,z) = (x\,y)\,x$.
(c) Distributivity: $x\,(y + z) = x\,y + x\,z$.

Proof of Theorem A.2: We show how to prove one of these properties, the proof for the rest is similar. Let's see the commutativity of multiplication. Given the complex numbers $x = (a, b)$ and $y = (c, d)$ we have

$$x\,y = (a, b)(c, d) = \big((ac - bd), (ad + bc)\big)$$

and

$$y\,x = (c, d)(a, b) = \big((ca - db), (cb + da)\big)$$

therefore we get that $x\,y = y\,x$. The rest of the properties can be proven in a similar way. This establishes the Theorem. □

We now mention a few more properties of complex numbers which are straightforward from the definitions above. For all complex numbers (a, b) we have that

$$(0, 0) + (a, b) = (a, b)$$
$$(-a, -b) + (a, b) = (0, 0)$$
$$(a, b)(1, 0) = (a, b).$$

From the first equation above the complex number $(0, 0)$ is called the *zero* complex number. From the second equation above the complex number $(-a, -b)$ is called the *negative* of (a, b), and we write

$$-(a, b) = (-a, -b).$$

From the last equation above the complex number $(1, 0)$ is called the *identity* for the multiplication.

The *inverse* of a complex number (a, b), denoted as $(a, b)^{-1}$, is the complex number satisfying

$$(a, b)\,(a, b)^{-1} = (1, 0).$$

Since the inverse of a complex number is itself a complex number, it can be written as

$$(a, b)^{-1} = (c, d)$$

for appropriate components c and d. The next result gives us a formula for these components. The next result says that every nonzero complex number has an inverse.

Theorem A.3. *The inverse of (a, b), with either $a \neq 0$ or $b \neq 0$, is*

$$(a, b)^{-1} = \Big(\frac{a}{(a^2 + b^2)}, \frac{-b}{(a^2 + b^2)}\Big). \tag{A.3}$$

Proof of Theorem A.3: A complex number $(a, b)^{-1}$ is the inverse of (a, b) iff

$$(a, b)\,(a, b)^{-1} = (1, 0).$$

When we write $(a, b)^{-1} = (c, d)$, the equation above is

$$(a, b)(c, d) = (1, 0).$$

If we compute explicitly the left-hand side above we get

$$\big((ac - bd), (ad + bc)\big) = (1, 0).$$

The equation above implies two equations for real numbers,

$$ac - bd = 1, \qquad ad + bc = 0.$$

In the case that either $a \neq 0$ or $b \neq 0$, the solution to the equations above is

$$c = \frac{a}{(a^2 + b^2)}, \qquad d = \frac{-b}{(a^2 + b^2)}.$$

Therefore, the inverse of (a, b) is

$$(a, b)^{-1} = \Big(\frac{a}{(a^2 + b^2)}, \frac{-b}{(a^2 + b^2)}\Big).$$

This establishes the Theorem. □

Example A.1. Find the inverse of $(2, 3)$. Then, verify your result.

Solution: The formula above says that $(2, 3)^{-1}$ is given by

$$(2, 3)^{-1} = \Big(\frac{2}{(2^2 + 3^2)}, \frac{-3}{(2^2 + 3^2)}\Big) \quad \Rightarrow \quad (2, 3)^{-1} = \Big(\frac{2}{13}, \frac{-3}{13}\Big).$$

This is correct, since

$$\begin{aligned}(2, 3)\Big(\frac{2}{13}, \frac{-3}{13}\Big) &= \Big(\big(\frac{4}{13} - \frac{(-9)}{13}\big), \big(\frac{-6}{13} + \frac{6}{13}\big)\Big)\\ &= \Big(\frac{13}{13}, \frac{0}{13}\Big)\\ &= (1, 0).\end{aligned}$$

◁

A.1. Extending the Real Numbers. The set of all complex numbers of the form $(a, 0)$ satisfy the same properties as the set of all real numbers a. Indeed, for all a, c reals holds

$$(a, 0) + (c, 0) = (a + c, 0), \qquad (a, 0)(c, 0) = (ac, 0).$$

We also have that

$$-(a, 0) = (-a, 0),$$

and the formula above for the inverse of a complex number says that

$$(a, 0)^{-1} = \Big(\frac{1}{a}, 0\Big).$$

From here it is natural to identify a complex number $(a, 0)$ with the real number a, that is,

$$(a, 0) \quad \longleftrightarrow \quad a.$$

This identification suggests the following definition.

