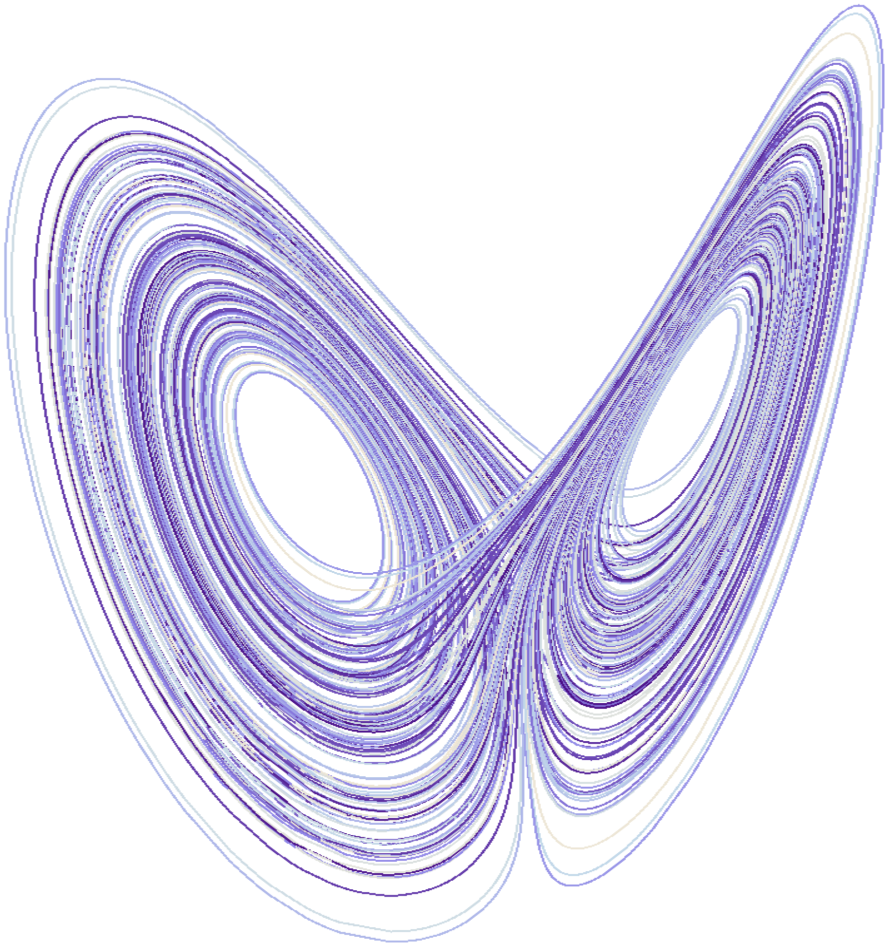


Ordinary Differential Equations

Evan Randles

Preface

These are the course notes for Ordinary Differential Equations (MA 311). Many of the ideas presented in these notes are not original and, at least in part, have been influenced by a number of excellent texts on the subject, including *Elementary Differential Equations and Boundary Value Problems* by William E. Boyce and Richard C. DiPrima [3], *Differential Equations with Applications and Historical Notes* by George F. Simmons [6], and *Differential Equations and Their Applications* by Martin Braun [4]. As these notes represent an active working draft, please update/download them frequently as I will often make corrections and changes without explicit warning. Any block of text or word in red is just a note for me (one I'm leaving to myself while editing); these should be disregarded. Words in blue are often hyperlinks to an interesting reference, usually a video, and you should click on them if you're reading along digitally. Also, if you find or suspect an error or typo – no matter how trivial – please email me to let me know!

Acknowledgment:

These course notes have grown and evolved significantly since their original use in Ordinary Differential Equations (MA311) in FA2018 at Colby College. I owe a great deal of gratitude to my students, who have served as test subjects for these notes as well as editors and proofreaders. In particular, I would like to thank Shabab Ahmed, William Bae, Yilei (Jerry) Bao, Huan Bui, Ben Capodanno, Julia Chahine, Tim (TC) Clifford, Henry Doud, Caroline Dunsby, Qidong He, Ja'Sean Holmes, Paula Jaramillo, Michelle Ling, Cara Moynihan, Lilly Naimie, Yiheng Su, Terry Surrette, Audrey Vaver, Matt Welch, Simon Xu, Trevaughn Wright-Reynolds, and Cathy Zhao. Of course, I am responsible for the errors that inevitably remain.

Contents

Chapter 1

Introduction: The game and the players

1.1 Four Differential Equations

A differential equation is, generally speaking, an equation which relates a function to its derivatives. This course is dedicated to the study of an important class of differential equations, called *ordinary differential equations*. Before laying out precise definitions (which is done in the next section) and beginning our study in earnest, the goal of this introductory section is to get a feel for the subject and build some intuition about where we are and where we're going. To this end, we survey four differential equations, the first of which comes from population ecology.

Example 1: Exponential Population Growth

Given a constant r, consider the differential equation

$$\frac{dP}{dt} = rP. \tag{1.1}$$

This equation gives a relation between some function P and its derivative with respect to an independent variable t. This differential equation comes from population ecology and is used to model population growth of a population (in an extremely simple setting). Here, $P = P(t)$ represents the population of a community at time t and r is a parameter which measures the difference between the birth and death rates of the population.

Observe that the function

$$P(t) = e^{rt} \tag{1.2}$$

solves the differential equation (1.1) above because

$$\frac{dP}{dt} = \frac{d}{dt}e^{rt} = re^{rt} = rP(t)$$

which holds for all real numbers t (also written "for all $t \in \mathbb{R}$"). As you can easily see, multiplying the function (1.2) by any constant, e.g., $5e^{rt}$ or $0 = 0e^{rt}$, produces another solution to Equation (1.1). This tells us, in particular, that differential equations can have multiple solutions. In fact, as we will see later in the semester, every solution to Equation (1.1) is of the form

$$P(t) = Ce^{rt}$$

for $t \in \mathbb{R}$ where C is a constant. Such a solution (or collections of solutions) is called a *general solution* because it allows one to capture all possible solutions by simply varying a constant or finite collection of constants.

In practice, one seeks to model the growth/decay of a population governed by (1.1) while knowing the value of the population at a fixed time t_0. For example, suppose that at time $t_0 = 0$ the population of a certain species has the value P_0, a fixed number. Then to model the population for all time $t > t_0 = 0$, one wants to find a function $P = P(t)$ which solves the differential equation (1.1) and also has the property that $P(t_0) = P(0) = P_0$. This is called an initial value problem and is often expressed in symbols as

$$\begin{cases} \dfrac{dP}{dt} = rP, & P(0) = P_0 \end{cases} .$$

The equation $P(0) = P_0$ is called an initial condition. You should verify that $P(t) = P_0 e^{rt}$ satisfies both the differential equation and the initial condition. Though it's not presently obvious, this is the only such solution.

Example 2: Newton's Second Law

Consider the differential equation

$$m\ddot{x} = F(t, \dot{x}, x) \tag{1.3}$$

where m is a fixed positive number and $F : \mathbb{R}^3 \to \mathbb{R}$ is a specified function. Equation (1.3) describes the motion of a particle (or massive body) of inertial mass m undergoing a force F. Here, $x = x(t)$ is the position of the particle at time t and $\dot{x} = \frac{dx}{dt}$ and $\ddot{x} = \frac{d^2x}{dt^2}$ are its velocity and acceleration, respectively[a]. Under fairly general circumstances, it is physically reasonable to assume that the external force F on a particle depends not only time but also on the particles position x and velocity[b] $\dot{x}$.

Of course, you might recognize Equation (1.3) better as "$F = ma$" or Newton's Second Law of Motion. This equation was first understood by Isaac Newton and originally presented in his 1687 treatise on mechanics and gravitation, *Philosopia Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy). Together with his eponymous law of gravity, Newton was able to use Equation (1.3) to show that the planets move in ellipses with the sun at one focus. This was a great triumph of physics (and mathematics) in that it furnished a theoretical description of the empirical law of Johannes Kepler (based on the observations of Tycho Brahe) known as Kepler's Second Law of Planetary Motion, published originally in 1609.

Stepping away from gravity, let's consider a particularly simple form of F in Eq. (1.3) which is used to describe spring-mass systems (and pendula). This is the situation in which a particle of mass m at position x is pulled toward the equilibrium position $x = 0$ by a force proportional to its displacement from 0. It is modeled by the force function $F(t, \dot{x}, x) = -kx$ where $k > 0$ is constant and is determined by the physical nature of the spring. In this case, (1.3) can be written in the form

$$m\ddot{x} = -kx.$$

or, equivalently,

$$\ddot{x} + \frac{k}{m}x = 0. \tag{1.4}$$

In the Chapter 2, we will see how to solve (1.4). For now, however, we can easily see that the function

$$x(t) = 5\cos\left(\left(\frac{k}{m}\right)^{1/2} t\right)$$

solves Eq. (1.4). To see this, it is customary to write

$$x(t) = 5\cos(\omega t)$$

where $\omega = \sqrt{k/m}$ and then observe that

$$\ddot{x}(t) = \frac{d^2}{dt^2}\left(5\cos(\omega t)\right) = -5\omega^2\cos(\omega t) = -\omega^2 5\cos(\omega t) = -\omega^2 x(t)$$

for all $t \in \mathbb{R}$. Since[c] $\omega^2 = k/m$, we have

$$\ddot{x}(t) + \frac{k}{m}x(t) = -\omega^2 x(t) + \frac{k}{m}x(t) = 0$$

for all $t \in \mathbb{R}$. Hence $x(t) = 5\cos(\omega t)$ satisfies Eq. (1.4) when $\omega = \sqrt{k/m}$. We will see, in fact, that (1.4) has the general solution

$$x(t) = C_1\cos(\omega t) + C_2\sin(\omega t)$$

where C_1 and C_2 are constant.

For a nice demonstration of the motion of a spring-mass system, modeled by Eq. (1.4), I encourage you to watch the Spray Paint Oscillator video from MIT's Physics Department [1] (Click here). The video illustrates a spring-mass system where a displaced mass (a can of spray paint) is subjected to a vertical force from a spring. The oscillatory behavior the mass exhibits is then captured in time as the spray can paints a (co)sinusoidal wave on a moving roll of paper. Of course, this oscillatory behavior is nicely predicted by the general solution above.

[a] As is customary, I will often use several different notation for (ordinary and partial) derivatives. Some common notations for derivatives of a function $y = y(t)$ are y', dy/dt, $\dot{y}$, Dy and $D_t y$.

[b] The dependence on velocity often comes from the resistive force (friction) a particle experiences while traveling through a fluid (air, water, etc).

[c] Of course, we designed ω precisely so that $\omega^2 = k/m$.

Exercise 1: Warm up: Playing around with trigonometric identities

In this exercise, you will play around with an alternative way of writing the general solution to (1.4). Please do the following.

1. Given $A \geq 0$ and $\delta \in \mathbb{R}$, show that the function

$$x(t) = A\cos(\omega t + \delta)$$

defined for $t \in \mathbb{R}$ solves Eq. (1.4), i.e., has the property that

$$\ddot{x}(t) + \frac{k}{m}x(t) = 0$$

for all $t \in \mathbb{R}$, where, as before, $\omega = \sqrt{k/m}$. Hint: If you're not used to doing things with arbitrary constants, first try things when, e.g., $A = 3$, $\omega = 4$ and $\delta = 5$. Keeping close track of the numbers $3, 4$ and 5 in your calculation, replace these numbers with the symbols A, ω and δ and make sure things work out as they should.

2. Show that

$$x(t) = C_1\cos(\omega t) + C_2\sin(\omega t)$$

can be equivalently expressed as

$$x(t) = A\cos(\omega t + \delta)$$

for $t \in \mathbb{R}$. More precisely, show that, given any real numbers C_1 and C_2, there are constants $A \geq 0$ and $\delta \in \mathbb{R}$ for which

$$C_1\cos(\omega t) + C_2\sin(\omega t) = A\cos(\omega t + \delta)$$

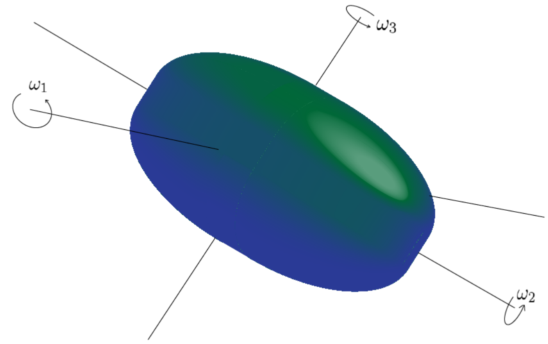
for all $t \in \mathbb{R}$. Please give the exact formulae for the constants C_1 and C_2 in terms of A and δ. Correspondingly, give the exact formulae for the constants A and δ in terms of C_1 and C_2. Hint: You might want to use the trigonometric identity $\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$. If you're very careful, you'll see that multiple δs will work.

Example 3: Rigid Body Dynamics

Given positive constants I_1, I_2 and I_3, consider the system[a] of differential equations

$$\begin{aligned} I_1\dot{\omega}_1 &= (I_2 - I_3)\omega_2\omega_3 \\ I_2\dot{\omega}_2 &= (I_3 - I_1)\omega_1\omega_3 \\ I_3\dot{\omega}_2 &= (I_1 - I_2)\omega_1\omega_2. \end{aligned} \tag{1.5}$$

Here, ω_1, ω_2 and ω_3 are functions of time t and $\dot{\omega}_1, \dot{\omega}_2$ and $\dot{\omega}_3$ are their derivatives in time, respectively. These equations are known as the Euler's Equations of rigid body dynamics and Eq. (1.5), specifically, gives the simplified torque-free version. This system of equations describes the evolution of the angular velocities[b] ω_1, ω_2 and ω_3 of a three-dimensional rigid body where I_1, I_2 and I_3 are the body's principal moments of inertia. In slightly simpler language, this equation describes how a three-dimensional rigid body rotates about its center of mass. This situation is illustrated in the figure below.



As it turns out, one can (through much effort) write down solutions to Eq. (1.5) in terms of elliptic functions[c] [7]. We won't however have such a need for writing down solutions. We will instead develop some very power methods to analyze these equations and deduce precise qualitative behavior of solutions through dynamical system analysis.

In the situation that $I_1 > I_2 > I_3 > 0$ or, equivalently that the mass profile of the rigid body is different in each of its three principle directions, Russian Cosmonaut Vladimir Dzhanibekov observed (in space) that the rigid body will rotate stably about its smallest and largest inertial axes (those corresponding to ω_1 and ω_2) and unstably about its intermediate axis (corresponding to ω_2). Essentially, this means that an object can be spun along two of its principal axes and it will continue to hold that spin; however, the object cannot be reliably spun around its intermediate axis. This phenomenon, known as the Dzhanibekov effect, can be completely understood by analyzing Eq. 1.5 and the corresponding theoretical explanation is aptly called the "Intermediate Axis Theorem". This result, also called the "Tennis Racket Theorem", was originally explained by M. Ashbaugh, C. Chicone and R. Cushman [2]. You can demonstrate this phenomenon yourself by flipping a book or a deck of cards. A nice illustration of the Dzhanibekov effect can also be seen here and here. By the end of the course, we will be able to explain the Intermediate Axis Theorem!

[a]We call this a *system of differential equations* because the three equations, considered together, interrelate three functions and their derivatives. Later we shall see that this distinction (between equations and systems of equations) is unnecessary once we allow solutions of "differential equations" to be vector-valued (and not simply real-valued).

[b]These are measured in radians per second.

[c]For those of you who have taken complex analysis, these are special types of meromorphic functions.

Example 4: Black-Scholes, Heat, and Schrödinger

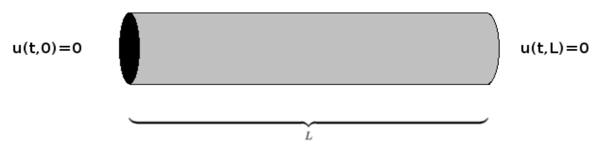
Given a (possibly complex) function $A = A(x,t)$, consider the differential equation

$$\frac{\partial u}{\partial t} + A\frac{\partial^2 u}{\partial x^2} = 0 \tag{1.6}$$

where $u = u(x,t)$ is a function of x and t. This equation, which relates partial derivatives of u in x and t, has an incredible number of applications, ranging from economics to physics to probability.

When $A(x,t) = \sigma^2 x^2/2 > 0$ for a real constant σ, Equation (1.6) is a special case of the Black-Scholes equation and it arises in the study of mathematical finance. In this case, u represents the price of an option governed by European call, a rule under which stocks are purchased, t represents time and x represents the price of a stock on which the option depends. This Black-Scholes model for financial markets was named by economist Robert C. Merton to give credit to Fischer Black and Myron Scholes, two economists who first began to understand the model and its applicability. The 1997 Nobel Prize in Economic Sciences was awarded to Merton and Scholes for their analysis and discovery of the so-called risk-neutral argument.

When $A = -\alpha^2 < 0$ for a real constant α, Equation (1.6) is called the heat equation and it arises in the study of heat conduction and thermodynamics. To explain how this situation arises, consider a rod of length L made of a thermally homogeneous material. Suppose that at time $t = 0$, the initial temperature of the rod is known and is given by $u(0, x) = u_0(x)$ for all $0 \leq x \leq L$. Think of the function $u_0 : [0, L] \to \mathbb{R}$ as some function that is known experimentally at $t = 0$. Let's assume that, for $t > 0$, the ends of the rod are attached to something (a thermal bath) that keeps their temperature fixed at 0 for all $t > 0$, i.e., $u(t, 0) = u(t, L) = 0$ for all $t > 0$. This set-up is depicted in the figure below.



A big question in the study of heat conduction is the following: *If we know the thermal properties of the rod, can we find the temperature of the rod, $u(t, x)$, for all $t > 0$ and all $0 \leq x \leq L$?* It is known, both experimentally and theoretically (see [9] for a clear derivation from first principles), that the temperature u in the rod satisfies Equation (1.6) for all $t > 0$ and $0 < x < L$ where α is a constant which depends on the rod's thermal conductivity, density and specific heat. Solving this problem (and answering the question posed above) was one of the main goals of Joseph Fourier in his treatise *Théorie analytique de la chaleur* (Analytic theory of heat) and his solution began an entire field of mathematics which we now call *Fourier analysis.*

When A is a purely imaginary number, say $A = i\hbar/2\mu$ where $\hbar$ is Planck's constant, Equation (1.6) is called Schrödinger's equation and explains the quantum mechanical evolution of a "free-particle" of (reduced) mass μ. In this case, u is called a wave function (usually denoted instead by ψ) and its square $|u|^2 = |\psi(t, x)|^2$ gives the probability of finding the quantum particle at position x and time t.

Now we've seen four examples of differential equations. The first three equations are known as ordinary differential equations because they only involve derivatives with respect to one variable. Ordinary differential equations are the subject of this course. The last equation is an example of a partial differential equation and is often the subject of our topics course, Mathematics 411.

1.2 What is an ordinary differential equation?

As stated in the previous section, this course is about ordinary differential equations. Let's give a definition.

Definition 1.2.1. *Given a function $F : \mathbb{R}^{n+2} \to \mathbb{R}$, an equation of the form*

$$F(t, y^{(n)}, y^{(n-1)}, \dots, y', y) = 0 \tag{1.7_0}$$

is called an ordinary differential equation (ODE). An analytic solution $y = y(t)$ of (1.7_0) is a sufficiently differentiable function (n-times differentiable with continuous n^{th} derivative) such that

$$F(t, y^{(n)}(t), y^{(n-1)}(t), \dots, y'(t), y(t)) = 0$$

for all t in a domain (usually an interval) on which $F(t, \cdot, \cdot, \dots)$ is defined and well-behaved.

What we mean by $F(t, \cdot, \cdot, \cdots)$ being defined and well-behaved will be made clear shortly (when we discuss existence and uniqueness of ordinary differential equations). For now, you can think of this as the domain in which F makes sense and is nice as a function of t. Also, we will often refer to an analytic solution simply as a solution. There is a small distinction between these two terms, but that distinction won't be important to us.

Example 5

Consider the function $F : \mathbb{R}^3 \to \mathbb{R}$ defined by the rule

$$F(a, b, c) = 2ac - b$$

for real numbers a, b, and c (or equivalently, for $(a, b, c) \in \mathbb{R}^3$). This function gives rise to the differential equation

$$F(t, y', y) = 2ty - y' = 0.$$

As the function F is defined and differentiable in t for all $t \in \mathbb{R}$, an analytic solution to this differential equation is a function $y = y(t)$ which is once continuously differentiable in t and has

$$2ty(t) - y'(t) = 0$$

for all $t \in \mathbb{R}$. You should check that the function $y(t) = 3e^{t^2}$ is a solution.

The *order* of an ordinary differential equation is the order of the highest-order derivative appearing in a non-trivial way. In the definition above, the order of (1.7_0) is n, as long as the derivative $y^{(n)}$ actually appears in the $F(t, y^{(n)}, y^{(n-1)}, \dots, y', y)$. This happens, in particular, if the partial derivative of F with respect to its second argument is non-zero. Pertaining to the example above,

$$2ty - y' = 0$$

is a first-order differential equation because the highest order derivative appearing in it is 1.

We say that an n^{th}-order ordinary differential equations is in *standard form* if it written as

$$y^{(n)} = f(t, y^{(n-1)}, y^{(n-2)}, \dots, y', y) \tag{1.7}$$

for some function $f : \mathbb{R}^{n+1} \to \mathbb{R}$. Many of the differential equations in this course will be written in standard form (or able to be equivalently expressed in standard form).

Example 6

The equation

$$y' = 2ty$$

is a first-order ordinary differential equation in standard form. This can be expressed as

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

where $f : \mathbb{R}^2 \to \mathbb{R}$ is given by $f(a, b) = 2ab$. It is easy to see that this equation is equivalent to the differential equation $2ty = y'$ considered in the preceding example. By contrast, the equation

$$\sin(y') + e^{y'} - ty = 0$$

is an ordinary differential equation but its not expressed in standard form. If you think about it carefully, it doesn't seem like it would be easy, if possible at all, to express this equation in standard form, i.e., to solve for y'. One, in fact, needs a very deep theorem from multivariate calculus, called the implicit function theorem, to place it (locally) in standard form. We shall learn about the implicit function theorem later in these notes.

Let's now isolate an important class of ordinary differential equations which will later be of great importance.

Definition 1.2.2. *A n^{th}-order differential equation is said to be linear if it is of the form*

$$a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + a_{n-2}(t)y^{(n-2)} + \cdots + a_1(t)y' + a_0(t)y = g(t) \tag{1.8}$$

where $a_0, a_1, \ldots, a_n$ and g are real-valued functions of t defined on an interval $I = (\alpha, \beta) \subseteq \mathbb{R}$ called coefficients; we shall assume further that $a_n(t) \neq 0$ for all $t \in I$. If it happens to be the case that $g = 0$, i.e., (1.8) is equivalently

$$a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + a_{n-2}(t)y^{(n-2)} + \cdots + a_0(t)y = 0,$$

then the differential equation is said to be a linear homogeneous ordinary differential equation or simply homogeneous. Otherwise, the equation is said to be inhomogeneous.

Observe that the familiar equation

$$y' - 2ty = 0$$

is a first-order linear homogeneous differential equation. Here $a_1(t) = 1$ and $a_0(t) = -2t$. By contrast, the equation

$$y' + (y)^2 = 0$$

is not a linear equation. We shall take up the study of linear differential equations in the chapters to come.

1.3 What is an initial value problem?

Consider an n^{th}-order differential equation

$$y^{(n)} = f(t, y^{(n-1)}, \ldots, y', y). \tag{1.9}$$

An *initial value problem* for the differential equation (1.9) comes by specifying a time t_0 and n numbers $y_0^{(n-1)}, y_0^{(n-2)}, \ldots, y_0', y_0$ and asking that a solution y=y(t) to (1.9) also satisfies the constraints

$$y^{(n-1)}(t_0) = y_0^{(n-1)}, \quad y^{(n-2)}(t_0) = y_0^{(n-2)}, \quad \cdots \quad y'(t_0) = y_0' \text{ and } y(t_0) = y_0. \tag{1.10}$$

The numbers $y_0^{(n-1)}, y_0^{(n-2)}, \ldots, y_0', y_0$ are called *initial values* and the n equations in (1.10) are called *initial conditions.* Such an initial value problem is often written in the form

$$\begin{cases} y^{(n)} = f(t, y^{(n-1)}, \ldots, y', y) \\ \\ \quad y^{(n-1)}(t_0) = y_0^{(n-1)}, \\ \quad y^{(n-2)}(t_0) = y_0^{(n-2)} \\ \qquad \vdots \\ \quad y'(t_0) = y_0' \\ \quad y(t_0) = y_0 \end{cases} \tag{1.11}$$

or, equivalently, in the form

$$\left\{ y^{(n)} = f(t, y^{(n-1)}, y^{(n-2)}, \ldots, y', y), \quad y^{(n-1)}(t_0) = y_0^{(n-1)}, \ldots, y'(t_0) = y_0', \, y(t_0) = y_0. \right.$$

A solution y to this initial value problem is, by definition, a function $y = y(t)$ solving the differential equation (1.9) and the n-initial conditions (1.10).

For example, consider the initial value problem

$$\begin{cases} y' = 2ty \\ \\ y(1) = 5. \end{cases} \tag{1.12}$$

We notice that, because this is a first-order differential equation, the initial value problem consists of only one initial condition, $y(1) = 5$, where $t_0 = 1$ and $y_0 = 5$. You should verify that the function $y_1(t) = (5/e)e^{t^2} = 5e^{t^2-1}$ satisfies both the differential equation $y' = 2ty$ and the initial condition $y(1) = 5$ (Check it!). On the other hand, the function $y_2(t) = 2e^{t^2}$ solves the differential equation $y' = 2ty$ but not the initial condition $y(1) = 5$. Therefore y_1 is a solution to the initial value problem (1.12) but y_2 is not.

Though there are a number of interesting mathematical reasons to discuss initial value problems, the posing of an initial value problem is often motivated directly by application. For example, consider the equation for a spring-mass system

$$\ddot{x} + 4x = 0.$$

As we discussed in the introductory section, this differential equation describes an object with mass $m = 1$ (kg) being displaced from equilibrium by a spring with spring constant $k = 4$ (kg/s^2); here $\omega = \sqrt{k/m} = \sqrt{4} = 2$. A solution to this differential equation is a function $x = x(t)$ which gives the displacement of the object from equilibrium at any future time $t \geq 0$. When it comes to describing the movement of an (actual) object in time, we could expect to know the position $x(0) = x_0$ and velocity $\dot{x}(0) = \dot{x}_0$ of the object at the starting point $t_0 = 0$. For example, suppose that the object was initially displaced from equilibrium 0.1 (in meters) and had zero initial velocity. In this case, we would seek a solution to $\ddot{x} + 4x = 0$ which satisfied the initial conditions $x(0) = 0.1$ and $\ddot{x} = 0$. In other words, to describe the evolution of the spring which started displaced by 0.1 meters and had zero initial velocity, we would seek a solution to the initial value problem

$$\{\ddot{x} + 4x = 0 \qquad \dot{x}(0) = 0,\, x(0) = 0.1. \tag{1.13}$$

Observe that the above initial value problem represents a 2^{nd}-order differential equation and, as we required, it comes with two initial conditions. If you take for granted the earlier claim that any such solution to $\ddot{x} + 4x = 0$ is of the form

$$x(t) = C_1 \cos(2t) + C_2 \sin(2t)$$

where C_1 and C_2 are (to be determined) constants, we can attempt to solve the initial value problem (1.13). To this end, let's plug in our general solution to the initial first initial condition. We have

$$0.1 = x(0) = C_1 \cos(2 \cdot 0) + C_2 \sin(2 \cdot 0) = C_1$$

and therefore $C_1 = 0.1$. For the second initial condition, we obtain

$$0 = \dot{x}(0) = C_1(-2\sin(2 \cdot 0)) + C_2(2\cos(2 \cdot 0)) = 2C_2$$

and therefore $C_2 = 0$. Putting this information together, we suspect that

$$x(t) = 0.1 \cos(2t)$$

solves the initial value problem (1.13). You should check for yourself that our suspicion is correct, i.e., this does solve the initial value problem.

As we've seen already, differential equations themselves admit multiple solutions. By pairing a differential equation with enough initial conditions, one hopes to isolate a single solution satisfying both the differential equation and the initial conditions. For example, in the mass-spring system above, we started with a collection of solutions $x(t) = C_1 \cos(2t) + C_2 \sin(2t)$ and, by appealing to initial conditions, isolated one particular solution $x(t) = 0.1 \cos(2t)$. This is, of course, a physically reasonable thing to expect. By knowing the initial position and velocity of an object, one hopes that the evolution of the object is then determined (at least approximately) by physics (here, Newton's Second Law). In the next chapter, we will study a major result called the Picard-Lindelöff theorem. For a given differential equation and corresponding initial value problem, this theorem, in particular, gives sufficient conditions under which a single (unique!) solution can be found. The following exercise will hlep you think about this issue of uniqueness and isolating solutions.

Exercise 2

Let k be a non-zero constant and consider the ordinary differential equation

$$\frac{dy}{dt} = ky. \tag{1.14}$$

1. Show that all (analytic) solutions to (1.14) are of the form $y(t) = Ce^{kt}$ where C is a constant. *Note: As will be true throughout this course, the exact phrasing of this question is important. This question is not asking you to show that Ce^{kt} is a solution (which is easy to verify). This question is asking you to show that, if you're given* ***any*** *solution to (1.14), then it* ***must*** *be of the form Ce^{kt}. Hint: Assume that $w = w(t)$ is a solution and use methods from single-variable calculus to show that $w(t)/e^{kt}$ is necessarily a constant.*

2. Given a (fixed) constant y_0, use your previous result to show that the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = ky \\ y(0) = y_0 \end{cases}$$

has one and only one solution.

Chapter 2

First-order ODEs

The focus of this chapter is to understand first-order differential equations and their initial value problems. We begin this chapter by considering a few special classes of first-order differential equations for which there are methods to produce analytic (closed-form) solutions. Our aim here is to begin to understand how the structure of the differential equation affects the quantitative and qualitative behavior of its solutions. Another aim is to look at some applications of these equations. We shall quickly find, however, that most first-order differential equations are not solvable by such simple methods yet we still would like to understand their solutions. That leads to an important question: If we have no method for solving a differential equation or initial value problem, how do we know a solution exists? Further, if we were able to find a solution to an initial value problem, how would we know it was the only one? Though these questions seem abstract, they are essential questions to ask if one wants to solve differential equations in the real world. To answer these questions, we will then turn to the major existence/uniqueness theorem for differential equations, the Picard-Lindelöff theorem. The theorem will tell us conditions under which a given initial value problem has a solution and when that solution is unique. This will then be helpful to us as we turn our focus to qualitative (dynamical system) analysis and methods for numerical approximations of solutions. Let's begin with the easy stuff.

2.1 First Method: Separable Equations

The first class of differential equations that we can solve are called separable equations. These are defined as follows.

Definition 2.1.1. *Given continuous real-valued functions g and h of a single real variable, a first-order separable differential equation is an equation of the form*

$$\frac{dy}{dt} = \frac{g(t)}{h(y)}. \tag{2.1}$$

In view of the definition above, we will generally assume that g and h are continuous functions on an interval (or a union of intervals). Further, we shall assume that $h(y)$ is non-zero on this domain so that the quotient $g(t)/h(y)$ is a continuous function of t and y. Though you don't have to worry about it now, these domains of continuity, especially that for which $h(y)$ is non-zero, will turn out to be important when we study existence and uniqueness of solutions.

Example 1

1. The differential equation
$$\frac{dy}{dt} = \frac{\sin(t)}{y}$$
is separable. For this equation, $g(t) = \sin(t)$ and $h(y) = y$. Here $g(t)$ is continuous for all $y \in \mathbb{R}$ and y is continuous and nonzero on the intervals $(-\infty, 0)$ and $(0, \infty)$.

2. The differential equation
$$\frac{dy}{dt} = \sin(t + y)$$
is not separable.

Let's now outline a method for producing solutions of separable equations. To this end, suppose that $y = y(t)$ is a solution to (2.1). Then, for all t for which y is defined, continuous and has continuous derivative $dy/dt = y'(t)$, we have

$$h(y(t))y'(t) = g(t).$$

Computing antiderivatives of both sides of this equation gives

$$\int h(y(t))y'(t)\,dt = \int g(t)\,dt + C.$$

where we are using the fact that any two antiderivatives differ by, at most, a constant C. Focusing on the left hand side and noting that $dy = \frac{dy}{dt}dt = y'(t)dt$, we obtain

$$\int h(y(t))y'(t)\,dt = \int h(y)dy;$$

this is really u-substitution where we have replaced the function $y(t)$ with the independent variable y. Consequently we have

$$\int h(y)\,dy = \int g(t)\,dt + C$$

or, equivalently,

$$H(y) = G(t) + C$$

where H is an antiderivative of h and G is an antiderivative of g, i.e., $H' = h$ and $G' = g$. Then, provided that the function H is invertible (so it has an inverse), a solution to the differential equation (2.1) is given by

$$y(t) = y = H^{-1}(G(t) + C)$$

where C is some constant of integration. If the differential equation (2.1) comes paired with an initial condition, i.e., to form an initial value problem, we can subject our solution to the initial condition. In practice, this will often allow you to determine an appropriate constant C which will solve the initial value problem.

Example 2

Consider the separable differential equation

$$\frac{dy}{dt} = \frac{2t}{3y^2}.$$

Here $h(y) = 3y^2$ and $g(t) = 2t$. We separate variables to find

$$\int 3y^2\,dy = \int h(y)\,dy = \int g(t)\,dt + C = \int 2t\,dt + C.$$

In other words

$$y^3 = H(y) = G(t) + C = t^2 + C$$

because $H'(y) = 3y^2$ and $G'(t) = 2t$. As $H(y) = y^3$ is an invertible function of y, we can solve the above equation to obtain the family of solutions

$$y(t) = H^{-1}(G(t) + C) = (t^2 + C)^{1/3}.$$

You should verify that this, in fact, is a solution to the differential equation.

Caution: Where the C is placed in the above solution is important! In fact, though it looks incredibly similar to the solution we found, the function

$$\tilde{y}(t) = (t^2)^{1/3} + C$$

does not solve the differential equation for any $C \neq 0$.

In the general procedure above, we used a u-substitution to treat y as an independent variable (in the integration) and not as a function of t (dependent variable). In practice, it's not necessary to worry about this step and instead move directly from (2.1) to the identity

$$\int h(y)\,dy = \int g(t)\,dt + C.$$

In practice, to solve a first-order separable ordinary differential equation of the form

$$\frac{dy}{dt} = \frac{g(t)}{h(y)},$$

you should follow three steps:

Step 1. Formally, by treating the differentials dy and dx as independent quantities, move everything involving y in the differential equation to the left-hand side and everything involving t to the right-hand side. This yields

$$h(y)\,dy = g(t)\,dt.$$

Step 2. Place integration signs on both sides of the equation above to obtain

$$\int h(y)\,dy = \int g(t)\,dt + C$$

where C is a (necessary) constant of integration. Compute both indefinite integrals (noting that you've already accounted for the additive constant C).

Step 3. Solve your remaining equation for $y = y(t)$.

As Step 1 of the above procedure isolates the (independent variable) t from the (dependent) variable y, we call this method *separation of variables*. Let's see this method in practice.

Example 3

Let's use separation of variables to solve the initial value problem

$$\begin{cases} y' = 2ty^2 \\ y(1) = -\frac{1}{2}. \end{cases}$$

Upon separating variables we find,

$$\int \frac{1}{y^2}\, dy = \int 2t\, dt + C$$

or, equivalently,

$$\frac{-1}{y} = t^2 + C.$$

Let's solve this equation to find

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

To solve the initial value problem, we now consider the initial condition $y(1) = -1/2$. Subject to this initial condition, we have

$$-\frac{1}{2} = y(1) = \frac{-1}{1^2 + C}$$

and so $C = 1$. Thus, the method of separation of variables produces the solution

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

You should, in fact, check directly that this solves the initial value problem.

Example 4

It will often be the case that the equation

$$H(y(t)) = G(t) + C \tag{2.2}$$

cannot be easily solved for $y(t)$, at least in terms of functions you know. This is precisely the situation in which $H(y)$ does not have an inverse you can write down. In this case, solutions can be left in the *implicit form* (2.2). For example, consider the differential equation.

$$\frac{dy}{dt} = \frac{2t}{y^2 + 1}$$

is separable. In this case, $g(t) = 2t$ and $h(y) = y^2 + 1$. Observe that $g(t)$ is a continuous for all $t \in \mathbb{R}$ and $h(y)$ is a continuous (and strictly positive) function for all $y \in \mathbb{R}$. Let's separate variables to find

$$\int (1+y^2)\,dy = \int h(y)\,dy = \int g(t)\,dt + C = \int 2t\,dt + C = t^2 + C.$$

Since $y + y^3/3$ is an antiderivative of $1+y^2$, we have

$$H(y(t)) = y(t) + \frac{y(t)^3}{3} = t^2 + C.$$

Solving this cubic equation in y, though possible, is not necessarily easy. In cases like this, leaving things in implicit form is sufficient.

Exercise 3

1. Determine whether or not the following first-order ordinary differential equations are separable. If so, give the general solution.

   (a)
   $$\frac{dy}{dt} = (1+2t)(1+y)$$

   (b)
   $$\frac{dy}{dt} = 1 - t + y^2 - ty^2$$

   (c)
   $$\frac{dy}{dt} = 4y + t$$

   (d)
   $$\frac{dx}{dt} = te^{x-t^2+3}$$

2. Solve the following initial value problems.

   i.
   $$\begin{cases} \dfrac{dy}{dt} = \dfrac{t^2}{y+t^3y} \\ y(0) = -2 \end{cases}$$

   ii.
   $$\begin{cases} \dfrac{dz}{dt} = 2tz^2 + 3t^2z^2 \\ z(1) = -1 \end{cases}$$

Example 5

As it turns out, the method of separation of variables can often miss solutions. Consider, for instance, the differential equation

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

Using the method of separation of variables, we have

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

On one hand

$$H(y) = \int \frac{1}{(y-1)^2}\,dy = \frac{-1}{(y-1)}$$

and on the other hand

$$G(t) = \int 2t\,dt = t^2.$$

Consequently, the method of separation of variables produces solutions $y = y(t)$ for which

$$\frac{-1}{y-1} = H(y) = G(t) + C = t^2 + C$$

or, equivalently,

$$y(t) = 1 - \frac{1}{t^2 + C} \tag{2.3}$$

where C is a constant. It is easy to see (and you should check it yourself) that this gives a solution to the given differential equation for every real number C.

Consider now the initial value problem

$$\{y' = 2t(y-1)^2 \qquad y(0) = 1\ .$$

Assuming that our solution is of the form (2.3), we seek a constant C for which

$$1 = y(0) = 1 - \frac{1}{0^2 + C} = 1 - \frac{1}{C}$$

or, equivalently, a constant C for which $0 = 1/C$. Of course, no such constant exists. For this reason, (2.3) cannot be a general solution to the given differential equation because not ever initial value problem can be solved by simply specifying the constant C. This is a limitation of the method of separation of variables. Let us notice, however, that the function which is identically one, i.e., $y(t) = 1$ for all $t \in \mathbb{R}$, solves the initial value problem. To see this, simply observe that $y(0) = 1$ and

$$\frac{dy}{dt} = \frac{d}{dt}1 = 0 = 2t(1-1)^2 = 2t(y(t)-1)^2$$

for all t. Hence the constant function $y(t) = 1$ does indeed solve the initial value problem.

2.1.1 An application: Newton's Law of Cooling

Consider an object with initial temperature T_0 placed in a thermal bath of ambient temperature T_a. We want to understand how the temperature of the object, $T(t)$, changes over time from its initial temperature T_0. For example, if a hot cup of coffee at temperature $T_0 = 200°F$ is placed in a room of constant temperature $T_a = 70°F$, we would like to know the temperature $T(t)$ as a function of time $t > 0$.

A good model for this situation is to assume that the temperature $T(t)$ changes in time in a way that is proportional to the difference between $T(t)$ and the ambient temperature T_a. As a differential equation, this is

$$\frac{dT}{dt} = -k(T - T_a),$$

known as Newton's Law of Cooling. Assuming that the constant k is positive, let's make note of a few things which make this model physically plausible:

1. In the case that the temperature of the object T is equal to the ambient temperature T_a, the right-hand side of the differential equation is zero and hence $dT/dt = 0$. In this case we expect the temperature to remain constant.

2. In the case that the object is warmer than the ambient temperature, i.e., $T > T_a$, the right-hand side of the differential equation is negative (because $k > 0$) and so $dT/dt < 0$ which predicts that the object will cool over time.

3. Finally, in the case that the object is cooler than the ambient temperature, e.g., you place a cold can of soda outside on a hot summer day, the same argument predicts that $dT/dt > 0$ and hence the object will warm in time.

Though the above observations provide a good heuristic account of why Newton's Law of Cooling makes some sense, the real justification can only come empirically. As it turns out, this model does provide a fairly reasonable description of cooling/warming for many materials away from phase transitions [9]. The constant k can often be measured for a given object and is a function of the chemical makeup of the object and the object's shape (especially its ratio of the surface area to volume).

The corresponding initial value problem for this model is

$$\begin{cases} \dfrac{dT}{dt} = -k(T - T_a) \\ \\ \quad T(0) = T_0 \end{cases}$$

where, again, k is a constant, T_a is the ambient temperature of the thermal bath (environment) and T_0 is the initial temperature of the object. To solve the differential equation, we separate variables to obtain

$$\frac{dT}{T - T_a} = -k\,dt$$

and so

$$\log(T - T_a) = \int \frac{dT}{T - T_a} = \int -k\,dt + C = -kt + C.$$

Exponentiating both sides give

$$T(t) - T_a = T - T_a = e^{-kt+C} = e^C e^{-kt} = Ke^{-kt}$$

where $e^C = K$ is a constant. Thus

$$T(t) = Ke^{-kt} + T_a.$$

where K is constant. Combining this with the initial condition $T(0) = T_0$, we obtain a solution to the initial value problem,

$$T(t) = (T_0 - T_a)e^{-kt} + T_a$$

defined for $t \geq 0$.

Remark 2.1.2. Had we been more careful in the above computations, we would have written

$$\int \frac{dT}{T - T_a} = \log|T - T_a|,$$

accounting for the possibility that $T < T_a$, and also made mention that the constant $K = e^C$ should necessarily be positive. In being so careful, one then has to really think hard about the signs of things while solving for $T(t)$. As it turns out, all of this worry is unnecessary and these two wrongs (missing the absolute value and not restricting K to be positive) cancel each other out and still give

$$T(t) = (T_0 - T_a)e^{-kt} + T_a$$

as the correct solution. Getting the correct answer by being cavalier about absolute values won't always be a good idea and so sometimes the worrying is good practice.

Example 6: Cooling coffee

Suppose that a cup of coffee, initially at $200°F$ is placed in a room of constant ambient temperature $T_a = 60°F$. Due to the thermal conductivity of the coffee (and its vessel), suppose that it is known that Newton's Law of Cooling is approximately valid with $k = \log(2)/10 \approx 0.0693\ldots(\mathrm{min})^{-1}$. How long will it take for the coffee to cool to the drinkable temperature of $130°F$? Does the temperature of the coffee equilibriate in the long run?

To answer these questions, we appeal to the solution to the initial value problem obtained above. We have

$$\begin{aligned} T(t) &= (T_0 - T_a)e^{-kt} + T_a = (200 - 60)e^{-\log(2)t/10} + 60 \\ &= 140(e^{-\log(2)})^{(t/10)} + 60 = \frac{140}{2^{(t/10)}} + 60 \end{aligned}$$

for $t \geq 0$. To answer the first question, we ask: For which t is

$$T(t) = \frac{140}{2^{(t/10)}} + 60 = 130$$

or, equivalently, for which t is

$$2^{(t/10)} = 2?$$

In this case, we see that $t = 10$ minutes.

Exercise 4: A warming can of Moxie

On a hot day in Waterville (ambient temperature is 90° F), you place a cold can of Moxie (Maine's official soft drink), which is initially at 40°, on a picnic table. Five minutes later, the Moxie has warmed to 50°. If the outside temperature remains at a constant 90°, what will happen to the temperature of the Moxie after it remains on the table for 20 minutes? How many minutes will it take for the Moxie to warm to 80°? What happens to the temperature of the Moxie in the long run, i.e., as $t \to \infty$?