Definition A.4. *The* ***real part*** *of $z = (a, b)$ is a and—then it is natural to call—the* ***imaginary part*** *of z is b. We also use the notation*

$$a = \mathsf{Re}(z), \qquad b = \mathsf{Im}(z).$$

A.2. The Imaginary Unit. We understood complex numbers of the form $(a, 0)$. They are no more than the real numbers. Now we study complex numbers of the form $(0, b)$—complex numbers with no real part. In particular, we focus on the complex number $(0, 1)$, which we call the *imaginary unit*. Let us compute its square,

$$(0, 1)^2 = (0, 1)(0, 1) = (-1, 0) = -(1, 0) \quad \Rightarrow \quad (0, 1)^2 = -(1, 0).$$

Within the complex numbers we do have a number whose square is negative one, and that number is the imaginary unit $(0, 1)$. Actually, there are two complex numbers whose square is negative one, one is $(0, 1)$ and the other is $-(0, 1)$, because

$$(0, -1)^2 = (0, -1)(0, -1) = (0 - (-1)(-1), 0 + 0) = (-1, 0) = -(1, 0).$$

So, in the set of complex numbers we do have solutions for the $\sqrt{-(1,0)}$, given by

$$\sqrt{-(1,0)} = \pm(0,1).$$

Notice that $\sqrt{-1}$ has no solutions, but $\sqrt{-(1,0)}$ has two solutions. This is the origin of the confusion with del Ferro's calculation. Most of his calculation used numbers of the form $(a,0)$—written as a— except at one tiny spot where a number $(0,1)$ shows up and later on cancels out. Del Ferro's calculation makes perfect sense in the complex realm, and almost all of it can be reproduced with real numbers, but not all.

A.3. Standard Notation. We can now relate the ordered pair notation we have been using for complex numbers with the notation used by the early mathematicians. We start noticing that

$$(a,b) = (a,0) + (0,b) = (a,0) + (b,0)(0,1).$$

Therefore, if we write a for $(a,0)$, b for $(b,0)$, and we use $i = (0,1)$, we get that every complex number (a,b) can be written as

$$(a,b) = a + bi.$$

Recall, a and b are the real and imaginary parts of (a,b). And the equation

$$(0,1)^2 = -(1,0)$$

in the new notation is

$$i^2 = -1.$$

This notation $(a+bi)$ is useful to manipulate formulas involving addition and multiplication. If we multiply $(a+bi)$ by $(c+di)$ and use the distributive and associative properties we get

$$(a+bi)(c+di) = ac + adi + cbi + bdi^2,$$

and if we recall that $i^2 = -1$ and we reorder terms, we get

$$(a+bi)(c+di) = ac - bd + (ad+bc)i.$$

So, we do not need to remember the formula for the product of two complex numbers. With the new notation, this formula comes from the distributive and associative properties. Similarly, to compute the inverse of a complex number $a+bi$ we may write

$$\begin{aligned}\frac{1}{a+bi} &= \frac{1}{(a+bi)}\,\frac{(a-bi)}{(a-bi)}\\ &= \frac{(a-bi)}{(a+bi)(a-bi)}.\end{aligned}$$

Notice that

$$(a+bi)(a-bi) = a^2 + b^2,$$

which has only a real part. Then we can write

$$\frac{1}{a+bi} = \frac{a-bi}{(a^2+b^2)} \quad\Rightarrow\quad \frac{1}{a+bi} = \frac{a}{(a^2+b^2)} - \frac{b}{(a^2+b^2)}\,i$$

which agrees with the formula we got in Theorem A.3.

A.4. Useful Formulas. The powers of i can have only four possible results.