2.2 Linear First-Order Equations

A linear first-order ordinary differential equation is an equation of the form

$$\frac{dy}{dt} + a(t)y = b(t) \tag{2.4}$$

where $a(t)$ and $b(t)$ are real-valued functions. We shall assume that the functions a and b are continuous on some interval $I = (\alpha, \beta)$. In comparing this form to Definition 1.2.2, the above equation can be seen equivalent by setting $a(t) = a_0(t)/a_1(t)$ and $b(t) = g(t)/a_0(t)$. We say that the equation (2.4) is *homogeneous* if $b(t) = 0$, i.e., if (2.4) is of the form

$$\frac{dy}{dt} + a(t)y = 0. \tag{2.5}$$

We observe that a homogeneous first-order equation is separable. To find an analytic solution $y = y(t)$, we separate variables to find

$$\frac{1}{y}\frac{dy}{dt} = -a(t)$$

and therefore

$$\ln|y(t)| = -\int a(t)\,dt + c_1$$

Upon taking exponentials of both sides, we obtain

$$|y(t)| = \exp\left(-\int a(t)\,dt + c_1\right).$$

This can be written equivalently as

$$\left|y(t)\exp\left(\int a(t)\,dt\right)\right| = e^{c_1} = k \tag{2.6}$$

which holds for all $t \in I$, i.e., for all $\alpha < t < \beta$. At this point, we'd like to drop the absolute values and so we'll need to worry a little bit. The following exercise will help in this.

Exercise 5

Let f be a real-valued function on $I = (\alpha, \beta)$. Suppose that, for some constant k,

$$|f(t)| = k \tag{2.7}$$

for all $\alpha < t < \beta$.

1. Suppose that f is a continuous function on I. Using the intermediate value theorem (from Calculus 1), argue why f must be identically constant on I, i.e., there is some real number C for which $f(t) = C$ for all $\alpha < t < \beta$. What is C?

2. Give an example of a (possibly discontinuous) function f which satisfies (2.7) but for which f is not constant.

Observe that, as $a = a(t)$ is a continuous function on I, any antiderivative $\int a(t)\,dt$ will also be continuous. Also, as we seek an analytic solution, $y(t)$ is necessarily continuous. Consequently, the expression

$$y(t)\exp\left(\int a(t)\,dt\right)$$

is continuous and, in view of (2.6), its absolute value is constant. Using the result from Item 1 of the exercise above, we conclude that

$$y(t)\exp\left(\int a(t)\,dt\right) = C$$

for some constant C or, equivalently,

$$y(t) = C \exp\left(-\int a(t)\,dt\right) = Ce^{-A(t)}. \tag{2.8}$$

where $A(t) = \int a(t)\,dt$ is an antiderivative of a on I. Since any two antiderivatives of $a(t)$ on I differ by, at most, an additive constant[1], a replacement of one antiderivative in (2.8) with another will result to simply changing the multiplicative constant C. Equation (2.8) is said to be the general solution of the homogeneous equation, (2.5). Let's verify directly that (2.8) solves (2.5). As $A'(t) = a(t)$, we have, for any constant C,

$$\frac{d}{dt}y(t) = \frac{d}{dt}Ce^{-A(t)} = -CA'(t)e^{-A(t)} = -a(t)Ce^{-A(t)} = -a(t)y(t)$$

and therefore

$$\frac{dy}{dt} + a(t)y(t) = 0$$

for all $\alpha < t < \beta$.

Example 7

Consider the linear homogeneous differential equation

$$\frac{dy}{dt} + \frac{t}{\sqrt{1+t^2}}y = 0.$$

To find a solution to the equation, we must find an antiderivative $A(t)$ of $a(t) = t(1+t^2)^{-1/2}$. Upon making a u-substitution, $u = t^2$, we find

$$A(t) = \int \frac{t}{\sqrt{1+t^2}}\,dt = \sqrt{1+t^2}.$$

Therefore, the method outlined above yields the solution

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

for each $C \in \mathbb{R}$. You should verify that this solves the given differential equation.

We've extensively treated the homogeneous case. Let's now deal with the non-homogeneous case

$$\frac{dy}{dt} + a(t)y = b(t).$$

To solve this equation, we need a "trick". This trick is found by asking the following question: Could we multiply this non-homogeneous equation by a function $\mu(t)$ which would make the left hand side equivalent to the derivative of the product of μ and y, i.e., we want to find a suitable function μ for which

$$\frac{d}{dt}(\mu(t)y(t)) = \mu(t)\frac{dy}{dt} + a(t)\mu(t)y(t).$$

In studying the solution we found for the homogeneous equation, we suspect that $\mu(t) = e^{A(t)}$ where $A(t)$ is an antiderivative of $a(t)$. If this is the case, we make use of the product

[1] Remember this from your single-variable calculus course?

rule to see that

$$\begin{aligned}
\frac{d}{dt}\left(\mu(t)y(t)\right) &= \frac{d}{dt}\left(e^{A(t)}y(t)\right) = e^{A(t)}\frac{dy}{dt} + \frac{d}{dt}\left(e^{A(t)}\right)y(t) \\
&= e^{A(t)}\frac{dy}{dt} + A'(t)e^{A(t)}y(t) \\
&= e^{A(t)}\frac{dy}{dt} + a(t)e^{A(t)}y(t);
\end{aligned}$$

this is just as we desired. For this reason, the function

$$\mu(t) = e^{A(t)} = \exp\left(\int a(t)\,dt\right)$$

is called an integrating factor[2]. Let's multiply the integrating factor into the inhomogenous equation (1.8) and make use of the product property. We have

$$\frac{d}{dt}\left(\mu(t)y(t)\right) = \mu(t)\frac{dy}{dt} + a(t)\mu(t)y(t) = \mu(t)b(t).$$

In view of the fundamental theorem of calculus

$$\mu(t)y(t) = \int \frac{d}{dt}\left(\mu(t)y(t)\right) = \int \mu(t)b(t)\,dt + C$$

and so

$$y(t) = \frac{1}{\mu(t)}\int \mu(t)b(t)\,dt + \frac{C}{\mu(t)}.$$

Let's verify that this, in fact, solves the inhomogeneous differential equation. We have

$$\begin{aligned}
\frac{dy}{dt} &= \frac{d}{dt}\left(\frac{1}{\mu(t)}\int \mu(t)b(t)\,dt\right) + \frac{d}{dt}\left(\frac{C}{\mu(t)}\right) \\
&= \frac{d}{dt}\left(e^{-A(t)}\int e^{A(t)}b(t)\,dt\right) + \frac{d}{dt}\left(Ce^{-A(t)}\right) \\
&= -A'(t)e^{-A(t)}\left(\int e^{A(t)}b(t)\,dt\right) + e^{-A(t)}\frac{d}{dt}\left(\int e^{A(t)}b(t)\,dt\right) - A'(t)Ce^{-A(t)} \\
&= -a(t)e^{-A(t)}\int e^{A(t)}b(t)\,dt + e^{-A(t)}\left(e^{A(t)}b(t)\right) - a(t)Ce^{-A(t)} \\
&= b(t) - a(t)\left(e^{-A(t)}\int e^{A(t)}b(t)\,dt + Ce^{-A(t)}\right) \\
&= b(t) - a(t)\left(\frac{1}{\mu(t)}\int \mu(t)b(t)\,dt + \frac{C}{\mu(t)}\right) \\
&= b(t) - a(t)y(t)
\end{aligned}$$

and so

$$\frac{dy}{dt} + a(t)y(t) = b(t)$$

as desired. As it turns out, the solution $y(t)$ above is a general solution to the differential equation (1.8) which means that **every** solution to (1.8) is of this form. We state this as a theorem and postpone the proof until the end of the chapter.

[2]Note, we are using a new notation for the exponential function, $e^x = \exp(x)$.

Theorem 2.2.1. *Let $a(t)$ and $b(t)$ be continuous functions on an interval $I = (\alpha, \beta)$ and let $A(t)$ be an antiderivative of $a(t)$. Then the differential equation*

$$\frac{dy}{dt} + a(t)y = b(t)$$

has infinitely many solutions, all of which are given by

$$y(t) = \frac{1}{\mu(t)} \int \mu(t)b(t)\,dt + \frac{C}{\mu(t)}$$

for $t \in I$ where C is a constant and $\mu(t) = e^{A(t)}$.

Remark 2.2.2. In practice, it's often not of great use to memorize the formula above for $y(t)$. If you understand the basic idea of how this solution was found, it's straightforward to reproduce it.

Example 8

Consider the inhomogeneous differential equation

$$\frac{dy}{dt} + y = 10t.$$

This is a linear inhomogeneous first-order ordinary differential equation; it is not separable. Given that $a(t) = 1$, to solve this differential equation we introduce the integrating factor

$$\mu(t) = \exp\left(\int a(t)\,dt\right) = e^t$$

making use of the antiderivative $A(t) = t$ of $a(t)$. Multiplying both sides of the differential equation by $\mu(t)$ we obtain

$$e^t \frac{dy}{dt} + e^t y(t) = 10te^t.$$

As designed, we recognize the left-hand side by the derivative of the product $e^t y(t)$ and, in this way, we obtain

$$\frac{d}{dt}\left(e^t y(t)\right) = 10te^t$$

and therefore

$$e^t y(t) = \int \frac{d}{dt}\left(e^t y(t)\right)\,dt + C = \int 10te^t\,dt + C = 10\int te^t\,dt + C.$$

To compute this integral, it's useful to integrate by parts where $u = t$ and $dv = e^t$. This gives

$$\int te^t\,dt = uv - \int v\,du = te^t - \int e^t \cdot 1\,dt = te^t - e^t.$$

Putting everything together yields

$$e^t y(t) = 10(te^t - e^t) + C$$

from which we obtain the general solution

$$y(t) = 10t - 10 + Ce^{-t}.$$

Example 9

Consider the initial value problem

$$\begin{cases} \dfrac{dy}{dt} + 4y = e^{-3t} \\ \quad y(0) = 2. \end{cases}$$

Let's first solve the inhomogenous differential equation. Given that $a(t) = 4$, we multiply both sides by the integrating factor

$$\mu(t) = \exp\left(\int 4\,dt\right) = e^{4t}$$

which yields

$$\frac{d}{dt}\left(e^{4t}y(t)\right) = e^{4t}\frac{dy}{dt} + 4e^{4t}y(t) = e^{4t}e^{-3t} = e^{t}.$$

Consequently,

$$\left((e^{4t}y(t)\right) = \int e^{t}\,dt + C = e^{t} + C.$$

Thus, our general solution is

$$y(t) = \frac{1}{e^{4t}}\left(e^{t} + C\right) = e^{-3t} + Ce^{-4t}.$$

Let's subject this general solution to the initial condition $y(0) = 2$. We have

$$2 = y(0) = e^{-3\cdot 0} + Ce^{-4\cdot 0} = 1 + C$$

and hence $C = 1$. Plugging this C into the general solution yields

$$y(t) = e^{-3t} + e^{-4t}.$$

This is, in fact, a solution to the initial value problem. I encourage you to verify this directly, i.e., verify that this y solves the differential equation and also satisfies the initial condition.

Now, it's your turn.

Exercise 6

1. Solve the following linear equations.

   (a)
   $$\frac{dy}{dt} + \frac{2t}{1+t^2}y = \frac{1}{1+t^2}$$

(b)

$$\frac{dz}{dt} + z = te^t$$

(c)

$$\dot{x} - 2x = \sin(2t)$$

(d)

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

2. Solve the following initial value problems.

(a)

$$\begin{cases} \frac{dy}{dt} + \frac{y}{1+t} = 2 \\ y(0) = 1. \end{cases}$$

(b)

$$\begin{cases} \frac{dy}{dt} + 2ty = t \\ y(0) = 3/2 \end{cases}$$

Let's study a nice application of linear first-order differential equations.

Exercise 7

Radioactivity is a property of substances whose constituent atoms undergo spontaneous decomposition. This process of spontaneous decomposition is known as radioactive decay. This decay usually occurs at some constant rate and can be measured by a Geiger counter. Generally speaking, the more radioactive the material, the faster the decay. Based on work by Ernest Rutherford and his contemporaries (including Marie Curie), it was discovered that radioactive decay was well modeled by a first-order ordinary differential equation of the form

$$\frac{dN}{dt} + \lambda N = 0$$

where N is the number of atoms of a substance at time t and λ is the rate of decay, a constant. If substance is known to have N_0 atoms at time t_0, the number of atoms for $t > t_0$ satisfies the initial value problem

$$\begin{cases} \frac{dN}{dt} + \lambda N = 0 \\ N(t_0) = N_0. \end{cases}$$

1. The half-life of a radioactive substance is defined to be the amount of time $t - t_0$

at which the number of atoms left $N(t)$ is exactly half of the initial number of atoms, N_0. If the half-life of a certain radioactive sample of carbon-14 is 5570 years, and the decay is described by the initial value problem above, find the rate of decay λ.

2. Given your result from the previous item, find the amount of time it will take for this sample of carbon-14 to decay to 1/10 of its original mass.

2.3 Qualitative analysis: slope fields and solutions to initital value problems

In this section, we study first-order differential equations from a qualitative perspective. We hope our study here builds a strong intuition and informs upon the long term behavior of solutions of first-order ODEs. To this end, we consider the general first-order ordinary differential equation

$$\frac{dy}{dt} = f(t, y).$$

We ask:

What does this equation tell us graphically?

To answer this question, we first introduce the notion of a slope field. Given a function $f(t, y)$, a slope field is produced by selecting points $(t, y) \in \mathbb{R}^2$ in the domain of f and plotting, at the point (t, y), a small line segment with slope $f(t, y)$. An example of a slope field for a function $f(t, y)$ is plotted in Figure 2.1. The small line segments drawn are called *mini tangent lines* – you can disregard the arrowheads drawn as the only essential feature is the slope of these lines (and not their direction).

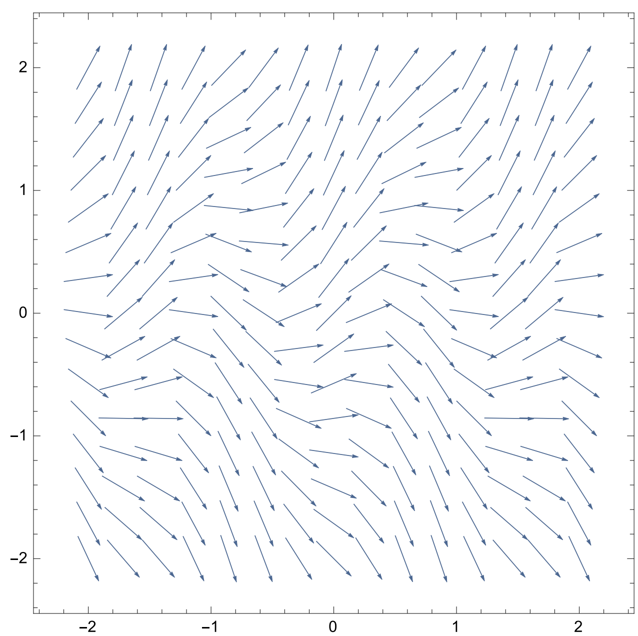


Figure 2.1: A slope field

Given a function $f(t, y)$ and its corresponding slope field, we say that $\mathcal{C}$ is an *integral curve* for this slope field if $\mathcal{C}$ is a smooth curve (continuous with well-defined tangent lines) such that, at each point (t, y) through which $\mathcal{C}$ passes, the tangent line (or direction vector) of $\mathcal{C}$ is parallel with the mini tangent lines of the slope field. Loosely speaking, an integral curve $\mathcal{C}$ is a path whose direction is determined by the mini tangent lines of the slope field. Figure 2.2 shows three integral curves for the slope field shown in Figure 2.1.

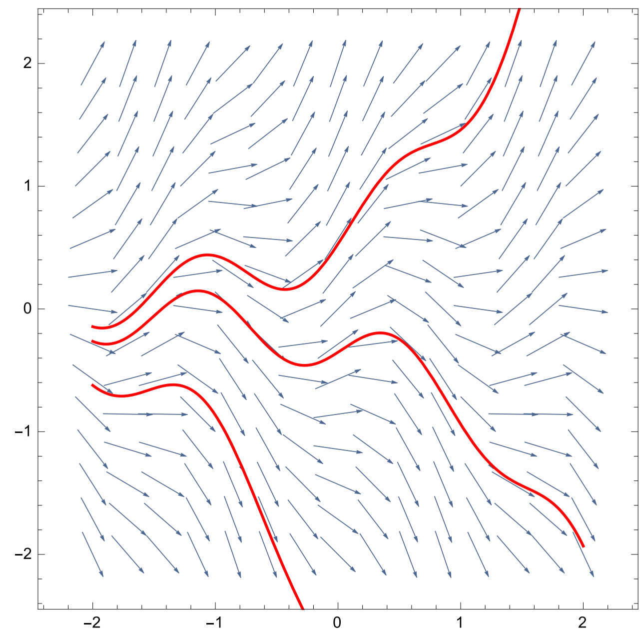


Figure 2.2: Three integral curves

Let's suppose now that y is a continuously differentiable function. As the derivative of y represents the tangent slope of the graph of $y = y(t)$ in the t-y plane, for y to be a solution to the differential equation $dy/dt = f(t, y)$, we want the graph of y to be parallel to the mini tangent lines of the slope field defined by $f(t, y)$ at each point (t, y). More precisely, we want the graph of y to be an integral curve for this slope field. Correspondingly, the three integral curves depicted in Figure 2.2 analogously depict three solutions to the differential equation $dy/dt = f(t, y)$. Let's summarize this observation:

The graph of a solution $y(t)$ to the differential equation

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

is an integral curve $\mathcal{C}$ for the slope field corresponding to $f(t, y)$.

To get some practice drawing slope fields and visualizing integral curves and solutions, it is illustrative to consider a few example.
Note here.

Example 10

Consider the differential equation

$$\frac{dy}{dt} = ty^2.$$

The slope field for $f(t,y) = ty^2$ is illustrated in Figure 2.3 for $-4 \leq t \leq 4$ and $-4 \leq y \leq 4$.

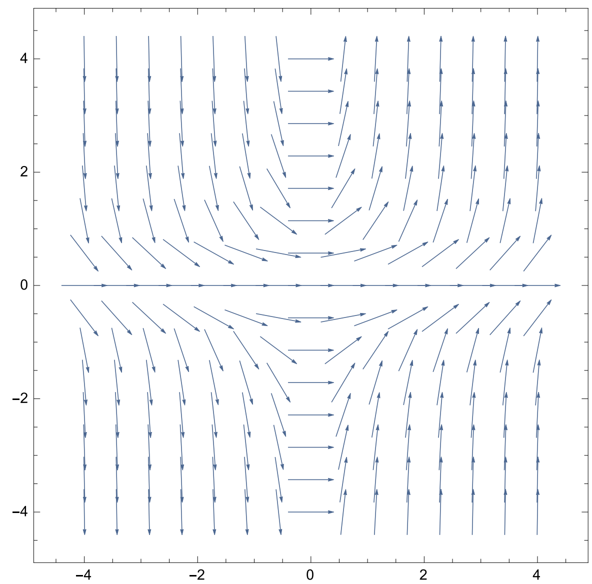


Figure 2.3: A slope field for $f(t,y) = ty^2$

In looking at the figure, we can get a good idea of what the integral curves should look like. Luckily, as this differential equation is separable, we can also find some solutions explicitly by the method of separation of variables. Let's do this.

We have

$$\int \frac{1}{y^2}\, dy = \int t\, dt$$

or, equivalently,

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

for some constant C. Solving this equation for $y = y(t)$, we obtain the family of solutions of the form

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

The method of separation of variables also misses a rather obvious solution: The constant solution $y(t) = 0$ for all t (which you should check directly is a solution). This solution is interesting for a couple of reasons. First, it cannot be found using the method of separation of variables – this gives credence to calling separation of variables a method and nothing more. Secondly, it is what's known as an equilibrium solution. Such solutions will be seen to be extremely important from a qualitative/macroscopic perspective. Figure 2.4, illustrates the equilibrium solution $y = 0$ along with the solutions

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

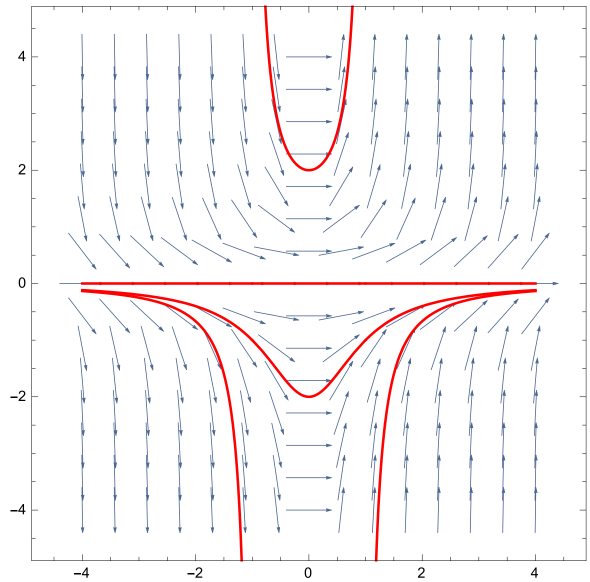


Figure 2.4: A slope field for $f(t, y) = ty^2$

In studying these solutions, it should be noted that, for all three solutions illustrated,

$$\lim_{t \to \infty} y(t) = 0.$$

In other words, the three solutions tend to the equilibrium solution $y = 0$ in as t becomes large; in fact, this is true for all solutions to this differential equation. This "drift toward equilibrium" is an incredibly important concept.

Example 11

Consider the differential equation

$$\frac{dy}{dt} = y - t.$$

The slope field for $f(t, y) = y - t$ is illustrated in Figure 2.5 for $-2 \leq t \leq 2$ and $-2 \leq y \leq 4$.

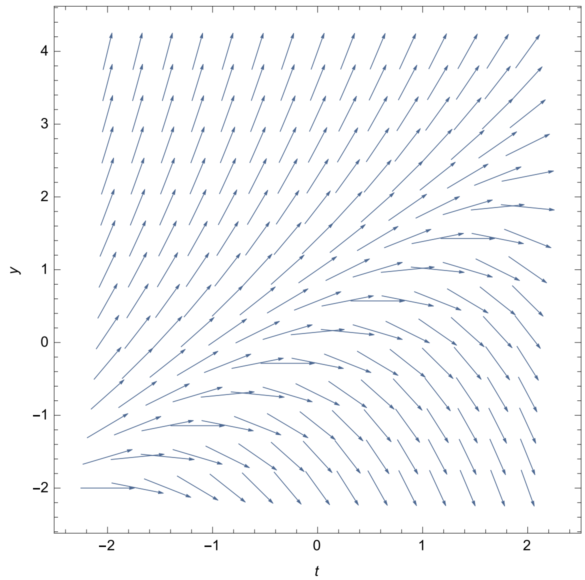


Figure 2.5: A slope field for $f(t, y) = y - t$

In looking at the figure, we can get a good picture of what we expect for our integral curves. As this differential equation is linear we can find all solutions, and hence integral curves, explicitly. Let's rewrite the equations as

$$\frac{dy}{dt} - y = -t,$$

and compute the integrating factor

$$\mu(t) = \exp\left(\int -1\,dt\right) = e^{-t}.$$

In view of Theorem 2.2.1, the general solution is

$$\begin{aligned} y(t) &= \frac{1}{e^{-t}} \int (-t)e^{-t}\,dt + \frac{C}{e^{-t}} \\ &= e^{-t}\left(te^{-t} + e^{-t}\right) + Ce^{t} \\ &= 1 + t + Ce^{t}. \end{aligned}$$

Figure (2.6) illustrates several solutions to this differential equation. It is interesting to note that most (all except that for which $C = 0$) diverge from the line $t \mapsto t+1$ as $t \to \infty$ and converge to this line as $t \to -\infty$. This is another example of "equilibrium behavior".

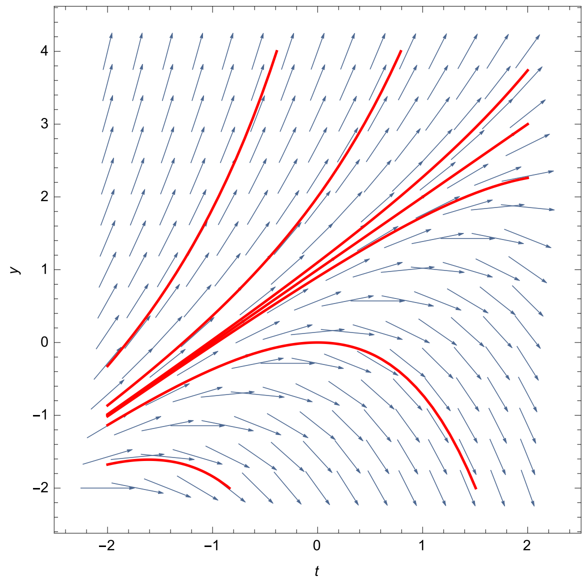


Figure 2.6: A slope field for $f(t, y) = y - t$ and some integral curves

Now, it's your turn.

Exercise 8

For each of the three differential equations below, do the following:

a. Sketch a slope field on the grid $-4 \leq t, y \leq 4$. Please draw the most accurate sketch you can and include at least 20 distinct mini tangent lines.

b. Sketch a number (≥ 4) of integral curves/graphs of solutions. Please do this on top of the slope field and use different colors.

c. Consider, in particular, the solution $y = y(t)$ whose graph passes through $(0, 1)$. Without explicitly solving the equation, describe the behavior of this solution as $t \to \infty$.

1.
$$\frac{dy}{dt} = y(4 - 2y)$$

2.
$$\frac{dy}{dt} = \sin(y)$$

3.
$$\frac{dy}{dt} = t - y$$

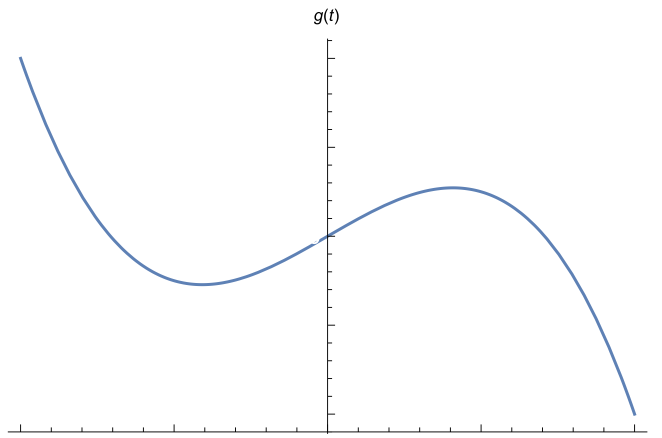
The following is a useful exercise. It will help you to think about how exactly slope fields are generated from the function $f(t, y)$. The exercise makes you think about the cases in which $f(t, y)$ is only a function of t or only a function of y. In doing this exercises, it is important to remember that the value of $f(t, y)$ determines the slope of the mini tangent lines in a slope field at (t, y).

Exercise 9

1. Consider a differential equation of the form

$$\frac{dy}{dt} = g(t)$$

where $g(t)$ is graphed in the following figure.

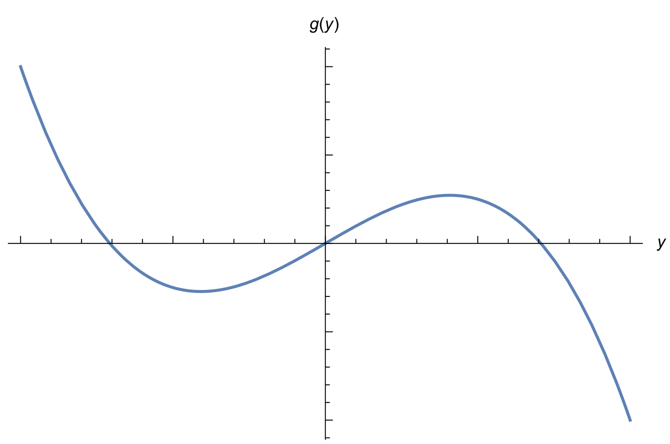


Sketch the slope field corresponding to this differential equation and illustrate some possible solutions. Please label the slope field and solutions in different colors.

2. Consider a differential equation of the form

$$\frac{dy}{dt} = g(y)$$

where $g(y)$ is graphed in the following figure.



Sketch the slope field corresponding to this differential equation and illustrate some possible solutions. Please label the slope field and solutions in different colors.

In the context of slope fields and integral curves, we are now ready to understand initial value problems graphically. Consider the differential equation

$$\frac{dy}{dt} = f(t,y) \tag{2.9}$$

and the associated initial value problem

$$\begin{cases} \dfrac{dy}{dt} = f(t,y) \\ \\ y(t_0) = y_0. \end{cases} \tag{2.10}$$

Recall, a solution to (2.10) is, by definition, a once continuously differentiable function $y = y(t)$ which solves the differential equation 2.9 and also satisfies the initial condition $y(t_0) = y_0$. In other words, the graph of a solution $y = y(t)$ to the initial value problem is an integral curve passing through the point (t_0, y_0). In this context, (t_0, y_0) is called an *initial point*. This is illustrated in the Figure 2.7.

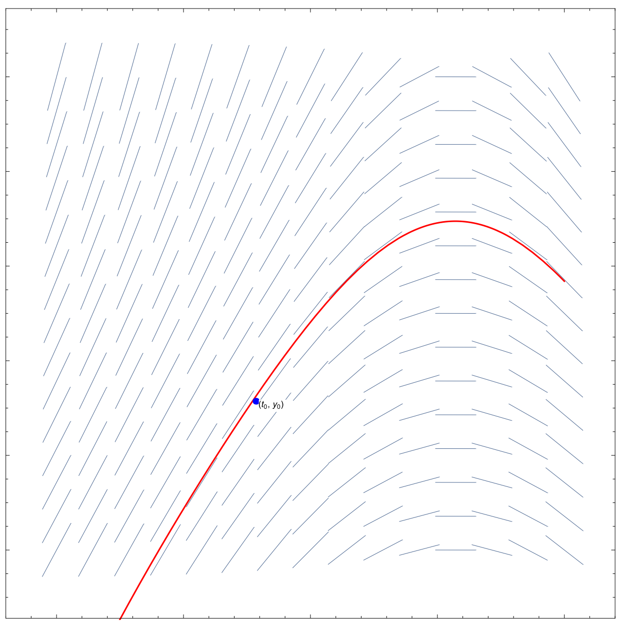


Figure 2.7: Integral curve through the initial point (t_0, y_0)

In view of the connection between solutions and integral curves, to solve an initial value problem of the form (2.10), one seeks an integral curve $\mathcal{C} = \mathcal{C}_{(t_0,y_0)}$ passing through the initial point (t_0, y_0); this integral curve will be, at least locally, the graph of a solution $y = y(t)$ to the initial value problem.

Exercise 10

In this exercise, we study the correspondence between integral curves passing through initial points and solutions to initial value problems. To this end, consider the differential equation

$$\frac{dy}{dt} = -\frac{t}{y}. \tag{2.11}$$

1. Carefully sketch the slope field for (2.11) on the domain $\mathcal{D} = \{(t, y) : -2 \leq t \leq 2, -2 \leq y \leq 2\}$. *Note: By convention, you should draw vertical mini tangent lines at points for which* $y = 0$. *Also, think about why this convention would make sense.*

2. In your slope field, sketch an integral curve $\mathcal{C} = \mathcal{C}_{(t_0,y_0)}$ through the initial point $(t_0, y_0) = (0, 1)$. *Hint: Your answer should be a closed curve passing through* $(1, 0)$, $(0, -1)$ *and* $(-1, 0)$.

3. In light of your answer to the previous question, is it possible for there to be a single[a] function $y = y(t)$ whose graph is (all of) $\mathcal{C}$? Why or Why not?

4. Can $\mathcal{C}$ be described by an algebraic equation of the form

$$p(y, t) = C$$

where p is a polynomial in the variables y and t and C is a constant? If so, what is this algebraic equation?

5. Use separation of variables to find a solution $y = y(t)$ to the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = -\dfrac{t}{y} & y(0) = 1 \end{cases}$$

and answer the following.

   (a) Does the graph of this solution also pass through $(0, -1)$? Why or why not?

   (b) Does the graph of your solution pass through $(1, 0)$? Further, discuss the behavior of $y(t)$ and $y'(t)$ as $t \nearrow 1$, i.e., as t approaches 1 from the left.

   (c) In what sense does your solution $y = y(t)$ agree or differ from your algebraic equation $p(y, t) = C$?

The moral of the example is this: While the graphs of solutions to first-order ordinary differential equations give rise to integral curves, they often only describe portions of more "complete" integral curves. Hence, not every integral curve can be gotten as the graph of a solution.

[a] In other words, does there exists a single-valued function?

2.3.1 Equilibrium Solutions

Through out our studies in this section, we have observed a number of examples of differential equations whose solutions exhibited "equilibrium behavior". To understand this behavior, it is useful to have the following vocabulary.

Definition 2.3.1 (Equilibrium Solution). *An equilibrium solution to the first-order differential equation*

$$\frac{dy}{dt} = f(t, y) \tag{2.12}$$

is, by definition, a solution which is identically constant. In other words, a solution $y(t)$ *to* (2.12) *is an equilibrium solution to* (2.12) *if* $y(t) = \gamma$ *for all* t *where* γ *is a constant.*

Example 12

Consider the differential equation

$$\frac{dy}{dt} = ty(1-y).$$

Observe that the constant function y defined by $y(t) = 1$ for all $t \in \mathbb{R}$ satisfies the differential equation because

$$\frac{dy}{dt} = \frac{d}{dt}(1) = 0 = t \cdot 1(1-1) = t \cdot y(t)(1-y(t))$$

for all $t \in \mathbb{R}$. Hence $y(t) = \gamma = 1$ is an equilibrium solution. By precisely the same argument, we have that the zero function, $y(t) = \gamma = 0$ for $t \in \mathbb{R}$, is also an equilibrium solution.

In general, for $y(t) = \gamma$ to be an equilibrium solution, one must always have that

$$0 = \frac{d}{dt}\gamma = \frac{dy}{dt} = f(t, y(t)) = f(t, \gamma)$$

for all t. Conversely, if γ is a real number for which $f(t, \gamma) = 0$ for all $t \in \mathbb{R}$, we can easily see that this number defines an equilibrium solution, y, given by $y(t) = \gamma$ for $t \in \mathbb{R}$. For this reason, the numbers γ which solve the equation $f(t, \gamma) = 0$ for all t are called *equilibrium values* for the differential equation (or for the function $f(t, y)$). We summarize these observations as follows.

Proposition 2.3.2. *For the differential equation*

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

the constant function $y(t) = \gamma$ is an equilibrium solution if and only if the number γ is an equilibrium value.

Of course, in looking at the previous example, the equilibrium solutions $y(t) = 0$ and $y(t) = 1$ corresponded to the equilibrium values 0 and 1 for the function $f(t, y) = ty(1-y)$. The take-away of the proposition above is that, to search for equilibrium solutions, it suffices to search for equilibrium values for the function $f(t, y)$. As we have seen in previous examples in this section, and is often the case, equilibrium solutions will often be seen to have the property that they attract or repel solutions nearby solutions in the $t \to \infty$ limit.

Exercise 11

Suppose that the constant function $y(t) = 1$ is a solution to the differential equation

$$\frac{dy}{dt} = f(t, y) \tag{2.13}$$

where $f(t, y)$ is a continuous function on $\mathbb{R}^2$.

1. Based only on this information, what can you say about the function $f(t, y)$?

2. What can you say about the slope field for (2.13)? How much of it can you sketch?

3. What can you say about solutions to the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = f(t,y) & y(0) = 1\,? \end{cases}$$

4. Can anything be said about solutions to the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = f(t,y) & y(0) = y_0 \end{cases}$$

where $y_0 \neq 1$?

Exercise 12

In this exercise, we revisit the differential equations

$$\frac{dy}{dt} = ty^2 \tag{2.14}$$

and

$$\frac{dy}{dt} = y - t \tag{2.15}$$

which we considered in our introduction of slope fields and integral curves.

1. Find all equilibrium values and equilibrium solutions to (2.14).
2. In looking back at our slope field analysis for (2.14), describe the long-time behavior of integral curves near the equilibrium solutions. Be as precise as possible. For instance, you could say "For each integral curve above/below the graph of the equilibrium solution with value y_0, the curve drifts toward/away from the equilibrium solution at $t \to \pm\infty$".
3. Are there equilibrium solutions for (2.15)? If so, describe them. If not, explain why none exist.
4. Regardless of your answer to the previous item, in Example 11 we discussed the nature of integral curves for (2.15) drifting toward the graph of the function $t \mapsto t + 1$. To see this as a equilibrium behavior, we set

$$y(t) = (t+1) + u(t)$$

where y and u are unknown functions. Find a differential equation[a] for u (which will be linear and homogeneous) with the property that u solves this differential equation if and only if y solves (2.15).

5. Given your differential equation in u found in the previous item, what are its equilibrium values/solutions? Explain roughly why this shows $y(t) = t + 1$ should be interpreted as a (quasi) equilibrium solution.

[a] In fact, transforming one differential equation into a simpler one is a handy technique. In this case, finding all solutions for the differential equation in u will yield all solutions for the differential equation (2.15). Do you see why? Would these solutions in y be consistent with those found in Example 11?

2.3.2 Autonomous equations and the classification of equilibria

In this subsection we turn our attention to a special class of differential equations called autonomous differential equations. A first-order ordinary differential equations is said to be autonomous if it is of the form

$$\frac{dy}{dt} = h(y) \tag{2.16}$$

where h is only a function of y. In other words, an autonomous first-order differential equation is one which does not explicitly depend on time. You should observe that autonomous first-order equations are separable. Though these equations are relatively easy to solve, their generalizations, autonomous systems, which we shall study later in these notes are complicated objects whose solution are elusive. For this reason, it's instructive to develop some machinery to understand qualitative behavior of solutions to autonomous equations – this machinery will extend into the realm of autonomous systems. Let's motivate this machinery by looking at a specific example.

Example 13

Consider the autonomous differential equation

$$\frac{dy}{dt} = y(y-1)(y+1)^2.$$

Figures 2.8 and 2.9 below illustrates the function $h(y) = y(y-1)(y+1)^2$ and the slope field for this differential equation[a].

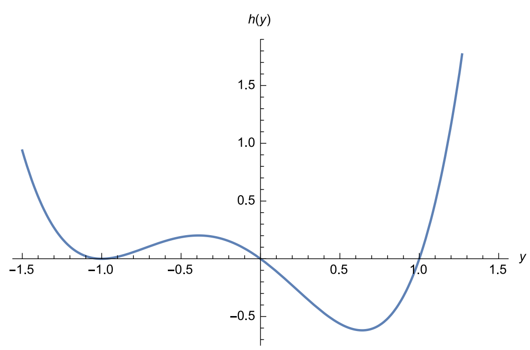


Figure 2.8: $h(y)$ vs y

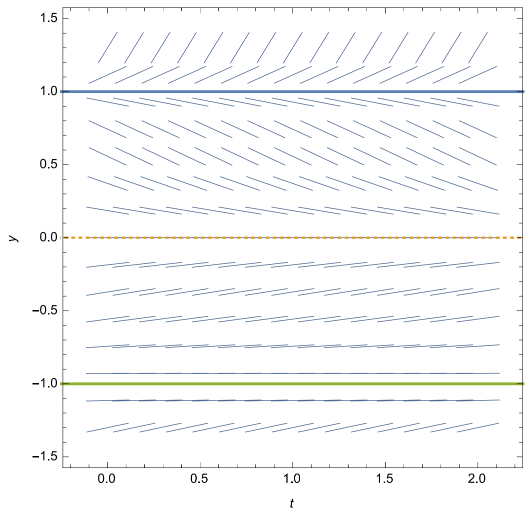


Figure 2.9: The slope field for $dy/dt = h(y)$

In studying the function $h(y) = y(y-1)(y+1)^2$, we immediately see that there are three equilibrium values, $y = -1, 0$ and 1. These equilibrium values can be seen by looking at Figure 2.8 and the corresponding equilibrium solutions are clearly seen in Figure 2.9. Taking for granted that there is one and only one solution to each initial value problem[b], we see immediately that, if $y_0 = -1, 0$ or 1, the solution $y = y(t)$ to the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = y(y-1)(y+1)^2 \\ \\ \qquad y(0) = y_0 \end{cases} \tag{2.17}$$

has the property that $\lim_{t\to\infty} y(t) = y_0$. This is, of course, a simplistic observation: any solution beginning at an equilibrium value is an equilibrium solution and will therefore take the same value for all time. It is therefore natural to ask: What is the large-time behavior of solutions whose initial values are close to (but not exactly) equilibrium values? In other words, what happens asymptotically to integral curves which come close to the equilibria? To address this question, let's consider the solution to (2.17) where y_0 takes values near (but not equal to) $-1, 0$ and 1. The figure below illustrates six solutions (only for $t \geq 0$) corresponding to $y_0 = -1.1, -0.9, -0.1, 0.1, 0.9$ and 1.1.

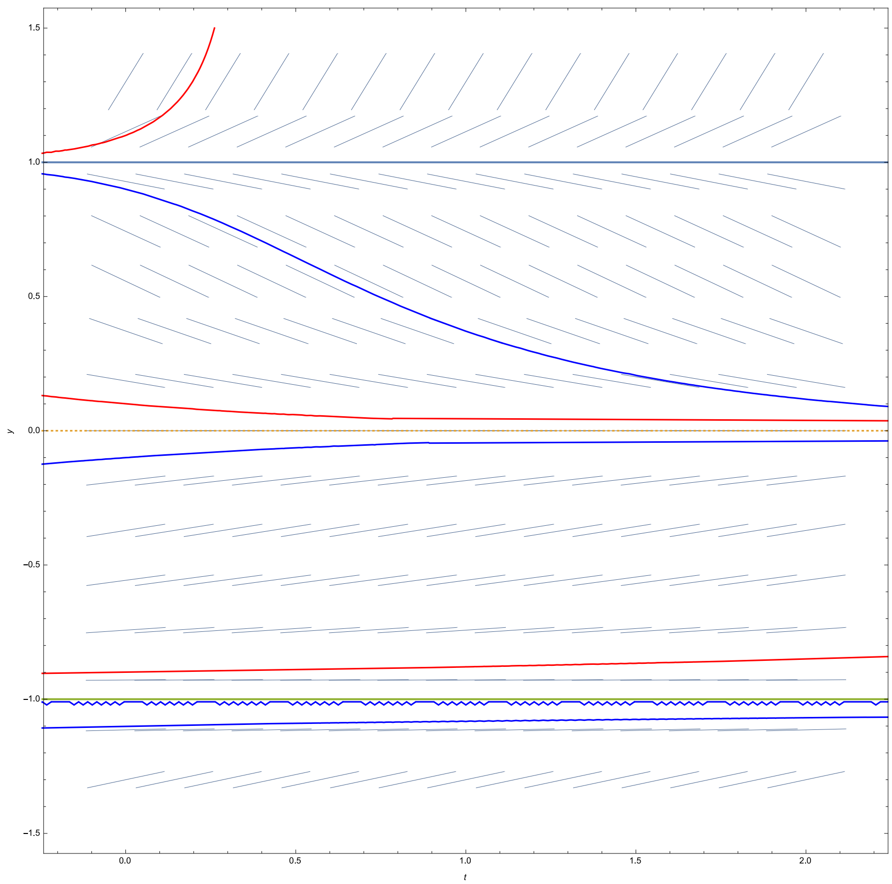


Figure 2.10: Six integral curves

Observe that, for the solutions $y(t)$ starting near the equilibrium at $y = 0$, i.e., those for which $y(0) = \pm 0.1$, we have $\lim_{t\to\infty} y(t) = 0$. In other words, the solutions close to the equilibrium value 0, converge on this equilibrium value in large time. By contrast, the solutions starting near the equilibrium value 1 drift away from this value in large time. In particular, for the solution with initial condition $y(0) = 1.1$, we have $\lim_{t\to\infty} y(t) = \infty$ and, for the solution with initial condition $y(0) = 0.9$, we have $\lim_{t\to\infty} y(t) = 0$. Finally, and stranger still, the solutions starting just slightly below $y = -1$ converge to -1 in large time while solutions starting just above -1, drift away from -1 in large time. In fact, the following is true: If $y(t)$ solves the initial value problem (2.17), then

$$\lim_{t\to\infty} y(t) = \begin{cases} \infty & y(0) > 1 \\ 1 & y(0) = 1 \\ 0 & -1 < y(0) < 1 \\ -1 & y(0) \le -1. \end{cases}$$

This equilibrium behavior can be easily seen by studying Figure 2.10.

[a]Observe that the slope field has the property that its mini tangent lines are parallel along any horizontal line – this is characteristic of autonomous equations.

[b]This is true in this case. We will have to tools to show it in the next section

The preceding example illustrates the vastly different character of equilibria. Studying the situation closely, in particular Figure 2.8 it is apparent the the sign (and change thereof) of $h(y)$ near equilibria was the determining factor in this long time behavior. This leads us to the following definition.

Definition 2.3.3. *Let y_0 be an equilibrium value for the autonomous differential equation*

$$\frac{dy}{dt} = h(y)$$

where h is a continuous function near y_0.

1. *If, on some neighborhood[3] of y_0, $h(y) > 0$ for $y < y_0$ and $h(y) < 0$ for $y > y_0$, we say that y_0 is a sink.*
2. *If, on some neighborhood of y_0, $h(y) < 0$ for $y < y_0$ and $h(y) > 0$ for $y > y_0$, we say that y_0 is a source.*
3. *If y_0 is neither a sink nor a source, we say y_0 is a node.*

Sinks are said to be stable equilibria, while sources and nodes are said to be unstable.

Though the above definition seems somewhat complicated, what it's saying is this: If, at an equilibrium value y_0 at which h has a downcrossing, i.e., $h(y)$ changes sign from positive to negative as y increases, then y_0 is a sink. If h has an upcrossing at y_0, i.e., $h(y)$ changes sign from negative to positive as y increases, then y_0 is a source. If it isn't a sink or a source, it's a node. In looking to the preceding example, we see that 0 is a sink, $y = 1$ is a source and $y = -1$ is a node.

Of course, if $h(y)$ is a differentiable function, the derivative of h at an equilibrium point y_0 can help used to indicate the way in which h's graph crosses the horizontal axis (provided it does cross it) and so it can help indicate if the equilibrium value is a source or sink. This is captured by the following proposition.

Proposition 2.3.4. *Let y_0 be an equilibrium value for the autonomous differential equation*

$$\frac{dy}{dt} = h(y).$$

and suppose that $h(y)$ is differentiable at y_0.

1. *If $h'(y_0) < 0$, then y_0 is a sink.*
2. *If $h'(y_0) > 0$ then y_0 is a source.*
3. *If $h'(y_0) = 0$, nothing can be said and further analysis is required.*

Example 14: Revisiting the previous example

For the differential equation (2.17), we found equilibrium values $-1, 0$ and 1. Let's apply the proposition above to see if we can classify these equilibria and confirm what we already understand. First, the function $h(y) = y(y-1)(y+1)^2$ is a polynomial and is therefore differentiable at the equilibrium values. We have

$$\begin{aligned} \frac{\partial h}{\partial y}(y) &= (y-1)(y+1)^2 + y(y+1)^2 + 2y(y-1)(y+1) \\ &= (y+1)(4y^2 - y - 1) \end{aligned}$$

[3] By a neighborhood of y_0, we mean an interval of the form $I = (y_0 - \delta, y_0 + \delta)$ for some positive number δ. Thus, by saying that y_0 is a sink, we mean there is some positive number δ for which $h(y) > 0$ whenever $y_0 - \delta < y < y_0$ and $h(y) < 0$ whenever $y_0 < y < y_0 + \delta$.

for $y \in \mathbb{R}$. At $y = 0$, we see that

$$h'(0) = (0+1)(4 \cdot 0^2 - 0 - 1) = -1 < 0$$

and so, in view of the preceding proposition, we confirm that 0 is a sink. At $y = 1$, we see that

$$h'(1) = (1+1)(4 \cdot 1^2 - 1 - 1) = 4 > 0$$

and so, as we already knew, 1 is a source. Finally, at $y = -1$, we see that

$$h'(-1) = (-1+1)(4 \cdot (-1)^2 - (-1) - 1) = 0$$

and so the above proposition isn't able to classify this equilibrium value. Of course, $y = -1$ is easily seen to be a node.

Example 15: A skydiver in freefall

During freefall, a skydiver who is accelerated toward earth by the gravity also encounters air resistance. Close to earth, the gravitational force can be modeled by $F_g = mg$ where m is the mass of the skydiver (in kg) and $g = 9.81\, m/s^2$. Knowing that air resistance opposes motion of the skydiver and is stronger at higher speeds, it is modeled by $F_a = -kv$ where v is the velocity of the skydiver and $k > 0$ is a constant depending on the shape of the skydiver[a]; k has units of kg/s. In principal, if the skydiver presents a large shape/cross section to the air in freefall, then the value of k will be large. Appealing the Newton's second law, we find

$$ma = m\dot{v} = F_g + F_a = mg - kv$$

and so we obtain

$$\dot{v} = g - (k/m)v$$

where v is the velocity of the skydiver and $\dot{v} = dv/dt$ is the acceleration. This equation is an autonomous first-order differential equation. Let's study the equilibrium behavior. We put

$$h(v) = g - (k/m)v = 0$$

and obtain a single equilibrium value, $v_T = v = mg/k$. Of course, h is differentiable at v_T and we find

$$h'(v_T) = 0 - (k/m) \cdot 1 = -k/m < 0$$

from which we conclude that v_0 is a sink in view or Proposition 2.3.4. Applying our techniques of qualitative analysis, we see that, regardless of the skydivers initial (vertical) velocity relative to Earth, the velocity of the skydiver will approach the value $v_T = mg/k$ in time, i.e.,

$$\lim_{t \to \infty} v(t) = v_T = mg/k.$$

For this reason, v_T is called the *terminal velocity* of the skydiver. It is dependent, in particular, on the skydiver's mass and shape and the gravitational constant g.

We note that the above differential equation can be solved explicitly, as it's a linear first-order inhomogeneous equation. It's easy to verify (and you should) that

$$v(t) = mg/k + (v_0 - mg/k)\, e^{-kt/m}$$

solves the initial value problem

$$\begin{cases}\dot{v} = g - (k/m)v, \quad v(0) = v_0\end{cases}$$

where v_0 is the skydiver's initial velocity. From this solution, it's easy to see that $\lim_{t\to\infty} v(t) = mg/k$ as our qualitative analysis predicted.

[a]Of course, k should also depend on the density of air, which is a function of altitude, and therefore cannot truly be constant provided the skydiver falls a significant distance.

Exercise 13

Given the following autonomous differential equations, do the following:

1. Find all equilibrium values and their corresponding equilibrium solutions.
2. For each equilibrium value, determine if it's a sink, source or node.
3. Give a rough sketch of the corresponding slope field and integral curves (enough to include the equilibria) and make comment to the behavior of solutions with initial values near the equilibrium values.

a.
$$\frac{dy}{dt} = y(1-y)$$

b.
$$\frac{dy}{dt} = 1 - y^2$$

c.
$$\frac{dy}{dt} = y^3$$

d.
$$\frac{dy}{dt} = \begin{cases} y(1-y) & y > 0 \\ \frac{1}{2}y & y \le 0. \end{cases}$$

Upon looking back at all of our work in this subsection, there seems to be one glaring omission (and you should see if you can find it before reading further). To pinpoint this omission, let's recap what we've done. In general, we spent a lot of effort studying autonomous first-order differential equations of the form (2.22) and their equilibrium values. In particular, we studied the behavior of the function $h(y)$ near these equilibrium values and gave names (sink, source, node) to them based on h's behavior. This study and our vocabulary for it were essentially motivated by our observations in one single example. In that example, we observed that solutions starting out near a sink happened to converge to that equilibrium value, i.e., solutions with initial values near 0 tended to 0 in large time. We also observed that solutions starting near a source (in that case $y = 1$) tended away from the source in large time. Beyond this example we'd like to infer that, in some generality, this large time behavior of solutions with initial values near equilibrium values is determined by the type (sink, source, node) of equilibrium[4]. Pushing beyond our intuition, the following proposition provides a partial result in that direction.

[4]For otherwise, it would have been a little silly to name them.

Proposition 2.3.5. *Let γ be a sink for the autonomous first-order equation*

$$\frac{dy}{dt} = h(y)$$

and let $\mathcal{O} = (\kappa, \rho)$ be an interval containing γ for which $h(y) > 0$ for all $\kappa < y < \gamma$ and $h(y) < 0$ for all $\gamma < y < \rho$. Given $y_0 \in \mathcal{O}$ and real numbers $\alpha < t_0$, suppose that $y = y(t) \in C^1(\alpha, \infty)$ (meaning y is once continuously differentiable on the interval (α, ∞)) is a solution to the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = h(y) & y(t_0) = y_0. \end{cases}$$

Then $y(t)$ converges to the equilibrium value γ as $t \to \infty$, i.e.,

$$\lim_{t\to\infty} y(t) = \gamma.$$

Before proving the proposition, it's helpful to first treat the following lemma[5].

Lemma 2.3.6. *Assume the hypotheses and notation of Proposition 2.3.5. If, for any $t_0 \in [t_0, \infty)$, $y(t_1) = \gamma$, then $y(t) = \gamma$ for all $t \geq t_1$ and so*

$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} \gamma = \gamma.$$

In other words, if the solution $y(t)$ coincides with the equilibrium (sink) value γ at any point, it remains there forever.

Proof. Let us assume, to reach a contradiction, that $y(t_1) = \gamma$ but $y(t_2) \neq \gamma$ for some $t_2 > t_1$. By the continuity of y, we may assume that t_2 is chosen so that $y(t_2) \in \mathcal{O}$. In the case that $\gamma = y(t_1) < y(t_2) < \rho$, set $\tau_1 = \max\{t \in [t_1, t_2] : y(t) = \gamma\}$, that is, τ_1 is the largest[6] element in the interval $[t_1, t_2]$ for which $y(t) = \gamma$. By construction (and the intermediate value theorem), it follows that $\gamma = y(\tau_1) < y(t)$ for all $t \in (\tau, t_2]$. Now, let $\tau_2 = \min\{t \in (\tau_1, t_2] : y(t) \geq y(t_2)\}$. By construction, the continuity of y, and the intermediate value theorem, it follows that $y(\tau_2) = y(t_2) < \rho$ and $\gamma = y(\tau_1) < y(t) < y(\tau_2) = y(t_2) < \rho$ for all $t \in [\tau_1, \tau_2]$. An appeal to the mean value theorem gives a number $\xi \in [\tau_1, \tau_2$ for which

$$y(\tau_2) - y(\tau_1) = y'(\xi)(\tau_2 - \tau_1).$$

Since $y(\tau_2) > y(\tau_1)$ we find that $y'(\xi) > 0$. However, since $\xi \in [\tau_1, \tau_2]$,$\gamma < y(\xi) < \rho$ and therefore $0 > h(y(\xi)) = y'(\xi)$, our desired contradiction. In the case that $\kappa < y(t_2) < y(t_1)$, an analogous argument also yields a contradiction. Consequently, no such t_2 exists and so $y(t) = \gamma$ for all $t \geq t_1$. □

Proof of Proposition 2.3.5. We shall consider the situation in which $\kappa < y_0 \leq \gamma$; the proof is similar when $\gamma < y_0 < \rho$. If $y_0 = y(t_0) = \gamma$, the desired result follows immediately by the preceding lemma. We therefore assume that $\kappa < y_0 < \gamma$. In this situation, we have that $y'(t_0) = h(y(t_0)) = h(y_0) > 0$ and, by the continuity of y, there must be an interval of the form $I = [t_0, \beta)$ on which $y(t)$ is strictly increasing. From this, there are two possibilities.

[5] A "lemma" in mathematics is a factual statement which is (usually) used in the proof of a theorem or proposition (which are themselves also factual statement).

[6] In fact, τ should be first defined as a supremum. The continuity of y then guarantees that it is indeed a maximum.

Possibility 1: $y(t)$ is a strictly increasing function for all $t > t_0$. In this case, we necessarily have $\kappa < y_0 = y(t_0) < y(t)$ for all t. If there is some t_1 for which $y(t_1) \geq \gamma$, then the intermediate value theorem ensures that $y(t)$ coincided with γ and so the desired result follows directly from the lemma. We therefore assume that $\kappa < y_0 = y(t_0) < y(t) < \gamma$ for all t. Given that y is strictly increasing, we have

$$\kappa < y_0 < \lim_{t\to\infty} y(t) =: \eta \leq \gamma,$$

in particular, this limit exists. Our goal now becomes to show that $\eta = \gamma$ and we will do this by appealing to the mean value theorem. Observe that, for each $t \in [t_0, \infty)$, the mean value theorem guarantees $c_t \in [t, t+1]$ for which

$$y(t+1) - y(t) = y'(c_t)(t+1-t) = y'(c_t) = h(y(c_t)).$$

Observe that, since $t < c_t < t+1$ for all $t <$, as $t \to \infty$, $c_t \to \infty$. Consequently,

$$0 = \eta - \eta = \lim_{t\to\infty}(y(t+1) - y(t)) = \lim_{t\to\infty} h(y(t_c)) = h(\lim_{t\to\infty} y(t_c)) = h(\eta)$$

in view of the continuity of h. Recalling that $\eta \in (\kappa, \gamma] \in \mathcal{O}$, it follows that $\eta = \gamma$ (for otherwise $h(\eta) > 0$) and so

$$\lim_{t\to\infty} y(t) = \gamma.$$

Possibility 2: $y(t)$ is not strictly increasing on $[t_0, \infty)$. If $y(t) \geq \gamma$ for some $t \geq t_0$, the solution $y(t)$ necessarily coincides with γ and so the result follows from the preceding lemma. It remains to rule out the case that $y(t) < \gamma$ for all $t > t_0$.

To this end, we assume, to reach a contradiction, that $y(t) < \gamma$ for all $t > t_0$. The assumption that y is not strictly increasing guarantees that there is a number $t_1 > t_0$ such that $y(t_1) \leq y_0 = y(t_0)$. Given that y is necessarily strictly increasing initially, i.e., strictly increasing on an interval of the form $[t_0, \beta)$, the inequality $y(t_1) \leq y(t_0)$ implies that y has a local maximum between t_0 and t_1; this follows from the mean value theorem. If $t \in [t_0, t_1]$ makes $y(t)$ a local maximum, then $y'(t) = 0$ and we set τ to be the minimum of such values, i.e., $\tau = \min\{t \in [t_1, t_2] : y'(t) = 0\}$. By the continuity of y', we have $h(y(\tau)) = y'(\tau) = 0$. Given that $y(t)$ is strictly increasing on $[t_0, \beta)$, it must be the case that $y_0 = y(t_0) < y(t) \leq y(\tau) < \gamma$ for all $t \in [t_0, \tau]$; any other arrangement would produce a local extremum between t_0 and τ thus contradicting the definition of τ. However, this is impossible for it would imply that $h(y(\tau)) > 0$ because $\kappa < y_0 < y(\tau) < \gamma$; this is our desired contradiction. $\square$

2.4 Existence and uniqueness: The Picard-Lindelöff theorem

In this section, we investigate the existence and uniqueness of differential equations and their corresponding initial value problems. Precisely, given a first-order ordinary differential equation

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

we ask the following two questions:

*Q*1. Given an initial value y_0 and time t_0, when does there exist a solution $y(t)$ (which is, at least, continuously differentiable near t_0) to the corresponding initial value problem

$$\begin{cases} \dfrac{dy}{dt} = f(t, y) \\ \\ y(t_0) = y_0? \end{cases}$$

$Q2$. If the answer to the previous question $Q1$ is "yes", when is $y(t)$ the only solution satisfying the initial condition $y(t_0) = y_0$?

From a graphical perspective (in terms of slope fields and integral curves), these two questions can be rephrased as follows:

$Q1_g$ Given an initial point (t_0, y_0) in the slope field defined by $f(t, y)$, when does there exist an integral curve for this slope field passing through the point (t_0, y_0)?

$Q2_g$ If there is an integral curve passing through this initial point (t_0, y_0) in the slope field for $f(t, y)$, when is this the only such integral curve?

As we shall see, these questions and their answers makeup a rather delicate and technical business. To sort things out satisfactorily, we will need to worry about rather technical aspects of functions (domains, domains/regions of continuity and differentiability, etc.). First, to illustrate why the answer to these questions isn't completely straightforward, we consider two examples.

Example 16: When $Q1$ goes wrong

Consider the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = \dfrac{y}{t} \\ \\ y(0) = 1. \end{cases} \tag{2.18}$$

We see that the function $f(t, y) = y/t$ blows up near the initial time $t_0 = 0$. Thus, for a solution $y(t)$ to the differential equation, its derivative must grow without bound as t nears 0 and we suspect that, if a function satisfying the initial condition $y(0) = 1$ exists, it has to be pretty wild, perhaps too wild to be differentiable. Further, in looking at the slope field for $f(t, y) = y/t$, Figure 2.11, we suspect that fitting an integral curve through $(0, 1)$ might be difficult.

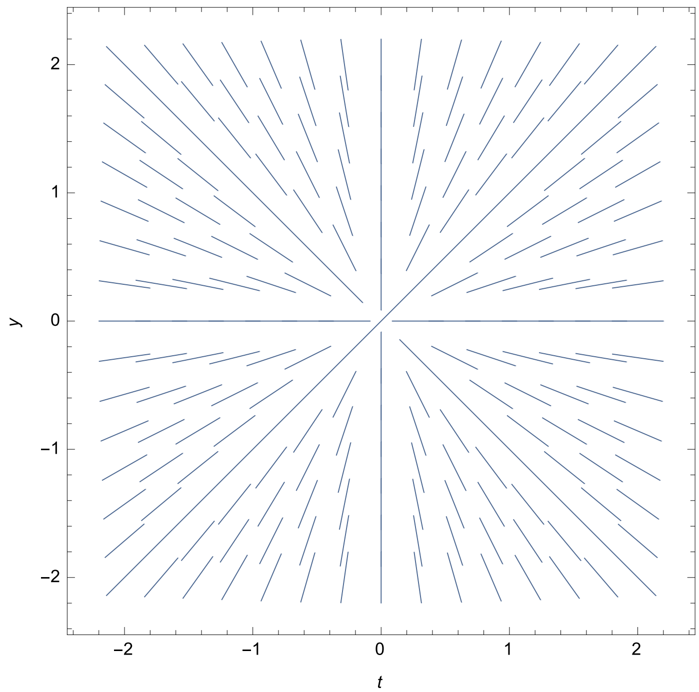


Figure 2.11: Slope Field for $f(t, y) = y/t$

As the following proposition shows, it's impossible.

Proposition 2.4.1. *There is no analytic solution to the initial value problem* (2.18).

Proof. Let's assume, to reach a contradiction that there exists an interval I containing $t_0 = 0$ and a function $y \in C^1(I)$ which satisfies the initial value problem (2.18). Correspondingly, for all non-zero $t \in I$, we have

$$ty'(t) = t\frac{dy}{dt}(t) = y(t)$$

Given that $y' = dy/dt$ and y are continuous functions on I (because $y \in C^1(I)$), we have

$$y(0) = \lim_{t \to 0} y(t) = \lim_{t \to 0} ty'(t) = 0 \cdot y'(0) = 0 \neq 1,$$

our desired contradiction. □

Example 17: When $Q2$ goes wrong

Consider the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = 3y^{2/3} \\ \\ y(0) = 0. \end{cases}$$

In seeking solutions to this initial value problem, we immediately observe that the equilibrium solution, $y(t) = 0$ for $t \in \mathbb{R}$ does the trick, i.e., it satisfies the differential equation for all time and also the initial condition $y(0) = 0$. So, in particular, $Q1$

has an affirmative answer. To investigate $Q2$, we ask: Are there any other solutions to this initial value problem? An application of separation of variables, in fact, gives one immediately: The function $\tilde{y}$ defined by $\tilde{y}(t) = t^3$ is also a solution and satisfies $\tilde{y}(0) = 0$. As $y(t) \neq \tilde{y}(t)$ for all t, these are distinct functions and hence distinct solutions to the initial value problem. These solutions are both illustrated in the Figure 2.12 below.

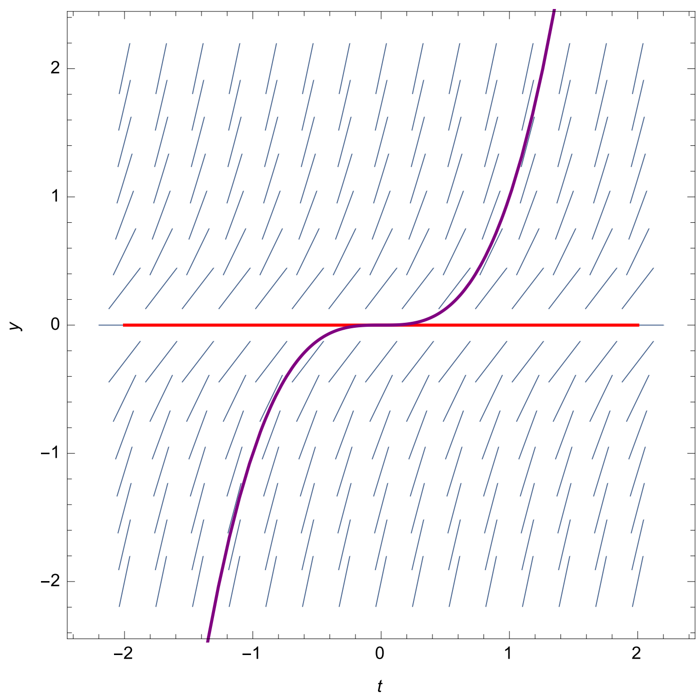


Figure 2.12: The solutions y and $\tilde{y}$

The two preceding examples should bother you; they certainly bother me. Suppose, for example, that the differential equation $dy/dt = 3y^{2/3}$ describes the position of a particle undergoing a force and starting at position 0 at time 0. Does the existence of multiple solutions say something about nature? Does the particle experience two laws of nature and two possible dynamical behaviors, both encoded in a single initial value problem? This would present us with a philosophical conundrum (perhaps paralleled by those of quantum mechanics). Mathematically, however, the problem with both differential equations is that the functions y/t and $3y^{2/3}$ are badly behaved near the given initial points (t_0, y_0); the first fails to be continuous and the second fails to be differentiable. The following theorem shows that, so as long as $f(t, y)$ isn't too badly behaved near the initial point (t_0, y_0), Questions $Q1$ and $Q2$ have affirmative answers.

Theorem 2.4.2 (Picard-Lindelöff)**.** *Let $f = f(t, y)$ be defined, continuous and have continuous partial derivative $\partial f/\partial y$ on the rectangle*

$$R = \{(t, y) : \alpha < t < \beta, \gamma < y < \kappa\} = (\alpha, \beta) \times (\gamma, \kappa);$$

here $\alpha, \beta, \gamma, \kappa$ are such that $\alpha < \beta$ and $\gamma < \kappa$. If $(t_0, y_0) \in R$, i.e., $\alpha < t_0 < \beta$ and $\gamma < y_0 < \kappa$, then the differential equation

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

has a unique solution $y(t)$ passing through the initial point (t_0, y_0). More precisely, there is one and only one function $y(t)$ satisfying the initial value problem

$$\begin{cases} \frac{d}{dt} y(t) = f(t, y(t)) & y(t_0) = y_0 \end{cases}$$

for all $t \in I = (t_0 - \delta, t_0 + \delta) \subseteq (\alpha, \beta)$; here, δ is some positive number and $y(t)$ is necessarily once-continuously differentiable[7] *on the interval I.*

In reference to Theorem 2.4.2, we say that f satisfies the theorem's hypotheses at the point (t_0, y_0) if there is a rectangle R containing (t_0, y_0) and on which f satisfies the hypotheses of the theorem, i.e., on this rectangle R, f is defined, continuous and has continuous partial derivative $\partial f/\partial y$.

Example 18

Returning to our first example of this section, i.e., that for which

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

We observe that, given any rectangle R containing the initial point $(0, 1)$ in its interior, the function $f(t, y) = y/t$ is necessarily discontinuous on this rectangle. Hence f does not satisfy the hypotheses of the Picard-Lindelöff theorem at $(0, 1)$. Of course, this is no surprise as we demonstrated that no solution to this initial value problem exists and hence the theorem's conclusion is false.

Pertaining to our second example, i.e., that for which

$$f(t, y) = 3y^{2/3},$$

we observe that f is continuous on $\mathbb{R}^2$ and hence it's continuous on every rectangle R containing the initial point $(0, 0)$. Computing the partial derivative of f with respect to y, we obtain

$$\frac{\partial f}{\partial y}(t, y) = \frac{2}{y^{1/3}}.$$

This function is clearly discontinuous on any rectangle R containing the initial point $(0, 0)$. Consequently, f does not satisfy the hypotheses of the Picard-Lindelöff theorem at $(0, 0)$ and hence the theorem cannot guarantee the existence of unique solutions for this initial point. As we observed, in fact, solutions to this corresponding initial value problem are not unique.

In both of the above cases, the initial points for which $y_0 = 0$ were problematic. Both functions however do satisfy the hypotheses of Theorem 2.4.2 at any initial point (t_0, y_0) for which $y_0 \neq 0$. Thus, in view of the theorem, there exists unique solutions to these differential equations through any such initial point.

In many situation we've studied so far (separable and linear first-order initial differential equations), we've been able to establish directly that solutions do exists for many initial values. As these are simply methods for producing solutions, and nothing more, they cannot be used to show if/when a given solution is unique. For this task, we must appeal to the Picard-Lindelöff theorem. Along these lines, we revisit Theorem 2.2.1 and complete

[7] This means that $y'(t)$ exists and is continuous on the interval $I = (t_0 - \delta, t_0 + \delta)$. In this case, we write $y \in C^1(I)$.

its proof. To remind you, we left off by establishing a general form solutions to first-order linear differential equations. It remains to show that, in fact, all solutions are of this form.

Proof. Consider the linear first-order differential equation

$$\frac{dy}{dt} + a(t)y = b(t) \tag{2.19}$$

where $a(t)$ and $b(t)$ are continuous on an interval $I = (\alpha, \beta)$. As we saw previously, for any constant C, the function

$$y(t) = \frac{1}{\mu(t)} \int \mu(t)b(t)\,dt + \frac{C}{\mu(t)}$$

is a solution to the differential equation (2.19) where $\mu(t) = e^{A(t)}$ is the associated integrating factor and A is an antiderivative of a. For simplicity, we shall write this as

$$y(t) = y_p(t) + \frac{C}{\mu(t)} \tag{2.20}$$

where

$$y_p(t) = \frac{1}{\mu(t)} \int \mu(t)b(t)\,dt$$

is a differentiable function on the interval I. As stated above, to complete the proof of the theorem, we must show that all solutions are of the form (2.20).

To this end, let's first observe that (2.19) can be rewritten in the form

$$\frac{dy}{dt} = f(t,y) = b(t) - a(t)y.$$

Given that $a(t)$ and $b(t)$ are continuous on the interval $I = (\alpha, \beta)$, we see the function $f(t,y)$ is continuous on the rectangle

$$R = \{(t,y) : \alpha < t < \beta, y \in \mathbb{R}\}.$$

Moreover,

$$\frac{\partial f}{\partial y}(t,y) = \frac{\partial}{\partial y}\left(b(t) - a(t)y\right) = -a(t)$$

is also continuous on R. By an appeal to the Picard-Lindelöff theorem, given any initial point $(t_0, y_0) \in R$, there exists a unique solution whose graph passes through (t_0, y_0). Equivalently, for any real number y_0 and time $\alpha < t_0 < \beta$, there exists a unique solution to the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = b(t) - a(t)y \\ \\ y(t_0) = y_0. \end{cases}$$

Let's use this fact to complete the proof. Let $\tilde{y}(t)$ be an arbitrary solution to (2.19). By definition, $\tilde{y}(t)$ must be a continuously differentiable function on some subinterval J of I. Taking any element $t_0 \in J$, we set $y_0 = \tilde{y}(t_0)$ and so trivially, this function $\tilde{y}$ satisfies the above initial value problem for this y_0. Observe that, by choosing $C = \mu(t_0)(y_0 - y_p(t_0)$ in (2.20),

$$y(t_0) = y_p(t_0) + \frac{C}{\mu(t_0)} = y_p(t_0) + \frac{\mu(t_0)(y_0 - y_p(t_0))}{\mu(t_0)} = y_0$$

and hence $y(t) = y_p(t) + C/\mu(t)$ for this particular C also satisfies this initial value problem. In view of the Picard-Lindeöff theorem, the solutions y and $\tilde{y}$ must be the same, i.e., our arbitrary solution $\tilde{y}$ is given by

$$\tilde{y}(t) = y_p(t) + \frac{C}{\mu(t)} = \frac{1}{\mu(t)} \int \mu(t) b(t)\, dt + \frac{C}{\mu(t)},$$

as desired. □

The above proof completes our general theory concerning first-order linear differential equations. We may now rightly call the given solutions *general solutions*. This theory can be generalized slightly thereby obtaining general solutions under less restrictive hypotheses on a and b; however, this is unnecessarily general for our purposes. We now widen our focus and study equations which are not linear nor separable and so, at least at present, we don't have analytic methods for finding solution. In this context, the Picard-Lindelöff theorem is still useful for it guarantees the existence and uniqueness of solutions without actually needing to find them. To this end, please do the following exercise.

Exercise 14

Use Theorem 2.4.2 to show that, given an arbitrary initial point (t_0, x_0) in the t-x plane, the initial value problem

$$\begin{cases} \dfrac{dx}{dt} = x^3 - t & x(t_0) = x_0 \end{cases}$$

has a unique solution[a].

[a] You should note that this equation is not separable nor linear and so, by the methods we've learned so far, we don't know how to come up with a solution. In fact, a Mathematica query does not yield a closed-form solution.

Exercise 15

This exercise will help to explain how $y(t)$, guaranteed by Theorem 2.4.2, is only defined on $(t_0 - \delta, t_0 + \delta)$ and not necessarily on all of $\mathbb{R}$. For the given initial value problems, do the following:

1. Solve the initial value problem.
2. Find the natural domain of your solution $y(t)$ and describe what happens to the solution $y(t)$ as t approaches the boundary/limits of its natural domain.
3. Explain why your observation in the above item is consistent with the conclusion of the Picard-Lindelöff theorem.

a.

$$\begin{cases} \dfrac{dy}{dt} = y^5 \\ y(0) = 1 \end{cases}$$

b.

$$\begin{cases} \dfrac{dy}{dt} = \dfrac{1}{(t+1)^2} \\ \\ y(0) = 1 \end{cases}$$

Remark 2.4.3. Let's make some final remarks about the Picard-Lindelöff Theorem and what it guarantees.

1. The Picard-Lindelöff theorem tells us when our search for a solution to a differential equation (or initial value problem) isn't fruitless.

2. The theorem puts our slope field analysis on a correct and rigourous footing. It tells us when there is one and only one integral curve passing though an initial point.

3. The theorem also gives credence to a new technique: guessing. That is, if in looking for a solution to an initial value problem you somehow stumble on a solution, the Picard-Lindelöff Theorem tells you (under certain conditions) that you can stop looking.

4. Though the Picard-Lindelöff theorem guarantees when a solution exists, it does not give a general method for finding it. In some cases, the theorem will ensure that a solution exists though you could never write it down.

5. In view of the theorem, we can start saying "the" solution instead of "a" solution when talking about initial value problems.

Exercise 16: Solution Curves

Given $f : \mathbb{R}^2 \to \mathbb{R}$, we consider the differential equation

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

and its corresponding slope field drawn in the t-y plane. For an ordered pair (t_0, y_0) in the t-y plane, solving the initial value problem

$$\begin{cases} \dfrac{dy}{dt} = f(t, y) \qquad y(t_0) = y_0 \end{cases} \tag{2.21}$$

is equivalent to fitting a "smooth" curve through the point (t_0, y_0) whose tangent lines match the slope field's minitangent lines at every point. Such a curve through the point (t_0, x_0) is called a *solution curve* and (t_0, y_0) is called an *initial point*. In this context, the solution curve through the initial point (t_0, y_0) is the graph of the solution $y = y(t)$ to the initial value problem (2.21).

1. Suppose that f satisfies the hypotheses of Theorem 2.4.2 at the point (t_0, y_0). How many solution curves can pass through the initial point (t_0, y_0)? Give a one-sentence explanation of your answer.

2. Suppose now that f satisfies the hypotheses of Theorem 2.4.2 at every point (t, y) in the t-y plane. Can any two-solution curves in the t-y plane intersect? Give a one-sentence explanation of your answer.

3. Let's now consider the particular first-order autonomous differential equation

$$\frac{dy}{dt} = 2\sqrt{|y|},$$

i.e., where $f(t, y) = 2\sqrt{|y|}$ for $(t, y) \in \mathbb{R}^2$. Does this f meet the hypotheses of Theorem 2.4.2 at every point in the t-y plane? Please explain your answer, i.e., show how each hypothesis is true or false.

4. Use separation of variables to solve the corresponding initial value problem

$$\begin{cases} \frac{dy}{dt} = 2\sqrt{|y|} & y(0) = 0. \end{cases}$$

If you're worried about the absolute value, you can assume that if $y(0) = 0$, $y(t) \geq 0$ for all t for which y is defined[a].

5. It is easy to check (and you should) that the zero function $y(t) = 0$ for all t is an equilibrium solution to the preceding initial value problem. In light of this and your work from the previous part, are solutions to this initial value problem unique? Does this contradict the statement of Theorem 2.4.2?

6. Draw the slope field for the differential equation

$$\frac{dy}{dt} = 2\sqrt{|y|}.$$

Does this slope field allow for intersecting solution curves? If so, draw them.

[a]It's also fun to think about starting at an initial value $y(0) = y_0 < 0$. Here, you'll have to be a little more careful about the absolute value in the integrand.

Exercise 17

(Is it possible to cross equilibrium solution curves?) Consider the differential equation

$$\frac{dy}{dt} = y(1 - y). \tag{2.22}$$

Note that 0 and 1 are equilibrium values for this equation and so this equation admits the equilibrium solutions defined[a] by $y^{(0)}(t) = 0$ and $y^{(1)}(t) = 1$ for all $t \in \mathbb{R}$.

a. Use Theorem 2.4.2 to show that, given any solution $y(t)$ to (2.22) such that $y(0) > 0$, we must have $y(t) > 0$ for all t for which the solution is defined. Hint: The desired inequality can also be written as $y(t) > y^{(0)}(t)$ for all t.

b. Is it true that, given any solution $y(t)$ to (2.22) such that $0 < y(0) < 1$, we have $0 < y(t) < 1$ for all t (for which the solution is defined)? Explain your answer.

c. Given $f : \mathbb{R}^2 \to \mathbb{R}$ satisfying the hypotheses of Theorem 2.4.2 (on any/every rectangle in the t-y plane) and the corresponding differential equation

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

what can be said about solution curves crossing (or the inability for them to cross) equilibrium values? Can you formulate a precise statement about this[b]?

[a] The superscripts are just to give tractable notation for the equilibrium solutions (they do not represent derivatives or anything like that).

[b] You are essentially being asked to generalize the results of Parts a and b. To this end, you should think about what is really going on concerning uniqueness.

2.5 Exact Equations

In this section, we focus on another method for (analytically) solving (certain) first-order differential equations. The differential equations we study here are called *exact equation* and, to develop the theory surrounding these equations and their solutions, it is useful to have a sturdy grasp on some multivariable calculus, especially the chain rule. The following formulation of the chain rule will be useful to us.

Theorem 2.5.1 (A version of the chain rule)**.** *Let $\mathcal{D} \subseteq \mathbb{R}^2$ be an open set and let $f : \mathcal{D} \to \mathbb{R}$ have continuous first-order partial derivatives at every point $(x, y) \in \mathcal{D}$. Also, let $I \subseteq \mathbb{R}$ be an open interval and let $x(t)$ and $y(t)$ be continuously differentiable functions on I, i.e., $x, y \in C^1(I)$, which satisfy $(x(t), y(t)) \in \mathcal{D}$ for every $t \in I$. Then the function $h : I \to \mathbb{R}$ defined by $h(t) = f(x(t), y(t))$ is continuously differentiable on I and*

$$h'(t) = \frac{\partial f}{\partial x}(x(t), y(t))x'(t) + \frac{\partial f}{\partial y}(x(t), y(t))y'(t) \tag{2.23}$$

for all $t \in I$.

In the statement of the above theorem, we have been precise about our hypotheses and explicit about where the partial derivatives of f are being evaluated. Often, the above formula for h' is written simply as

$$h' = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

where it is implied that $\partial f/\partial x$ and $\partial f/\partial y$ are evaluated at $(x(t), y(t))$ and $dx/dt = x'$, $dy/dt = y'$, and h' are evaluated at $t \in I$. Further, the hypotheses concerning f having continuous first-order partial derivatives is actually stronger than really needed – one only needs that f is differentiable in the sense of a two-variable function. For some background on differentiability of multivariable functions and a "uniform" perspective of the chain rule of which Theorem 2.5.1 is a special case, the reader is referred to Appendix B.

Exercise 18

Consider the function $f(x, y) = x^2 + y^2$ which is defined and has continuous first-order partial derivatives on all of $\mathcal{D} = \mathbb{R}^2$. For the following pairs $(x(t), y(t))$ of functions, use the chain rule to compute $h'(t)$ where $h(t) = f(x(t), y(t))$ for $t \in \mathbb{R}$.

1. $(x(t), y(t)) = (t^2, t)$
2. $(x(t), y(t)) = (e^t, e^{-t})$
3. $(x(t), y(t)) = (\cos(t), \sin(t))$

Also, by first computing h (and simplifying), explain why your answer for Item 3 makes sense.

Exercise 19

Let $\psi : \mathbb{R}^2 \to \mathbb{R}$ have continuous first-order partial derivatives everywhere (on the whole of $\mathbb{R}^2$) and let $y : \mathbb{R} \to \mathbb{R}$ be differentiable everywhere (on the whole of $\mathbb{R}$). We consider the function $h : \mathbb{R} \to \mathbb{R}$ defined by

$$h(x) = \psi(x, y(x)) \qquad \text{for } x \in \mathbb{R}.$$

1. Use the chain rule to show that

$$\frac{dh}{dx}(x) = \frac{\partial \psi}{\partial x}(x, y(x)) + \frac{\partial \psi}{\partial y}(x, y(x))\frac{dy}{dx}(x) \qquad \text{for } x \in \mathbb{R}.$$

2. Suppose that, for some constant C,

$$\psi(x, y(x)) = C \qquad \text{for all } x \in \mathbb{R}.$$

   Conclude that y satisfies the ordinary differential equation

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

   where

$$M(x, y) = \frac{\partial \psi}{\partial x} \qquad \text{and} \qquad N(x, y) = \frac{\partial \psi}{\partial y}.$$

Now that we have the chain rule at our fingertips, we are ready to introduce a new class of first-order differential equations which we are able to "solve". As is customary in the literature, we will denote our independent variable by x and the dependent variable by $y = y(x)$.

Definition 2.5.2. *Let $M(x, y)$ and $N(x, y)$ be continuous functions on some (open) subset $\mathcal{D}$ of $\mathbb{R}^2$ and consider the differential equation*

$$M(x, y) + N(x, y)\frac{dy}{dx} = 0. \tag{2.24}$$

This differential equation is said to be exact, if there exists a once-differentiable function $\psi : \mathcal{D} \to \mathbb{R}$ for which

$$M(x, y) = \psi_x(x, y) = \frac{\partial \psi}{\partial x}(x, y) \qquad N(x, y) = \psi_y(x, y) = \frac{\partial \psi}{\partial y}(x, y)$$

for $(x, y) \in \mathcal{D}$. Often, the function ψ is said to be a primitive of (2.24).

Given a primitive $\psi(x, y)$ of an exact equation (2.24), the constraint equation

$$\psi(x, y) = C$$

where C is a constant, forces (under mild conditions[8]) y to be a continuously differentiable function of x (at least locally). Correspondingly, $y(x)$ satisfies

$$\psi(x, y(x)) = C$$

[8] The condition that $\phi_y = N \neq 0$ suffices in view of the implicit function theorem (See [14]).

for all x for which y is continuously differentiable. On this domain, the chain rule Theorem 2.5.1 guarantees that

$$\begin{aligned} 0 = \frac{d}{dx}C = \frac{d}{dx}\psi(x, y(x)) &= \frac{\partial \psi}{\partial x}(x, y(x)) + \frac{\partial \psi}{\partial y}(x, y(x))y'(x) \\ &= M(x, y(x)) + N(x, y(x))\frac{dy}{dx}(x) \end{aligned}$$

and thus $y = y(x)$ solves the differential equation 2.24. Thus, to solve an exact equation with primitive ψ, one simply writes down the constraint equation

$$\psi(x, y) = C$$

and solves for y as a function of x.

Example 19

Consider the differential equation

$$(3x^2 + y^3) + 3xy^2\frac{dy}{dx} = 0. \tag{2.25}$$

Here $M(x, y) = 3x^2 + y^3$ and $N(x, y) = 3xy^2$; both are continuous on $\mathbb{R}^2$. To seek a primitive ψ, thereby determining if (2.25) is exact, we must have

$$\psi(x, y) = \int \psi_x \, dx = \int M(x, y) \, dx = \int (3x^2 + y^3) \, dx = x^3 + xy^3 + h_1(y)$$

where $h_1(y)$ is an unknown function of y. Here, computing the (partial) indefinite integral $\int M \, dx$ means to find a function ψ whose partial derivative in x is M; this is precisely why the function $h_1(y)$ appears as any such function would be "killed" by partial differentiation in x. Similarly

$$\psi(x, y) = \int \psi_y \, dy = \int N(x, y) \, dy = \int 3xy^2 \, dy = xy^3 + h_2(x)$$

where $h_2(x)$ is an unknown function of x. For the equation to be exact, these two computations must be consistent, i.e., we should be able to identify functions $h_1(y)$ and $h_2(x)$ for which

$$xy^3 + h_2(x) = \psi(x, y) = x^3 + xy^3 + h_1(y).$$

By inspection, a choice of $h_1(y) = 0$ and $h_2(x) = x^3$ does the job. Hence, the differential equation (2.25) and has primitive

$$\psi(x, y) = x^3 + xy^3$$

which is differentiable on $\mathbb{R}^2$. It's straightforward to check (as you should to double check your computations) that $\psi_x = M$ and $\psi_y = N$. Given any constant C, we then see that the constraint equation

$$C = \psi(x, y) = x^3 + xy^3$$

implicitly defines a solution $y = y(x)$ to (2.25) . When $x \neq 0$ (which rules out $C = 0$), we can solve for y as a function of x. This gives the solution

$$y(x) = y = \left(\frac{C - x^3}{x}\right)^{1/3} = \left(\frac{C}{x} - x^2\right)^{1/3}$$

defined for all $x \neq 0$. You should verify that this does indeed solve (2.25).

Remark 2.5.3. Given the constraint equation $\psi(x, y) = C$ to an exact equation, it's not always easy (if possible) to solve explicitly for y as a function of x without appealing to special functions of numerical methods. For example, given the primitive

$$\psi(x, y) = xe^y + y^2$$

to the exact differential equation

$$e^y + (xe^y + 2y)\frac{dy}{dx} = 0,$$

it isn't possible by hand to solve the equation

$$xe^y + y^2 = C$$

for y as a function of x (You should try it). In such cases, it is okay to leave the solution in the so-called implicit form,

$$\psi(x, y) = C.$$

Remark 2.5.4. Unsurprisingly, solving initial value problems for exact differential equations amounts to determining the constant C in the constraint equation $\psi(x, y) = C$. Whether or not the solution $y(x)$ is found explicitly or left in implicit form, it is always easiest to determine the constant C directly from the implicit form, i.e., before you solve for y. This is done by simply plugging in the initial point (x_0, y_0) corresponding to the initial condition $y(x_0) = y_0$ and hence

$$C = \psi(x_0, y_0).$$

Example 20

Consider the differential equation

$$(3x^2 + y^3) + 6xy^2\frac{dy}{dx} = 0.$$

Following the same procedure in the previous example, we seek a primative ψ by partial integration. We have

$$\int M(x, y)\, dx = \int (3x^2 + y^3)\, dx = x^3 + xy^3 + h_1(y)$$

and

$$\int N(x, y)\, dy = \int 6xy^2\, dy = 2xy^3 + h_2(x).$$

For a primitive to exist, these computations must be consistent and so we seek functions $h_1(y)$ and $h_2(x)$ for which

$$x^3 + xy^3 + h_1(y) = 2xy^3 + h_2(x).$$

Upon simplifying, this is the equation

$$x^3 - xy^3 = h_2(x) - h_1(y).$$

As the polynomial $x^3 - xy^3$ cannot be expressed as the difference of function $h_2(x)$ and $h_1(y)$ which only depend on x and y respectively, we conclude that no choice of h_1 and h_2 will work. Thus, no primitive exists and we conclude that the given equation is not exact.