Theorem A.5. *The integer powers of i can have only four results: 1, i, -1, and $-i$.*

Proof of Theorem A.5: We just show that this is the case for the first powers. Dy definition of a power zero and power one we know that

$$i^0 = 1,$$
$$i^1 = i.$$

We also know that

$$i^2 = -i.$$

We can compute the next powers, using that $(a+bi)^{m+n} = (a+bi)^m (a+bi)^n$, so we get

$$\begin{aligned}
i^3 &= i^2\, i = (-1)\, i = -i \\
i^4 &= i^3\, i = -i\, i = -i^2 = 1 \\
i^5 &= i^4\, i = (1)\, i = i \\
i^6 &= i^5\, i = i\, i = -1 \\
i^7 &= i^6\, i = (-1)\, i = -i \\
&\vdots
\end{aligned}$$

An argument using induction would proof this Theorem. □

The *conjugate* of a complex number $a+bi$ is the complex number

$$\overline{a+bi} = a - bi.$$

For example,

$$\overline{1+2i} = 1-2i, \qquad \overline{a} = a, \qquad \overline{i} = -i, \qquad \overline{4i} = -4i.$$

If we conjugate twice we get the original complex number, that is $\overline{\overline{a+bi}} = a+bi$.

The *modulus* or *absolute value* of a complex number $a+bi$ is the real number

$$|a+bi| = \sqrt{a^2+b^2}.$$

For example

$$|3+4i| = \sqrt{9+16} = \sqrt{25} = 5, \qquad |a+0i| = |a|, \qquad |i| = 1, \qquad |1+i| = \sqrt{2}.$$

Using these definitions is simple to see that

$$(a+bi)\,\overline{(a+bi)} = (a+bi)(a-bi) = (a^2+b^2) = |a+bi|^2.$$

Using these definitions we can rewrite the formula in Eq. (A.3) for the inverse of a complex number as follows,

$$\frac{1}{(a+bi)} = \frac{1}{(a^2+b^2)}(a-bi).$$

If we call $z = a+bi$, then the formula for z^{-1} reduces to

$$z^{-1} = \frac{\overline{z}}{|z|^2}.$$

Example A.2. Write $\dfrac{1}{(3+4i)}$ in the form $c+di$.

Solution: You multiply numerator and denominator by $3-4i$,

$$\begin{aligned}\frac{1}{(3+4i)} &= \frac{1}{(3+4i)}\frac{(3-4i)}{(3-4i)}\\ &= \frac{(3-4i)}{(3^2+4^2)}\\ &= \frac{3-4i}{25}\\ &= \frac{3}{25}-\frac{4}{25}\,i.\end{aligned}$$

So, we have found that the inverse of $(3+4i)$ is $\Big(\dfrac{3}{25}-\dfrac{4}{25}\,i\Big)$. ◁

The absolute value of complex numbers satisfy the triangle inequality.

Theorem A.6. *For all complex numbers z_1, z_2 holds $|z_1+z_2| \leqslant |z_1|+|z_2|$.*

Remark: The idea of the Proof of Theorem A.6 is to use the graphical representations of complex numbers as vectors on a plane. Then $|z_1|$ is the length of the vector given by z_1, and the same holds for the vectors associated to z_2 and z_1+z_2, the latter being the diagonal in the parallelogram formed by z_1 and z_2. Then it is clear that the triangle inequality holds.

The absolute value of a complex number also satisfies the following properties.

Theorem A.7. *For all complex numbers z_1, z_2 holds $|z_1z_2| = |z_1|\,|z_2|$.*

Proof of Theorem A.7: For an arbitrary complex numbers $z_1 = a+bi$ and $z_2 = c+di$, we have

$$z_1z_2 = (ac-bd)+(ad+bc)i,$$

therefore,

$$\begin{aligned}|z_1z_2|^2 &= (ac-bd)^2+(ad+bc)^2\\ &= (ac)^2+(bd)^2-2acbd+(ad)^2+(bc)^2+2adbc\\ &= a^2c^2+b^2d^2+a^2d^2+b^2c^2\\ &= a^2(c^2+d^2)+b^2(d^2+c^2)\\ &= (a^2+b^2)(d^2+c^2)\\ &= |z_1|^2\,|z_2|^2,\end{aligned}$$

Taking a square root we get

$$|z_1z_2| = |z_1|\,|z_2|.$$

This establishes the Theorem. □

Theorem A.8. *Every complex number z satisfies that $|z^n| = |z|^n$, for all integer n.*