In studying the preceding two examples, we see that it can be somewhat involved to determine whether or not a differential equation of the form (2.24) is exact. Fortunately, there is an easy condition to check. If (2.24) is exact with primitive ψ, by virtue of Clairaut's theorem (the equality of mixed partials), we have

$$\frac{\partial N}{\partial x} = \frac{\partial}{\partial x}\left(\frac{\partial \psi}{\partial y}\right) = \frac{\partial^2 \psi}{\partial x \partial y} = \frac{\partial^2 \psi}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial \psi}{\partial x}\right) = \frac{\partial M}{\partial y}.$$

provided that the partial derivatives $\partial N/\partial x$ and $\partial M/\partial y$ exist and are continuous. Thus, for the equation (2.24) to be exact, it is necessary to have

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}.$$

provided these partial derivatives exist and are continuous. Fortunately, under certain conditions[9] on the common domain of M and N, the above conditions is also sufficient. The following theorem captures a particularly simple case.

Theorem 2.5.5. *Let the function M, N, $\partial M/\partial y$ and $\partial N/\partial x$ be continuous on the rectangle*

$$R = \{(x,y) \in \mathbb{R}^2 : a < x < b, c < y < d\}$$

where a, b, c and d are such that $a < b$ and $c < d$ (we allow the possibility that $a = c = -\infty$ and $b = d = \infty$). Then (2.24) *is an exact equation (in R) if and only if*

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

at each point in R.

Let us check our results of the previous two examples against the theorem above. For the first example, we have

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}(3x^2 + y^3) = 3y^2 = \frac{\partial}{\partial x}(3xy^2) = \frac{\partial N}{\partial x}$$

for all $(x,y) \in \mathbb{R}^2$ and so, by virtue of the theorem, the corresponding differential equation is exact. Of course, this is consistent with our result as we found the equation to have primative $\psi(x,y) = x^3 + xy^3$. For the second example,

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}(3x^2 + y^3) = 3y^2 \neq 6y^2 = \frac{\partial}{\partial x}6xy^2 = \frac{\partial N}{\partial x}$$

and so the equation is not exact; this is the conclusion we found previously.

[9]This depends on a certain topological property of the domain of M and N which asks that the domain has "no holes". The technical term is called *simply connected*.

Exercise 20

Using Theorem 2.5.5, determine whether or not each of the following equations is exact or can be rewritten as an exact equation. If exact, find the solution[a] (which should involve an arbitrary constant C).

1. $$(2x+3)+(2y-2)\frac{dy}{dx}=0$$

2. $$(2x+4y)+(2x-2y)\frac{dy}{dx}=0$$

3. $$(e^x \sin y - 2y\sin x)dx + (e^x \cos y + 2\cos x)dy = 0$$

4. $$(9x^2+y-1)-(4y-x)\frac{dy}{dx}=0$$

5. $$\frac{dy}{dx}=\frac{-2xy}{1+x^2+y^2}$$

[a] If you can solve for y explicitly, please do it. If not, you can leave your solution in implicit form.

Exercise 21

Given continuously differentiable functions M and N, each mapping $\mathbb{R}$ into $\mathbb{R}$, use Theorem 2.5.5 to show that

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

is exact. Conclude that separable equations are also exact.

2.6 A Brief Look at Numerics: Euler's Method

In the course of this chapter, we have studied three methods for analytically producing solutions to certain types of ordinary differential equations, namely separable, linear and exact equations. There are many other interesting and powerful methods for producing solutions which we do not treat here. Generally, our prescription has been to classify a given differential equation as a certain type and then apply the known method adapted to solve equations of that type to produce a formula for the solution. Most differential equations, especially those modeling real-life situations, do not fit into the types for which analytic solution techniques are known. Often in practice, you will be given a differential equation and corresponding initial value problem which you can verify has a unique solution via the Picard-Lindelöff theorem but have no method for producing the solution. We ask: How can we gain information about a solution without being able to write it down? The answer is in approximation.

In this section, we discuss one method for numerically approximating solutions to initial value problems. The method we treat here, called Euler's method, is fantastically simple and is perhaps the oldest known method for approximating solutions. It should be noted that Euler's method is rarely used in practice as it is suboptimal and computationally expensive compared to more modern methods, such as Runge-Kutta. It is however an illustrative of the flavor of numerical methods.
Consider and initial value problem of the form

$$\begin{cases} \dfrac{dy}{dt} = f(t,y), \quad y(t_0) = y_0. \end{cases} \tag{2.26}$$

We assume that $f(t,y)$ satisfies the hypotheses of the Picard-Lindelöff theorem near the point (t_0, y_0) and therefore a unique solution $y = y(t)$ exists and is continuously differentiable on an interval of the form $(t_0 - \delta, t_0 + \delta)$. We shall assume additionally that y is twice differentiable on this interval, i.e., that the second derivative y'' exists at all points in $(t_0 - \delta, t_0 + \delta)$. In principle, we will know nothing additional about y. The goal of Euler's method is to produce a numerical scheme to approximate y on a subinterval of the form $I = [t_0, T] \subseteq (t_0 - \delta, t_0 + \delta)$. We want this numerical scheme to approximate $y(t)$ at a number of equally spaced times $t_0, t_1, \ldots, t_N = T$ in the interval I. To this end, we fix an integer N and set

$$t_k = t_{k-1} + h_N$$

for $k = 1, 2, 3, \ldots N$ where $h_N = (T - t_0)/N$ is the so-called time step associated with this partition of the interval I. Equivalently,

$$t_k = t_{k-1} + h_N = t_0 + kh_N = t_0 + k\left(\frac{T - t_0}{N}\right)$$

for $k = 1, 2, \ldots, N$. With these points fixed, our goal is approximate the numbers $y(t_0), y(t_1), \ldots, y(t_k), \ldots, y(t_N)$, i.e., the solution y evaluated at the times $t_0, t_1, \ldots, t_N$, by $N + 1$ numbers $y_0, y_1, \ldots, y_N$. Figure 2.13 illustrates this situation (with $N = 5$).

Obviously, we should choose $y_0 = y(t_0)$ as we know y passes through the initial point (t_0, y_0). To make a choice of y_1, we make an appeal to Taylor's theorem from calculus. As y is twice differentiable on an interval containing t_0 and $t_1 = t_0 + h_N$, Taylor's theorem guarantees that

$$y(t_1) = y(t_0 + h_N) = y(t_0) + y'(t_0)h_N + \frac{y''(\xi)}{2!}h_N^2 \tag{2.27}$$

for some $\xi \in (t_0, t_1)$. We note that this formula is exact and the existence of ξ is due to the mean value theorem. Now, $y(t_0) = y_0$ and, because y solves the differential equation on the interval $(t_0 - \delta, t_0 + \delta)$, we know that

$$y'(t_0) = f(t_0, y(t_0)) = f(t_0, y_0).$$

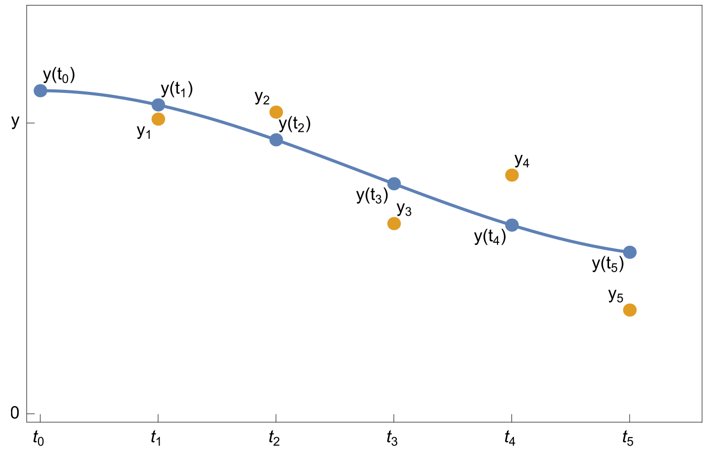


Figure 2.13: The function $y(t)$ and the discrete approximants $y_1, y_2, \ldots, y_5$.

With these observations in mind, (2.27) simplifies to

$$y(t_1) = y_0 + f(t_0, y_0)h_N + \frac{y''(\xi)}{2}h_N^2$$

In the case that $0 < h_N < 1$ is a small number, which can be produced by choosing N to be a large integer, h_N^2 even smaller and we would therefore expect[10] the term

$$\frac{y''(\xi)}{2}h_N^2$$

to be small compared to $y_0 + f(t_0, y_0)h_N$. We therefore set

$$y_1 = y_0 + f(t_0, y_0)h_N$$

to be our approximation for $y(t_1)$. Continuing in this manner we iteratively define

$$y_k = y_{k-1} + f(t_{k-1}, y_{k-1})h_N \tag{2.28}$$

for $k = 1, 2 \ldots, N$ thus producing our collection of so-called approximants $y_0, y_1, \ldots, y_N$ of the actual values of $y(t)$ at $t_0, t_1, \ldots, t_N$. Given our discussion above, we heuristically expect this approximation scheme to get better as $h_N \to 0$ or, equivalently, as $N \to \infty$. We summarize this approximation scheme as follows.

[10]At present, this is more of a “hope” than an expectation. We shall study this in detail in the next section.

Euler's method

To approximate the solution $y(t)$ to the initial value problem

$$\begin{cases} y' = f(t,y), \quad y(t_0) = y_0 \end{cases}$$

on the interval $I = [t_0, T]$, do the following:

1. Select an integer N.
2. Set $h_N = \frac{T-t_0}{N}$ and put
$$t_k = t_0 + kh_N$$
for $k = 1, 2, \ldots$.
3. Iteratively define
$$y_k = y_{k-1} + f(t_{k-1}, y_{k-1})h_N$$
for $k = 1, 2 \ldots, N$ where t_0 and y_0 are those given in the initial condition.

This scheme approximates the values $y(t_0), y(t_1), \ldots, y(t_N)$ by the numbers $y_0, y_1, \ldots, y_N$.

Example 21

To illustrate Euler's method, it's instructive to apply it to an initial value problem whose solution we know and understand well. To this end, we consider the initial value problem

$$\begin{cases} y' = y, \quad y(0) = 1. \end{cases}$$

Here, $t_0 = 0$, $y_0 = 1$ and $f(t,y) = y$. Of course, this is a linear homogeneous differential equation. A moment's thought shows that the unique solution is $y(t) = e^t$; this is the function to which we will compare our linear approximation.

Let's use Euler's method to approximate y on the interval $[0,1] = [t_0, T]$. For a natural number N, we have $h_N = 1/N$ and

$$t_k = t_0 + kh_N = \frac{k}{N}$$

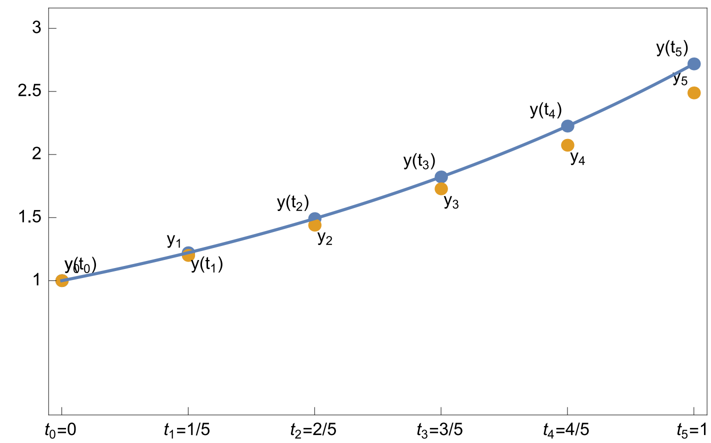
for $k = 1, 2, \ldots, N$. Following the approximation scheme, we have

$$y_k = y_{k-1} + f(t_{k-1}, y_{k-1})h_N = y_{k-1} + y_{k-1}h_N = \left(1 + \frac{1}{N}\right) y_{k-1}$$

for $k = 1, 2, \ldots, N$. Thus each y_k is gotten from multiplying the previous iterate y_{k-1} by $(1 + 1/N)$. From this we obtain the formula

$$y_k = \left(1 + \frac{1}{N}\right)^k y_0 = \left(1 + \frac{1}{N}\right)^k$$

for $k = 0, 1, \ldots, N$. This figure below illustrates this approximation for $N = 5$.



We observe, in this case, the approximants $y_1, y_2, \ldots, y_5$ are an underestimate for the solution $y(t) = e^t$ evaluated at the points $t_k = k/5$ for $k = 0, 1, \ldots, 5$. As we discussed, we expect this approximation to get better as $N \to \infty$, which is the case. In particular, because $t_N = N/N = 1$ for all N, we should expect

$$\lim_{N\to\infty} y_N = y(1) = e.$$

Of course, you might remember from calculus, this is the famous sequential approximation

$$e = \lim_{N\to\infty} \left(1 + \frac{1}{N}\right)^N.$$

Exercise 22: Euler's Method applied to Ricatti's Equation

Consider the initial value problem

$$\begin{cases} y' = t^2 - y^2 & y(0) = 1. \end{cases} \tag{2.29}$$

It is easy to verify that the initial value problem (2.29) has a unique solution $y(t)$. In this exercise, you use Euler's method to approximate this solution on the interval $I = [0, 1]$. To this end, do the following:

1. Verify that the differential equation $y' = t^2 - y^2$ is not separable, linear or exact. Use Wolfram Alpha, or any computational software you want, to find an analytic solution to this initial value problem; this should involve the so-called Bessel function of the first kind, a function defined via power series. You don't need to write down the solution, which is complicated, but please include a copy of the software's output.

2. Run Euler's method for $N = 4$ for (2.29) on the interval $I = [0, 1]$, by hand. Please list your answer in a table of the form:

t_0	t_1	t_2	t_3	t_4
y_0	y_1	y_2	y_3	y_4

3. On the course website, I have provided the Matlab m-file I used to run Euler's method for the initial value problem of the previous example. This m-file also runs one of Matlab's relatively sophisticated numerical approximation schemes, "ode45", for comparison. Adjust the m-file to the initial value problem (2.29) and run it for $N = 4, 10$ and 100. Print your adjusted m-file and the plots.

2.6.1 The Error in Euler's Method

For the initial step of Euler's method, we approximated $y(t_1)$ by

$$y_1 = y(t_0) + y'(t_0)(t_1 - t_0) = y_0 + f(t_0, y_0)h_N$$

where $h_N = t_1 - t_0 = (T - t_0)/N$. Looking carefully, we see that this is nothing more than a linear approximation of y centered at the initial point (t_0, y_0) and the approximant y_1 is simply the value gotten by evaluating this linear approximation at time t_1. The absolute error between y_1 and $y(t_1)$ is then simply a measure of how much this linear approximation differs from the function at t_1. As we discussed, given that y is twice differentiable[11], the mean value theorem guarantees a number ξ_1 between t_0 and t_1 for which

$$y_1 = y(t_1) = y(t_0) + f(t_0, y_0)h_N + \frac{y''(\xi_1)}{2}h_N^2 = y_1 + \frac{y''(\xi_1)}{2}h_N^2$$

and so the absolute error between $y(t_1)$ and y_1 is

$$\mathcal{E}_1 = |y(t_1) - y_1| = \frac{|y''(\xi)|h_N^2}{2}.$$

Provided that we can get some information concerning y'', this error can be made arbitrarily small by taking N sufficiently large.

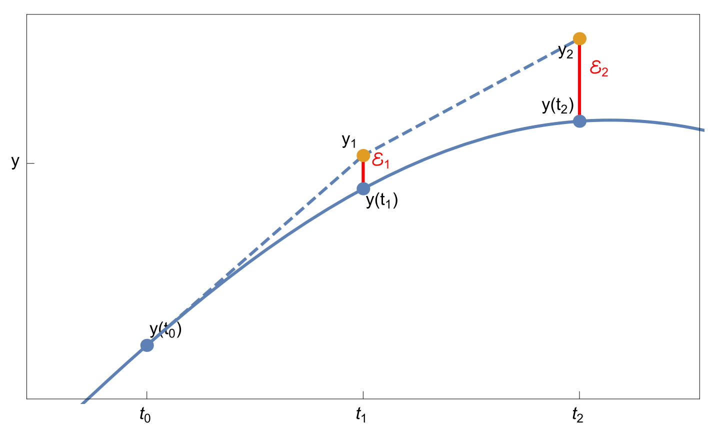


Figure 2.14: Aggregating Error in Euler's Method

For the second iteration, getting a handle on the error $\mathcal{E}_2 = |y(t_2) - y_2|$ is somewhat more tricky. We recall

$$y_2 = y_1 + f(t_1, y_1)h_N$$

which is, like y_1, gotten by a linear approximation. In contrast to the first iteration, the initial point used in this linear approximation is (t_1, y_1) which, unlike (t_0, y_0), does not

[11] See Exercise 24

(usually) sit on the graph of y. Of course, since we don't know the value of $y(t_1)$ exactly, we can't center our approximation at $(t_1, y(t_1))$ and the best (using Euler's method) we can do is center our approximation at the preceding approximant (t_1, y_1). This is illustrated in Figure 2.14. For this reason, Euler's method aggregates error: the absolute error for each iteration $\mathcal{E}_k$ depends not only on the error produced by the linearization but also on the previous absolute errors $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_{k-1}$. The following lemma helps us to quantify this error.

Lemma 2.6.1. *Consider the initial value problem*

$$\begin{cases} \frac{dy}{dt} = f(t,y) & y(t_0) = y_0 \end{cases}$$

where we shall assume that f, $\partial f/\partial y$ and $\partial f/\partial t$ are bounded and continuous functions on the whole of $\mathbb{R}^2$. Let M_1 and M_2 be positive constants for which

$$\left|\frac{\partial f}{\partial y}(t,y)\right| \leq M_1 \qquad \left|\frac{\partial f}{\partial t}(t,y) + \left(\frac{\partial f}{\partial y}(t,y)\right) f(t,y)\right| \leq M_2$$

for all $(t,y) \in \mathbb{R}^2$. Following Euler's method, let $T > t_0$, N be a natural number and set $h_N = (T - t_0)/N$. Taking t_0 and y_0 as given by the initial value problem, for each $k = 1, 2, \ldots, N$, define

$$t_k = t_0 + kh_N,$$

and

$$y_k = y_{k-1} + f(t_{k-1}, y_{t_{k-1}})h_N.$$

Then, for each $k = 1, 2, \ldots, N$, the absolute error $\mathcal{E}_k = |y(t_k) - y_k|$ between the solution $y(t)$ at time t_k and its approximation y_k satisfies

$$\mathcal{E}_k \leq (1 + M_1 h_N)\mathcal{E}_{k-1} + \frac{M_2 h_N^2}{2} \tag{2.30}$$

where $\mathcal{E}_0 = |y(t_0) - y_0| = 0$.

The inequality (2.30) essentially says that, at worst, the error in the kth approximation depends on the error of the $(k-1)$th approximation, the step size h_N, and certain bounds concerning f and its first-order partial derivatives. Before giving the proof, which you are strongly encouraged to read, it should be noted that the hypotheses of the result (concerning continuity and boundedness on all of $\mathbb{R}^2$) are overly restrictive and can be weakened significantly. The result is stated as is to give you the essential idea of what's going on while avoiding getting bogged down in non-essential details.

Proof. In view of the Picard-Lindelöff theorem, let y be the unique solution to the initial value problem. This function is necessarily once continuously differentiable on an interval[12] of the form $I = (t_0 - \delta, t_0 + \delta)$ where $\delta > 0$ and we shall assume that $[t_0, T] \subseteq [t_0, t_0 + \delta)$. By virtue of Exercise 24 below, $y(t)$ is twice continuously differentiable on I and, for all such t,

$$y''(t) = \frac{\partial f}{\partial t}(t, y(t)) + \frac{\partial f}{\partial y}(t, y(t))f(t, y(t)). \tag{2.31}$$

The case for $k = 1$ is slightly easier than the others and so we will treat it separately: As discussed previously, the mean value theorem guarantees a number ξ_1 between t_0 and

[12] In fact, the lemma's hypotheses actually guarantee (by a stronger version of the Picard-Lindelöff theorem) that the solution exists and is continuously differentiable on all of $\mathbb{R}$.

t_1 for which

$$\begin{aligned} y(t_1) &= y(t_0) + y'(t_0)(t_1 - t_0) + \frac{y''(\xi_1)(t_1 - t_0)^2}{2} \\ &= y_0 + f(t_0, y_0)h_N + \frac{y''(\xi_1)h_N^2}{2} \\ &= y_1 + \frac{y''(\xi_1)h_N^2}{2} \end{aligned}$$

where we have used the fact that $y(t_0) = y_0$ and $y'(t_0) = f(t_0, y(t_0)) = f(t_0, y_0)$. By virtue of (2.31), we have

$$\mathcal{E}_1 = |y(t_1) - y_1| = |y''(\xi_1)|\frac{h_N^2}{2} = \left|\frac{\partial f}{\partial t}(\xi_1, y(\xi_1)) + \frac{\partial f}{\partial y}(\xi_1, y(\xi_1))f(\xi_1, y(\xi_1))\right|\frac{h_N^2}{2}.$$

In view of the lemmas hypotheses, we have

$$\left|\frac{\partial f}{\partial t}(\xi_1, y(\xi_1)) + \frac{\partial f}{\partial y}(\xi_1, y(\xi_1))f(\xi_1, y(\xi_1))\right| \leq M_2$$

and therefore

$$\mathcal{E}_1 \leq \frac{M_2 h_N^2}{2} = \frac{M_1^2 h_N^2}{2}\mathcal{E}_0 + \frac{M_2 h_N^2}{2}.$$

Let us now focus on $\mathcal{E}_k$ where $k = 2, 3, \ldots, N$. Upon fixing one such k, the mean value theorem, there is ξ_k between t_{k-1} and t_k for which

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

and so

$$\begin{aligned} y(t_k) - y_k &= y(t_{k-1}) + f(t_{k-1}, y(t_{k-1}))h_N + y''(\xi_k)\frac{h_N^2}{2} - (y_{k-1} + f(t_{k-1}, y_{k-1})h_N) \\ &= (y(t_{k-1}) - y_{k-1}) + (f(t_{k-1}, y(t_{k-1})) - f(t_{k-1}, y_{k-1}))h_N + y''(\xi_k)\frac{h_N^2}{2} \end{aligned}$$

By applying the mean value theorem to the function $y \mapsto f(t_{k-1}, y)$, the mean value theorem guarantees a $\tilde{y}$ between $y(t_{k-1})$ and y_{k-1} for which

$$f(t_{k-1}, y(t_{k-1})) - f(t_{k-1}, y_{k-1}) = \frac{\partial f}{\partial y}(t_{k-1}, \tilde{y})(y(t_{k-1}) - y_{k-1}).$$

Combining these two identities yields

$$\begin{aligned} y(t_k) - y_k &= (y(t_{k-1}) - y_{k-1}) + \frac{\partial f}{\partial y}(t_{k-1}, \tilde{y})(y(t_{k-1}) - y_{k-1})h_N + y''(\xi_k)\frac{h_N^2}{2} \\ &= (y(t_{k-1}) - y_{k-1})\left(1 + \frac{\partial f}{\partial y}(t_{k-1}, \tilde{y})h_N\right) + y''(\xi_k)\frac{h_N^2}{2} \end{aligned}$$

and so

$$\mathcal{E}_k = |y(t_k) - y_k| = |y(t_{k-1}) - y_k|\left|1 + \frac{\partial f}{\partial y}(t_{k-1}, \tilde{y})h_N\right| + \frac{|y''(\xi_k)|h_N^2}{2}.$$

By an appeal to (2.31) the lemma's hypotheses,

$$\left|1 + \frac{\partial f}{\partial y}(t_{k-1}, \tilde{y})h_N\right| \leq 1 + \left|\frac{\partial f}{\partial y}(t_{k-1}, \tilde{y})\right| h_N \leq 1 + M_1 h_N$$

and

$$|y''(\xi_k)| = \left|\frac{\partial f}{\partial t}(\xi_k, y(\xi_k)) + \frac{\partial f}{\partial y}(\xi_k, y(\xi_k))f(\xi_k, y(\xi_k))\right| \leq M_2.$$

Consequently,

$$\mathcal{E}_k \leq |y(t_{k-1}) - y_{k-1}|(1 + M_1 h_N) + \frac{M_2 h_N^2}{2} = \mathcal{E}_{k-1}(1 + M_1 h_N) + \frac{M_2 h_N^2}{2},$$

as desired. □

Exercise 23: Regularity by the bootstrap

Consider the differential equation

$$\frac{dy}{dt} = f(t, y).$$

where we assume that, given an open interval $J = (\alpha, \beta)$, the function $f(t, y)$ is continuous on the rectangle $R := J \times \mathbb{R} = \{(t, y) \in \mathbb{R}^2 : t \in J \text{ and } y \in \mathbb{R}\}$. Given a subinterval $I = (a, b) \in J$, assume that $y : I \to \mathbb{R}$ is a function which is differentiable for all $t \in I$ and satisfies

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

for all $t \in I$.

1. Show that $y'(t)$ is necessarily a continuous function on the interval I. In this case we say that y is once continuously differentiable on I and write $y \in C^1(I)$. Conclude that y is a *bona fide* (analytic) solution to the given differential equation on I.

2. Assume additionally that $\partial f/\partial t$ and $\partial f/\partial y$ exist and are continuous function on the rectangle R. Use the chain rule to show that

$$y''(t) = \frac{\partial f}{\partial t}(t, y(t)) + \frac{\partial f}{\partial y}(t, y(t))f(t, y(t))$$

   for $t \in I$. Conclude that y is twice differentiable on I and $y''(t)$ is continuous on I. In this case, we say that y is *twice continuously differentiable on* I and write $y \in C^2(I)$.

3. Assume now that f has (both pure and mixed) partial derivatives of all orders which are continuous on R. Conclude that y can be differentiated *ad infinitum* (as many times as you wish) and all of its derivatives are continuous on I. In this case, y is said to be *smooth* on I and we write $y \in C^\infty(I)$.

As it turns out, the previous lemma helps us to provide a uniform estimate on the errors $\mathcal{E}_k$. Here, we use the word uniform to mean that the upper bound doesn't depend on k, but only on f, $T - t_0$ and N. Such an estimate is extremely useful because it allows us to tell exactly how large N should be to get the global error, i.e., the maximum of the $\mathcal{E}_k$'s, within any desired tolerance. For example, if you only need to estimate your solution $y(t)$ within a desired global tolerance on a given interval, you can figure out, *a priori*, how many computations need to be done to obtain this desired error. Provided you're using computational software to approximate the solution, this means that you will know ahead of time how long the computer will need to run to get within this tolerance. We formulate the precise result in the following theorem. As is true for the lemma, the hypotheses of the theorem can be significantly weakened and the result of the theorem still holds; for this, we refer the reader to [8].

Theorem 2.6.2 (The Error of Euler's Method)**.** *Consider the initial value problem*

$$\begin{cases} \frac{dy}{dt} = f(t,y) \quad y(t_0) = y_0 \end{cases}$$

where we shall assume that f, $\partial f/\partial y$ and $\partial f/\partial t$ are bounded and continuous functions on the whole of $\mathbb{R}^2$. Let M_1 and M_2 be positive constants for which

$$\left|\frac{\partial f}{\partial y}(t,y)\right| \leq M_1 \qquad \left|\frac{\partial f}{\partial t}(t,y) + \left(\frac{\partial f}{\partial y}(t,y)\right) f(t,y)\right| \leq M_2$$

for all $(t,y) \in \mathbb{R}^2$. Following Euler's method, let $T > t_0$, N be a natural number and set $h_N = (T - t_0)/N$. Taking t_0 and y_0 as given by the initial value problem, for each $k = 1, 2, \ldots, N$, define

$$t_k = t_0 + kh_N,$$

and

$$y_k = y_{k-1} + f(t_{k-1}, y_{t_{k-1}})h_N.$$

Then the absolute error $\mathcal{E}_k = |y(t_k) - y_k|$ satisfies

$$\mathcal{E}_k \leq \frac{M_2 h_N}{2M_1}\left(e^{M_1(T-t_0)} - 1\right) \tag{2.32}$$

for all $k = 0, 1, 2 \ldots, N$.

Proof. In view of the lemma, we have

$$\mathcal{E}_k \leq A\mathcal{E}_{k-1} + B$$

for each $k = 1, 2, \ldots, N$ where $A = 1 + M_1 h_N > 1$ and $B = M_2 h_N^2/2 > 0$. For $k \geq 2$, combining the inequalities for k and $k-1$, we easily obtain

$$\mathcal{E}_k \leq A(A\mathcal{E}_{k-2} + B) + B = A^2\mathcal{E}_{k-2} + AB + B.$$

Continuing in this manner (or by rigorously using mathematical induction), it follows that

$$\mathcal{E}_k \leq A^k\mathcal{E}_0 + A^{k-1}B + A^{k-2}B + \cdots AB + B$$

and since $\mathcal{E}_0 = 0$,

$$\mathcal{E}_k \leq B(A^{k-1} + A^{k-2} + \cdots + A + 1)$$

which holds for $k \geq 1$. For each such k, note that

$$(A-1)(A^{k-1}+A^{k-2}+\cdots+A+1) = A^k+A^{k-1}-A^{k-1}+A^{k-2}-A^{k-2}+\cdots+A-A-1) = A^k-1$$

and because $A > 1$, we have

$$\frac{A^k - 1}{A - 1} = (A^{k-1} + A^{k-2} + \cdots + A + 1)$$

and therefore

$$\mathcal{E}_k \leq B\frac{A^k - 1}{A - 1} = \frac{B}{A-1}(A^k - 1) = \frac{M_2 h_N^2/2}{M_1 h_N}((1 + M_1 h_N)^k - 1). \tag{2.33}$$

Given that $M_1 h_N > 0$, observe that

$$\begin{aligned} e^{M_1 h_N} &= \frac{(M_1 h_N)^0}{0!} + \frac{(M_1 h_N)^1}{1!} + \frac{(M_1 h_N)^2}{2!} + \frac{(M_1 h_N)^3}{3!} + \cdots \\ &= 1 + M_1 h_N + \text{"positive stuff"} \end{aligned}$$

and therefore

$$(1 + M_1 h_N)^k \le (e^{M_1 h_N})^k = e^{M_1 k h_N} \tag{2.34}$$

for each $k \ge 1$. Combining (2.33) and (2.34) yields

$$\mathcal{E}_k \le \frac{M_2 h_N}{2M_1}\left(e^{M_1 k h_N} - 1\right)$$

for each $k \ge 1$. Finally, upon noting that $k h_N = t_k - t_0 \le T - T_0$, we obtain

$$\mathcal{E}_k \le \frac{M_2 h_N}{2M_1}\left(e^{M_1(T-t_0)} - 1\right) = \frac{M_2(T-t_0)}{2M_1 N}\left(e^{M_1(T-T_0)} - 1\right)$$

for $k = 1, 2, \ldots N$, as desired. □

Note Here

Example 22

Consider the initial value problem

$$\begin{cases} \frac{dy}{dt} = \cos^2(y) & y(0) = 0. \end{cases}$$

For this equation, $f(t,y) = \cos^2(y)$ is continuous on all of $\mathbb{R}^2$ and we have

$$\frac{\partial f}{\partial t}(t,y) = 0 \quad \text{and} \quad \frac{\partial f}{\partial y}(t,y) = -2\cos(y)\sin(y) = -\sin(2y),$$

both of which are continuous on all of $\mathbb{R}^2$. It is easy to see that, for all $(t,y) \in \mathbb{R}^2$,

$$\left|\frac{\partial f}{\partial y}(t,y)\right| = |-\sin(y)| \le 1 := M_1$$

and

$$\left|\frac{\partial f}{\partial t}(t,y) + \left(\frac{\partial f}{\partial y}(t,y)\right) f(t,y)\right| = \left|-\sin(2y)\cos^2(y)\right| \le 1 := M_2.$$

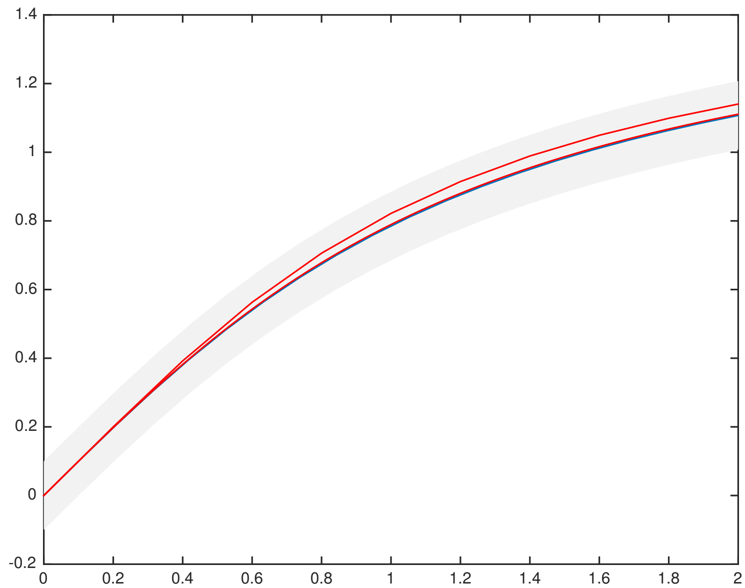
In using Euler's method to numerically approximate the solution to this initial value problem on, say, the interval $[0,2]$, we see that our situation is ripe for the application of Theorem 2.6.2 and we can therefore obtain so-called *a priori* bounds on the global approximation error $\mathcal{E} = \max_{k=1,2,\ldots,N} \mathcal{E}_k$. To this end, let's fix $T = 2 > t_0 = 0$ and a natural number N. Running Euler's method to approximate the solution $y = y(t)$ to this initial value problem on the interval $[0,2] = [t_0, T]$ yields approximants defined iteratively by $y_0 = 0$ and

$$y_k = y_{k-1} + f(t_{k-1}, y_{k-1}) h_N = y_{k-1} + \cos^2(y_{k-1}) \frac{2}{N}$$

for $k = 1, 2, \ldots, N$ where $t_k = t_0 + k h_N = 2k/N$. In view of Theorem 2.6.2, we have

$$\mathcal{E}_k \le \frac{M_2(2-0)}{2M_1 N}\left(e^{M_1(2-0)} - 1\right) = \frac{1}{N}\left(e^2 - 1\right) \le \frac{10-1}{N} = \frac{9}{N}$$

for $k = 1, 2, \ldots, N$. Thus, if we wish to get our global approximation error within $0.1 = 10^{-1}$, it is sufficient to choose $N = 90$. In the figure below, I have plotted the solution[a] $y(t)$ to this initial value problem along with the Euler approximation for $N = 10$ and 90.



The graph of the solution is shown in blue with a gray error band of radius 0.1 surrounding it. The approximants given by Euler's method are both plotted in red for $N = 10$ and $N = 90$, the latter being almost visually indistinguishable from the known solution. In studying this example, you should observe that the upper bound

$$\frac{M_2(T - t_0)}{2M_1 N}\left(e^{M_1(T-t_0)} - 1\right)$$

is a significant over estimation of the error, this is frequently the case.

[a]This initial value problem is easily solvable using the technique of separation of variables. The solution is $y(t) = \arctan(t)$.

Exercise 24

To get a handle on the preceding theorem, consider the initial value problem

$$\begin{cases} \frac{dy}{dt} = \frac{y}{2(1+y^2)} & y(0) = 1. \end{cases}$$

Here, $f(t, y) = y/(2(1 + y^2))$ which is evidently continuous on $\mathbb{R}^2$.

1. Compute the partial derivatives $\partial f/\partial t$ and $\partial f/\partial y$ and verify that they are also continuous on all of $\mathbb{R}^2$.

2. In fact, by inspection, you should see that f, $\partial f/\partial t$ and $\partial f/\partial y$ are also bounded

on all of $\mathbb{R}^2$. Use a method of your choosing to obtain positive constants M_1 and M_2 for which

$$\left|\frac{\partial f}{\partial y}(t,y)\right| \leq M_1 \qquad \text{and} \left|\frac{\partial f}{\partial t}(t,y) + \left(\frac{\partial f}{\partial y}(t,y)\right) f(t,y)\right| \leq M_2$$

for all $(t,y) \in \mathbb{R}^2$. Hint: It's straightforward to show that M_1 can be taken to be $1/2$ and M_2 to be $1/4$. If you can show this, great! If you can do better, i.e., get smaller values of M_1 and M_2, please do it! Note: If you use computational software, please include your source code.

3. Let's now run Euler's method. Set $t_0 = 0$, $T = 1$ and let $N = 3$, calculate the approximants y_1, y_2 and y_3 and put your answers in a table of the form

t_0	t_1	t_2	t_3
y_0	y_1	y_2	y_3

4. As the differential equation $y' = y/2(1+y^2)$ is separable, you can obtain a solution to the given initial value problem; however, the solution will need to be left in implicit form. Please give this solution in implicit form.

5. Using the implicit solution obtained in the previous part, I used Mathematica to appromate $y(t_1)$, $y(t_2)$ and $y(t_3)$. Putting the computer precision to 20 decimal places, I obtain:

$$\begin{aligned} y(t_1) &= 1.0832428248405900569 \\ y(t_2) &= 1.1659882614021548341 \\ y(t_3) &= 1.2478564015933930436 \end{aligned}$$

Assuming these are the actual values (which they aren't, but good enough for our purposes), compute the errors $\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_3$.

6. Let's now check these errors against the theorem. Using your values of M_1 and M_2 and $T = 1$, $t_0 = 0$ and $N = 3$, compute the upper bound in (2.32). Verify that your errors satisfy the inequality (2.32).

7. Finally, using (2.32) and your computed value of the upper bound in the present situation, find how large N would need to be so that the global approximation error yielded by Euler's method satisfies

$$\mathcal{E} = \max_{k=1,2,\ldots,N} \{\mathcal{E}_k\} \leq 10^{-10}.$$

Chapter 3

Higher Order Ordinary Differential Equations

In this chapter, we study higher[1] order ordinary differential equations and their initial value problems. Much of our focus, however, will be spent on second-order equations and this is essentially for two reasons. First, the methods we develop here to solve linear second-order equations generalizes straightforwardly to n^{th}-order equations for $n > 2$. Secondly, the vast majority of differential equations found in nature (and elsewhere) are second-order equations, e.g., Newton's Second Law, Bessel's Equation (which helps to explain the vibration of the head of a drum), the equation for the Arrow-Pratt measure of absolute risk aversion (in finance/economics).

3.1 Second-order Equations: Existence and Uniqueness

We recall, a second-order ordinary differential equations is an equation of the form

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

where $f : \mathcal{D} \to \mathbb{R}$ where $\mathcal{D} \subseteq \mathbb{R}^3$. We recall from Section 1.3 that second-order differential equations are paired with two initial conditions, one involving y and the other involving y'. Given a time t_0 and two numbers y_0 and y_0', we consider the corresponding initial value problem

$$\begin{cases} y'' = f(t, y', y) \\ \\ \quad y'(t_0) = y_0' \\ \quad y(t_0) = y_0. \end{cases} \tag{3.1}$$

A solution for this initial value problem is, by definition, a function y which is twice continuously differentiable[2] on an open interval J containing t_0 which satisfies both initial conditions $y(t_0) = y_0$ and $y'(t_0) = y_0'$ and has

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

for all $t \in J$.

[1] Of order $n > 1$

[2] This means that y is twice differentiable on J and its second derivative y'' is itself continuous on J.

Since we're reasonably comfortable now with questions of existence and uniqueness (and because I want to start writing "the solution" instead of "a solution"), let's state the Picard-Lindelöff theorem for second-order equations.

Theorem 3.1.1 (Picard-Lindelöff Theorem for Second-order Equations)**.** *Let t_0, y_0' and $y_0 \in \mathbb{R}$ and consider the initial value problem* (3.1) *where $f : \mathcal{D} \to \mathbb{R}$ for $\mathcal{D} \subseteq \mathbb{R}^3$. If there is an open region $R \subseteq \mathcal{D}$ (perhaps an open prism of the form $R = (a,b) \times (c,d) \times (e,f)$) containing the initial point $(t_0, y_0', y_0) \in \mathbb{R}^3$ on which $f = f(t,a,b)$, $\partial f/\partial a$ and $\partial f/\partial b$ are all continuous, then the initial value problem* (3.1) *has a unique solution y.*

In the next chapter, we will see that this theorem is a simple consequence of an analogous Picard-Lindelöff type theorem for first-order systems of differential equations. Though it is an interesting study, we shall not here analyze situations in which sufficiently "bad" functions f yield initial value problems for which the conclusions of this theorem are not true (as we did for first-order equations).

As was true in one dimension, initial value problems are often solved by producing a general solution – when it exists – to the differential equation $y'' = f(t, y', y)$ and then subjecting this general solution to initial conditions. Thus, as there are two initial conditions here, we expect general solutions to have two constants of integration, usually denoted by C_1 and C_2. As was the case for first-order differential equations, a class of equations which always have general solutions are linear equations. In some sense, these are the easiest differential equations to solve and, as we will see, they are framed within a beautiful theory which is tightly tied to linear algebra. This is the theory to which we now turn.

3.2 Linear Second-Order Ordinary Differential Equations: the general theory

In looking back to Definition 1.2.2, a second-order linear differential equation is an equation of the form

$$a_2(t)y'' + a_1(t)y' + a_0(t)y = g(t) \tag{3.2}$$

where a_0, a_1, a_2 and g are real-valued functions defined on an open interval I; the equation is said to be homogeneous provided g is the zero function. For example, two important second-order linear homogeneous equations are

$$(1-t^2)y'' - 2ty' + \alpha(\alpha-1)y = 0$$

and

$$t^2y'' + ty' + (t^2 - \nu^2)y = 0$$

where α and ν are (fixed) natural numbers. The first equation is called Legendre's equation[3] and appears in the study of gravitation and electrostatics [12]. The second equation is called Bessel's equation[4] and arises in a number of places, including the study of the vibration of a drum.

We shall primarily study the situation in which the "coefficient" $a_2 = 1$ or, equivalently, the second-order linear differential equation

$$y'' + p(t)y' + q(t)y = r(t) \tag{3.3}$$

[3]Adrien Marie Legendre (1752-1833) was a French mathematician best known for his work on elliptic integrals and the method of least squares.

[4]Friedrich Wilhelm Bessel (1784-1846) was a German astronomer and applied mathematician who first observed stellar parallax and used it to get one of the first accurate measurements of the "size" of the universe.

where p, q and r are real-valued functions on I (and can be gotten from (3.2) by multiplying though by $1/a_2$). To investigate the conditions under which solutions to (3.3) are unique (subject to initial conditions), we rewrite this equation as

$$y'' = r(t) - p(t)y' - q(t)y$$

so as to employ the machinery of Theorem 3.1.1. Here,

$$f(t, a, b) = r(t) - p(t)a - q(t)b$$

and, for f to satisfy the hypotheses of the theorem, we can see easily that we should require $r(t)$, $p(t)$ and $q(t)$ to be continuous functions (Do you see why?). More precisely, we have the following theorem.

Theorem 3.2.1. *Let I be an open interval and let p, q and r be continuous functions on I. Then, given any $t_0 \in I$ and any y_0 and $y_0' \in \mathbb{R}$, the initial value problem*

$$\begin{cases} y'' + p(t)y' + q(t)y = r(t) \\ \\ y(t_0) = y_0 \\ y'(t_0) = y_0' \end{cases}$$

has a unique solution y. Furthermore, this solution is twice continuously differentiable on the entire interval I.

Proof. We shall use Theorem 3.1.1 to show that a unique solution y exists. The statement about y being twice continuously differentiable on the entire interval I (and not just a subinterval J containing t_0) makes use of a more refined version of the Picard-Lindelöff theorem than we have stated; the proof of this fact is omitted (Note Here). Upon writing

$$f(t, a, b) = r(t) - p(t)a - q(t)b,$$

we have

$$\frac{\partial f}{\partial a} = -p(t) \qquad \text{and} \qquad \frac{\partial f}{\partial b} = -q(t)$$

for $t \in I$ and $a, b \in \mathbb{R}$. We observe that, given any $t_0 \in I$ and any two real numbers y_0 and y_0', the initial point (t_0, y_0, y_0') is contained in the rectangle

$$R = I \times \mathbb{R} \times \mathbb{R} = \{(t, a, b) : t \in I \text{ and } -\infty < a, b < \infty\}$$

on which f, $\partial f/\partial a$ and $\partial f/\partial b$ are continuous. The result now follows immediately by an application of Theorem 3.1.1. □

Example 1

Consider the initial value problem

$$\begin{cases} y'' + y = 0 \\ \\ y(0) = 1 \\ y'(0) = 0. \end{cases}$$

We observe that the functions $q(t) = 1$ and $p(t) = r(t) = 0$ are continuous everywhere and so the theorem above guarantees the existence and uniqueness of solutions to this initial value problem (and, in fact, any initial value problem). By guessing (perhaps), we observe that

$$y(t) = \cos(t),$$

is twice continuously differentiable, solves the differential equation and has

$$y(0) = \cos(0) = 1 \qquad \text{and} \qquad y'(0) = -\sin(0) = 0.$$

Thus, in view of the theorem, this is the one and only solution.

Exercise 25: Existence/uniqueness for Legendre's equation

Fix an integer α and consider Legendre's equation, that is, the homogeneous second-order linear equation

$$(1-t^2)y'' - 2ty' + \alpha(\alpha-1)y = 0.$$

Determine the intervals I on which any initial value problem corresponding to $t_0 \in I$ can be solved (and uniquely). Your answer should list three intervals (and have reasoning to illustrate your claim).

3.3 Homogeneous Equations and General Solutions

In this section, we focus our attention on general linear homogeneous second-order differential equations of the form

$$y'' + p(t)y' + q(t)y = 0 \tag{3.4}$$

where p and q are continuous functions on an interval I. The focus of this section is to understand how and when we can produce any/every solution to (3.4) given that (somehow) we already know a couple of solutions. For example, let's return to the differential equation

$$y'' + y = 0$$

discussed in the last section. As we saw, the function $y_1(t) = \cos(t)$, in particular, solves this differential equation. As you can probably easily guess, $y_2(t) = \sin(t)$ is also a solution to this differential equation. In this section, we discuss how we can use these two solutions to produce any other solution and, equivalently, solve any initial value problem. The following proposition is an important step in this direction; it's called the principle of superposition.

Proposition 3.3.1 (The Principle of Superposition)**.** *Let y_1 and y_2 be two solutions to* (3.4)*. Then, for any constants C_1 and C_2,*

$$y(t) = C_1y_1(t) + C_2y_2(t)$$

is a solution to (3.4)*.*

Proof. We have

$$y'(t) = C_1y_1'(t) + C_2y_2'(t) \qquad \text{and} \qquad y''(t) = C_1y_1''(t) + C_2y_2''(t)$$

for all $t \in I$. Consequently,

$$\begin{aligned} &y''(t) + p(t)y'(t) + q(t)y(t) \\ &\quad= (C_1y_1''(t) + C_2y_2''(t)) + p(t)(C_1y_1'(t) + C_2y_2'(t)) + q(t)(C_1y_1(t) + C_2y_2(t)) \\ &\quad= C_1y_1''(t) + C_2y_2''(t) + C_1p(t)y_1'(t) + C_2p(t)y_2'(t) + C_1q(t)y_1(t) + C_2q(t)y_2(t) \\ &\quad= C_1(y_1''(t) + p(t)y_1'(t) + q(t)y_1(t)) + C_2(y_2''(t) + p(t)y_2'(t) + q(t)y_2(t)) \qquad (3.5) \end{aligned}$$

for $t \in I$. However, since y_1 and y_2 are solutions to the differential equation (3.4), for all $t \in I$, we have

$$y_1''(t) + p(t)y_1'(t) + q(t)y_1(t) = 0 \quad \text{and} \quad y_2''(t) + p(t)y_2'(t) + q(t)y_2(t) = 0 \tag{3.6}$$

and so, by combining (3.5) and (3.6), we obtain

$$y''(t) + p(t)y'(t) + q(t)y(t) = 0$$

for all $t \in I$ and therefore $y = C_1y_1 + C_2y_2$ is a solution to (3.4). □

Remark 3.3.2. Proposition 3.3.1 says that the linear combination of two solutions is another solution. This guarantees, in particular, that the zero function is always a solution. It also guarantees that a constant multiple of any solution is another solution.

Let's now suppose that we have (however they've been found) two solutions y_1 and y_2 to the homogeneous equation (3.4) on the interval I. As the proposition shows, we can produce more solutions to (3.4) by simply taking linear combinations of y_1 and y_2. In this way, a natural question arises: Can every solution to (3.4) be gotten by a linear combination of y_1 and y_2? In discussing this question, the following vocabulary is helpful.

Definition 3.3.3. *Let $y_1(t)$ and $y_2(t)$ be solutions to* (3.4) *on the interval I. If every solution to* (3.4) *is given by*

$$y(t) = C_1y_1(t) + C_2y_2(t) \tag{3.7}$$

by specifying constants C_1 and C_2, we say that (3.7) *is a general solution to* (3.4).

As every solution to an initial value problem for the differential equation (3.4) is itself a solution to (3.4) and, conversely, every solution to (3.4) solves some initial value problem, we immediately obtain the following characterization of general solutions.

Proposition 3.3.4. *Let $y_1(t)$ and $y_2(t)$ be solutions to the differential equation* (3.4) *on the interval I. Then* (3.7) *is a general solution to* (3.4) *if and only if, given any $t_0 \in I$, $y_0, y_0' \in \mathbb{R}$, there are constants $C_1, C_2 \in \mathbb{R}$ for which* (3.7) *solves the initial value problem*

$$\begin{cases} y'' + py' + qy = 0 \\ \\ y'(t_0) = y_0' \\ y(t_0) = y_0 \end{cases}$$

If either of these equivalent conditions is satisfied, we call the pair $\{y_1, y_2\}$ a fundamental generating set of solutions to (3.4).

Example 2

Consider the second-order linear homogeneous differential equation

$$y'' - y' - 2y = 0. \tag{3.8}$$

One can easily verify that $y_1(t) = 2e^{2t}$ and $y_2(t) = -e^{2t}$ are solutions to (3.8). We ask: Is $y(t) = C_1y_1(t) + C_2y_2(t)$ a fundamental solution? To answer this question, let's consider the associated initial value problem

$$\begin{cases} y'' - y' - 2y = 0 \\ \\ y'(0) = 1 \\ y'(0) = -1. \end{cases} \tag{3.9}$$

If $y(t) = C_1y_1(t)+C_2y_2(t) = C_1\left(2e^{2t}\right)+C_2\left(-e^{2t}\right) = (2C_1-C_2)e^{2t}$ solves the initial value problem (3.9), we must have

$$1 = y(0) = (2C_1 - C_2)e^0 = 2C_1 - C_2$$

and

$$-1 = y'(0) = 2(2C_1 - C_2)e^0 = 2(2C_1 - C_2).$$

It is easy to see, however, that there is no choice of constants C_1 and C_2 for which these equations can hold simultaneously. This can be seen by substitution or elimination (or whatever linear solution method you use). Consequently, $y(t) = C_1((2e^{2t})+C_2\left(-e^{2t}\right)$ is not a general solution to (3.8) and hence $\{2e^{2t}, -e^{2t}\}$ is not a fundamental generating set of solutions.

Under further investigation, observe that $y(t) = e^{-t}$ solves (3.8) and, in fact, solves the initial value problem (3.9). As e^{-t} is sufficiently different in character from $2e^{2t}$, we might wonder if the pair $2e^{2t}$ and e^{-t} form a fundamental generating set of solution to (3.8). In fact, $\{2e^{2t}, e^{-t}\}$ is a fundamental generating set of solutions (and you should think about how you might show this). This fact will be confirmed in a subsequent example after we develop a little more machinery.

We return to the general picture and suppose that $y_1(t)$ and $y_2(t)$ are solutions to the differential equation (3.4). At present, sorting out whether or not y_1 and y_2 form a fundamental generating set of solutions seems like an arduous task. Let us instead, momentarily, focus on a more simple question: Given some fixed $t_0 \in I$ and $y_0, y_0' \in \mathbb{R}$, can the initial value problem

$$\begin{cases} y'' + py' + qy = 0 \\ \\ \quad y'(t_0) = y_0' \\ \quad y(t_0) = y_0 \end{cases} \tag{3.10}$$

be solved by (3.7) upon specifying constants C_1 and C_2? Here we are asking if (3.7) can be used to solve an initial value problem corresponding to a prespecified initial time t_0. We are not (yet) asking if this can be done for all t_0, a question whose affirmative answer would characterize (3.7) as a general solution. To this end, we must see if (3.7) can be used to satisfy the initial conditions at the initial time $t_0 \in I$. We want

$$y(t_0) = C_1y_1(t_0) + C_2y_2(t_0) = y_0$$

and

$$y'(t_0) = C_1y_1'(t_0) + C_2y_2'(t_0) = y_0'.$$

As the numbers y_0, y_0', $y_1(t_0)$, $y_2(t_0)$, $y_1'(t_0)$ and $y_2'(t_0)$ are all known, this is a 2×2 linear system in the variables C_1 and C_2. We can write this equivalently as the matrix equation

$$\begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} C_1y_1(t_0) + C_2y_2(t_0) \\ C_1y_1'(t_0) + C_2y_2'(t_0) \end{pmatrix} = \begin{pmatrix} y_0 \\ y_0' \end{pmatrix}.$$

We recall from linear algebra, that this system is solvable for C_1 and C_2 provided the matrix

$$W_{y_1,y_2}(t_0) = \begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix}$$

is invertible and this happens if and only if the determinant

$$w_{y_1,y_2}(t_0) = \det(W_{y_1,y_2}(t_0)) = y_1(t_0)y_2'(t_0) - y_2(t_0)y_1'(t_0) \neq 0.$$

The matrix $W_{y_1,y_2}(t_0)$ is called the *Wronskian matrix* for the solutions y_1 and y_2 evaluated at t_0 and $w_{y_1,y_2}(t_0)$ is called the *Wronskian determinant* evaluated at t_0; both objects are named in honor of Polish mathematician Józef Hoene-Wroński. In the case that the Wronskian determinant at t_0 is non-zero, we have

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix}^{-1} \begin{pmatrix} y_0 \\ y_0' \end{pmatrix} = (W_{y_1,y_2}(t_0))^{-1} \begin{pmatrix} y_0 \\ y_0' \end{pmatrix}$$

and so the inivial value problem (3.10) is solved by putting $y = C_1 y_1 + C_2 y_2$. We state this as a theorem.

Theorem 3.3.5. *Let y_1 and y_2 be solutions to* (3.4)*. Given $t_0 \in I$, if*

$$w_{y_1,y_2}(t_0) = y_1(t_0)y_2'(t_0) - y_2(t_0)y_1'(t_0) \neq 0,$$

or equivalently, if the Wronskian matrix at t_0 is invertible, then, for any real numbers y_0 and y_0', the initial value problem (3.10) *is solved by*

$$y(t) = C_1 y_1(t) + C_2 y_2(t)$$

where

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix}^{-1} \begin{pmatrix} y_0 \\ y_0' \end{pmatrix} = (W_{y_1,y_2}(t_0))^{-1} \begin{pmatrix} y_0 \\ y_0' \end{pmatrix}.$$

Example 3

Consider the homogeneous linear second-order differential equation

$$y'' + 4y = 0$$

where $p(t) = 0$ and $q(t) = 4$, both continuous on the whole of the real line $\mathbb{R}$. Also, consider

$$y_1(t) = \cos(2t) \qquad \text{and} \qquad y_2(t) = \sin(2t)$$

defined for $t \in \mathbb{R}$. We observe that

$$y_1''(t) + 4y_1(t) = (\cos(2t))'' + 4\cos(2t) = -4\cos(2t) + 4\cos(2t) = 0$$

and

$$y_2''(t) + 4y_2(t) = (\sin(2t))'' + 4\sin(2t) = -4\sin(2t) + 4\sin(2t) = 0$$

for all $t \in \mathbb{R}$ and therefore y_1 and y_2 are solutions to the second order linear homogeneous differential equation $y'' + 4y = 0$. Let's now consider the initial value problem

$$\begin{cases} y'' + 4y = 0 \\ \\ y(\pi/4) = 2 \\ y'(\pi/4) = -3 \end{cases}$$

where, by comparison to (3.10), $t_0 = \pi/4$, $y_0 = 2$ and $y_0' = -3$. Given y_1 and y_2, the Wronskian matrix at $t_0 = \pi/4$ is

$$\begin{aligned} W_{y_1,y_2}(\pi/4) &= \begin{pmatrix} y_1(\pi/4) & y_2(\pi/4) \\ y_1'(\pi/4) & y_2'(\pi/4) \end{pmatrix} \\ &= \begin{pmatrix} \cos(2\cdot\pi/4) & \sin(2\cdot\pi/4) \\ -2\sin(2\cdot\pi/4) & 2\cos(2\cdot\pi/4) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\pi/2) & \sin(\pi/2) \\ -2\sin(\pi/2) & 2\cos(\pi/2) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}. \end{aligned}$$

Consequently,

$$w_{y_1,y_2}(\pi/4) = \det\left(W_{y_1,y_2}(\pi/4)\right) = \det\begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix} = 2 \neq 0.$$

Thus, in view of Theorem 3.3.5, the above initial value problem is solved by a linear combination of y_1 and y_2. To find the coefficients, we look to the theorem and put

$$\begin{aligned} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} &= \left(W_{y_1,y_2}(\pi/4)\right)^{-1}\begin{pmatrix} y_0 \\ y_0' \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}^{-1}\begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 & -1/2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 2 \end{pmatrix} \end{aligned}$$

where we have used the fact that the inverse of a 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is given by

$$A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\det(A)}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

provided $\det(A) = ad - bc \neq 0$. Thus, in view of Theorem 3.3.5,

$$y(t) = C_1 y_1(t) + C_2 y_2(t) = \frac{3}{2}\cos(2t) + 2\sin(2t)$$

defined for $t \in \mathbb{R}$ solves the initial value problem above. This is easily verified (and you should always do this type of verification) by first observing that any linear combination of solutions is also a solution (Proposition 3.3.1) and further

$$y(\pi/4) = \frac{3}{2}\cos(2\cdot\pi/4) + 2\sin(2\cdot\pi/4) = 0 + 2 = 2$$

and

$$y'(\pi/4) = \frac{3}{2}(-2)\sin(2\cdot\pi/4) + 2(2)\cos(2\cdot\pi/4) = -3 + 0 = -3$$

as required. Finally, in view of Theorem 3.2.1, we know that this is the unique solution to this initial value problem.

Remark 3.3.6. Theorem 3.3.5 shows that, proviced $w_{y_1,y_2}(t_0) \neq 0$, we can solve the initial value problem (3.10) using the linear combination

$$y(t) = C_1 y_1(t) + C_2 y_2(t)$$

and it specifies exactly that constants C_1 and C_2 in terms of the inverse of the matrix $W_{y_1,y_2}(t_0)$, which is itself in terms of $y_1(t_0), y_2(t_0), y_1'(t_0), y_2'(t_0), y_0$ and y_0'. The essential thing for you to understand about this theorem is this: The condition that $w_{y_1,y_2}(t_0) \neq 0$ is sufficient for $y = C_1 y_1 + C_2 y_2$ to solve the initial value problem at hand – all you need to do is specify the constants C_1 and C_2. Though Theorem 3.3.5 does give you an explicit method for finding these constants, you can also simply solve the linear system

$$\begin{cases} C_1 y_1(t_0) + C_2 y_2(t_0) = y_0 \\ C_1 y_1'(t_0) + C_2 y_2'(t_0) = y_0' \end{cases}$$

directly for the constants C_1 and C_2. It is up to you whether you want to solve using the prescription given in Theorem 3.3.5 (in terms of matrices) or solve the above linear system directly.

Example 4

Consider the differential equation

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

and the corresponding initial value problem

$$\begin{cases} y'' - \frac{2}{t^2} y = 0 \\ y(1) = 1 \\ y'(1) = 0. \end{cases}$$

It is straightforward to see (and you should check it for yourself) that

$$y_1(t) = \frac{1}{t} \quad \text{and} \quad y_2(t) = t^2$$

defined for $t > 0$ are solutions to the differential equation. As $t_0 = 1$, we have $y_1(1) = 1/1 = 1$, $y_1'(1) = -1/1^2 = -1$, $y_2(1) = 1^2 = 1$ and $y_2'(1) = 2(1) = 2$. Therefore

$$w_{y_1,y_2}(t_0) = w_{y_1,y_2}(1) = (1)(2) - (1)(-1) = 3 \neq 0$$

and so Theorem 3.3.5 guarantees that the initial value problem is solved by

$$y(t) = \frac{C_1}{t} + C_2 t^2$$

for $t > 0$. Instead of constructing the Wronskian matrix, let's simply subject this general solution to the given initial conditions. We have

$$y(1) = \frac{C_1}{1} + C_2 1 = C_1 + C_2 = 1$$

and

$$y'(1) = C_1 \frac{-1}{1^2} + C_2(2)(1) = -C_1 + 2C_2 = 0.$$

By substitution (or addition), we find $1 = C_2 + 2C_2 = 3C_1$ and therefore $C_2 = 1/3$ and $C_1 = 2/3$. Thus our solution is

$$y(t) = \frac{2}{3t} + \frac{1}{3}t^2$$

defined for $t > 0$. You should verify directly that this solves the initial value problem.

Exercise 26

Consider the homogeneous linear second-order differential equation

$$y'' - 4y = 0$$

and notice that the coefficients $p = 0$, $q = -4$ are continuous on $I = \mathbb{R}$.

1. Consider $y_1(t) = e^{2t}$ and $y_2(t) = e^{-2t}$ defined for all real numbers t. Verify that y_1 and y_2 are solutions to this differential equation on $I = \mathbb{R}$.

2. Compute the Wronskian matrix $W_{y_1,y_2}(t_0)$ and Wronskian determinant $w_{y_1,y_2}(t_0)$ at an arbitrary point t_0. Is $w_{y_1,y_2}(t_0) = 0$ for any t_0?

3. Use the preceding theorem to solve the initial value problem

$$\begin{cases} y'' - 4y = 0 \\ \\ y(0) = 3 \\ y'(0) = -1. \end{cases}$$

Let's take stock of our progress thus far. We started with a second-order linear homogeneous differential equation (3.4) and assumed that we (somehow) knew two solutions y_1 and y_2. Using ideas from linear algebra, we observed that any linear combination of y_1 and y_2 was also a solution to (3.4). We asked: Given an initial value problem (which involved three fixed parameters $t_0 \in I$, y_0 and y_0') when could we solve the associated initial value problem (3.10) by a linear combination of y_1 and y_2. Throughout this investigation, we were led to the consideration of a matrix

$$W_{y_1,y_2}(t) = \begin{pmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{pmatrix}$$

called the Wronskian matrix and its determinant

$$w_{y_1,y_2}(t) = \det\left((W_{y_1,y_2}(t)\right) = y_1(t)y_2'(t) - y_2(t)y_1'(t)$$

called the Wronskian determinant[5]. We found that if the Wronskian matrix at t_0, $W_{y_1,y_2}(t_0)$, was invertible or, equivalently, if $w_{y_1,y_2}(t_0) \neq 0$, then **any** initial value problem posed at $t_0 \in I$ for (3.4) could be solved by a simple linear combination of y_1 and y_2 where the coefficients were given by a formula involving the inverse of $W_{y_1,y_2}(t_0)$ and the corresponding initial values; this was Theorem 3.3.5.

Our theory thus far has been fruitful and we've established a sufficient condition under which initial value problems at $t_0 \in I$ can be solved. In looking back to Proposition

[5] Caution: In many textbooks on ordinary differential equations, the "Wronskian" or "W" refer only to the determinant. The matrix W_{y_1,y_2} is often left unnamed.

3.3.4, Example 2 and the discussion which follows it, we'd like to understand how this theory depends on $t_0 \in I$. If we ask that $w_{y_1,y_2}(t) \neq 0$ for all $t \in I$, then Theorem 3.3.5 guarantees that all initial value problems for all initial times can be solved and hence, in view of Proposition 3.3.4, (3.7) is a general solution to (3.4). As the following theorem shows, a weaker condition will suffice: If the Wronskian determinant is non-zero at some (any!) time t, then (3.7) is a general solution.

Theorem 3.3.7. *Consider the homogeneous linear second-order ordinary differential equation* (3.4) *where we assume p and q are continuous functions on an interval I. Suppose that y_1 and y_2 solve the differential equation* (3.4)*. If*

$$w_{y_1,y_2}(t) = y_1(t)y_2'(t) - y_2(t)y_1'(t) \neq 0 \quad \text{for some } t \in I \tag{3.11}$$

(any t will do!), then (3.7) *is a general solution to* (3.4)*. More precisely, if the condition* (3.11) *is satisfied (that is for any t whatsoever in I), then the Wronskian matrix $W_{y_1,y_2}(t)$ is invertible for all $t \in I$ and so any initial value problem of the form*

$$\begin{cases} y'' + p(t)y' + q(t)y = 0 \\ \\ \qquad y(t_0) = y_0 \\ \qquad y'(t_0) = y_0' \end{cases} \tag{3.12}$$

where $t_0 \in I$ and $y_0, y_0' \in \mathbb{R}$ can be solved by (3.7) *by simply putting*

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = (W_{y_1,y_2}(t))^{-1} \begin{pmatrix} y_0 \\ y_0' \end{pmatrix} = \begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix}^{-1} \begin{pmatrix} y_0 \\ y_0' \end{pmatrix}.$$

We begin the proof of this theorem by first treating the following wonderful little lemma which is due to Abel.

Lemma 3.3.8. *Let y_1, y_2 be solutions to* (3.4)*. Then the Wronskian determinant w_{y_1,y_2} solves the first-order differential equation*

$$w' + p(t)w = 0$$

on the interval I.

Proof. As y_1 and y_2 are twice continuously differentiable on I, we observe that

$$w_{1,2}(t) := w_{y_1,y_2}(t) = y_1(t)y_2'(t) - y_2(t)y_1'(t)$$

is necessarily once continuously differentiable on I. Using the linearity and product rules for derivatives

$$\begin{aligned} w_{1,2}'(t) &= \frac{d}{dt}\left(y_1(t)y_2'(t) - y_2(t)y_1'(t)\right) \\ &= \frac{d}{dt}(y_1(t)y_2'(t)) - \frac{d}{dt}(y_2(t)y_1'(t)) \\ &= (y_1'(t)y_2'(t) + y_1(t)y_2''(t)) - (y_2'(t)y_1'(t) + y_2(t)y_1''(t)) \\ &= y_1(t)y_2''(t) - y_2(t)y_1''(t) + y_1'(t)y_2'(t) - y_2'(t)y_1'(t) \\ &= y_1(t)y_2''(t) - y_2(t)y_1''(t) + 0 \\ &= y_1(t)y_2''(t) - y_2(t)y_1''(t) \end{aligned} \tag{3.13}$$

for $t \in I$. By our assumption that y_1 and y_2 solve the differential equation (3.4), we can write

$$y_1''(t) = -p(t)y_1'(t) - q(t)y_1(t) \qquad \text{and} \qquad y_2''(t) = -p(t)y_2'(t) - q(t)y_2(t)$$

for $t \in I$. Substituting these identities for y_1'' and y_2'' into (3.13) yields

$$\begin{aligned}
w_{1,2}'(t) &= y_1(t)y_2''(t) - y_2(t)y_1''(t) \\
&= y_1(t)\left(-p(t)y_2'(t) - q(t)y_2(t)\right) - y_2(t)\left(-p(t)y_1'(t) - q(t)y_1(t)\right) \\
&= -p(t)y_1(t)y_2'(t) + p(t)y_2(t)y_1'(t) - q(t)y_1(t)y_2(t) + q(t)y_2(t)y_1(t) \\
&= -p(t)(y_1(t)y_2'(t) - y_2(t)y_1'(t)) - q(t)(y_1(t)y_2(t) - y_2(t)y_1(t))
\end{aligned}$$

for all $t \in I$. We see easily that the q term above is simply zero but what's perhaps more surprising is the the first term is exactly $-p(t)w_{1,2}(t)$! We have therefore shown that, provided y_1 and y_2 satisfy the differential equation (3.4), we have

$$w_{1,2}'(t) + p(t)w_{1,2}(t) = 0$$

for all $t \in I$. □

From this lemma, we immediately obtain a corollary. As you'll see, the proof of the corollary relies on our theory for first-order differential equations we developed in Chapter 1.

Corollary 3.3.9. *Let y_1 and y_2 be solutions to the differential equation* (3.4). *Then Wronskian determinant $w_{y_1,y_2} = y_1y_2' - y_2y_1'$ is either identically zero or it's never zero. In particular, if there is a single $t \in I$ for which $w_{y_1,y_2}(t) \neq 0$, then $w_{y_1,y_2}(t) \neq 0$ for all $t \in I$ and consequently, the Wronskian matrix $W_{y_1,y_2}(t)$ is invertible for all $t \in I$.*

Proof. Given that p is continuous on I, the first-order differential equation

$$w' + p(t)w = 0$$

satisfies the hypotheses of the first-order Picard-Lindelöff theorem (Theorem 2.4.2) on the rectangle $R = I \times \mathbb{R}$. Consequently, no two integral curves for this differential equation can intersect in view of the results shown in Exercise 17 (2.4).

Observe now that $w = 0$ is an equilibrium solution to this equation. By the previous lemma the Wronskian determinant w_{y_1,y_2} is also a solution. The inability for integral curves to cross then gives us the following dichotomy: Either w_{y_1,y_2} is the zero (equilibrium) solution or w_{y_1,y_2} is non-zero and hence can never take the value zero. □

Let's now put all of these pieces together to prove Theorem 3.3.7.

Proof of Theorem 3.3.7. Let y_1 and y_2 be solutions of (3.4) and suppose that Condition (3.11) is satisfied, that is, $w_{y_1,y_2}(t) \neq 0$ for some t. Then, in view of Corollary 3.3.9, the Wronskian matrix $W_{y_1,y_2}(t_0)$ is invertible for all initial times $t_0 \in I$ and so, by virtue of Theorem 3.3.5, any initial value problem of the form (3.12) can be solved by a linear combination of y_1 and y_2. In other words, (3.7) is a general solution to (3.4). □

Example 5: Examples 2 and 4 revisited

1. Consider (again) the differential equation

$$y'' - y' - 2y = 0$$

and the solutions $y_1(t) = 2e^{2t}$ and $y_2(t) = e^{-t}$. Here, $p(t) = -1$ and $q(t) = -2$ are continuous functions on $I = \mathbb{R}$. We have

$$w_{y_1,y_2}(t) = y_1(t)y_2'(t) - y_2(t)y_1'(t) = \left(2e^{2t}\right)\left(-e^{-t}\right) - \left(e^{-t}\right)\left(4e^{2t}\right) = -6e^t$$

for $t \in \mathbb{R}$. We observe that, in particular, at $t = 0$, $w_{y_1,y_2}(0) = -6 \neq 0$ and by virtue of Theorem 3.3.7 we may conclude that $\{y_1, y_2\}$ is a fundamental generating set of solutions. Of course, it is easy to see the conclusion of Corollary 3.3.9 here as $w_{y_1,y_2}(t) = -6e^t \neq 0$ for all $t \in \mathbb{R}$. Observe further that, as guaranteed by Lemma 3.3.8, $w_{y_1,y_2}(t) = -6e^t$ solves the first-order linear differential equation

$$w' + p(t)w = w' - w = 0.$$

2. Consider the differential equation

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

and the solutions $y_1(t) = 1/t$ and $y_2(t) = t^2$ on the interval $I = (0, \infty)$; here, $p(t) = 0$ and $q(t) = -2/t^2$ are continuous functions on $I = (0, \infty)$. We have

$$w_{y_1,y_2}(t) = \left(\frac{1}{t}\right)(2t) - (t^2)\left(-\frac{1}{t^2}\right) = 3.$$

In particular, $w_{y_1,y_2}(1) = 3 \neq 0$ and so Theorem 3.3.7 guarantees that $\{1/t, t^2\}$ is a fundamental generating set of solutions. The conclusion to Corollary 3.3.9 is easily seen in this example and furthermore $w'_{y_1,y_2}(t) = d(3)/dt = 0$ and so $w_{y_1,y_2}(t)$ solves

$$w' + p(t)w = w' + 0w = 0$$

as required by Lemma 3.3.8.

Exercise 27

Consider the linear homogeneous second-order differential equation

$$y'' + \frac{1}{t}y' - \frac{4}{t^2}y = 0.$$

1. What are the functions $p(t)$ and $q(t)$ in this differential equation? Verify (just a one-sentence explanation) that p and q are continuous on $I = (0, \infty)$.

2. Verify that

$$y_1(t) = t^2 \qquad \text{and} \qquad y_2(t) = \frac{1}{t^2}$$

solve this differential equation on $I = (0, \infty)$.

3. Compute the Wronskian matrix $W_{y_1,y_2}(t)$ and Wronskian determinant $w_{y_1,y_2}(t)$ for $t \in I$.

4. Confirm the conclusion of Lemma 3.3.8 by showing that w_{y_1,y_2} satisfies $w' + p(t)w = 0$ for $t \in I$.

5. Confirm the conclusion of Corollary 3.3.9, i.e., show that w_{y_1,y_2} falls on one side of this dichotomy.

6. Finally, verify that the hypotheses of Theorem 3.3.7 have been met for y_1 and

y_2 and use them to solve the initial value problem

$$\begin{cases} y'' + \frac{1}{t}y' - \frac{4}{t^2}y = 0 \\ \\ y(1) = 2 \\ y'(1) = -1 \end{cases}$$

on the interval $I = (0, \infty)$.

This section developed the theory of linear homogeneous second-order differential equations under the assumption that two solutions y_1 and y_2 were known. To this end, we put down an easily checkable condition under which $y = C_1 y_1 + C_2 y_2$ was a general solution; this condition was simply that the Wronskian determinant was non-zero somewhere. Beyond the result of the previous exercise (which is really nice!), two nagging questions remain:

1. How do we know two such solutions y_1 and y_2 (those for which the conclusion of Theorem 3.3.7 holds) exist?
2. If they exist, how do we find them?

The following theorem gives us a satisfactory answer to the first question; it is unfortunately purely existential. We'll have to postpone the second question for two more sections.

Theorem 3.3.10. *Consider the linear homogeneous second-order differential equation*

$$y'' + p(t)y' + q(t)y = 0$$

where p and q are continuous functions on an interval I. There exists two solutions y_1 and y_2 for which the conclusion of Theorem 3.3.7 holds. In other words, there are two functions y_1 and y_2 such that

$$y(t) = C_1 y_1(t) + C_2 y_2(t)$$

is a general solution to the above differential equation.

Proof. Let $t_0 \in I$ (any element will do) and consider the initial value problems

$$\begin{cases} y'' + p(t)y' + q(t)y = 0 \\ \\ y(t_0) = 1 \\ y'(t_0) = 0 \end{cases}$$

and

$$\begin{cases} y'' + p(t)y' + q(t)y = 0 \\ \\ y(t_0) = 0 \\ y'(t_0) = 1 \end{cases}.$$

In view of Theorem 3.2.1, both of these initial value problems have unique solutions which we denote by y_1 and y_2 respectively. Necessarily, $y_1(t_0) = 1$, $y_1'(t_0) = 0$, $y_2(t_0) = 0$ and $y_2'(t_0) = 0$ and since they satisfy distinct initial conditions, these solutions are necessarily distinct. We observe that at this chosen time t_0,

$$w_{y_1,y_2}(t_0) = \det(W_{y_1,y_2}(t_0)) = \begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \neq 0.$$

The result now follows by an appeal to Theorem 3.3.7. ☐

3.3.1 Abel's Identity and a useful application

In the proof of Corollary 3.3.9, we used only the uniqueness aspect of the Picard-Lindelöff theorem to make our conclusion. In that the proof, we readily observed that for two solutions y_1 and y_2 of (3.4), the Wronskian determinant satisfies $w' + p(t)w = 0$. Since the equation $w' + p(t)w = 0$ is a linear first-order equation, Theorem 2.2.1 says that all solutions (and so w_{y_1,y_2} in particular) are of the form

$$w(t) = Ce^{-P(t)}$$

where $P(t)$ is an antiderivative of $p(t)$. In other words, to each pair of solutions y_1 and y_2 to (3.4), there is a constant $C_{y_1,y_2} \in \mathbb{R}$ for which

$$w_{y_1,y_2}(t) = C_{y_1,y_2}e^{-P(t)}. \tag{3.14}$$

This provides us with another way to see that either w_{y_1,y_2} is never zero, i.e., when $C_{y_1,y_2} \neq 0$, or w_{y_1,y_2} is always zero, i.e., when $C_{y_1,y_2} = 0$. In view of our preceding results, we see that y_1 and y_2 form a fundamental generating set of solutions if and only if C_{y_1,y_2} is non-zero. The identity (3.14) is called Abel's identity and is due to N. Abel.

Among being interesting in its own right, Abel's identity gives a way of producing solutions to (3.4). Specifically, if one solution y_1 of (3.4) is known, Abel's identity can be used to produce another solution y_2 for which the pair $\{y_1, y_2\}$ forms a fundamental generating set of solutions. Let's explore this application.

Suppose that y_1 is a known non-zero solution to (3.4) and we seek another solution $y = y_2$ to (3.4). If such a solution y is to be found for which the pair $\{y_1, y\}$ to form a fundamental generating set of solutions, it is necessary that

$$y_1(t)y'(t) - y_1'(t)y(t) = w_{y_1,y}(t) = C_{y_1,y}e^{-P(t)}$$

for a non-zero constant $C_{y_1,y}$. Observe that, in this equation, y_1', y_1 and $e^{-P(t)}$ are known; the unknowns are the constant $C_{y_1,y}$ and the function y along with its derivative y'. Of course, this itself gives us a first order linear differential equation for y, provided we can specify the constant $C_{y_1,y}$. However, as any constant multiple of the desired solution y is another solution, we can simply demand that $C_{y_1,y} = 1$ and seek a solution y for which

$$y_1y' - y_1'y = e^{-P(t)}$$

for $t \in I$. Though we made this argument on the basis of necessity, it is not too terribly hard to show that y being a solution to the above equation is also sufficient for y to be a solution to (3.4) and have that the pair $\{y_1, y\}$ form a fundamental generating set of solutions. We state this as a proposition and leave its proof to the interested reader.

Proposition 3.3.11. *Let y_1 be a non-zero solution to* (3.4) *and let y solve the first-order linear differential equation*

$$y_1(t)y' - y_1'(t)y = e^{-P(t)}$$

for $t \in I$. Then $y_2 = y$ also solves (3.4) *(and so is twice continuously differentiable on I) and furthermore the pair $\{y_1, y_2\}$ is a fundamental generating set of solutions for* (3.4).

We make use of this proposition through the following steps:

Producing solutions via Abel's identity

Given a non-zero solution y_1 to the second-order linear homogeneous differential equation

$$y'' + p(t)y' + q(t)y = 0,$$

to find a second solution y_2 for which $\{y_1, y_2\}$ is a fundamental generating set of solutions, do the following:

1. Find an antiderivative $P(t)$ of p.
2. Solve the first order differential equation
$$y' - \frac{y_1'(t)}{y_1(t)}y = \frac{e^{-P(t)}}{y_1(t)} \tag{3.15}$$
for a non-zero solution y. This can be done using Theorem 2.2.1 or by any method you desire.
3. Set $y_2 = y$ and verify that y_2 solves the second-order differential equation and that $\{y_1, y_2\}$ is a fundamental generating set of solutions.

Example 6

Consider the differential equation

$$y'' + \frac{1}{t}y' - \frac{1}{t^2}y = 0$$

on the interval $I = (0, \infty)$. It is easy to see (and you might have already guessed) that $y_1(t) = t$ is a solution. Let's use the method above to produce another solution.

First, we have $p(t) = \frac{1}{t}$ and and we can choose $P(t) = \ln(t)$ as an antiderivative of P. Second, using the fact that $y_1(t) = t$ and therefore $y_1'(t) = 1$, the first-order differential equation (3.15) given by Abel's identity is

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

To solve this first-order equation, we first obtain the integrating factor $\mu(t) = 1/t$ and multiply through to find

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

Thus

$$\frac{y}{t} = \int \frac{1}{t^3}\,dt = -\frac{1}{2t^2} + C.$$

As any such (non-zero) solution will do, we choose $C = 0$ and multiply by t to obtain

$$y(t) = -\frac{1}{2t}$$

for $t \in I$. Finally, we set $y_2(t) = y(t) = -1/2t$.

Let's verify that this, in fact, gave us what we wanted. We have $y_2'(t) = 1/2t^2$ and $y_2''(t) = -1/t^3$ and therefore

$$\begin{aligned} & y_2''(t) + \frac{1}{t}y_1'(t) - \frac{1}{t^2}y_2(t) \\ = & \quad -\frac{1}{t^3} + \frac{1}{t}\frac{1}{2t^2} - \frac{1}{t^2}\left(\frac{-1}{2t}\right) \\ = & \quad \left(-1 + \frac{1}{2} + \frac{1}{2}\right)\frac{1}{t^3} = 0 \end{aligned}$$

for $t \in I$ and so y_2 is indeed a solutions. To see that $\{y_1, y_2\}$ is a fundamental generating set, we observe that

$$w_{y_1,y_2}(t) = y_1(t)y_2'(t) - y_1'(t)y_2(t) = t\left(\frac{1}{2t^2}\right) - 1\left(\frac{-1}{2t}\right) = \frac{1}{t} \neq 0$$

for $t \in I$. Of course, this result should come at no surprise for, by forcing Abel's identity to hold with $C = 1$, we designed y_2 in such a way that

$$w_{y_1,y_2}(t) = e^{-P(t)} = e^{-\ln(t)} = \frac{1}{t}.$$

for all $t \in I$.

Exercise 28

Consider the differential equation

$$y'' - 4y' + 4y = 0 \tag{3.16}$$

and the function $y_1(t) = e^{2t}$.

1. Verify that y_1 is a solution to (3.16).
2. By following the procedure outlined above, find a second solution y_2 to (3.16).
3. Verify directly that your function y_2 is a solution and that $\{y_1, y_2\}$ is a fundamental generating set of solutions.

Exercise 29

In using Abel's identity to produce a second solution $y = y_2$ to the differential equation

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

given that one solution y_1 is known, you obtain the first-order linear differential equation

$$y' - \frac{y_1'(t)}{y_1(t)}y = \frac{e^{-P(t)}}{y_1(t)}$$

which is of the form

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

where $a(t) = -y_1'(t)/y_1(t)$ and $b(t) = e^{-P(t)}/y_1(t)$.

1. In noting that an antiderivative of $a(t)$ can be readily found (try u-substitution), use Theorem 2.2.1 to obtain a formula for $y = y_2$. In this formula, feel free to choose $C = 0$. Your answer should contain an indefinite integral and be in terms of y_1 and P.
2. Conclude that $y = y_2$ is a product[a] of y_1 and another function $u(t)$, i.e.,

$$y(t) = y_1(t)u(t)$$

for $t \in I$. What is the function $u(t)$?

3. In looking back to the previous exercise, i.e., where $y_1(t) = e^{2t}$ and $P(t) = -4t$, show that the formula you obtain in Item 1 yields the y_2 you found previously.

[a]The observation that y_2 is obtained by multiplying y_1 by another function u is the basis for a technique called *reduction of order* found in the literature.

3.4 Homogeneous Equations: A Linear Algebraic Perspective

In this section, we study second-order linear homogeneous differential equations through the lens of linear algebra. As it turns out, our perspective here will help us understand the theory of the preceding section in a more complete way. To phrase things in these terms, we first need to think about vector spaces for, as you know, linear algebra is the study of vector spaces and linear maps between them[6]. To this end, let's introduce the vector spaces relevant to our study of linear differential equations. Given an open interval $I = (a, b)$, we will denote the set of continuous real-valued functions on I by $C^0(I)$. In other words,

$$C^0(I) = \{f : I \to \mathbb{R} : f \text{ is continuous on } I\}.$$

For any two elements $f, g \in C^0(I)$, i.e., continuous functions f and g on I, we can add them to form another function $f + g$ and defined by

$$(f + g)(t) = f(t) + g(t) \tag{3.17}$$

for $t \in I$. Also, given a real number α and a function $f \in C^0(I)$, we can create another function $\alpha \cdot f$ defined by

$$(\alpha \cdot f)(t) = \alpha f(t) \tag{3.18}$$

for $t \in I$; we shall abbreviate $\alpha \cdot f$ by αf. These two operations on functions are called addition and scalar multiplication, respectively. Though you might not have seen them put in this way, these operations are exactly what you've been working with all along, i.e., this is the normal way you add functions and multiply functions by scalars. Recalling the rules of limits from introductory calculus, it is evident that, because $f, g \in C^0(I)$, $f + g$ and αf are continuous functions on I and are therefore members of $C^0(I)$ themselves. From here it is not difficult to see that, in fact, equipped with these operations of function addition and scalar multiplication, $C^0(I)$ forms a vector space; the zero vector in $C^0(I)$ is simply the zero function, i.e., the function that assigns the number 0 to each $t \in I$. We state this as a proposition; verifying the vector space axioms[7] is left for you.

[6]This is a good time to brush up on your linear algebra skills if you feel rusty. For this, I encourage you to see your old linear algebra textbook and the appendix of this text.

[7]See Appendix C.

Proposition 3.4.1. *Consider the interval*[8] $I = (a, b)$ *and the set*

$$C^0(I) = \{f : I \to \mathbb{R} : f \text{ is continuous on } I\}$$

of continuous functions on I. Equipped with the addition $(+)$ and scalar multiplication $(\cdot)$ defined by (3.17) *and* (3.18)*, respectively, the set $C^0(I)$ forms a vector space over $\mathbb{R}$. It is called the vector space of continuous functions on I.*

There is a lot of deep mathematics connected to the space $C^0(I)$, which you will study if you take a course in mathematical analysis. In this section, we will focus our attention on certain subspaces of $C^0(I)$. We define

$$C^1(I) = \{f : I \to \mathbb{R} : f \text{ is differentiable and its derivative } f' \text{ is continuous on } I\}$$

and

$$C^2(I) = \{f : I \to \mathbb{R} : f \text{ is twice differentiable and } f'' \text{ is continuous on } I\}.$$

Exercise 30

Show the following:

1. Show that, as sets,
$$C^2(I) \subseteq C^1(I) \subseteq C^0(I).$$

2. Show that these are all "proper" containments, i.e., find a function $f \in C^0(I)$ which is not a member of $C^1(I)$. Also, find a function $g \in C^1(I)$ which is not a member of $C^2(I)$. Here, you can let $I = (-1, 1)$ or any open interval you want.

3. For a general I, show that $C^1(I)$ is a subspace of $C^0(I)$. Show that $C^2(I)$ is a subspace of $C^0(I)$. Is $C^2(I)$ a subspace of $C^1(I)$?

As it turns out, the spaces $C^0(I)$, $C^1(I)$ and $C^2(I)$ are huge. In contrast to familiar vector spaces like $\mathbb{R}^n$ which is n-dimensional (finite-dimensional), they are infinite dimensional spaces and so much of our intuition about vector spaces gathered from our knowledge of $\mathbb{R}^n$ breaks down. If this is of interest to you, I encourage you to do the exercises in Appendix C.

In view of the preceding exercise, we observe that

$$C^1(I) = \left\{f \in C^0(I) : f' = \frac{d}{dt}f \text{ exists on } I \text{ and } f' \in C^0(I)\right\}$$

and

$$C^2(I) = \left\{f \in C^1(I) : f'' = \frac{d}{dt}f' \text{ exists on } I \text{ and } f'' \in C^0(I)\right\}.$$

Continuing inductively, for each $n \geq 1$, we define

$$C^n(I) = \left\{f \in C^{n-1}(I) : f^{(n)} = \frac{d}{dt}f^{(n-1)} \text{ exits on } I \text{ and } f^{(n)} \in C^0(I)\right\};$$

[8] We shall allow for the possibilities that $a = -\infty$ or $b = \infty$

here, the symbol $f^{(k)}$ denotes the kth derivative of f. It is evident that[9], for all $0 \leq n \leq m$, $C^m(I)$ is a subspace of $C^n(I)$. In particular, for each $n \geq 0$, $C^n(I)$ is a vector space; it is called the space of *n-times continuously differentiable functions on I*. Another vector space which will be on interest is the space of *smooth functions on I* defined by

$$C^\infty(I) = \bigcap_{n\geq 0} C^n(I)$$

which is, equivalently, the set of function on I that have derivatives of all orders. Naturally, the functions in $C^\infty(I)$ are said to be *smooth*. Many of the functions you know (and know well) are smooth functions. For example, $e^t, \cos(t)$, $\sin(t)$ and all polynomials are smooth functions. In fact, as you might remember from calculus, any function having a convergent power series representation on I is smooth. What is perhaps more interesting is that there are smooth functions on I which cannot be represented by power series.