Proof of Theorem A.8: One proof uses the previous theorem A.7 and induction in n. For $n = 2$ it is proven by the theorem above,

$$|z^2| = |zz| = |z|\,|z| = |z|^2.$$

Now, suppose the theorem is true for $n-1$, so $|z^{n-1}| = |z|^{n-1}$. Then

$$|z^n| = |z^{n-1}z| = |z^{n-1}|\,|z|$$

where we used the previous theorem A.7. But in the first factor we use the inductive hypothesis,

$$|z^{n-1}|\,|z| = |z|^{n-1}\,|z| = |z|^n.$$

So we have proven that $|z^n| = |z|^n$. This establishes the Theorem. □

Remark: A second proof, independent of the previous theorem is that, for an arbitrary non-negative integer n we have,

$$|z^n| = \sqrt{z^n\overline{z^n}} = \sqrt{z^n(\overline{z})^n} = \sqrt{(z\overline{z})^n} = \left(\sqrt{z\overline{z}}\right)^n = |z|^n$$

Example A.3. Verify the result in Theorem A.8 for $n = 3$ and $z = 3 + 4i$.

Solution: First we compute $|z|$ and then its cube,

$$|z| = |3 + 4i| = \sqrt{9 + 16} = 5 \quad \Rightarrow \quad |z|^3 = 125.$$

We now compute z^3, and then its absolute value,

$$z^3 = (3 + 4i)(3 + 4i)(3 + 4i) = -117 + 44\,i \quad \Rightarrow \quad |z^3| = \sqrt{117^2 + 44^2} = 125.$$

Therefore, $|z|^3 = |z^3|$. As an extra bonus, we found another perfect triple, besides the famous $3^2 + 4^2 = 5^2$, which is

$$44^2 + 117^2 = 125^2.$$

◁

A.5. Complex Functions.

We know how to add and multiply complex numbers

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$
$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

This means we know how to extend any real-valued function defined on real numbers having a Taylor series expansion. We use the function Taylor series as the definition of the function for complex numbers. For example, the real-valued exponential function has the Taylor series expansion

$$e^{at} = \sum_{n=0}^{\infty} \frac{a^n t^n}{n!}.$$

Therefore, we define the complex-valued exponential as follows.

Definition A.9. *The* ***complex-valued exponential function*** *is given by*

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}. \tag{A.4}$$

Remark: We are particularly interested in the case that the argument of the exponential function is of the form $z = (a \pm bi)t$, where $r_\pm = a \pm bi$ are the roots of the characteristic polynomial of a second order linear differential equation with constant coefficients. In this case, the exponential function has the form

$$e^{(a+bi)t} = \sum_{n=0}^{\infty} \frac{(a+bi)^n t^n}{n!}.$$

The infinite sum on the right-hand side in equation (A.4) makes sense, since we know how to multiply—hence compute powers—of complex numbers, and we know how to add complex numbers. Furthermore, one can prove that the infinite series above converges, because the series converges in absolute value, which implies that the series itself converges. Also important, the name we chose for the function above, the exponential, is well chosen, because this function satisfies the exponential property.

Theorem A.10 (Exp. Property)**.** *For all complex numbers z_1, z_2 holds $e^{z_1+z_2} = e^{z_1}\, e^{z_2}$.*

Proof of Theorem A.10: A straightforward calculation using the binomial formula implies

$$\begin{aligned} e^{z_1+z_2} &= \sum_{n=0}^{\infty} \frac{(z_1+z_2)^n}{n!} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n}{k} \frac{z_1^k\, z_2^{n-k}}{n!} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{z_1^k\, z_2^{n-k}}{k!(n-k)!}, \end{aligned}$$

where we used the notation $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$. This double sum is over the triangular region in the nk space given by

$$0 \leqslant n \leqslant \infty \qquad 0 \leqslant k \leqslant n.$$

We now interchange the order of the sums, the indices be given by

$$0 \leqslant k \leqslant \infty \qquad k \leqslant n \leqslant \infty,$$

so we get

$$\sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{z_1^k\, z_2^{n-k}}{k!(n-k)!} = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \frac{z_1^k\, z_2^{n-k}}{k!(n-k)!}.$$

If we introduce the variable $m = n - k$ we get that

$$\begin{aligned} \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \frac{z_1^k\, z_2^{n-k}}{k!(n-k)!} &= \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{z_1^k\, z_2^m}{k!m!} \\ &= \Big(\sum_{k=0}^{\infty} \frac{z_1^k}{k!}\Big)\Big(\sum_{m=0}^{\infty} \frac{z_2^m}{m!}\Big) \\ &= e^{z_1}\, e^{z_2}. \end{aligned}$$

So we have shown that $e^{z_1+z_2} = e^{z_1}\, e^{z_2}$. This Establishes the Theorem. □

The exponential property in the case that the exponent is $z = (a + bi)t$ has the form

$$e^{(a+bi)t} = e^{at}\, e^{ibt}.$$

The first factor on the right-hand side above is a real exponential, which—for a given value of $a \neq 0$—it is either a decreasing ($a < 0$) or increasing ($a > 0$) function of t. The second factor above is an exponential of a pure imaginary exponent. These exponentials can be summed in a closed form.

Theorem A.11 (Euler Formula)**.** *For any real number θ holds that $e^{i\theta} = \cos(\theta) + i\sin(\theta)$.*

Proof of Theorem A.11: Recall that i^n can have only four results, 1, i, -1, $-i$. This result can be summarized as

$$i^{2n} = (-1)^n \quad \Rightarrow \quad i^{2n+1} = (-1)^n\, i.$$

If we split the sum in the definition of the exponential into even and odd terms in the sum index, we get

$$e^{i\theta} = \sum_{n=0}^{\infty} \frac{i^n \theta^n}{n!} = \sum_{n=0}^{\infty} \frac{i^{2n}\theta^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{i^{2n+1}\theta^{2n+1}}{(2n+1)!},$$

and using the property above on the powers of i we get

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

Recall that Taylor series expansions of the sine and cosine functions

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

Therefore, we have shown that

$$e^{i\theta} = \cos(\theta) + i\sin(\theta).$$

This establishes the Theorem. □

A.6. Complex Vectors. We can extend the notion of vectors with real components to vectors with complex components. For example, complex-valued vectors on a plane are vectors of the form

$$\boldsymbol{v} = \langle a + bi, c + di\rangle,$$

where a, b, c, d are real numbers. We can add two complex-valued vectors component-wise. So, given

$$\boldsymbol{v}_1 = \langle a_1 + b_1 i, c_1 + d_1 i\rangle, \qquad \boldsymbol{v}_2 = \langle a_2 + b_2 i, c_2 + d_2 i\rangle,$$

we have that

$$\boldsymbol{v}_1 + \boldsymbol{v}_2 = \big\langle (a_1 + a_2) + (b_1 + b_2)i, (c_1 + c_2) + (d_1 + d_2)i\big\rangle.$$

For example

$$\langle 2 + 3i, 4 + 5i\rangle + \langle 6 + 7i, 8 + 9i\rangle = \langle 8 + 10i, 12 + 14i\rangle.$$

We can also multiply a complex-valued vector by a scalar, which now is a complex number. So, given $\boldsymbol{v} = \langle a + bi, c + di\rangle$ and $z = z_1 + z_2 i$, then

$$z\boldsymbol{v} = (z_1 + z_2 i)\langle a + bi, c + di\rangle = \big\langle (z_1 + z_2 i)(a + bi), (z_1 + z_2 i)(c + di)\big\rangle.$$

For example

$$\begin{aligned} i\,\langle 2 + 3i, 4 + 5i\rangle &= \langle 2i - 3, 4i - 5\rangle \\ &= \langle -3 + 2i, -5 + 4i\rangle. \end{aligned}$$

The only non-intuitive calculation with complex-valued vectors is how to find the length of a complex vector. Recall that in the case of a real-valued vector $\boldsymbol{v} = \langle a, b \rangle$, the length of the vector is defined as