The focus of the remainder of this section is to consider an important type of map between the vector space $C^2(I)$ and $C^0(I)$ which is relevant to our study of differential equations. To this end, let $p, q \in C^0(I)$ and consider the function $L : C^2(I) \to C^0(I)$ defined by taking a function $y \in C^2(I)$ and sending it to a function $L[y] \in C^0(I)$ by the rule

$$L[y](t) = y''(t) + p(t)y'(t) + q(t)y(t)$$

for $t \in I$. Let's observe an important property of this map[10]. Given any functions y_1 and y_2 in $C^2(I)$, observe that $(y_1 + y_2)'(t) = y_1'(t) + y_2'(t)$ and $(y_1 + y_2)''(t) = y_1''(t) + y_2''(t)$ for $t \in I$ by the rules of calculus (the derivative of the sum is the sum of derivatives). Consequently,

$$\begin{aligned}
L[y_1 + y_2](t) &= (y_1 + y_2)''(t) + p(t)(y_1 + y_2)'(t) + q(t)(y_1 + y_2)(t)\\
&= y_1''(t) + y_2''(t) + p(t)(y_1'(t) + y_2'(t)) + q(t)(y_1(t) + y_2(t))\\
&= y_1''(t) + y_2''(t) + p(t)y_1'(t) + p(t)y_2'(t) + q(t)y_1(t) + q(t)y_2(t)\\
&= y_1''(t) + p(t)y_1'(t) + q(t)y_1(t) + y_2''(t) + p(t)y_2'(t) + q(t)y_2(t)\\
&= L[y_1](t) + L[y_2](t)
\end{aligned}$$

for all $t \in I$. Thus the map L has the property that $L[y_1 + y_2] = L[y_1] + L[y_2]$ for $y_1, y_2 \in C^2(I)$. By a similar argument, it isn't hard to see that $L[\alpha y] = \alpha L[y]$ for all $\alpha \in \mathbb{R}$ and $y \in C^2(I)$. We can therefore conclude that $L : C^2(I) \to C^0(I)$ is a linear transformation (also called a linear operator) from the vector space $C^2(I)$ into the vector space $C^0(I)$.

We have established that L is a linear map. You might say "okay, so what"? If you recall from linear algebra, the kernel of each linear map is an important object; it measures the degree to which the map is injective (one-to-one). From the perspective of linear differential equations, the kernel of our linear transformation L is of paramount importance. Recalling that the zero function is the zero vector in $C^0(I)$, we have

$$\begin{aligned}
\ker(L) &= \{y \in C^2(I) : L[y] \text{ is the zero function}\}\\
&= \{y \in C^2(I) : L[y](t) = y''(t) + p(t)y'(t) + q(t)y(t) = 0 \text{ for all } t \in I\}
\end{aligned}$$

From this we see trivially that $y \in \ker(L)$ if and only if y solves the linear homogeneous differential equation (3.4). We state all of this together in the following proposition.

[9] Precisely the same argument you used in the preceding exercise will generalize here.

[10] Before we do, you should verify yourself that, indeed, $L[y] \in C^0(I)$ provided $y \in C^2(I)$.

Proposition 3.4.2. *The map $L : C^2(I) \to C^0(I)$ is a linear operator from the vector space $C^2(I)$ of twice continuously differentiable functions into the vector space of continuous functions on the interval I. The kernel of this operator* $\ker(L)$ *is precisely the set of solutions to the homogeneous differential equation* (3.4).

In view of the proposition above, our goal is to understand $\ker(L)$ which characterizes all solutions of (3.4). From linear algebra, you should recall that the $\ker(L)$ is a subspace of $C^2(I)$. Consequently, we have the property that $\ker(L)$ is closed under linear combinations. That is, for any $y_1, y_2 \in \ker(L)$ and $C_1, C_2 \in \mathbb{R}$, we have $C_1y_1 + C_2y_2 \in \ker(L)$. Using the previous proposition to translate this into a statement about solutions to (3.4), we conclude that, for all y_1, y_2 solving (3.4) and constants C_1 and C_2,

$$y(t) = C_1y_1(t) + C_2y_2(t)$$

is a solution to (3.4). Of course, this should be no surprise for it is precisely the principle of superposition, Proposition 3.3.1!

In view of the paragraph above, we get the sense that this linear algebraic perspective is useful and that $\ker(L)$ is indeed an important object. To this end, let's ask about things like linear independence, span, basis and dimension. We first focus our attention on linear independence. In looking back to our knowledge of linear algebra, we recall that vectors $v_1, v_2, \dots, v_k$ are said to be linearly independent if the equation

$$C_1v_2 + C_2v_2 + \cdots + C_kv_k = 0$$

can only hold provided that the constants $C_1, C_2 \dots, C_k$ are all zero. Consequently, two functions y_1 and y_2 in $\ker(L) \subseteq C^2(I)$ are linearly independent if the only constants C_1 and C_2 for which

$$C_1y_1(t) + C_2y_2(t) = 0$$

for all $t \in I$ are the constants $C_1 = C_2 = 0$. As the following result shows, the question of linear independence in $\ker(L)$ is closely related to the Wronskian determinant – an object which arose for us in the completely different context of aiming to solve initial value problems.

Proposition 3.4.3. *Let $y_1, y_2 \in \ker(L)$. Then y_1 and y_2 are linearly independent if and only if*

$$w_{y_1,y_2}(t) \neq 0$$

for some (and hence all) $t \in I$.

Proof. Suppose that C_1 and C_2 are numbers for which

$$C_1y_1(t) + C_2y_2(t) = 0$$

for all $t \in I$; this means the linear combination of y_1 and y_2 is equal to the zero function. As both y_1 and y_2 are differentiable, the previous equation gives

$$C_1y_1'(t) + C_2y_2'(t) = 0$$

for all $t \in I$. In combining these two equations in terms of C_1, C_2 and y_1, y_2 and their derivatives, we obtain the matrix equation.

$$\begin{pmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Of course, we immediately see the Wronskian matrix rear its head. This is equivalently the equation

$$W_{y_1,y_2}(t)\begin{pmatrix}C_1\\C_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}.$$

If, for some $t \in I$, $w_{y_1,y_2}(t) \neq 0$, then the Wronskian matrix is invertible and hence the only solution to the equation above is the trivial (zero) solution, i.e., $C_1 = C_2 = 0$. We therefore see that y_1 and y_2 must be linearly independent.

Conversely, let y_1 and y_2 be linearly independent members of $\ker(L)$. We assume, to reach a contradiction, that $w_{y_1,y_2}(t_0) = 0$ for some $t_0 \in I$. In this case, the Wronskian matrix $W_{y_1,y_2}(t_0)$ is not invertible and hence there is a non-trivial solution (in C_1 and C_2) to the system

$$W_{y_1,y_2}(t_0)\begin{pmatrix}C_1\\C_1\end{pmatrix}=\begin{pmatrix}y_1(t_0) & y_2(t_0)\\ y_1'(t_0) & y_2'(t_0)\end{pmatrix}\begin{pmatrix}C_1\\C_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}.$$

Correspondingly, for this fixed pair C_1 and C_2 (not both zero), $y(t) = C_1y_1(t) + C_2y_2(t)$ solves the initial value problem

$$\begin{cases} L[y] = y'' + py' + qy = 0 \\ \\ y(t_0) = 0 \\ y'(t_0) = 0 \end{cases}.$$

Of course, this initial value problem is also satisfied by the zero solution (the zero function) and so, by the uniqueness property guaranteed by Theorem 3.2.1, $y(t)$ must be the zero function, i.e. $y(t) = C_1y_1(t) + C_2y_2(t) = 0$ for all $t \in I$. But this cannot be true, for we have assumed that y_1 and y_2 are linearly independent and so we have obtained a contradiction. Hence $w_{y_1,y_2}(t) \neq 0$ for all $t \in I$. □

Remark 3.4.4. Upon reading carefully through the preceding proof, you'll see that validity of the forward statement (that $w_{y_1,y_2}(t) \neq 0$ for some t implies that y_1 and y_2 are linearly independent) doesn't actually require that $y_1, y_2 \in \ker(L)$. The converse statement must however make use of this assumption for, in 1889, G. Peano proved that it is possible to have two linearly independent functions whose Wronskian determinant is always zero [13].

Let's now turn our focus to basis and dimension. Though the following result is essentially one we've seen before (Theorem 3.3.10), it takes on new life when phrased through our new linear-algebraic lens. In this context, the theorem shows, in particular, that the $\ker(L)$ is a 2-dimensional subspace of $C^2(I)$, which is itself an infinite dimensional vector space.

Theorem 3.4.5. *Let $p, q \in C^0(I)$ and define $L : C^2(I) \to C^0(I)$ by*

$$L[y] = y'' + py' + qy.$$

Then there exist functions $y_1, y_2 \in \ker(L)$ for which $\{y_1, y_2\}$ is a basis of $\ker(L)$. In other words, $\ker(L)$ is a two-dimensional vector space.

Proof. This proof is essentially that given for Theorem 3.3.10 but it is included here as it is insightful to see the argument phrased in the language of linear algebra. To this end, pick any $t_0 \in I$ and consider the initial value problems

$$\begin{cases} L[y] = y'' + py' + qy = 0 \\ y(t_0) = 1 \\ y'(t_0) = 0 \end{cases} \quad \text{and} \quad \begin{cases} L[y] = y'' + py' + qy = 0 \\ y(t_0) = 0 \\ y'(t_0) = 1 \end{cases}.$$

In view of Theorem 3.2.1, these initial value problems have solutions y_1 and y_2, i.e., there are functions $y_1, y_2 \in \ker(L)$ for which $y_1(t_0) = 1$, $y_1'(t_0) = 0$, $y_2(t_0) = 0$ and $y_2(t_0) = 1$. I claim that y_1 and y_2 form a basis for $\ker(L)$. To see this, first observe that

$$w_{y_1,y_2}(t_0) = \det \begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix} = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \neq 0.$$

In view of the preceding proposition, y_1 and y_2 are linearly independent. It remains to show that y_1 and y_2 span $\ker(L)$. To this end, let $y \in \ker(L)$, i.e., $y \in C^2(I)$ solves (3.4). Given that the Wronskian determinant for y_1, y_2 is non-zero, $\{y_1, y_2\}$ is a fundamental generating set of solutions and therefore

$$y = C_1 y_1 + C_2 y_2$$

which shows that y is a linear combination of y_1 and y_2. Consequently $\{y_1, y_2\}$ is indeed a basis. Since all bases (provided they exist) have the same number of elements, we conclude that $\ker(L)$ is a two-dimensional vector space. □

In studying the theorem above and thinking carefully, we see that two functions y_1 and y_2 form a basis for the $\ker(L)$ if and only if they form a fundamental generating set of solutions to (3.4). Thus the notions (of bases and fundamental generating sets) are equivalent. In this language, let us end this section by stating a theorem which aggregates all of this information.

Theorem 3.4.6. *Given any two elements $y_1, y_2 \in \ker(L)$, the following are equivalent.*

1. *y_1 and y_2 form a fundamental generating set of solutions to* (3.4).
2. *y_1 and y_2 form a basis for* $\ker(L)$.
3. *y_1 and y_2 are linearly independent.*
4. *$w_{y_1,y_2}(t) \neq 0$ for some $t \in I$.*
5. *$w_{y_1,y_2}(t) \neq 0$ for all $t \in I$.*

To check your understanding, I encourage you to check all of the implications of the theorem and point to the relevant results coming from the previous section. It is straightforward to argue that (1) $\implies$ (2) $\implies$ (3) $\implies$ (4) $\implies$ (5) $\implies$ (1).

Exercise 31

Consider $D^2 : C^2(I) \to C^0(I)$ defined for $y \in C^2(I)$ by

$$D^2[y](t) = y''(t).$$

As D^2 is of the form $D^2[y] = y'' + py' + qy$ where $p = q = 0$, D^2 is indeed a second order differential operator. Let's denote by $\mathcal{P}_1(I)$ the vector space of first degree polynomial function on the interval I, i.e.,

$$\mathcal{P}_1(I) = \{f : I \to \mathbb{R} : f(t) = a_0 + a_1 t \text{ for } t \in I \text{ where } a_0, a_1 \in \mathbb{R}\}.$$

1. Show that any $f \in \mathcal{P}_1(I)$ satisfies $D^2[f] = 0$. In other words, show that

$$\mathcal{P}_1(I) \subseteq \ker(D^2).$$

2. Using the fact that $f'(t) = 0$ on I if and only if f is a constant function, show

that

$$\ker(D^2) \subseteq \mathcal{P}_1(I).$$

In this way, your have shown that

$$\ker(D^2) = \mathcal{P}_1(I)$$

or, equivalently, every solution to the second order differential equation $y'' = 0$ is of the form $y(t) = a_1 t + a_0$.

3.5 Producing Solutions

In the preceding sections, we developed a general theory of second-order linear homogeneous ordinary differential equations. Almost all of the results in this theory started with the assumption that, somehow, two solutions y_1 and y_2 to the equation

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

were known. Using Abel's identity, we were able to further to reduce our efforts by being able to obtain a second linearly independent solution y_2 given that we knew a single non-zero solution y_1 – this however still involved considerable effort. The question remains: *How can we produce linearly independent solutions y_1 and y_2 or, at least, a non-zero solution y_1?* Though there is a fairly general theory for producing such solutions, it is fairly complicated and involves the matrix exponential. In this section, we shall restrict our attention to two special classes of linear second order differential equations for which finding solutions y_1 and y_2 is easy.
Note Here

3.5.1 Second-order equations with constant coefficients

Consider a second-order linear homogeneous equation of the form

$$L[y] = y'' + by' + cy = 0 \tag{3.19}$$

where $b, c \in \mathbb{R}$. Such an equation is said to have constant coefficients, i.e., where $p(t) = b$ and $q(t) = c$ are simply constant functions on $I = \mathbb{R}$. To produce solutions to this equation, we put forth the following idea: As a solution y to (3.19) must have the property that a (constant) linear combination of y'', y' and y must sum to zero, we expect that the derivatives of the function y must themselves look like y, i.e., be constant multiples of themselves. As the exponential function is the canonical function having this property, it is reasonable to expect a solution to (3.19) to be of the form $y(t) = e^{rt}$ for some constant r. Taking this as an educated guess, often called an *ansatz*, we substitute $y(t) = e^{rt}$ into (3.19) to find

$$0 = y''(t) + by'(t) + cy(t) = r^2 e^{rt} + bre^{rt} + ce^{rt} = (r^2 + br + c)e^{rt}$$

For this identity hold for all $t \in \mathbb{R}$, it is necessary and sufficient that the number r is such that

$$r^2 + br + c = 0. \tag{3.20}$$

This quadratic equation (3.20) is called the *characteristic equation* for the differential equation 3.19. Let's work through an example.

Example 7

Consider the constant-coefficient equation

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

In view of the observations above, or by making the ansatz that a solution is of the form e^{rt}, we obtain the corresponding characteristic equation

$$r^2 - r - 6 = 0.$$

Thinking back to your training in introductory algebra, we can solve this polynomial equation by factoring, which is easy in this case, or by using the quadratic formula. By the first method, we have

$$(r-3)(r+2) = r^2 - r - 6 = 0$$

thus obtaining two distinct solutions $r = 3$ and $r = -2$. This is excellent for, in view of our arguments preceding this example, this gives us candidates for two solutions, $y_1(t) = e^{3t}$ and $y_2(t) = e^{-2t}$. Let's verify that $y_1(t)$ is indeed a solution. We have $y_1'(t) = 3e^{3t}$ and $y_1''(t) = 9e^{3t}$ and so

$$y_1''(t) - y_1'(t) - 6y_1(t) = 9e^{3t} - 3e^{3t} - 6e^{3t} = 0e^{3t} = 0$$

for all $t \in \mathbb{R}$ and hence y_1 solves the equation. You should verify for yourself that $y_2(t) = e^{-2t}$ also works. Observe that

$$w_{y_1,y_2}(t) = \det\begin{pmatrix} e^{3t} & e^{-2t} \\ 3e^{3t} & -2e^{-2t} \end{pmatrix} = -2e^{3t}e^{-2t} - 3e^{3t}e^{-2t} = -5e^t \neq 0$$

and therefore y_1 and y_2 form a fundamental generating set of solutions to the given equation.

In generalizing the example above, we have the following proposition.

Proposition 3.5.1. *If the characteristic equation* (3.20) *has distinct real roots* r_1 *and* r_2*, then* $y_1(t) = e^{r_1 t}$ *and* $y_2(t) = e^{r_2 t}$ *form a fundamental generating set of solutions to* (3.19)*.*

Proof. In view of our previous arguments, it is evident that $y_1(t) = e^{r_1 t}$ and $y_2(t) = e^{r_2 t}$ solve the differential equation 3.19. It remains to show that they are linearly independent. To see this, we observe that

$$w_{y_1,y_2}(t) = \det\begin{pmatrix} e^{r_1 t} & e^{r_2 t} \\ r_1 e^{r_1 t} & r_2 e^{r_2 t} \end{pmatrix} = r_2 e^{r_1 t} e^{r_2 t} - r_1 e^{r_1 t} e^{r_2 t} = (r_2 - r_1)e^{(r_1+r_2)t}$$

for all t and this non-zero in view of our assumption that r_1 and r_2 are distinct. $\square$

As you likely remember from your introductory algebra courses, there are other possibilities for solving polynomial equations beyond the case of distinct real roots. In fact, there are three possibilities for solutions to the (quadratic) characteristic equation (3.20) given $b, c \in \mathbb{R}$:

i. It has distinct real roots r_1 and r_2.

ii. It has a single real root r of multiplicity 2.

iii. It has complex roots which are complex conjugates.

This trichotomy can be seen as a result of the quadratic formula

$$r = \frac{-b \pm \sqrt{b^2 - 4c}}{2},$$

applied to the (3.20). If $b^2 - 4c > 0$, (3.20) has the distinct real roots

$$r_1 = \frac{-b + \sqrt{b^2 - 4c}}{2} \quad \text{and} \quad r_2 = \frac{-b - \sqrt{b^2 - 4c}}{2}$$

and, in this case, Proposition 3.5.1 applies. In the case that $b^2 - 4c = 0$, (3.20) has a single root $r = -b/2$ of multiplicity two and the corresponding polynomial is easily seen to be a perfect square, i.e.,

$$r^2 + br + c = (r - a)^2 = 0$$

where $a = -b/2$. Correspondingly, the differential equations takes the form

$$L[y] = y'' - 2ay' + a^2 y = 0. \tag{3.21}$$

Given that $r = a$ solves the characteristic equation, we observe that $y_1(t) = e^{at}$ is a solution to (3.21). In contrast to the case with distinct real roots, the trouble here is to find a second linearly independent solution. Such a solution can be produced using Abel's identity and you should pursue this if you're curious. A second linearly independent solution can also be found using linear algebra, as the following two exercises show.

Exercise 32: Similarity Transformation

Let V be a vector space and let L and M be linear operators from V to V, i.e. mapping V into itself. Assume that M is invertible (an isomorphism) with inverse M^{-1} and set $S = M^{-1} \circ L \circ M$ and note that $M \circ S = L \circ M$ (and so they say that L and S are similar).

1. Given $v \in \ker(S)$, show that $Mv \in \ker(L)$.

2. If $\{v_1, v_2, \dots, v_d\}$ is a basis for $\ker(S)$, show that $\{M(v_1), M(v_2), \dots, M(v_d)\}$ is a basis for $\ker(L)$. Hint: We've already discussed how isomorphisms preserve linear independence, so necessarily $\{M(v_1), M(v_2), \dots, M(v_d)\}$ is a linearly independent set. Your job is to show that this collection spans $\ker(L)$.

Exercise 33

In this exercise we seek a linearly independent pair of solutions to the differential equation

$$L[y] = y'' - 2ay' + a^2 y = 0.$$

Equivalently, we seek a basis for $\ker(L)$. To achieve our goal, we take motivation for the previous exercise and define a linear operator M_a defined by

$$M_a[f](t) = e^{at} f(t)$$

where f is a function (which can be taken in $C^2(I)$ or $C^0(I)$).

1. Show that the inverse of M_a (as an operator from $C^2(I)$ onto $C^2(I)$) is given by
$$(M_a)^{-1}[f](t) = e^{-at}f(t) = M_{-a}[f](t)$$

2. For any $y \in C^2(I)$, show that
$$((M_a)^{-1} \circ L \circ M_a)[y] = D^2[y] \tag{3.22}$$
where D^2 is the operator defined by $D^2[y] = y''$ for $y \in C^2(I)$. In this way you show that L and D^2 are similar in the sense of the previous exercise.

3. As we saw in Exercise 31, $\ker(D^2) = \mathcal{P}_1(I)$ and so a basis for the $\ker(D^2)$ is $\{1, t\}$ (the constant function and the identity function). Use this fact and the result of the previous exercise to find a basis for $\ker(L)$.

4. Verify directly that the functions you obtained, $y_1(t)$ and $y_2(t)$, satisfy the differential equation $L[y] = 0$.

5. Using the Wronskian determinant, verify that these solutions are linearly independent.

Example 8

Consider the constant-coefficient second-order linear homogeneous differential equation
$$y'' - 2y' + y = 0.$$
Its associated characteristic equation is
$$r^2 - 2r + 1 = (r-1)^2 = 0$$
and so we obtain $r = 1$ as a solution of multiplicity 2. Immediately, we have $y_1(t) = e^t$ as a solution and, in view of the previous exercise, we expect $y_2(t) = te^t$ to also be a solution. Let's verify this. We have
$$\begin{aligned} y_2''(t) - 2y_2'(t) + y_2(t) &= (te^t)'' - 2(te^t)' + (te^t) \\ &= (e^t + te^t)' - 2(e^t + te^t) + te^t \\ &= (e^t + e^t + te^t) - 2e^t - 2te^t + te^t \\ &= 2e^t - 2e^t + te^t - 2te^t - te^t \\ &= 0 \end{aligned}$$
which holds for all $t \in \mathbb{R}$. Hence, $y_2(t) = te^t$ satisfies the differential equation. Observe now that
$$w_{y_1,y_2}(t) = y_1'(t)y_2(t) - y_1(t)y_2'(t) = e^t(te^t) - e^t(e^t + te^t) = -e^{2t} \neq 0$$
for all $t \in \mathbb{R}$ and therefore $\{y_1, y_2\}$ form a fundamental generating set of solutions to the given differential equation.

In the final case that $b^2 - 4c < 0$, (3.20) has complex roots

$$\begin{aligned} r &= \frac{-b \pm \sqrt{b^2 - 4c}}{2} = -\frac{b}{2} \pm \frac{\sqrt{(4c - b^2)(-1)}}{2} = -\frac{b}{2} \pm \frac{\sqrt{4c - b^2}\sqrt{-1}}{2} \\ &= \left(-\frac{b}{2}\right) \pm \left(\frac{\sqrt{4c - b^2}}{2}\right) i \end{aligned}$$

where $i = \sqrt{-1}$ is the imaginary unit. We write this as

$$r = \alpha \pm \beta i$$

where $\alpha = -b/2$ and $\beta = \sqrt{4c - b^2}/2$. In following our standard prescription for producing solutions, we consider the functions,

$$y_1(t) = e^{(\alpha + i\beta)t} = e^{\alpha t + i\beta t} \qquad \text{and} \qquad y_2(t) = e^{(\alpha - i\beta)t} = e^{\alpha t - i\beta t}.$$

defined[11] for $t \in \mathbb{R}$. These complex-valued functions are, in fact, solutions to (3.19) as you will show directly in Exercise 35. To this end, it will be necessary to have a notion of the "derivative" of a complex-valued function of a real variable t which you will develop in that exercise. For now, however, we shall not worry about the complex nature of things and manipulate these functions as if they were the real-valued function with which we are familiar. As a consequence of Euler's identity[12], i.e., that

$$e^{i\theta} = \cos(\theta) + i\sin(\theta) \tag{3.23}$$

for all θ, we have

$$\begin{aligned} y_1(t) &= e^{(\alpha + i\beta)t} = e^{\alpha t} e^{i\beta t} = e^{\alpha t}(\cos(\beta t) + i\sin(\beta t)) \\ &= e^{\alpha t}\cos(\beta t) + ie^{\alpha t}\sin(\beta t) \end{aligned} \tag{3.24}$$

and

$$\begin{aligned} y_2(t) &= e^{(\alpha - i\beta)t} = e^{\alpha t} e^{-i\beta t} = e^{\alpha t}(\cos(-\beta t) + i\sin(-\beta t)) \\ &= e^{\alpha t}(\cos(\beta t) - i\sin(\beta t)) \\ &= e^{\alpha t}\cos(\beta t) - ie^{\alpha t}\sin(\beta t) \end{aligned}$$

for $t \in \mathbb{R}$. Upon recalling that $\ker(L)$ is a vector space and, as we're taking for granted that $y_1, y_2 \in \ker(L)$,

$$\tilde{y}_1(t) = \frac{1}{2}\left(y_1(t) + y_2(t)\right) = \frac{1}{2}\left(2e^{\alpha t}\cos(\beta t)\right) = e^{\alpha t}\cos(\beta t) \in \ker(L),$$

i.e., $\tilde{y}_1(t) = e^{\alpha t}\cos(\beta t)$ is a solution to (3.19). By a similar argument[13], we have

$$\tilde{y}_2(t) = \frac{1}{2i}(y_1(t) - y_2(t)) = e^{\alpha t}\sin(\beta t) \in \ker(L),$$

i.e., $\tilde{y}_2(t) = e^{\alpha t}\sin(\beta t)$ is a solution to (3.19). For simplicity, let's rename these solutions as y_1 and y_2, that is, we set

$$y_1(t) = e^{\alpha t}\cos(\beta t) \qquad \text{and} \qquad y_2(t) = e^{\alpha t}\sin(\beta t) \tag{3.25}$$

[11] The complex exponential function $z \mapsto e^z$ is defined via power series and behaves just as you expect. For example, for any complex numbers z and w, $e^{z+w} = e^z e^w$.

[12] If you haven't seen Euler's identity before, it is easily derived by simplifying the Maclaurin series for $e^{i\theta}$. Further, evaluating the identity at $\theta = \pi$ yields what some call the most beautiful equation in mathematics, $e^{i\pi} + 1 = 0$.

[13] Here we are actually using the fact that $\ker(L)$ is a vector space over $\mathbb{C}$ and so we can scale by $1/2i$ without trouble

for $t \in \mathbb{R}$. As these solutions are real-valued, you'll likely find them more tractable to work with than the complex-valued functions from which they were derived[14]. For the solutions y_1 and y_2 to (3.19), observe that

$$\begin{aligned} w_{y_1,y_2}(t) &= \det\begin{pmatrix} e^{\alpha t}\cos(\beta t) & e^{\alpha t}\sin(\beta t) \\ \alpha e^{\alpha t}\cos(\beta t) - \beta e^{\alpha t}\sin(\beta t) & \alpha e^{\alpha t}\sin(\beta t) + \beta e^{\alpha t}\cos(\beta t) \end{pmatrix} \\ &= e^{2\alpha t}\left(\alpha\cos(\beta t)\sin(\beta t) + \beta\cos^2(\beta t) - \alpha\cos(\beta t)\sin(\beta t) + \beta\sin^2(\beta t)\right) \\ &= \beta e^{2\alpha t}(\cos^2(\beta t) + \sin^2(\beta t) \\ &= \beta e^{2\alpha t} \end{aligned}$$

for all $t \in \mathbb{R}$. This, of course, never vanishes because $\beta = \sqrt{4c-b^2}/2 \neq 0$ and we therefore conclude that $y_1(t) = e^{\alpha t}\cos(\beta t)$ and $y_2(t) = e^{\alpha t}\sin(\beta t)$ form a fundamental generating set of solutions to (3.19). We summarize our conclusions as follows.

Proposition 3.5.2. *Consider the second-order constant coefficient linear homogeneous differential equation* (3.19) *and its characteristic equation* (3.20). *If* $b^2 - 4c < 0$, (3.20) *has complex-roots* $\alpha \pm i\beta$ *where* $\alpha = -b/2$ *and* $\beta = \sqrt{4c-b^2}/2$. *In this case,* $\{y_1, y_2\}$, *defined by* (3.25), *is a fundamental generating set of solutions to* (3.19).

Remark 3.5.3. The above argument took for granted that all formal computations involving complex-valued functions worked just as they do for real-valued functions. In not taking these computations for granted, to truly prove the above theorem, one needs to verify directly that y_1 and y_2, defined by (3.25) are solutions to (3.19) provided $b^2-4c<0$.

Exercise 34

Carry out the instructions in the remark above. That is, directly verify that y_1 and y_2, defined by (3.25) are solutions to (3.19) provided $b^2 - 4c < 0$.

Exercise 35: Complex differentiation

As stated above, we can actually work directly with the complex exponential function and verify that $e^{(\alpha+i\beta)t}$ and $e^{(\alpha-i\beta)t}$ form a perfectly good fundamental generating set of solutions to (3.19) provided $b^2-4c<0$. Before we do this, we must understand what it means to differentiate a complex-valued function of a real variable. To this end, let $f:\mathbb{R}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$ and define $u:\mathbb{R}\to\mathbb{C}$ by

$$u(t) = f(t) + ig(t)$$

for $t \in \mathbb{R}$. We say that u is differentiable at t if f and g are differentiable at t and, in this case, the derivative of u at t is

$$u'(t) = f'(t) + ig'(t).$$

As with real-valued functions, differentiation of complex-valued functions is linear. For instance, given a complex number $r = \alpha + i\beta$ where $\alpha, \beta \in \mathbb{R}$,

$$\frac{d}{dt}(ru(t)) = ru'(t). \tag{3.26}$$

[14] It is my hope that, after doing Exercise 35, you will be just as comfortable working with the complex exponentials, which I personally find easier.

In this exercise, we study some properties of the complex exponential function and its derivatives.

1. Though it seems obvious, verify (3.26). Note that $ru(t) = (\alpha + i\beta)(f(t) + ig(t)) = \alpha f(t) - \beta g(t) + i(\beta f(t) + \alpha g(t))$ and $ru'(t) = (\alpha + i\beta)(f'(t) + ig'(t)) = \alpha f'(t) - \beta g'(t) + i(\beta f'(t) + \alpha g'(t))$

2. Using (3.24) and the above definition of differentiability of a complex-valued function, show that
$$\frac{d}{dt}e^{t(\alpha+i\beta)} = (\alpha + i\beta)e^{t(\alpha+i\beta)}$$
and
$$\frac{d^2}{dt^2}e^{t(\alpha+i\beta)} = (\alpha + i\beta)^2 e^{t(\alpha+i\beta)}$$
Hint: Confirm the first equality directly. To confirm the second, apply the first equality and (3.26).

3. Using your result above, applied to $\alpha + i\beta = 0 + i\delta$ for a real number $\delta \neq 0$, verify that $y_1(t) = e^{i\delta t}$ satisfies the differential equation
$$L[y] = y'' + \delta^2 y = 0.$$
Verify that $y_2(t) = e^{-i\delta t}$ also solves the above equation. Using the Wronskian, verify that y_1 and y_2 are linearly independent and hence form a fundamental generating set of solutions. Note: The Wronskian test is applied in the same way to complex-valued solutions as it is for real-valued ones.

4. At this point, I suspect that you might be worried about the complex-valued nature of things. You might ask: If the solution set I found above consists of complex-valued functions, how can I solve initial value problems with real-valued initial conditions given that my generating set of solutions is complex-valued? To address this, let's consider the initial value problem
$$\begin{cases} L[y] = y'' + \delta^2 y = 0 & y(0) = y_0 \text{ and } y'(0) = y_0' \end{cases}$$
where y_0 and y_0' are real numbers. As our theory states, the general solution to the above initial value problem is
$$y(t) = C_1 y_1(t) + C_2 y_2(t) = C_1 e^{i\delta t} + C_2 e^{-i\delta t}.$$
Using this general solution, solve the initial value problem directly. You will need to write the constants C_1 and C_2 in terms of y_0 and y_0'. Using the formulas for cosine and sine found in the previous exercises, simplify your result as much as possible and confirm that it is in fact real-valued.

In this section, we considered a method for producing solution to the second-order constant-coefficient linear homogeneous equation (3.19). In the course of this, we were asked to solve an auxiliary polynomial equation (3.20) called the characteristic equation. We found that the characteristic equation had three types of solutions, all found using the quadratic formula, and these solutions led us to three possibilities for to (3.19). We treated each case separately and wrote down the relevant result for that case; these were Proposition (3.5.1), Exercise 33 and Proposition (3.5.2). For your convenience, we aggregate all of our results into the following theorem.

Theorem 3.5.4. *Let $b, c \in \mathbb{R}$ and consider the constant-coefficient linear differential* *equation*

$$L[y] = y'' + by' + cy = 0 \tag{3.27}$$

and its corresponding polynomial equation

$$r^2 + br + c = 0. \tag{3.28}$$

1. *If* $b^2 > 4c$*, set*

$$r_1 = -\frac{b}{2} + \frac{\sqrt{b^2 - 4c}}{2} \quad \textit{and} \quad r_2 = -\frac{b}{2} - \frac{\sqrt{b^2 - 4c}}{2};$$

   these are distinct real solutions to (3.28)*. Then* $y_1(t) = e^{r_1 t}$ *and* $y_2(t) = e^{r_2 t}$ *form a fundamental generating set of solutions to* (3.27)*.*

2. *If* $b^2 = 4c$*, set* $r = -b/2$ *(the unique solution to* (3.28)*, a root of multiplicity* 2*). Then* $y_1(t) = e^{rt}$ *and* $y_2(t) = te^{rt}$ *form a fundamental generating set of solutions to* (3.27)*.*

3. *If* $b^2 < 4c$*, set*

$$\alpha = -\frac{b}{2} \quad \textit{and} \quad \beta = \frac{\sqrt{4c - b^2}}{2}.$$

   Then $r_1 = \alpha + i\beta$ *and* $r_2 = \alpha - i\beta$ *are distinct complex solutions to* (3.28) *and* $e^{r_1 t}$ *and* $e^{r_2 t}$ *form a fundamental generating set of solutions to* (3.27)*. Equivalently* $y_1(t) = e^{\alpha t}\cos(\beta t)$ *and* $y_2(t) = e^{\alpha t}\sin(\beta t)$ *form a fundamental generating set of solutions to* (3.27)*.*

Exercise 36

For the three initial value problems below, use Theorem 3.5.4 to find a fundamental generating set of solutions $\{y_1, y_2\}$ to the differential equation, use the Wronskian determinant to verify that the solutions are linearly independent, and find constants C_1 and C_2 for which $y(t) = C_1 y_1(t) + C_2 y_2(t)$ solve the initial value problem.

1. $$\begin{cases} y'' - y' - 12y = 0, \quad y(0) = 7, \ \ y'(0) = 0 \\ \end{cases}$$

2. $$\begin{cases} y'' + 6y' + 9y = 0, \quad y(0) = 1, \ \ y'(0) = -2 \end{cases}$$

3. $$\begin{cases} y'' - 2y' + 2y = 0, \quad y(0) = 1, \ \ y'(0) = 1 \end{cases}$$

3.5.2 Power Series Solutions

In the previous subsection, we took the task of finding fundamental generating sets of solutions to the second-order constant-coefficient linear homogeneous equation 3.19. As you expect, the coefficients being constant made the problem extremely tractable and we were able to establish a general theory for such cases, culminating in Theorem 3.5.4. Moving beyond the constant-coefficient realm, we here study another useful method for producing solutions. We shall generally consider second-order equations in the form

$$a_2(t)y'' + a_1(t)y' + a_0(t)y = 0 \tag{3.29}$$

with corresponding initial value problem

$$\begin{cases} a_2(t)y'' + a_1(t)y' + a_0(t)y = 0 \\ y(t_0) = y_0 \\ y'(t_0) = y_0'. \end{cases}$$

where a_0, a_1 and a_2 are continuous real-valued functions defined on an interval I and $t_0 \in I$. In fact, we shall assume much more than continuity for the functions a_0, a_1 and a_2: We shall assume that these functions are real-analytic on the interval I. This means that, for each $t_0 \in I$, the functions a_0, a_1 and a_2 have convergent power series representations centered at t_0 with positive radii of convergence.

To make things relatively simple, we shall only treat the cases in which a_0, a_1 and a_2 are polynomials and are centered at $t_0 = 0$, i.e., we shall assume that these functions are all of the form

$$b_0 + b_1 t + b_2 t^2 + \cdots b_n t^n$$

for given constants $b_0, b_1, \dots, b_n$. Having the coefficients centered at $t_0 = 0$ will make our computations relatively simple; it also makes our approach well-suited to solving initial value problems corresponding to initial time $t_0 = 0$. In view of these assumptions, to solve (3.29), we assume that solutions are themselves representable via power series centered at $t_0 = 0$. In other words, we seek solutions of the form

$$y(t) = \sum_{n=0}^{\infty} c_n t^n$$

where the coefficients $\{c_n\}$ are to be determined. Provided that this series has positive radius of convergence R, it is infinitely differentiable at $t_0 = 0$ and we have

$$y'(t) = \sum_{n=1}^{\infty} c_n n t^{n-1}$$

and

$$y''(t) = \sum_{n=2}^{\infty} c_n n(n-1) t^{n-2},$$

both of which have radius of convergence R. To determine the coefficients $\{c_n\}$ we can now substitute the power series for y, y' and y'' into the differential equation (3.29) and combine like terms. This, generally, will yield a recurrence relation for $\{c_n\}$, called an *indicial equation*. From it, we are able to determine the values of $\{c_n\}$ given that c_0 and c_1 are specified. The best way to see this is to start with an example.

Example 9

Consider the equation

$$y'' + 4y = 0.$$

Of course, we already know how to find a general solution to this equation using Theorem 3.5.4 (and you should do this before reading further). Let's however proceed by the method described above and seek solutions of the form

$$y(t) = \sum_{n=0}^{\infty} c_n t^n.$$

As we argued above, we have

$$y''(t) = \sum_{n=2}^{\infty} c_n n(n-1)t^{n-2}.$$

Consequently, for y to solve the given differential equation, we must have

$$0 = y''(t) + 4y(t) = \sum_{n=2}^{\infty} c_n n(n-1)t^{n-2} + 4\sum_{n=0}^{\infty} c_n t^n. \tag{3.30}$$

To combine like terms (like monomials), it is helpful to reindex the terms of both series so that they are written in terms of t^k. For the series for y'', we want $t^{n-2} = t^k$ and we therefore set $k = n-2$ to reindex this series. With this choice, the summation must start at $k = 0$ and each n gets replaced by $k+2$ and therefore[a]

$$\sum_{n=2}^{\infty} c_n n(n-1)t^{n-2} = \sum_{k=0}^{\infty} c_{k+2}(k+2)(k+1)t^k.$$

For the series for $y(t)$, we simply set $k = n$ and write

$$4\sum_{n=0}^{\infty} c_n t^n = 4\sum_{k=0}^{\infty} c_k t^k = \sum_{k=0}^{\infty} 4c_k t^k.$$

Upon inserting these results into (3.30) and combining the result into a single power series, we obtain

$$0 = \sum_{k=0}^{\infty} c_{k+2}(k+2)(k+1)t^k + \sum_{k=0}^{\infty} 4c_k t^k = \sum_{k=0}^{\infty}(c_{k+2}(k+2)(k+1) + 4c_k)t^k.$$

For this resultant series

$$\sum_{k=0}^{\infty}(c_{k+2}(k+2)(k+1) + 4c_k)t^k$$

to be identically zero, all of its coefficients must be zero. We therefore have

$$c_{k+2}(k+2)(k+1) + 4c_k = 0$$

for all $k = 0, 1, 2, \ldots$. This is the indicial equation; it shows, in particular, that each term (c_{k+2}) is determined by its penultimate term. We rewrite this as

$$c_{k+2} = -\frac{4}{(k+2)(k+1)}c_k$$

for $k \geq 0$ or, equivalently,

$$c_n = -\frac{4}{n(n-1)}c_{n-2} \tag{3.31}$$

for $n \geq 2$. As discussed above, we observe that all of the coefficients for the series for y are determined as as soon as the first two, c_0 and c_1 are specified. This amounts to specifying two initial conditions for y (and y'). To make things easy, let's make the choice $c_0 = 1$ and $c_1 = 0$. From (3.31), we immediately see that

$0 = c_1 = c_3 = c_5 = \cdots$, i.e., $c_k = 0$ for all odd indicies k. The terms with even indices are non-zero. We have

$$c_2 = -\frac{4}{2\cdot 1}\cdot 1, \qquad c_4 = -\frac{4}{4\cdot 3}c_2 = -\frac{4}{4\cdot 3}\left(-\frac{4}{2\cdot 1}\right) = \frac{(-1)^2 4^2}{4\cdot 3\cdot 2\cdot 1},$$

$$c_6 = -\frac{4}{6\cdot 5}\left(\frac{(-1)^2 4^2}{4\cdot 3\cdot 2\cdot 1}\right) = \frac{(-1)^3 4^3}{6!}$$

and continuing the pattern, we see that

$$c_n = \frac{(-1)^{n/2}(4)^{n/2}}{n!} = \frac{(-1)^{n/2}}{n!}2^n \tag{3.32}$$

for all even natural numbers n. Thus, our solution is

$$\begin{aligned} y(t) &= \sum_{n=0,2,4,6,\dots}^{\infty} \frac{(-1)^{n/2}}{n!}2^n t^n = \sum_{n=0,2,4,6,\dots}^{\infty} \frac{(-1)^{n/2}}{n!}(2t)^n \\ &= 1 - \frac{(2t)^2}{2!} + \frac{(2t)^4}{4!} - \frac{(2t)^6}{6!} + \frac{(2t)^8}{8!} - \frac{(2t)^{10}}{10!} + \cdots. \end{aligned}$$

In thinking back to your training to power series, I hope you recognize that this solution is precisely the function $y(t) = \cos(2t)$. If we had, instead, selected $c_0 = 0$ and $c_1 = 1$, we would have produced by an analogous argument the solution $y(t) = \sin(2t)$. Consequently this method of solution via power series has produced two solutions to the given differential equation:

$$y_1(t) = \cos(2t) \qquad \text{and} \qquad y_2(t) = \sin(2t)$$

which you can easily verify to be linearly independent.

[a] You should think carefully about why this works. To see things clearly, it is often helpful to write out the first few terms of the series to make sure your reindexing worked correctly.

You can think of the procedure outlined in the preceding example as a prescription. We can state it in the following way:

Power Series Solutions

Given real analytic functions $a_2(t)$, $a_1(t)$, and $a_0(t)$ (which can simply be polynomials) with convergent Maclaurin series expansions, we consider the second order linear homogeneous equation

$$a_2(t)y'' + a_1(t)y' + a_0(t)y = 0$$

defined on a symmetric interval $I = (-R, R)$ of $t_0 = 0$ on which the Maclaurin series expansions for $a_2(t)$, $a_1(t)$, and $a_0(t)$ are absolutely convergent. To produce linearly independent solutions, we assume first that a solution $y(t)$ can be expressed as a convergent Maclaurin series of the form

$$y(t) = \sum_{n=0}^{\infty} c_n t^n$$

which is convergent on some interval about $t_0 = 0$. We then do the following:

1. Compute $y'(t)$ and $y''(t)$.
2. Plug both into the differential equation and using the rules of power series manipulation, simplify each term to produce an equation asserting that the sum of (at most three) power series is zero.
3. By reindexing, express all power series in the form

$$\sum_{n}^{\infty} (\text{Coefficients})t^n.$$

4. By removing a finite number of lower order monomial terms, you should be able to write your result as

$$d_0 + d_1 t + \cdots + d_{n_0-1}t^{n_0-1} + \sum_{n=n_0}^{\infty} (\text{Coefficients})t^n = 0$$

5. Using linear independence, you can conclude that the polynomial coefficients are zero and that the coefficients (labeled "Coefficients") of the remaining series, which will depend on integers and the coefficients of y, p and q, must also be zero. This will lead to an "indicial equation" from which you can relate each coefficient c_n to $c_1, c_2, \ldots, c_{n-1}$.
6. Finally, you will have freedom to choose some coefficients of $y(t)$. Your choices will lead, though the indicial equation, to the unique determination of the corresponding power series. Your freedom of choice for these remained coefficients can be varied to obtain linearly independent solutions.

The essential ingredient in the prescription above is the assumption that, when the coefficients $a_2(t)$, $a_1(t)$, and $a_0(t)$ are real analytic (have convergent Taylor expansions), a solution $y(t)$ will a also be real analytic (and hence you can assume it is given by a power series). This highly non-trivial assumption is a consequence of a deep theorem in differential equations. The theorem is due to Augustin-Louis Cauchy and Sofya Kovalevskaya and is eponymously called the Cauchy-Kovalevskaya Theorem [5]. Though Cauchy initially proved a special case of the theorem, Kovalevskaya was the mathematician who obtain the full result in 1875 which is stated in the context of partial differential equations.

Example 10: Airy's Equation

For a non-zero real number k, consider the second order linear homogeneous ordinary differential equation

$$y'' + kty = 0 \tag{3.33}$$

on $I = \mathbb{R}$. This differential equation is called Airy's equation and is named after the mathematician and astronomer, George Biddell Airy. We see that the coefficients, $a_2(t) = 1$, $a_1(t) = 0$ and $a_0(t) = kt$ are both real analytic and so, in view of the preceding prescription, we assume that

$$y(t) = \sum_{n=0}^{\infty} c_n t^n$$

is a solution to (3.33) where the coefficients $\{c_n\}$ are to be determined. We compute

$$y'(t) = \sum_{n=1}^{\infty} c_n n t^{n-1} \quad \text{and} \quad y''(t) = \sum_{n=2}^{\infty} c_n n(n-1)t^{n-2}.$$

By plugging this expression for $y''(t)$ and $y(t)$ in (3.33), we obtain

$$\begin{aligned} 0 &= y''(t) + kty(t) \\ &= \sum_{n=2}^{\infty} c_n n(n-1)t^{n-2} + kt\sum_{n=0}^{\infty} c_n t^n \\ &= \sum_{n=2}^{\infty} c_n n(n-1)t^{n-2} + \sum_{n=0}^{\infty} c_n k t^{n+1}. \end{aligned}$$

Taking cues from Step 3 of the prescription outlined above, we reindex the first series by sending $n \to n+2$ so that

$$\sum_{n=2}^{\infty} c_n n(n-1)t^{n-2} = \sum_{n=0}^{\infty} c_{n+2}(n+2)(n+1)t^n.$$

To reindex the second series, we send $n \to n-1$ so that

$$\sum_{n=0}^{\infty} c_n k t^{n+1} = \sum_{n=1}^{\infty} c_{n-1} k t^n.$$

Consequently, the above identity is

$$0 = \sum_{n=0}^{\infty} c_{n+2}(n+2)(n+1)t^n + \sum_{n=1}^{\infty} c_{n-1} k t^n.$$

By following Step 4 in the above prescription, in order to combine the above series expressions into a single expression, both series must start at $n_0 = 1$. Upon recognizing that

$$\sum_{n=0}^{\infty} c_{n+2}(n+2)(n+1)t^n = 2c_2 + \sum_{n=1} c_{n+2}(n+2)(n+1)t^n,$$

we obtain the identity

$$\begin{aligned} 0 &= 2c_2 + \sum_{n=1}^{\infty} c_{n+2}(n+2)(n+1)t^n + \sum_{n=1}^{\infty} c_{n-1}kt^n \\ &= 2c_2 + \sum_{n=1}^{\infty} \left(c_{n+2}(n+2)(n+1) + c_{n-1}k\right) t^n \end{aligned}$$

from which, in view of Step 5, we conclude that $c_2 = 0$ and

$$c_{n+2}(n+2)(n+1) + c_{n-1}k = 0$$

for all $n = 1, 2, \ldots$ in order for $y(t)$ to satisfy the differential equation. This says that $c_2 = 0$ and

$$c_{n+2} = -\frac{k}{(n+2)(n+1)}c_{n-1}$$

for $n \geq 1$; the latter is an indicial equation which gives a relationship that determines the coefficients $\{c_n\}$ iteratively. A final reindexing gives $c_2 = 0$ and the final (equivalent) indicial equation

$$c_{n+3} = -\frac{k}{(n+3)(n+2)}c_n \tag{3.34}$$

for $n = 0, 1, 2, \ldots$. In view of Step 6, we make the following observations from the indicial equation (3.34).

1. If c_0 is known (or chosen), then (3.34) determines the coefficients c_3, $c_6, c_9, \ldots$.

2. If c_1 is known (or chosen), then (3.34) determines the coefficients c_4, $c_7, c_{10}, \ldots$.

3. Since c_2 must be zero, (3.34) forces $c_5 = c_8 = c_{11} = \cdots = 0$.

With these observation, it is apparent that we are free to choose c_0 and c_1 and, once chosen, all other coefficients are determined uniquely and so the solution $y(t)$ is determined uniquely. Our freedom to choose two constants c_0 and c_1 coincides exactly with the fact that initial value problems come equipped with two initial conditions and equivalently that the $\ker(L)$ is two-dimensional. In this way, if we make two linearly independent choices for c_0 and c_1, we will obtain two linearly independent solutions. To this end, let's choose $c_0 = 1$ and $c_1 = 0$. Then, we have

$$c_3 = -\frac{k}{3 \cdot 2}c_0 = \frac{-k}{3 \cdot 2}$$

$$c_6 = -\frac{k}{6 \cdot 5}c_3 = \frac{k^2}{6 \cdot 5 \cdot 3 \cdot 2}$$

and, in general,

$$c_{3n} = \frac{(-1)^n k^n}{(3n)(3n-1)(3n-3)(3n-4)\cdots(3)(2)}.$$

for $n = 0, 1, 2, \ldots$. Also, since $c_1 = 0$ and $c_2 = 0$, we see that $c_1 = c_2 = c_4 = c_5 = c_7 = \cdots = 0$. Based on this choice of c_0 and c_1, the corresponding solution is

$$\begin{aligned} y_1(t) &= \sum_{n=0}^{\infty} c_{3n} t^{3n} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n k^n}{(3n)(3n-1)(3n-3)(3n-4)\cdots(3)(2)} t^{3n} \\ &= 1 - \frac{k}{6}t^3 + \frac{k^2}{180}t^6 - \frac{k^3}{12960}t^9 + \cdots + \end{aligned}$$

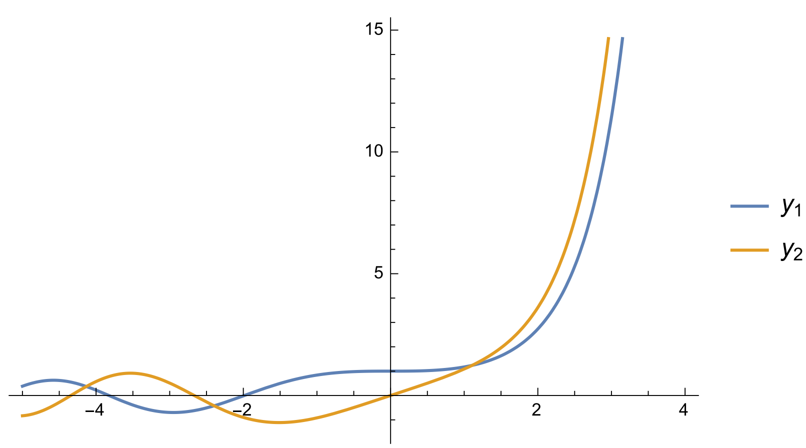
Making the choice $c_0 = 0$ and $c_1 = 1$, analogous reasoning gives us the solution

$$\begin{aligned} y_2(t) &= \sum_{n=0}^{\infty} \frac{(-1)^n k^n}{(3n+1)(3n)(3n-2)(3n-3)\cdots 4 \cdot 3} t^{3n+1} \\ &= t - \frac{k}{3 \cdot 4}t^4 + \frac{k^2}{504}t^7 - \frac{k^3}{54360}t^9 + \cdots \end{aligned}$$

It can be verified that both solutions y_1 and y_2 have infinite radii of convergence, i.e., they converge absolutely on $\mathbb{R}$. By design, we note that

$$w_{y_1, y_2}(0) = 1 \cdot 1 - 0 \cdot 0 = 1 \neq 0$$

and so our solutions constructed must form a fundamental generating sets of solutions to (3.33). The following figure illustrated the solutions $y_1(t)$ and $y_2(t)$ to Airy's equation.



Exercise 37

By assuming each equation below may be solved by a power series of the form $y(t) = \sum_{n=0}^{\infty} c_n t^n$, solve each equation given the specified values of c_0 and c_1. In the process, please identify the indicial equation and give a non-recursive formula (analogous to (3.32), which could be defined piece-wise) for the coefficients c_n. If the power series obtained is easily identifiable as a function you know (as was $\cos(2t)$, say what this function is.

1. $$y'' - 9y = 0 \qquad c_0 = 1,\, c_1 = -3$$

2. $$y'' - t^2 y = 0 \qquad c_0 = 1,\, c_1 = 0$$

Exercise 38

Consider Bessel's differential equation with parameter $\nu = 0$:

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

1. Explain why this equation does not satisfy the hypotheses of Theorem 3.1.1 on any open interval I containing $t_0 = 0$.

As expected by your conclusion above, this differential equation is somewhat poorly behaved on intervals containing zero. For example, we cannot expect to find two linearly independent solutions y_1 and y_2 with which we can solve every initial value problem with initial time $t_0 = 0$. Still however, this differential equation is incredibly important[a] and the method of power series does help produce a solution which is well-behaved at $t_0 = 0$. To this end, assume that

$$y(t) = \sum_{n=0}^{\infty} c_n t^n$$

solves the given differential equation.

2. By assuming that $c_0 = 1$, which turns out to be enough to determine all remaining coefficients, find the corresponding solution.

[a] It appears, for instance, in the study of vibrations of a drum.

3.6 The inhomogeneous problem

In this section, we turn our focus to linear inhomogeneous second-order differential equations. These, as you recall, are equations of the form

$$L[y] = y'' + py' + qy = r \tag{3.35}$$

where p, q and r are continuous functions on an open interval I, i.e., $p, q, r \in C^0(I)$. As we will see, it is also important to study the homogeneous equation

$$L[y] = y'' + py' + qy = 0 \tag{3.36}$$

in conjunction with the corresponding inhomogeneous equation (3.35). Let's make an observation in this direction: Suppose that, somehow, you know a solution y_p to (3.35), called a *particular solution* . Then, given any solution y_h to the homogeneous equation (3.36), we have

$$L[y_p + y_h] = L[y_p] + L[y_h] = r + 0 = r$$

by virtue of the linearity of L. In other words, $y = y_p + y_h$ also satisfies (3.35) given any solution y_h to (3.36). As the next proposition shows, all solutions to (3.35) are of this form.

Proposition 3.6.1. *Let y_p be a particular solution to the inhomogeneous equation* (3.35). *Then every solution to* (3.35) *is of the form*

$$y = y_p + y_h$$

where y_h solves the homogeneous equation (3.36).

The proposition above allows you to produce any solution to the inhomogeneous problem (3.35) by simply knowing a single solution y_p. This gives an easy recipe for producing general solutions thereby solving initial value problems. Before proving the proposition, we summarize this simple observation as the following corollary.

Corollary 3.6.2. *Let y_p be a particular solution to the inhomogeneous equation* (3.35) *and let y_1 and y_2 form a fundamental generating set of solutions to the homogeneous equation* (3.36)*, equivalently, $\{y_1, y_2\}$ is a basis for* $\ker(L)$*. Then*

$$y = y_p + C_1 y_1 + C_2 y_2$$

is a general solution to (3.35) *in the sense that every initial value problem for* (3.35) *can be solved by specifying constants C_1 and C_2.*

Proof of Proposition 3.6.1. Given an arbitrary solution y to (3.35), consider the difference $y - y_p$ where y_p is the given particular solution. Observe that

$$L[y - y_p] = L[y] - L[y_p] = r - r = 0$$

by virtue of the linearity of L. This shows that $y - y_p \in \ker(L)$ and is therefore equal to some y_h solving the homogeneous equation (3.36). Consequently $y = y_p + y - y_p = y_p + y_h$, as desired. □

Example 11

Consider the initial value problem

$$\begin{cases} y'' + y = t \quad y(0) = 0 \quad y'(0) = 2. \end{cases}$$

To apply our theory above, we first focus on the inhomogeneous differential equation

$$y'' + y = t. \tag{3.37}$$

By observation (we will do this systematically in the next section), we see (or guess) that

$$y_p(t) = t$$

is a solution to (3.37) because

$$(y_p(t))'' + y_p(t) = \frac{d^2}{dt^2} t + t = 0 + t = t$$

for all $t \in \mathbb{R}$. Let's note that y_p does not solve the given initial value problem because $y_p'(0) = 1 \neq 2$, so we have to keep looking. We note that $y_1(t) = \cos(t)$ and $y_2(t) = \sin(t)$ form a fundamental generating set of solution to the homogeneous equation

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

Thus, in view of Corollary 3.6.2,

$$y = y_p + C_1 y_1 + C_2 y_2 = t + C_1 \cos(t) + C_2 \sin(t)$$

is a general solution to (3.37). Let's find the C_1 and C_2 giving the solution to the given initial value problem. We want

$$0 = y(0) = 0 + C_1 \cos(0) + C_2 \sin(0) = C_1$$

and

$$2 = y'(0) = \frac{d}{dt}(t + C_1 \cos(t) + C_2 \sin(t))|_{t=0} = 1 - C_1 \sin(0) + C_2 \cos(0) = 1 + C_2$$

and therefore $C_2 = 1$. Thus the initial value problem has the unique solution

$$y(t) = t + \sin(t).$$

In this section, we developed a theory for solving linear inhomogeneous second-order equations. As we saw, this theory is closely tied to the theory for solving homogeneous equations. Essentially, what we learned is this: To produce all solutions (or solve any initial value problem) for an inhomogeneous linear differential equation $L[y] = r$, you simply need to know one solution y_p, which we've called a particular solution. All other solutions are produced by adding something in $\ker(L)$. This sets us on the quest to find a method for producing particular solutions. As you say in the previous example, sometimes particular solutions can be guessed. In fact, in the next section we will study an effective and systematic method for guessing called the *method of undetermined coefficients*. Beyond the method of undetermined coefficients, the following two exercises highlight methods for finding particular solutions to inhomogeneous equations. First exercise considers the very special case that the differential equation does not explicitly involve y, i.e., $q = 0$. The second exercise employs power series methods.

Exercise 39

Consider the second order linear inhomogeneous differential equation

$$L[y] = y'' + y' = \sin(t) \tag{3.38}$$

with homogeneous counterpart

$$L[y] = y'' + y' = 0. \tag{3.39}$$

To find a general solution to (3.38), in view of Corollary 3.6.2, we must find a particular solution, y_p, to (3.38) and a fundamental generating set of solutions $\{y_1, y_2\}$ to (3.39). To this end, please do the following.

1. To find a particular solution to (3.38), we set $u = y'$ and obtain a first order linear equation in u. What is this equation?

2. Solve the equation you found for u. As we're only looking for a single particular solution, you integration constant C can be taken to be zero. Once you have the solution u, integrate (and again, your constant of integration can be taken to be zero) to find a particular solution y_p.
3. Verify that the solution you found, y_p, is indeed a solution to (3.38).
4. Now, find a fundamental generating set of solution to (3.39).
5. In view of Corollary 3.6.2, write down a general solution to (3.38).
6. Use your general solution to solve the initial value problem

$$\begin{cases} y'' + y' = \sin(t) \\ y(0) = 1 \\ y'(0) = 1 \end{cases} .$$

Exercise 40

Find a particular solution y_p to the linear second order inhomogeneous equation

$$y'' - y = 2e^t$$

by first recognizing that

$$e^t = \sum_{n=0}^{\infty} \frac{1}{n!} t^n = 1 + t + \frac{1}{2}t^2 + \frac{1}{6}t^3 + \cdots .$$

To do this, you should assume that the particular solution $y = y_p$ has convergent power series representation of the form

$$y(t) = \sum_{n=0}^{\infty} c_n t^n$$

and plug it in to the inhomogeneous equation (where the power series for $2e^t$ has been put on the right-hand side) and follow the usual steps to identify the coefficients $\{c_n\}$. As with solving homogeneous equations, you will have some choice with the first couple of coefficients. Please choose $c_0 = 0$ and $c_1 = 1$.

3.7 Undetermined Coefficients

In this section, we focus on producing particular solutions to second order linear constant-coefficient inhomogeneous differential equations, where the inhomogeneous term r is of a certain familiar form. That is, we consider the inhomogeneous equation

$$y'' + by' + cy = r(t) \tag{3.40}$$

where $r(t)$ is a polynomial, exponential function, trigonometric function or a linear combination thereof. As derivatives of such functions are of a similar character to the original function, to produce a solution to the inhomogeneous equation, it makes sense to guess that a particular solution y_p looks like the function $r(t)$. Making this educated guess is at the heart of the method of undetermined coefficients. Let's turn to an example to illustrate this process.

Example 12

Consider the inhomogeneous equation

$$y'' + y' - 2y = 20\cos(2t) \tag{3.41}$$

for which we seek a particular solution. As we want to find a function y_p whose derivatives make the linear combination of the left hand side equal to $\cos(2t)$, it makes sense to to assume that y_p itself looks like $A\cos(2t)$ for some constant A. Unfortunately, this doesn't quite work because

$$\begin{aligned}
&(A\cos(2t))'' + (A\cos(2t))' - 2(A\cos(2t)) \\
&\qquad = -4A\cos(2t) - 2A\sin(2t) - 2A\cos(2t) \\
&\qquad = -6A\cos(2t) - 2A\sin(2t) \\
&\qquad \neq 20\cos(2t)
\end{aligned}$$

for any choice of A due to the term containing $\sin(2t)$. To get this term to cancel out, we adjust our guess by adding a term involving $\sin(2t)$. That is, we consider

$$y_p(t) = A\cos(2t) + B\sin(2t)$$

for some undetermined coefficients A and B which we hope to determine by plugging y_p into the differential equation. We have

$$\begin{aligned}
y_p''(t) + y_p'(t) - 2y_p(t) &= (A\cos(2t) + B\sin(2t))'' + (A\cos(2t) + B\sin(2t))' \\
&\qquad\qquad -2(A\cos(2t) + B\sin(2t)) \\
&= -4A\cos(2t) - 4B\sin(2t) + (-2A\sin(2t) + 2B\cos(2t)) \\
&\qquad\qquad -2A\cos(2t) - 2B\sin(2t) \\
&= -4A\cos(2t) - 4B\sin(2t) - 2A\sin(2t) + 2B\cos(2t) \\
&\qquad\qquad -2A\cos(2t) - 2\sin(2t) \\
&= (-6A + 2B)\cos(2t) + (-6B - 2A)\sin(2t)
\end{aligned}$$

for all $t \in \mathbb{R}$. Thus, for $y_p(t) = A\cos(2t) + B\sin(2t)$ to solve the inhomogeneous equation, we must have

$$(-6A + 2B)\cos(2t) + (-6B - 2A)\sin(2t) = 20\cos(2t)$$

for all $t \in \mathbb{R}$ and, due to the linear independence of $\sin(2t)$ and $\cos(2t)$, we must have

$$-6A + 2B = 20 \quad \text{and} \quad -6B - 2A = 0.$$

This is a 2×2 linear system in the variables A and B and we can solve it to find that

$$A = -3 \text{ and } B = 1.$$

Thus, our particular solution (which you should directly verify is a solution) is

$$y_p(t) = -3\cos(2t) + \sin(2t).$$

Noting that $y_1(t) = e^t$ and $y_2(t) = e^{-2t}$ form a fundamental generating set of solutions to the homogeneous equation,

$$y'' + y' - 2y = 0,$$

Corollary 3.6.2 gives

$$y(t) = C_1 e^t + C_2 e^{-2t} - 3\cos(2t) + \sin(2t)$$

as the general solution to the inhomogeneous equation (3.41).

In looking closely at our example above, the essential idea of the method of undetermined coefficients is to look at the function $r(t)$ in the giving inhomogeneous equation (3.40). If $r(t)$ is sine, cosine, polynomial, exponential or a linear combination thereof, your guess y_p should be a linear combination of functions (and their derivatives) in the same form as $r(t)$. The coefficients in your linear combination are, a priori, unknown and your goal is find them by plugging your guess y_p into the differential equation and find the correct coefficients to make the equation hold. For this reason this procedure is known as the *method of undetermined coefficients.* The following table will help in this task.

$r(t)$	$y_p(t)$
ae^{rt}	Ae^{rt}
$a\cos(\omega t) + b\sin(\omega t)$	$A\cos(\omega t) + B\sin(\omega t)$
$ae^{rt}\cos(\omega t) + be^{rt}\sin(\omega t)$	$Ae^{rt}\cos(\omega t) + Be^{rt}\sin(\omega t)$
$a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$	$A_0 + A_1 t + \cdots + A_n t^n$
Any sum of the items in this column	A sum of the items in this column

Table 3.1: A guess table for the method of undetermined coefficients

Example 13

Consider the inhomogeneous equation

$$y'' + y = e^{2t} + e^{-t} + t$$

to which we seek a particular solution y_p. As $r(t) = e^{2t} + e^{-t} + t$ is a sum of the exponentials e^{2t} and e^{-t} and the first-order polynomial $p(t) = t$, Table 3.1 suggests that we try the function

$$y_p(t) = Ae^{2t} + Be^{-t} + Ct + D$$

as a candidate for our solution. We have

$$y_p'(t) = 2Ae^{2t} - Be^{-t} + C$$

and

$$y_p''(t) = 4Ae^{2t} + Be^{-t}$$

and therefore

$$\begin{aligned} y_p''(t) + y_p(t) &= 4Ae^{2t} + Be^{-t} + Ae^{2t} + Be^{-t} + Ct + D \\ &= 5Ae^{2t} + 2Be^{-t} + Ct + D. \end{aligned}$$

For the inhomogeneous equation to be satisfied by y_p, we must have

$$5Ae^{2t} + 2Be^{-t} + Ct + D = e^{2t} + e^{-t} + t.$$

By the linear independence of the function e^{2t}, e^{-t}, t and 1, we conclude that this can hold if and only if the coefficients on both sides of this equation match. That is

$$5A = 1, \quad 2B = 1, \quad C = 1 \quad \text{and} \quad D = 0.$$

Consequently, $A = 1/5$, $B = 1/2$, $C = 1$, $D = 0$ and so our particular solution is

$$y_p(t) = \frac{1}{5}e^{2t} + \frac{1}{2}e^{-t} + t.$$

An appeal to Corollary 3.6.2 shows that

$$y(t) = C_1\cos(t) + C_2\sin(t) + \frac{1}{5}e^{2t} + \frac{1}{2}e^{-t} + t$$

is a general solution to the given inhomogeneous equation.

In the following example, we study a case in which Table 3.1, on its own, does not yield a particular solution.

Example 14

Consider the inhomogeneous equation

$$y'' + y = \cos(t).$$

Based on Table 3.1, our candidate for a particular solution should be of the form

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

In this case, we see that

$$\begin{aligned} y_p''(t) + y_p(t) &= (A\cos(t) + B\sin(t))'' + (A\cos(t) + B\sin(t)) \\ &= -A\cos(t) - B\sin(t) + A\cos(t) + B\sin(t) \\ &= (A - A)\cos(t) + (B - B)\sin(t) \\ &= 0. \end{aligned}$$

Through this, we see that we are unable to select constants A and B for which $y_p(t) = A\cos(t) + B\sin(t)$ solves the inhomogeneous equation because

$$y_p''(t) + y_p(t) = 0 \neq \cos(t).$$

The essential problem in the above example is that $r(t) = \cos(t)$ itself satisfies the corresponding homogeneous equation

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

Consequently, the guess $y_p(t) = A\cos(t) + B\sin(t)$, based on the form of $r(t)$, also satisfies the homogeneous equation and therefore cannot solve the given inhomogeneous equation. This observation pertains to the general picture: If $r(t)$ in (3.40) satisfies the corresponding homogeneous equation

$$y'' + by' + cy = 0 \tag{3.42}$$

then any guess y_p based on $r(t)$ coming from Table 3.1 will itself satisfy (3.42) and therefore cannot satisfy (3.40). Fortunately, there is a simple fix for this: simply multiply by the monomial t and let the product rule take care of the rest. Let's return to our example to give this a try.

Example 15

Again, we consider the inhomogeneous equation

$$y'' + y = \cos(t).$$

Multiplying our original guess $A\cos(t) + B\sin(t)$ by t, gives the new guess

$$y_p(t) = At\cos(t) + Bt\sin(t).$$

In view of the product rule, we have

$$y_p'(t) = A\cos(t) + B\sin(t) - At\sin(t) + Bt\cos(t)$$

and

$$\begin{aligned} y_p''(t) &= -A\sin(t) + B\cos(t) - A\sin(t) + B\cos(t) - At\cos(t) - Bt\sin(t) \\ &= -2A\sin(t) + 2B\cos(t) - At\cos(t) - Bt\sin(t). \end{aligned}$$

Consequently,

$$\begin{aligned} &y_p''(t) + y_p(t) \\ &\quad = -2A\sin(t) + 2B\cos(t) - At\cos(t) - Bt\sin(t) + At\cos(t) + Bt\sin(t) \\ &\quad = -2A\sin(t) + 2B\cos(t) \end{aligned}$$

and so, for y_p to solve the inhomogeneous equation, we identify

$$-2A\sin(t) + 2B\cos(t) = \cos(t).$$

Thus, $A = 0$ and $2B = 1$ from which we obtain

$$y_p(t) = \frac{1}{2}t\sin(t)$$

as a particular solution.

In general, we state this as a principle.

Principle: If any part of the initial guess y_p, based on Table 3.1, satisfies the homogeneous differential equation (3.42), replace this term with itself multiplied by t.

Exercise 41

Using the theory of inhomogeneous equations, find general solutions to the following inhomogeneous differential equations.

1. $$y'' + 3y' + 2y = \cos(t)$$

2. $$y'' + 4y = e^{2t}$$

3. $$y'' - 4y = e^{2t}$$

Use the theory of inhomogenous equations to solve the following initial value problems.

1. $$\begin{cases} y'' + 3y' + 2y = \cos(t) \\ \\ y(0) = 1 \\ y'(0) = 0 \end{cases}$$

2. $$\begin{cases} y'' - 4y = e^{2t} \\ \\ y(0) = 0 \\ y'(0) = 1 \end{cases}$$

3.8 Application: Damped, Undamped and Forced Oscillation

In this section, we focus on a particular application of linear second-order constant-coefficient differential equations. This is the study of harmonic oscillators and, in particular, we study the cases of undamped, damped, forced and unforced oscillation. To set up the general problem, consider an object of mass m, measured in kilograms (kg), which

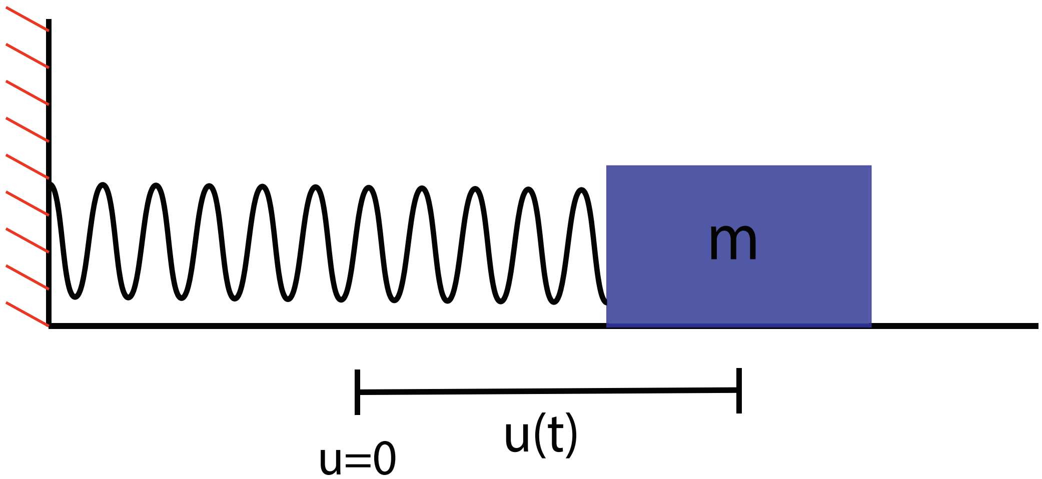


Figure 3.1: An object of mass m connected to a spring and sliding on a frictionless surface.

rests on a surface and is connected to a wall by means of a spring. This is illustrated in Figure 3.1. Let us denote by $u = u(t)$ the displacement of the object from the rest position ($u = 0$) as a function of time t; we shall measure $u(t)$ in meters and t in seconds (s). Assuming that the motion of the object is modeled by Hooke's law, the force done on the object by the spring, measured in Newtons (N), is given by

$$F_{\text{spring}} = -ku$$

where $k > 0$ is the so-called spring constant with units of Newton per meter. The spring force F_{spring} is also known as a restoring force as it pulls the object back to the equilibrium position with a magnitude proportional to the distance from equilibrium.

3.8.1 Free and undamped motion

Let us assume that the object attached to the spring, described above, undergoes frictionless and unforced motion. That is, we assume that there are no frictional/damping forces on the object nor any other external forces beyond those given by the spring. In this case, Newton's second law gives us the equation of motion

$$m\ddot{u} = \text{mass} \times \text{acceleration} = F_{\text{spring}} = -ku.$$

This is equivalently written as

$$\ddot{u} + \omega_0^2 u = 0 \tag{3.43}$$

and called the equation of free and undamped harmonic motion where the constant

$$\omega_0 := \sqrt{\frac{k}{m}} \tag{3.44}$$

is called the natural frequency of oscillation and has units $1/s$. This is a second order linear homogeneous constant coefficient differential equation. An appeal to our theory for such equations gives the general solution

$$u(t) = C_1 \cos(\omega_0 t) + C_2 \sin(\omega_0 t) \tag{3.45}$$

for $t \in \mathbb{R}$. Correspondingly, if the initial position and velocity of the object are known to be u_0 and u_0' respectively, the initial value problem

$$\begin{cases} \ddot{u} + \omega_0^2 u = 0 \\ u(0) = u_0 \\ \dot{u}(0) = u_0' \end{cases}$$

is solved by (uniquely) specifying C_1 and C_2 in (3.45). This concludes our treatment of free (unforced) undamped oscillation. To check your understanding of the situation, and the solution, I encourage you to return to the introductory section and read the example on Newton's second law and watch the correspoinding Spray Paint Oscillator video from MIT's Physics Department [1] (Click here). Does the experiment in the view coincide with what we've done here?
Note Here

3.8.2 Damped Free Oscillatory Motion

The physical situation described in the preceding subsection assumes that the moving object is not subject to frictional forces. It is therefore an idealized model and unrealistic. Any realistic model must account for frictional/damping forces, including the friction between the object and the surface on which it rests, the energy lost in the expansion and contraction of the crystalline structure which makes up the spring and air resistance on the object. As a first-order model, we can account for this friction by considering a force which is proportional to the velocity $\dot{u}$ and opposite in sign, i.e.,

$$F_{\text{damp}} = -c\dot{u}$$

where c is a non-negative real number called a damping coefficient. We note that the magnitude of this force increases and decreases along with the speed of the object and is always opposed to the direction of motion. You should think about why this makes $F_{\text{damp}} = -c\dot{u}$ a reasonable model for friction/damping. Accounting for this force, Newton's second law gives

$$m\ddot{u} = \text{mass} \times \text{acceleration} = F_{\text{damp}} + F_{\text{spring}} = -c\dot{u} - ku$$

or, equivalently,

$$\ddot{u} + \frac{c}{m}\dot{u} + \frac{k}{m}u = 0. \tag{3.46}$$

This is the equation of free (undriven) damped harmonic motion. Before we turn our focus to the study of its solutions, let's consider the following exercise which studies (3.46) from an energy perspective. The exercise leads to an interesting application: an alternate proof (in addition to the Picard-Lindelöff theorem) that solutions to to initial value problems for (3.46) are unique.

Exercise 42

Consider an object of mass m whose displacement $u \in C^2(\mathbb{R})$ satisfies the equation

$$L[u] = m\ddot{u} + c\dot{u} + ku = 0.$$

Throughout this exercise, we only assume that u is twice differentiable and satisfies the equation above[a]. To u we associate the following useful quantity: For $t \geq 0$, we define

$$E(t) = \frac{m}{2}(\dot{u}(t))^2 + \frac{k}{2}(u(t))^2 \tag{3.47}$$

called the *energy of u at time t*; note that $E(t) \geq 0$ for all t. By the fundamental theorem of calculus, we easily[b] observe that

$$E(t) = E(0) + \int_0^t \frac{dE}{dt'}\,dt' = E(0) + \int_0^t (m\dot{u}(t')\ddot{u}(t') + ku(t')\dot{u}(t'))\,dt' \tag{3.48}$$

1. Using (3.48) and the fact that $m\ddot{u} + c\dot{u} + ku = 0$, show that

   $$E(t) = E(0) - c\int_0^t (\dot{u}(t'))^2\,dt' \tag{3.49}$$

   for all $t > 0$.

2. Using the above result, conclude that, if $c = 0$, $E(t) = E(0)$ for all time t. The interpretation of this fact is this: Energy is conserved for an undamped free oscillator (this gives more justification that the constant c represents friction/damping).

3. In the general case that $c \geq 0$, use (3.49) to show that E is a non-increasing function of time.

4. In the next two items, we use the energy E to prove the uniqueness statement in the Picard-Lindelöf theorem. First, suppose that $u(0) = 0$ and $\dot{u}(0) = 0$, use (3.47) and the previous item to show that $E(t) = 0$ for all t. Conclude that $u(t) = 0$ for all t.

5. Now suppose that u_1 and u_2 solve the initial value problem

   $$\{L[u] = m\ddot{u} + c\dot{u} + ku = 0, \qquad u(0) = u_0 \text{ and } \dot{u}(0) = \dot{u}_0$$

   where u_0 and $\dot{u}_0$ are fixed constants. By considering $u = u_2 - u_1$ and applying your results from the previous item, show that $u_1(t) = u_2(t)$ for all t. In this way, you prove the uniqueness statement of the Picard-Lindelöf theorem.

6. What's the moral of this exercise?

[a] In other words, there is no reason to come up with an explicit formula for u.
[b] You should check this computation.

Let's now turn our focus to solving (3.46). To this end, we consider the characteristic polynomial

$$r^2 + \frac{c}{m}r + \frac{k}{m} = 0$$

which has solutions

$$r = -\frac{c}{2m} \pm \frac{1}{2m}\sqrt{c^2 - 4km}.$$

In looking at Theorem 3.5.4, we have three cases to consider:

1. **Underdamped Motion:** In the case that $c^2 < 4km$, meaning that the damping is small relative to the product of the spring constant and the mass of the object, the solutions to characteristic polynomial are complex-valued. Here,

   $$r = -\frac{c}{2m} \pm \omega i$$

where

$$\omega = \omega(c, k, m) = \frac{\sqrt{4km - c^2}}{2m}$$

and so

$$u(t) = C_1 e^{-\frac{c}{2m}t}\cos(\omega t) + C_2 e^{-\frac{c}{2m}t}\sin(\omega t)$$

is a general solution to (3.46) in the case that $c^2 < 4km$. This is so-called underdamped harmonic motion. We see that the solutions decay on the order of $e^{-ct/2m}$ in amplitude (which depends on the damping coefficient c) while oscillating ad infinitum with frequency $\omega = \omega(c, k, m)$. We note that this model recaptures our model of undamped harmonic motion when $c = 0$ and, in this case, $\omega(0, k, m) = \omega_0$.

2. **Overdamped Motion:** In the case that $c^2 > 4km$, i.e., the case in which the damping coefficient is large relative to the product of the spring coefficient and the mass of the oscillator, we obtain so-called overdamped motion. Here, the solutions to the characteristic polynomial are real valued and of the form

$$r_1 = -\frac{c}{2m} + \frac{\sqrt{c^2 - 4km}}{2m} \qquad \text{and} \qquad r_2 = -\frac{c}{2m} - \frac{\sqrt{c^2 - 4km}}{2m}$$

giving the general solution

$$u(t) = C_1 e^{r_1 t} + C_2 e^{r_2 t}$$

to (3.46) in the case of overdamped motion. We note that $0 < \sqrt{c^2 - 4km} < \sqrt{c^2} = c$ and so it follows that

$$r_1 = -\frac{c}{2m} + \frac{\sqrt{c^2 - 4km}}{2m} < 0$$

giving the inequality $r_2 < r_1 < 0$. This ensures that the non-oscillatory general solution $C_1 e^{r_1 t} + C_2 e^{r_2 t}$ decays exponentially to 0 at the rate $e^{-r_1 t}$ (i the slowest case).

3. **Critically Damped Motion:** The final possibility is the critical case in which $c^2 = 4km$; this is called critically damped motion. An appeal to Theorem 3.5.4 gives the general solution

$$u(t) = C_1 e^{-\frac{c}{2m}t} + C_2 t e^{-\frac{c}{2m}t} = (A + Bt)e^{-\frac{c}{2m}t}$$

which illustrates that the motion is non-oscillatory and decays exponentially to 0.

Note Here

Exercise 43

1. Show that, in the case of overdamped or critically damped motion, the mass can pass through its zero displacement position at most once, regardless of the initial conditions.

2. In the case of critically damped motion, show that the solution to the intial value problem with initial conditions $u(0) = u_0 > 0$ and $\dot{u}(0) = 0$ has the property that $u(t) \to 0$ as $t \to \infty$ but $u(t) \neq 0$ for any finite value of $t \geq 0$.

3. Again, in the case of critically damped motion such that $u(0) = u_0 > 0$, find conditions on $\dot{u}(0)$ that guarantee that u will pass though its zero displacement at some time t.

Note Here

Exercise 44

We now study forced undamped motion:

$$m\ddot{u} + ku = F(t) \tag{3.50}$$

where we set $\omega_0 = \sqrt{k/m}$, called the natural frequency of oscillation.

1. If $F(t) = \sin(\omega t)$ for a constant $\omega \neq \omega_0$, use the method of undetermined coefficients to find a general solution to (3.50). Use your general solution to solve the general initial value problem

   $$\{m\ddot{u} + ku = \sin(\omega t), \qquad u(0) = 0,\ \dot{u}(0) = 1.$$

   In the case that $m = k = 1$ and $\omega = 2$, plot your solution. Do you see resonance?

2. If $F(t) = \sin(\omega_0 t)$, use the method of undetermined coefficients to find a general solution to (3.50). Use your general solution to solve the general initial value problem

   $$\{m\ddot{u} + ku = \sin(\omega_0 t), \qquad u(0) = 0,\ \dot{u}(0) = 1.$$

   In the case that $m = k = 1$, plot your solution. Do you see resonance?

3. Analyze the case in which $F(t) = \sin(\omega t) + \sin(\omega_0 t)$. Does this example exhibit resonance?

here

Chapter 4

Systems

In this chapter, we study first-order coupled systems of ordinary differential equations. In the previous chapters, we've focused on the theory and applications of scalar equations. There the idea was to find a real-valued (scalar) function $y = y(t)$ which satisfied a scalar equation relating y and its derivatives, $y', y'', \ldots$. By contrast, a coupled first-order $n \times n$ system of differential equations is formed by specifying n ordinary differential equations which relate n functions $x_1(t), x_2(t), \ldots, x_n(t)$ and their first-order derivatives, $\dot{x}_1(t), \dot{x}_2(t), \ldots, \dot{x}_n(t)$.

To motivate our study of systems, we look to biology for a simple yet fascinating mathematical model, called the predator-prey model, of the population dynamics of two competing species. Suppose that the population of a species – the prey– is given by a function $P(t)$ and the population of another species – the predators – is modeled by a function $K(t)$. Making a number of assumptions and simplifications concerning the interaction between predator and prey, the Lotka-Volterra model says that the populations evolve according to the following equations,

$$\begin{aligned} \dot{P}(t) &= aP(t) - bP(t)K(t) \\ \dot{K}(t) &= cP(t)K(t) - dK(t), \end{aligned} \tag{4.1}$$

called the Lotka-Volterra equations or the predator-prey equations. Here $\dot{P} = dP/dt$, $\dot{K} = dK/dt$ and a, b, c, d are all positive numbers which depend on the dynamics of interaction of the two competing species. The first equation

$$\dot{P}(t) = aP(t) - bP(t)K(t),$$

called the prey equation, says that the population P of the prey species increases at a rate $aP(t)$ while decreasing at a rate $bP(t)K(t)$. This captures the idea that the prey species has unlimited food and constantly reproduces while simultaneously gets killed off at a rate which depends on both the predator and prey populations. By contrast, the predator equation

$$\dot{K}(t) = cP(t)K(t) - dK(t)$$

says that the population K of the predator species increases at a rate $cP(t)K(t)$ while simultaneously dying at a rate of $dK(t)$, due to natural death. In taking these two equations together, we see that the resulting 2×2 system (4.1) is truly coupled in the sense that changes in the one population is driven by the value of both that population and the population of its competitor.

In studying the Lotka-Volterra equations, we are interested to know the evolution of both predator and prey populations given that initial populations are known. In other words, we are generally interested in solving the the system (4.1) subject to the initial conditions

$$P(0) = P_0 \quad \text{and} \quad K(0) = K_0$$

where P_0 and K_0 are the known initial populations of predator and prey respectively. Though, in general, solving such initial value problems is not possible in terms of simple (trigonometric, exponential) functions, this system is relatively straightforward to analyze and we will do so with qualitative methods, analogous to those discussed in Chapter 1. NOTE HERE

4.1 First-order $n \times n$ systems and their initial value problems

Let's here begin to formalize precisely some of the terminology introduced in the introductory section. A $n \times n$ first-order system of ordinary differential equations is, by definition, a system of equations of the form

$$\begin{aligned} \dot{x}_1 &= F_1(t, x_1, x_2, \dots, x_n) \\ \dot{x}_2 &= F_2(t, x_1, x_2, \dots, x_n) \\ \vdots & \quad \vdots \\ \dot{x}_n &= F_n(t, x_1, x_2, \dots, x_n) \end{aligned} \tag{4.2}$$

where, for each $k = 1, 2, \dots, n$, F_k is a real-valued function of the $n+1$ variables, $t, x_1, x_2, \dots, x_n$. We shall assume that all function $F_1, F_2, \dots, F_k$ are defined and continuous on a common domain; more precisely, that there is some open interval I such that, for each $k = 1, 2, \dots, n$, $F_k : I \times \mathbb{R}^n \to \mathbb{R}$ is continuous[1]. To solve such a system means to find n once-continuously differentiable real-valued functions $x_1(t), x_2(t), \dots, x_n(t)$ for which

$$\begin{aligned} \dot{x}_1(t) &= F_1(t, x_1(t), x_2(t), \dots, x_n(t)) \\ \dot{x}_2(t) &= F_2(t, x_1(t), x_2(t), \dots, x_n(t)) \\ \vdots & \quad \vdots \\ \dot{x}_n(t) &= F_n(t, x_1(t), x_2(t), \dots, x_n(t)) \end{aligned}$$

for all t in a common (and non-trivial) subdomain J of I; here we use notation that, for each $k = 1, 2, \dots, n$, $\dot{x}_k(t) = dx_k/dt$. To simplify notation, it is useful to put these ideas in terms of vector-valued functions. To this end, we define $F : I \times \mathbb{R}^n \to \mathbb{R}^n$ by putting

$$F(t, \mathbf{x}) = F(t, x_1, x_2, \dots, x_n) = \begin{pmatrix} F_1(t, x_1, x_2, \dots, x_n) \\ F_2(t, x_1, x_2, \dots, x_n) \\ \vdots \\ F_n(t, x_1, x_2, \dots, x_n) \end{pmatrix}$$

for all $(t, \mathbf{x}) = (t, x_1, x_2, \dots, x_n) \in I \times \mathbb{R}^n$. In these terms, the $n \times n$ system (4.2) is equivalently written as,

$$\dot{\mathbf{x}} = F(t, \mathbf{x}) \tag{4.3}$$

[1] Here $I \times \mathbb{R}^n = \{(t, x_1, x_2, \dots, x_n) \in \mathbb{R}^{n+1} : t \in I \text{ and } x_1, x_2, \dots, x_n \in \mathbb{R}\}$.

where $\mathbf{x}$ is of the form

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} \quad \text{and} \quad \dot{\mathbf{x}}(t) = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}.$$

To talk about a "solution" to (4.3) (and hence (4.2)), it's useful to give a definition.

Definition 4.1.1. *Let I be an open interval and consider a vector-valued function $\mathbf{x} : I \to \mathbb{R}^n$ defined by*

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

where $x_1(t), x_2(t), \dots, x_n(t)$ are functions mapping from I to $\mathbb{R}$, called the components of $\mathbf{x}$. We say that $\mathbf{x}$ is continuous on I if its components are all continuous functions on I, i.e., $x_1, x_2, \dots, x_n \in C^0(I)$. The set of all such continuous functions is denoted by $C^0(I; \mathbb{R}^n)$. We say that $\mathbf{x}$ is once-continuously differentiable on I if its components are all once-continuously differentiable on I, i.e., $x_1, x_2, \dots, x_n \in C^1(I)$. In this case, the derivative of $\mathbf{x}$ is defined by

$$\dot{x}(t) = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}$$

for $t \in I$. The set of all such once-continuously differentiable is denoted by $C^1(I; \mathbb{R}^n)$.