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}} = \sqrt{a^2 + b^2},$$

where $\cdot$ is the dot product of vectors, that is, given the real-valued vectors $\boldsymbol{v}_1 = \langle a_1, b_1 \rangle$, $\boldsymbol{v}_2 = \langle a_2, b_2 \rangle$, their dot product is the real number

$$\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = a_1 a_2 + b_1 b_2.$$

We want to generalize the notion of length from real-valued vectors to complex-valued vectors. Notice that the *length of a vector*—real or complex—must be a *real number*. Unfortunately, in the case of a complex-valued vector $\boldsymbol{v} = \langle a + bi, c + di \rangle$ the formula $\sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$ is not always a real number, it may have a nonzero imaginary part. In order to get a real number for the length of a complex-valued vector we define

$$\|\boldsymbol{v}\| = \sqrt{\overline{\boldsymbol{v}} \cdot \boldsymbol{v}},$$

where the conjugate of a vector means to conjugate all its components, that is

$$\overline{\boldsymbol{v}} = \overline{\langle a + bi, c + di \rangle} = \langle a - bi, c - di \rangle.$$

We needed to introduce the conjugate in the first vector in the formula above so that the result is a real number. Indeed, we have the following result.

Theorem A.12. *The length of a complex-valued vector* $\boldsymbol{v} = \langle a + bi, c + di \rangle$ *is*

$$\|\boldsymbol{v}\| = \sqrt{\overline{\boldsymbol{v}} \cdot \boldsymbol{v}} = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

Proof of Theorem A.12: This is a straightforward calculation,

$$\begin{aligned} \|\boldsymbol{v}\|^2 &= \overline{\boldsymbol{v}} \cdot \boldsymbol{v} \\ &= \langle a - bi, c - di \rangle \cdot \langle a + bi, c + di \rangle \\ &= (a - bi)(a + bi) + (c - di)\,(c + di) \\ &= a^2 + b^2 + c^2 + d^2. \end{aligned}$$

So we get the formula

$$\|\boldsymbol{v}\| = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

This establishes the Theorem. □

Example A.4. Find the length of $\boldsymbol{v} = \langle 1 + 2i, 3 + 4i \rangle$

Solution: The length of this vector is

$$\|\boldsymbol{v}\| = \sqrt{1^2 + 2^2 + 3^2 + 4^2} = \sqrt{30}.$$

◁

A *unit vector* is a vector with length one, that is, $\boldsymbol{u}$ is a unit vector iff $\|\boldsymbol{u}\| = 1$. Sometimes one needs to find a unit vector parallel to some vector $\boldsymbol{v}$. For both real-valued and complex-valued vectors we have the same formula. A unit vector $\boldsymbol{u}$ parallel to $\boldsymbol{v} \neq \boldsymbol{0}$ is

$$\boldsymbol{u} = \frac{1}{\|\boldsymbol{v}\|}\, \boldsymbol{v}.$$

Example A.5. Find a unit vector in the direction of $\boldsymbol{v} = \langle 3+2i, 1-2i\rangle$.

Solution: First we check that $\boldsymbol{v}$ is not a unit vector. Indeed,

$$\begin{aligned}\|\boldsymbol{v}\|^2 &= \overline{\boldsymbol{v}} \cdot \boldsymbol{v}\\ &= \langle 3-2i, 1+2i\rangle \cdot \langle 3+2i, 1-2i\rangle\\ &= (3-2i)(3+2i) + (1+2i)(1-2i)\\ &= 3^2 + 2^2 + 1^2 + 2^2\\ &= 14.\end{aligned}$$

Since $\|\boldsymbol{v}\| = \sqrt{14}$, the vector $\boldsymbol{v}$ is not unit. A unit vector is

$$\boldsymbol{u} = \frac{1}{\sqrt{14}} \langle 3-2i, 1+2i\rangle$$

more explicitly,

$$\boldsymbol{u} = \Big\langle \Big(\frac{3}{\sqrt{14}} - \frac{2}{\sqrt{14}}\,i\Big), \Big(\frac{1}{\sqrt{14}} + \frac{2}{\sqrt{14}}\,i\Big) \Big\rangle$$

◁

Notes.

This appendix is inspired on Tom Apostol's overview of complex numbers given in his outstanding Calculus textbook, [1], Volume I, § 9.

B. Answers to Exercises

Chapter 1: First Order Equations

Section 1.1: Separable Equations

1.1.1.- Implicit form: $\dfrac{y^2}{2} = \dfrac{t^3}{3} + c$.

Explicit form: $y = \pm\sqrt{\dfrac{2t^3}{3} + 2c}$.

1.1.2.- $y^4 + y + t^3 - t = c$, with $c \in \mathbb{R}$.

1.1.3.- $y(t) = \dfrac{3}{3 - t^3}$.

1.1.4.- $y(t) = c\, e^{-\sqrt{1+t^2}}$.

1.1.5.- $y(t) = t\big(\ln(|t|) + c\big)$.

1.1.6.- $y^2(t) = 2t^2\big(\ln(|t|) + c\big)$.

1.1.7.- Implicit: $y^2 + ty - 2t = 0$.
Explicit: $y(t) = \dfrac{1}{2}\big(-t + \sqrt{t^2 + 8t}\big)$.

1.1.8.- Hint: Recall the definition of an Euler homogeneous equation and the remarks below that definition. Also recall that

$$y_1'(x) = f\big(x, y_1(x)\big),$$

for any independent variable x, for example for $x = kt$.

Section 1.2: Linear Variable Coefficient Equations

1.2.1.- $y(t) = c\, e^{2t^2}$.

1.2.2.- $y(t) = c\, e^{-t} - e^{-2t}$, with $c \in \mathbb{R}$.

1.2.3.- $y(t) = 2e^t + 2(t-1)\, e^{2t}$.

1.2.4.- $y(t) = \dfrac{\pi}{2t^2} - \dfrac{\cos(t)}{t^2}$.

1.2.5.- $y(t) = c\, e^{t^2(t^2+2)}$, with $c \in \mathbb{R}$.

1.2.6.- $y(t) = \dfrac{t^2}{n+2} + \dfrac{c}{t^n}$, with $c \in \mathbb{R}$.

1.2.7.- $y(t) = 3\, e^{t^2}$.

1.2.??.- $y(t) = c\, e^t + \sin(t) + \cos(t)$, for all $c \in \mathbb{R}$.

1.2.??.- $y(t) = -t^2 + t^2 \sin(4t)$.

1.2.??.- Define $v(t) = 1/y(t)$. The equation for v is $v' = tv - t$. Its solution is $v(t) = c\, e^{t^2/2} + 1$. Therefore,

$$y(t) = \frac{1}{c\, e^{t^2/2} + 1}, \quad c \in \mathbb{R}.$$

1.2.??.- $y(x) = \big(6 + c\, e^{-x^2/4}\big)^2$

1.2.??.- $y(x) = \big(4\, e^{3t} - 3\big)^{1/3}$

Bibliography

[1] T. Apostol. *Calculus*. John Wiley & Sons, New York, 1967. Volume I, Second edition.

[2] T. Apostol. *Calculus*. John Wiley & Sons, New York, 1969. Volume II, Second edition.

[3] G. Birkhoff and G. Rota. *Ordinary Differential Equations*. John Wiley and Sons, New York, 1989. 4th edition.

[4] W. Boyce and R. DiPrima. *Elementary differential equations and boundary value problems*. Wiley, New Jersey, 2012. 10th edition.

[5] R. Churchill. *Operational Mathematics*. McGraw-Hill, New york, 1958. Second Edition.

[6] S. Hassani. *Mathematical physics*. Springer, New York, 2006. Corrected second printing, 2000.

[7] Y. Pinchover and J. Rubinstein. *An Introduction to Partial Differential Equations*. Cambridge University Press, Cambridge, 2005.

[8] G. Simmons. *Differential equations with applications and historical notes*. McGraw-Hill, New York, 1991. 2nd edition.

[9] S. Strogatz. *Nonlinear Dynamics and Chaos*. Perseus Books Publishing, Cambridge, USA, 1994. Paperback printing, 2000.

[10] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*. American Mathematical Society, Providence, Rhode Island.

[11] E. Zeidler. *Nonlinear Functional Analysis and its Applications I, Fixed-Point Theorems*. Springer, New York, 1986.

[12] E. Zeidler. *Applied functional analysis: applications to mathematical physics*. Springer, New York, 1995.

[13] D. Zill and W. Wright. *Differential equations and boundary value problems*. Brooks/Cole, Boston, 2013. 8th edition.