Remark 4.1.2. As was true for scalar-valued functions, $C^1(I; \mathbb{R}^n) \subseteq C^0(I; \mathbb{R}^n)$. In fact, both of these sets are infinite-dimensional vector spaces and $C^1(I; \mathbb{R}^n)$ is a subspace of $C^0(I; \mathbf{R}^n)$. Not surprisingly, viewing these as vector spaces will be helpful to us when studying linear systems. Also, it should be noted that derivative $\dot{\mathbf{x}}$ of $\mathbf{x}$ coincides with the Jacobian derivative of $\mathbf{x}$.

By definition, a *solution* to the system (4.3) (equivalently (4.2)) is a vector-valued function $\mathbf{x} \in C^1(J; \mathbb{R}^n)$ which has

$$\dot{\mathbf{x}}(t) = F(t, \mathbf{x}(t))$$

for all $t \in J$ where J is some (non-trivial) subinterval of I. As was true for scalar equations, there are often multiple (an infinite number) of solutions to a given $n \times n$ system of differential equations. Of course, it is of interest to select a solution (among the many) which satisfies certain given initial conditions. To this end, introduce the concept of an initial value problem.

An initial value problem for (4.3) comes by specifying an initial time $t_0 \in I$ and a constant vector

$$\mathbf{x}_0 = \begin{pmatrix} x_{1,0} \\ x_{2,0} \\ \vdots \\ x_{n,0} \end{pmatrix} \in \mathbb{R}^n$$

and asking for solutions to (4.3) to satisfy

$$\mathbf{x}(t_0) = \mathbf{x}_0.$$

We write such an initial value problem as

$$\{\dot{\mathbf{x}}(t) = F(t, \mathbf{x}(t)), \qquad \mathbf{x}(t_0) = \mathbf{x_0}$$

or

$$\begin{cases} \dot{\mathbf{x}} = F(t, \mathbf{x}) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases}.$$

We are now ready to state the Picard-Lindelöff theorem for systems. As we shall see, this one theorem will capture all Picard-Lindelöff theorems seen previously.

Theorem 4.1.3 (The Picard-Lindelöf Theorem for $n \times n$ Systems)**.** *Let $I \subseteq \mathbb{R}$ be an interval and consider the $n \times n$ system*

$$\dot{\mathbf{x}} = F(t, \mathbf{x})$$

where $F : I \times \mathbb{R}^n \to \mathbb{R}^n$ has components $F_1, F_2, \ldots, F_n : I \times \mathbb{R}^n \to \mathbb{R}$. If the components $F_1, F_2, \ldots, F_n$ are all continuous on $I \times \mathbb{R}^n$ and the partial derivatives

$$\frac{\partial F_1}{\partial x_1}, \frac{\partial F_1}{\partial x_2}, \ldots, \frac{\partial F_1}{\partial x_n}; \frac{\partial F_2}{\partial x_1}, \frac{\partial F_2}{\partial x_2}, \ldots, \frac{\partial F_2}{\partial x_n}; \ldots; \ldots; \frac{\partial F_n}{\partial x_1}, \frac{\partial F_n}{\partial x_2}, \ldots, \frac{\partial F_n}{x_n}$$

are continuous on $I \times \mathbb{R}^n$, then, given any $t_0 \in I$ and fixed vector $\mathbf{x_0} \in \mathbb{R}^n$, the initial value problem

$$\{\dot{\mathbf{x}} = F(t, \mathbf{x}), \qquad \mathbf{x}(t_0) = \mathbf{x_0} \tag{4.4}$$

has a unique solution $\mathbf{x} \in C^1(J; \mathbb{R}^n)$ where J is a subinterval of I (possibly equal to I) containing t_0, i.e., $t_0 \in I_0 \subseteq I$.

Here

4.2 Linear Systems

A class of $n \times n$ systems which will be of great interest for us are linear systems. Understanding these systems, especially the so-called constant-coefficient linear systems will be fundamental to our understanding of non-linear systems, such as the Lotka-Volterra equations discussed in the introductory section. An $n \times n$ linear systems of first-order ordinary differential equations is a system of the form

$$\begin{aligned} \dot{x}_1 &= p_{11}(t)x_1 + p_{12}(t)x_2 + \cdots + p_{1n}(t)x_n + g_1(t) \\ \dot{x}_2 &= p_{21}(t)x_1 + p_{22}(t)x_2 + \cdots + p_{2n}(t)x_n + g_2(t) \\ \vdots \quad & \quad \vdots \\ \dot{x}_n &= p_{n1}(t)x_1 + p_{n2}(t)x_2 + \cdots + p_{nn}(t)x_n + g_n(t). \end{aligned}$$

This can be written in the equivalent form

$$\dot{\mathbf{x}} = P(t)\mathbf{x} + g(t) \tag{4.5}$$

where

$$P(t) = \begin{pmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1n}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}(t) & p_{n2}(t) & \cdots & p_{nn}(t) \end{pmatrix}$$

and

$$g(t) = \begin{pmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_n(t) \end{pmatrix}.$$

Specifically, P is a function from some interval I into the set $M_{n,n}(\mathbb{R})$ of $n \times n$ matrices with real entries, i.e., $P : I \to M_{n,n}(\mathbb{R})$, and by $P(t)\mathbf{x}$ we mean the matrix product of $P(t)$ with a vector-valued function $\mathbf{x}$. We say that (4.5) is *homogeneous* if the vector-valued function g is identically zero, i.e., the system (4.5) is

$$\dot{\mathbf{x}} = P(t)\mathbf{x}.$$

As was true for linear scalar equations, linear algebra is a great tool for understanding the solutions to $n \times n$ linear systems. Looking in this direction, for each interval I, we recognize that the set of vector valued functions $C^1(I;\mathbb{R}^n)$ is a real vector space when equipped with the (perhaps obvious) notion of addition and scalar multiplication: For $\mathbf{x}, \mathbf{y} \in C^1(I;\mathbb{R}^n)$ and $\alpha \in \mathbb{R}$,

$$(\mathbf{x}+\mathbf{y})(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} + \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} (x_1+y_1)(t) \\ (x_2+y_2)(t) \\ \vdots \\ (x_n+y_n)(t) \end{pmatrix} \in C^1(I;\mathbb{R}^n)$$

and

$$(\alpha\mathbf{x})(t) = \alpha \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} \alpha x_1(t) \\ \alpha x_2(t) \\ \vdots \\ \alpha x_n(t) \end{pmatrix} \in C^1(I;\mathbb{R}^n).$$

With precisely the same notions of addition and scalar multiplication, $C^0(I;\mathbb{R}^n)$ is a vector space and it is easy to see that $C^1(I;\mathbb{R}^n)$ is a subspace of $C^0(I;\mathbb{R}^n)$; they are both infinite dimensional vector spaces. These spaces can be recognized as *path spaces* in the sense that each $\mathbf{x} \in C^1(I;\mathbb{R}^n)$ traces a smooth curve (or path) in $\mathbb{R}^n$. For instance, $\mathbf{x} \in C^1(\mathbb{R};\mathbb{R}^2)$ defined by

$$\mathbf{x}(t) = \begin{pmatrix} e^{-t/10}\cos(t) \\ e^{-t/10}\sin(t) \end{pmatrix}$$

traces out the spiral illustrated in Figure 4.1. Note Here

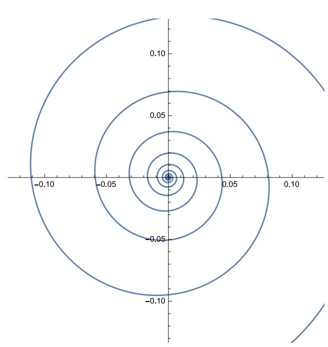


Figure 4.1: $\mathbf{x}(t) = (e^{-t/10}\cos(t), e^{-t/10}\sin(t))^\top$ for $-10 < t < 10$.

As the following theorem shows, the solutions of linear homogeneous systems form finite-dimensional subspaces of $C^1(I;\mathbb{R}^n)$.

Theorem 4.2.1. *Consider the matrix-valued function*

$$P(t) = \begin{pmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1n}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}(t) & p_{n2}(t) & \cdots & p_{nn}(t) \end{pmatrix} \tag{4.6}$$

where the components p_{jk} are continuous (real-valued) functions on an interval I of $\mathbb{R}$ and the associated $n \times n$ linear system associated to P,

$$\dot{\mathbf{x}}(t) = P(t)\mathbf{x}(t). \tag{4.7}$$

Then the set of solutions to (4.7) on the interval I,

$$S(I) = \left\{\mathbf{x} \in C^1(I;\mathbb{R}^n) : \dot{\mathbf{x}}(t) = P(t)\mathbf{x}\right\},$$

is an n-dimensional subspace of $C^1(I;\mathbb{R}^n)$. In particular, there exist n linearly independent elements $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)} \in S(I)$ such that, given any solution $\mathbf{x}$ to (4.7) on I, there exist constants $c_1, c_2, \ldots, c_n$ such that

$$\mathbf{x}(t) = c_1\mathbf{x}^{(1)}(t) + c_2\mathbf{x}^{(2)}(t) + \cdots + c_n\mathbf{x}^{(n)}(t)$$

for $t \in I$. Such an expression is called a general solution to the equation $\dot{\mathbf{x}} = P\mathbf{x}$.

The focus of the following exercise is to see how the linear structure of (4.7) (essentially) gives Theorem 4.2.1 as consequence of Theorem 4.1.3.

Exercise 45: Linear Systems

1. Given an interval I, assume that the entries of the matrix P, defined by (4.6), are continuous on I, i.e., $p_{j,k} \in C^0(I)$ for all $k, j = 1, 2, \ldots, n$. Show that the system
$$\dot{\mathbf{x}}(t) = P(t)\mathbf{x}(t)$$
satisfies the hypotheses of Theorem 4.1.3.

2. Suppose that $\mathbf{x}, \mathbf{y} \in C^1(I;\mathbb{R}^n)$ satisfy (4.7). Show that, for any constants $\alpha, \beta \in \mathbb{R}$, $\alpha\mathbf{x} + \beta\mathbf{y} \in C^1(I;\mathbb{R}^n)$ also satisfies (4.7). In this way, you show that the set of solutions to (4.7) is a subspace of $C^1(I;\mathbb{R}^n)$.

3. Assuming that the entries of P are continuous function on the interval $I = \mathbb{R}$ (which we will assume henceforth), use Theorem 4.1.3 to show that the set of solutions to (4.7) (near $t_0 = 0$) is n-dimensional, i.e., is an n-dimensional subspace of $C^1(I_0;\mathbb{R}^n)$ for some subinterval $I_0 \subseteq \mathbb{R}$. Hint: Take $t_0 = 0$ and apply Theorem 4.1.3 n times to the n initial vectors
$$\mathbf{x}_0^{(1)} = e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{x}_0^{(2)} = e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{x}_0^{(n)} = e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$
to obtain solutions $\mathbf{x}^{(1)}(t), \mathbf{x}^{(2)}(t), \ldots, \mathbf{x}^{(n)}(t)$. Show that every solution to (4.7) can be written as a linear combination of these solutions (this will use

the uniqueness statement of Theorem 4.1.3. Show that these solutions are also linearly independent (no need to use the Wronskian).

4. Given continuous functions $p(t)$ and $q(t)$, consider the 2×2 linear system

$$\dot{\mathbf{x}}(t) = \begin{pmatrix} \dot{x_1}(t) \\ \dot{x_2}(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -q(t) & -p(t) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

Show that the above 2×2 linear system is equivalent to the second-order linear differential equation

$$L[y] = y'' + p(t)y' + q(t)y = 0.$$

Does your result from the previous item (applied to this system) guarantee our result from earlier in the semester that $\ker(L)$ is 2-dimensional? How? Hint: Set $y = x_1$, then $y' = \dot{x}_1 = x_2$.

Suppose that we are given n solutions $\mathbf{x}^{(1)}(t), \mathbf{x}^{(2)}(t), \ldots, \mathbf{x}^{(n)} \in C^1(I; \mathbb{R}^n)$ to the $n \times n$ linear system

$$\dot{\mathbf{x}} = P\mathbf{x},$$

how could we verify that they formed a basis of the solution space $S(I)$? In light of the preceding theorem, we simply must verify they are linearly independent. This means, of course that, if there are constants $c_1, c_2, \ldots, c_n \in \mathbb{R}^n$ for which

$$c_1\mathbf{x}^{(1)}(t) + c_2\mathbf{x}^{(2)}(t) + \cdots + c_n\mathbf{x}^{(n)}(t) = 0 \tag{4.8}$$

for all $t \in I$ (here $0 = (0, 0, \ldots, 0)^\intercal$ is the zero vector), then it must be true that $c_1 = c_2 = \cdots c_n = 0$. We observe that equation (4.8) can be rewritten in the form

$$W\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}\right)(t) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where $W\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}\right)$ is the $n \times n$ matrix whose columns are the vector-valued functions $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$, i.e.,

$$\begin{aligned} &W\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}\right)(t) \\ &= \left(\mathbf{x}^{(1)}(t)|\mathbf{x}^{(2)}(t)|\ldots|\mathbf{x}^{(n)}(t)\right) = \begin{pmatrix} x_1^{(1)}(t) & x_1^{(2)}(t) & \ldots & x_1^{(n)}(t) \\ x_2^{(1)}(t) & x_2^{(2)}(t) & \ldots & x_2^{(n)}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)}(t) & x_n^{(2)}(t) & \ldots & x_n^{(n)}(t) \end{pmatrix}. \end{aligned}$$

Unsurprisingly, we call $W = W\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}\right)$ the *Wronskian matrix* associated to the vector-valued functions $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$. Its determinant,

$$w(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}) = \det(W(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)})$$

is called the *Wronskian determinant*. The following proposition gives a sufficient condition for the vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ to be linearly independent.

Proposition 4.2.2. *Let I be an interval and consider n vector-valued functions $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)} \in C^0(I;\mathbb{R}^n)$. If, for some $t_0 \in I$, $w\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}\right)(t_0) \neq 0$ (or equivalently, that $W\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}\right)(t_0)$ is an invertible matrix), then the vector-valued functions $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ are linearly independent.*

Exercise 46

1. Prove the proposition above.
2. Use the proposition to deduce the following fact: If $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(\mathbf{n})} \in C^1(I, \mathbb{R}^n)$ are solutions to the linear homogeneous system $\dot{\mathbf{x}} = P\mathbf{x}$ which satisfy $w\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}\right)(t_0) \neq 0$ for some $t_0 \in I$, then every solution to $\dot{\mathbf{x}} = P\mathbf{x}$ can be written in the form
$$\mathbf{x} = c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} + \cdots + c_n\mathbf{x}^{(n)}.$$
3. Suppose that y_1 and y_2 satisfy the linear differential equation
$$L[y] = y'' + p(t)y' + q(t)y = 0$$
where $p, q \in C^0(I)$. Use the equivalence of Item 4 of the preceding exercise to show that the Wonskian condition of Theorem 3.3.5, i.e., that
$$w_{y_1,y_2}(t_0) = y_1(t_0)y_2'(t_0) - y_2(t_0)y_1'(t_0) \neq 0$$
for some $t_0 \in I$, is equivalent to the condition obtained in the preceding item (pertaining to the Wronskian determinant for vector-valued functions). In other words, show that the Wronskian condition in Theorem 3.3.5 is a special case of the same condition for vector-valued functions.

4.3 Constant-coefficient linear systems

The theory discussed in the previous section gives the general structure of the theory of first-order linear systems. Analogous to our theory for higher-order linear (scalar) equations, the main idea is this: If a basis for the solution space of the linear system can be gotten, then every solution (and the solution to every initial value problem) can be expressed as a linear combination of these basis elements. Though bases always exist, producing them is far from trivial. However, in the special case that the system in question has constant coefficients, i.e., that $P(t) = A$ is a constant $n \times n$ matrix, producing a basis of the solution space essentially boils down to finding the eigenvalues and eigenvectors of A. Let's begin with a simple proposition.

Proposition 4.3.1. *Consider the $n \times n$ linear system*
$$\dot{\mathbf{x}} = A\mathbf{x} \tag{4.9}$$
where A is a (constant) $n \times n$ matrix. If $\lambda \in \mathbb{R}$ is an eigenvalue of A with corresponding eigenvector $v \in \mathbb{R}^n$, then
$$\mathbf{x}(t) = e^{\lambda t}v$$
is a solution to (4.9).

Proof. Given that λ is an eigenvalue of A with eigenvector

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix},$$

we have

$$\frac{d}{dt}\left(e^{\lambda t}v\right) = \frac{d}{dt}\begin{pmatrix} e^{\lambda t}v_1 \\ e^{\lambda t}v_2 \\ \vdots \\ e^{\lambda t}v_n \end{pmatrix} = \begin{pmatrix} e^{\lambda t}\lambda v_1 \\ e^{\lambda t}\lambda v_2 \\ \vdots \\ e^{\lambda t}\lambda v_n \end{pmatrix} = e^{\lambda t}\lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = e^{\lambda t}Av = A\left(e^{\lambda t}v\right)$$

for $t \in \mathbb{R}$. In other words,

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t)$$

for $t \in \mathbb{R}$, as desired. □

In the following exercise, you will use your background in linear algebra to develop a theory for the solutions to $n \times n$ constant-coefficient systems where the corresponding matrix A has n distinct real eigenvalues.

Exercise 47: Constant-coefficient linear systems

In what follows, A is an $n \times n$ constant matrix.

1. First, let's treat a purely algebraic result. If λ_1, λ_2 are distinct eigenvalues of A with corresponding eigenvectors v_1 and v_2, show that v_1 and v_2 are linearly independent.

2. Suppose now that λ_1, λ_2 and λ_3 are distinct eigenvalues of A with corresponding eigenvectors v_1, v_2 and v_3, show that v_1, v_2 and v_3 are linearly independent. Hint: You should use your result from the previous item.

3. In light of the previous two items, we will take the following fact for granted (which you should try to prove if you're curious):

   Fact 4.3.2. *Let A be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ and corresponding eigenvectors $v_1, v_2, \dots, v_n$. Then the vectors $v_1, v_2, \dots, v_n$ are linearly independent.*

   Under the hypotheses of the fact above, use properties of determinants to show that
   $$\det(v_1|v_2|\cdots|v_n) \neq 0.$$

4. As above, let A be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ and corresponding eigenvectors $v_1, v_2, \dots, v_n$. In view of the proposition,
   $$\mathbf{x}^{(1)}(t) = e^{\lambda_1 t}v_1,\ \mathbf{x}^{(2)}(t) = e^{\lambda_2 t}v_2\ \dots \mathbf{x}^{(n)}(t) = e^{\lambda_n t}v_n$$
   are solutions to the linear system (4.7). Thus, in view of the previous exercise, it is natural to ask if these n solutions span the solution space, i.e., if they are

linearly independent. Using properties of the determinant, show that

$$\begin{aligned} W\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}\right)(t) &= \det\left(\mathbf{x}^{(1)}(t)|\mathbf{x}^{(2)}(t)|\cdots|\mathbf{x}^{(n)}(t)\right) \\ &= e^{(\lambda_1+\lambda_2+\cdots+\lambda_n)t}\det(v_1|v_2|\cdots|v_n) \end{aligned}$$

for $t \in \mathbb{R}$. Using the previous results, conclude that $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}$ form a basis for the solutions space of (4.7).

Note here

Example 1

Consider the constant-coefficient linear system

$$\dot{\mathbf{x}} = A\mathbf{x},$$

where

$$A = \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix}$$

Our aim is to find a general solution for this system – equivalently, a basis for the solution space to the system. Taking cues from the preceding exercise, let's find the eigenvalues of A. We have

$$\begin{aligned} \det(A - \lambda I) &= \det\begin{pmatrix} 1-\lambda & -2 \\ 3 & -4-\lambda \end{pmatrix} \\ &= (1-\lambda)(4-\lambda)+6 \\ &= \lambda^2 + 3\lambda + 2 \\ &= (\lambda+1)(\lambda+2) \end{aligned}$$

and so λ is an eigenvalue for A if and only if

$$(\lambda+1)(\lambda+2) = \det(A - \lambda I) = 0.$$

Of course, this gives us eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -2$. Since these two eigenvalues are distinct, the result of the previous exercise shows that we can use them to find a basis to the solution space of the linear system. All we need to do is find associated eigenvectors.

Let's first find an eigenvector v_1 with eigenvalue $\lambda_1 = -1$. Necessarily, $v_1 = (x, y)$ is non-zero vector for which

$$(A - \lambda I)v_1 = \begin{pmatrix} 1-(-1) & -2 \\ 3 & -4-(-1) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x-2y \\ 3x-3y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and so a suitable choice is gotten by setting $x = y = 1$. This yields the eigenvector $v_1 = (1,1)^\intercal$ for A with eigenvalue $\lambda_1 = -1$. By a similar computation (which you should do), we find that $v_2 = (2,3)$ is an eigenvector for A with eigenvalue $\lambda_2 = -2$. To summarize,

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad v_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

are eigenvectors for A with eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -2$, respectively. As you proved in the preceding exercise, these eigenvectors are necessarily linearly independent and so it follows that

$$\mathbf{x}^{(1)}(t) = e^{\lambda_1} v_1 = e^{-t}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}e^{-t}\\e^{-t}\end{pmatrix} \quad \text{and} \quad \mathbf{x}^{(2)}(t) = e^{\lambda_2} v_2 = e^{-2t}\begin{pmatrix}2\\3\end{pmatrix} = \begin{pmatrix}2e^{-2t}\\3e^{-2t}\end{pmatrix}$$

form a basis for the solution space to $\dot{\mathbf{x}} = A\mathbf{x}$ and, in view of Theorem 4.2.1,

$$\mathbf{x}(t) = c_1\begin{pmatrix}e^{-t}\\e^{-t}\end{pmatrix} + c_2\begin{pmatrix}2e^{-2t}\\3e^{-2t}\end{pmatrix}$$

is a general solution to this linear system.

For general $n \times n$ constant coefficient linear systems of the form

$$\dot{\mathbf{x}} = A\mathbf{x},$$

which are the perhaps the simplest first-order systems we can analyze, producing general solutions is an involved task. When the eigenvalues of A are all real and distinct, the solution space to the system has a nice and simple form as you showed in the preceding exercise (and we confirmed in the example). Of course, there are many other possibilities for eigenvalues of an $n \times n$ matrix A and writing down a general prescription for solutions (to cover all possibilities) is possible, but it is perhaps too complicated for our needs. The main thing you should take away from this discussion is this: Analyzing linear systems comes down do spectral analysis (the analysis of eigenvalues/eigenvectors).

For simplicity, let's focus our attention on 2×2 constant coefficient systems where a full analysis is within reach. This result is captured by the following proposition.

Proposition 4.3.3. *Consider the 2×2 linear system*

$$\dot{\mathbf{x}}(t) = \begin{pmatrix}\dot{x}(t)\\\dot{y}(t)\end{pmatrix} = \begin{pmatrix}a & b\\c & d\end{pmatrix}\begin{pmatrix}x(t)\\y(t)\end{pmatrix} = A\mathbf{x}(t).$$

We have:

- *If the eigenvalues λ_1 and λ_2 of A are real and distinct with corresponding eigenvectors v_1 and v_2, then a general solution to the 2×2 system above is given by*

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2$$

  for $t \in \mathbb{R}$.

- *If A has complex eigenvalues $\lambda = \alpha \pm i\beta$ with (necessarily) complex eigenvectors $v = \vec{a} \pm i\vec{b}$ where $\vec{a}, \vec{b} \in \mathbb{R}^2$, a general solution to the 2×2 systems above is given by*

$$\mathbf{x}(t) = c_1 e^{\alpha t}(\cos(\beta t)\vec{a} - \sin(\beta t)\vec{b}) + c_2 e^{\alpha t}(\cos(\beta t)\vec{b} + \sin(\beta t)\vec{a})$$

  for $t \in \mathbb{R}$.

- *If A has only one eigenvalue (an eigenvalue of multiplicity 2) λ with corresponding eigenvector v, then a general solution to the 2×2 system above is given by*

$$\mathbf{x}(t) = c_1 e^{\lambda t} v + c_2 (t e^{\lambda t} v + e^{\lambda t} w)$$

  for $t \in \mathbb{R}$ where $w \in \mathbb{R}^2$ is such that

$$(A - \lambda I)w = v.$$

here

Looking back to the previous example, our 2×2 matrix

$$A = \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix}$$

had distinct real eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -2$. In that example, we produced a general solution to the system using the results of a previous exercise. You should confirm that we could have made the same conclusion for this system (to produce a general solution) using the first item of the proposition above. The following example treats the case in which eigenvalues are complex.

Example 2

Consider the 2×2 linear system
$$\dot{\mathbf{x}} = A\mathbf{x}$$
where
$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$
Let's use the preceding proposition to find a general solution to this system and use it to solve the initial value problem

$$\begin{cases} \dot{\mathbf{x}} = A\mathbf{x}, & \mathbf{x}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{cases}.$$

As the proposition necessitates, we will first find the eigenvalues of A. To this end, we solve

$$0 = \det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - (1)(-1) = \lambda^2 - 2\lambda + 2.$$

Using the quadratic formula, we find

$$\lambda = -\frac{-2}{2} \pm \frac{\sqrt{(-2)^2 - 4(1)(2)}}{2} = 1 \pm \frac{\sqrt{-4}}{2} = 1 \pm i$$

and so $\lambda = \alpha \pm i$ where $\alpha = 1$ and $\beta = 1$. In view of the proposition, our general solution will then be int terms of exponentials, sines and cosines. First, however, we must find eigenvectors associated to these eigenvalues and these will necessarily be of the form $v = \vec{a} \pm i\vec{b}$ – it's really our job to find $\vec{a}$ and $\vec{b}$. To this end, consider $\lambda = 1 + i$ and let's find a non-zero $v = (v_1, v_2)^\intercal \in ker(A - \lambda I)$, i.e., $v = (v_1, v_2)^\intercal$ for which

$$(A - \lambda I)v = \left(\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} - (1+i) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is equivalent to the linear system

$$\begin{cases} -iv_1 - v_2 & = 0 \\ v_1 - iv_2 & = 0 \end{cases}.$$

It should be noted that this system is consistent and underdetermined, i.e., the second equation must be a scalar multiple of the first (multiplying the second equation by $-i$ does the trick). This is by design: we selected our eigenvalue so that the

$\ker(A-\lambda I)$ was non-trivial. By putting $v_2 = 1$, we find that $v_1 = iv_2 = i$ yields a non-zero solution to the above system. In fact, all solutions must be scalar multiples of this one. An eigenvector for A with corresponding eigenvalue $\lambda = \alpha + i\beta = 1 + i$ is therefore

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + i\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \vec{a} + i\vec{b}$$

where

$$\vec{a} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \vec{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

As the proposition indicates, which is characteristic of complex eigenvector/eigenvalue pairs,

$$\vec{a} - i\vec{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - i\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

is necessarily an eigenvector of A with corresponding eigenvalue $\lambda = \alpha - i\beta = 1 - i$. This is quite a handy fact because it means that we needn't find another eigenvector for the remaining eigenvalue $\lambda = 1 - i$, as we would if the eigenvalues were real. Though you won't need to do this while solving linear systems with complex eigenvalues (we actually have all information needed already), let's verify this assertion for completeness. We have

$$A(\vec{a}-i\vec{b}) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -i \\ 1 \end{pmatrix} = \begin{pmatrix} -i-1 \\ -i+1 \end{pmatrix} = \begin{pmatrix} (1-i)(-i) \\ (1-i)(1) \end{pmatrix} = (1-i)\begin{pmatrix} -i \\ 1 \end{pmatrix} = \lambda(\vec{a}-i\vec{b})$$

and so $\vec{a} - i\vec{b}$ is an eigenvector for A with eigenvalue $\alpha - i\beta$, as was asserted.

Since we have identified α, β, $\vec{a}$, and $\vec{b}$, we are ready to write down a general solution to the linear system. In view of the proposition, a general solution is given by

$$\begin{aligned}
\mathbf{x}(t) &= c_1 e^{\alpha t}(\cos(\beta t)\vec{a} - \sin(\beta t)\vec{b}) + c_2 e^{\alpha t}(\cos(\beta t)\vec{b} + \sin(\beta t)\vec{a}) \\
&= c_1 e^t \left(\cos(t)\vec{a} - \sin(t)\vec{b}\right) + c_2 e^t \left(\cos(t)\vec{b} + \sin(t)\vec{a}\right) \\
&= c_1 e^t \left(\cos(t)\begin{pmatrix} 0 \\ 1 \end{pmatrix} - \sin(t)\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + c_2 e^t \left(\cos(t)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sin(t)\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\
&= c_1 \begin{pmatrix} -e^t\sin(t) \\ e^t\cos(t) \end{pmatrix} + c_2 \begin{pmatrix} e^t\cos(t) \\ e^t\sin(t) \end{pmatrix} \\
&= \begin{pmatrix} e^t(c_2\cos(t) - c_1\sin(t)) \\ e^t(c_1\cos(t) + c_2\sin(t)) \end{pmatrix}
\end{aligned}$$

for $t \in \mathbb{R}$. I strongly encourage you to verify that this is indeed a solution to the constant coefficient linear system $\dot{\mathbf{x}} = A\mathbf{x}$. Taking this for granted, let's use it to solve the given initial value problem.

We seek c_1 and c_2 for which $\mathbf{x}(0) = (1,1)^{\intercal}$. Specifically, we see c_1 and c_2 for which

$$\mathbf{x}(0) = \begin{pmatrix} e^0(c_2\cos(0) - c_1\sin(0)) \\ e^0(c_1\cos(0) + c_2\sin(0)) \end{pmatrix} = \begin{pmatrix} 1(c_2\cdot 1 - c_2\cdot 0) \\ 1(c_1\cdot 1 + c_2\cdot 0) \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Of course, choosing $c_1 = 1 = c_2$ gives the desired result and so our unique solution to the IVP is given by

$$\mathbf{x}(t) = \begin{pmatrix} e^t(\cos(t) - \sin(t)) \\ e^t(\cos(t) + \sin(t)) \end{pmatrix}.$$

Remark 4.3.4. In example above, we went out of our way to verify that $\vec{a} - i\vec{b}$ was an eigenvector of A with eigenvalue $\alpha - i\beta$ and we did this for exposition and completeness. To solve the given problem, by making use of the proposition, all that was really needed was α, β, $\vec{a}$ and $\vec{b}$. Thus, if you encounter complex eigenvalues (e.g., in the following exercise), it suffices to simply identify α, β, $\vec{a}$, and $\vec{b}$ and plug them in.

Exercise 48

Find general solutions to the following linear systems:

1. $$\dot{\mathbf{x}}(t) = A\mathbf{x} \qquad \text{where} \qquad A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

2. $$\dot{\mathbf{x}}(t) = A\mathbf{x} \qquad \text{where} \qquad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

3. $$\dot{\mathbf{x}}(t) = A\mathbf{x} \qquad \text{where} \qquad A = \begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}$$

If you're having fun with these calculations, try to do (but don't turn in) the case in which

$$A = \begin{pmatrix} -1 & -1 & 3 \\ 1 & 1 & -1 \\ -1 & -1 & 3 \end{pmatrix}.$$

4.4 The geometry of autonomous systems

A first-order autonomous system of ordinary differential equations is an $n \times n$ system of the form

$$\dot{\mathbf{x}} = F(\mathbf{x})$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$. Of course, this is an $n \times n$ system whose driving function F (present in Theorem 4.2.1) does not depend (explicitely) on time. For the autonomous linear systems studied in this section, we shall make the blanket assumption that F and its first-order partial derivatives are continuous wherever defined so, in particular, the hypotheses of Theorem 4.2.1 are met. Generally speaking, autonomous systems of differential equations model phenomena/games whose laws/rules are fixed in time and the evolution of their solutions depends only on these fixed laws rules and the starting point of the system. As we did for scalar autonomous equations, without loss of generality, we shall focus on systems and their initial value problems whose initial time is $t_0 = 0$.

The characterizations of equilibrium solutions (in terms of eigenvalues) given in class.

Proposition 4.4.1. *Consider the autonomous* 2×2 *system*

$$\dot{\mathbf{x}} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} F_1(x,y) \\ F_2(x,y) \end{pmatrix} = F(\mathbf{x})$$

where $F : \mathbb{R}^2 \to \mathbb{R}^2$ is differentiable. Assume that $\mathbf{x}_0 = (x_0, y_0)^\intercal \in \mathbb{R}^2$ is an equilibrium point for the system, i.e.,

$$F(\mathbf{x}_0) = \begin{pmatrix} F_1(x_0, y_0) \\ F_2(x_0, y_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and assume that the derivative matrix of F at $\mathbf{x}_0 = (x_0, y_0)^\intercal$,

$$DF(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{pmatrix},$$

is non-singular. Let λ_1 and λ_2 be the eigenvalues of $DF(\mathbf{x}_0)$.

1. *If λ_1 and λ_2 are real and negative, then $\mathbf{x}_0$ is a sink.*
2. *If λ_1 and λ_2 are real and positive, then $\mathbf{x}_0$ is a source.*
3. *If λ_1 and λ_2 are real and such that $\lambda_1 < 0 < \lambda_2$, then $\mathbf{x}_0$ is a saddle point.*
4. *If the eigenvalues λ_1, λ_2 are complex (and necessarily complex-conjugates) $\alpha \pm i\beta$ with $\alpha = 0$, then $\mathbf{x}_0$ is a center.*
5. *If the eigenvalues λ_1, λ_2 are complex (and necessarily complex-conjugates) $\alpha \pm i\beta$ with $\alpha > 0$, then $\mathbf{x}_0$ is a spiral source.*
6. *If the eigenvalues λ_1, λ_2 are complex (and necessarily complex-conjugates) $\alpha \pm i\beta$ with $\alpha < 0$, then $\mathbf{x}_0$ is a spiral sink.*

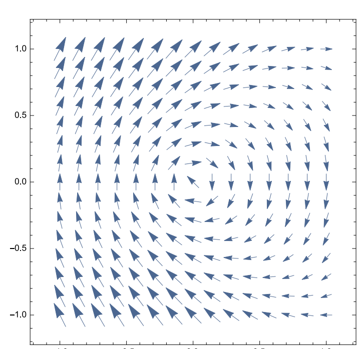
Given a (possibly non-linear) 2×2 system, you should be able to draw a phase portrait of the system. You should know how to find equilibrium values and characterize them as sources, sinks, saddle points, centers, spiral sources and spiral sinks) in line with Proposition 4.4.2 above. Though I covered this topic in lecture in the last week, you haven't been assigned homework on it. Correspondingly below, I've included two worked examples to aid your study.

Example 3

Consider the 2×2 system

$$\dot{\mathbf{x}} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} y \\ y^2 - x \end{pmatrix} = F(x, y)$$

A phase plane for this system is drawn in the figure below (you should try to sketch this by hand).



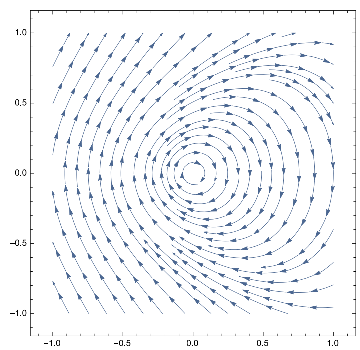
Let's seek equilibrium points for this system. To this end, we seek $(x, y)^\intercal$ such that

$$F(x, y) = \begin{pmatrix} y \\ y^2 - x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

From this we conclude that $y = 0$ and $y^2 - x = 0$ which has only one solution $(x, y) = (0, 0)$. Thus there is only one equilibrium value for the system above at the point $(0, 0)$. At this point, we compute

$$DF(0,0) = \begin{pmatrix} \frac{\partial}{\partial x} y & \frac{\partial}{\partial y} y \\ \frac{\partial}{\partial x}(y^2 - x) & \frac{\partial}{\partial y}(y^2 - x) \end{pmatrix} \Big|_{(x,y)=(0,0)} = \begin{pmatrix} 0 & 1 \\ -1 & 2y \end{pmatrix} \Big|_{(x,y)=(0,0)} = \begin{pmatrix} 0 & 1 \\ -1 & 0. \end{pmatrix}$$

The eigenvalues of this matrix $DF(0,0)$ are easily found to be $\lambda = \pm i$. Thus we classify the equilibrium point $(0, 0)$ as a center and so we expect (local) periodic solutions. This is not surprising given the phase diagram seen above. To confirm this, I have included a phase portrait illustrated in the following diagram.

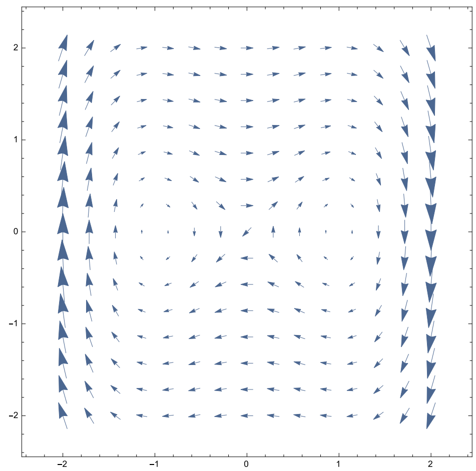


Example 4: Duffing's Equation

Though I gave this example in class, I think it might be useful to illustrate it here. Consider the following system (associated to Duffing's Equation)

$$\dot{\mathbf{x}} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} y \\ x - x^3 \end{pmatrix} = F(x, y).$$

I have plotted a phase diagram for Duffing's Equation below.



To find the equilibrium points of this system, we set $F(x, y) = (0, 0)^\intercal$, i.e., $y = 0$ and $x - x^3 = 0$. Thus the equilibrium points (x, y) have $y = 0$ and $x(1 - x^2) = 0$. The equation for x has three solutions $x = 0$ and $x = \pm 1$. From this we conclude that Duffing's system has three equilibrium points

$$(x, y) = (0, 0), \quad (x, y) = (1, 0) \quad \text{and} \quad (x, y) = (-1, 0).$$

To classify these points, we compute the derivative of F as follows:

$$DF(x, y) = \begin{pmatrix} \frac{\partial}{\partial x} y & \frac{\partial}{\partial y} y \\ \frac{\partial}{\partial x}(x - x^3) & \frac{\partial}{\partial y}(x - x^3) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 - 3x^2 & 0 \end{pmatrix}.$$

At $(x, y) = (0, 0)$, we have

$$DF(0, 0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

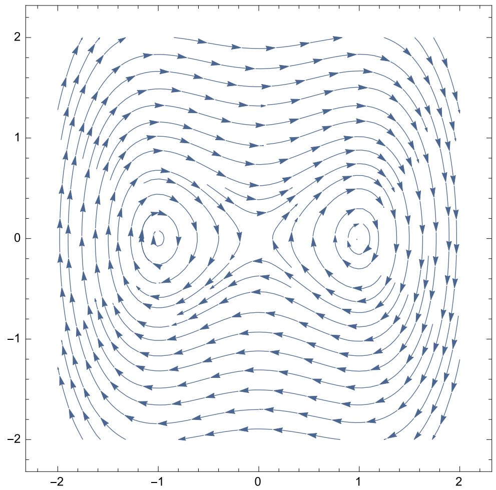
The eigenvalues of this matrix are (easily) computed to be $\lambda = \pm 1$. We may therefore classify $(x, y) = (0, 0)$ as a saddle point. At $(x, y) = (1, 0)$, we have

$$DF(1, 0) = \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}.$$

The eigenvalues of this matrix are (easily) computed to be $\lambda = \pm i\sqrt{2}$. We may therefore classify $(x, y) = (1, 0)$ as a center. Similarly at $(x, y) = (-1, 0)$, we have

$$DF(-1, 0) = \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}.$$

and this also has eigenvalues $\lambda = \pm i\sqrt{2}$. We may therefore classify $(x, y) = (-1, 0)$ as a center. The behavior of solutions near these equilibrium points are easily seen in the phase portrait illustrated below.



Proposition 4.4.2. *Consider the autonomous* 2×2 *system*

$$\dot{\mathbf{x}} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} F_1(x,y) \\ F_2(x,y) \end{pmatrix} = F(\mathbf{x})$$

where $F : \mathbb{R}^2 \to \mathbb{R}^2$ *is differentiable. Assume that* $\mathbf{x}_0 = (x_0, y_0)^\intercal \in \mathbb{R}^2$ *is an equilibrium point for the system, i.e.,*

$$F(\mathbf{x}_0) = \begin{pmatrix} F_1(x_0,y_0) \\ F_2(x_0,y_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and assume that the derivative matrix of F *at* $\mathbf{x}_0 = (x_0, y_0)^\intercal$,

$$DF(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{pmatrix},$$

is non-singular. Let λ_1 *and* λ_2 *be the eigenvalues of* $DF(\mathbf{x}_0)$.

1. *If* λ_1 *and* λ_2 *are real and negative, then* $\mathbf{x}_0$ *is a sink.*
2. *If* λ_1 *and* λ_2 *are real and positive, then* $\mathbf{x}_0$ *is a source.*
3. *If* λ_1 *and* λ_2 *are real and such that* $\lambda_1 < 0 < \lambda_2$, *then* $\mathbf{x}_0$ *is a saddle point.*
4. *If the eigenvalues* λ_1, λ_2 *are complex (and necessarily complex-conjugates)* $\alpha \pm i\beta$ *with* $\alpha = 0$, *then* $\mathbf{x}_0$ *is a center.*
5. *If the eigenvalues* λ_1, λ_2 *are complex (and necessarily complex-conjugates)* $\alpha \pm i\beta$ *with* $\alpha > 0$, *then* $\mathbf{x}_0$ *is a spiral source.*
6. *If the eigenvalues* λ_1, λ_2 *are complex (and necessarily complex-conjugates)* $\alpha \pm i\beta$ *with* $\alpha < 0$, *then* $\mathbf{x}_0$ *is a spiral sink.*

Exercise 49

Consider the 2×2 autonomous system

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} x - xy \\ xy - y \end{pmatrix}.$$

Find and (using the theorem above) classify the equilibrium points of this system.

Note that this is a special case of the Lotka-Volterra system.

Appendix A

Complex Numbers and Complex-Valued Functions of a Real Variable

We now move on to a discussion of complex roots and complex exponentials. As we discussed in class, for a complex number $z = a + ib$, we define

$$e^z = e^a(\cos(b) + i\sin(b)). \tag{A.1}$$

As we did in lecture, you can assume that the following properties hold:

1. For $z = a + ib$ and $w = c + id$,

$$e^{z+w} = e^z e^w.$$

2. For $z = a + ib$ (and so $-z = -a - ib = (-a) + i(-b)$),

$$e^{-z} = \frac{1}{e^z}.$$

Exercise 50: Warm-up (Don't turn this exercise in)

In this exercise, we study some basic properties of the complex exponential function.

1. Using only the properties above and (3.23), verify the angle addition formula for sine and cosine: For $\alpha, \beta \in \mathbb{R}$,

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$$

and

$$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \sin(\beta)\cos(\alpha).$$

Hint: Apply Property 1 in the case that $z = 0 + i\alpha$ and $w = 0 + i\beta$. Note that two complex numbers are equal if and only if their real and imaginary parts are equal.

2. Show that, for any $t \in \mathbb{R}$,

$$\cos(t) = \frac{e^{it} + e^{-it}}{2} \qquad \text{and} \qquad \sin(t) = \frac{e^{it} - e^{-it}}{2i}.$$

Appendix B

The Chain Rule

In this appendix, we present the multivariate chain rule.
Our first group of exercises will focus on some multivariate calculus and the multivariate chain rule. To this end we first focus on differentiability in euclidean space. We recall (from linear algebra) that n-dimensional euclidean space is the (real) vector space

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_k \in \mathbb{R} \text{ for } k = 1, 2, \ldots, n \right\}$$

equipped with the usual component-wise addition and scalar multiplication. Whether one uses column vectors (as above) or row vectors $(x_1, x_2, \ldots, x_n)$ to describe vectors in $\mathbb{R}^n$ is immaterial for most intents and purposes and the convention varies from textbook to textbook. I have chosen to use column vectors for my description because it makes easy the matrix calculations we will do in our description of differentiability and the chain rule. We shall take $\mathbb{R}^n$ to be equipped with the euclidean norm $\|\cdot\|$ defined by

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \qquad \text{for } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

We are interested in studying functions from n-dimensional euclidean space $\mathbb{R}^n$ to m-dimensional euclidean space $\mathbb{R}^m$. Such a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is given by

$$f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} \qquad \text{for } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

where, for $k = 1, 2, \ldots, m$, f_k is called the k-th component function of f and is a function from $\mathbb{R}^n$ to $\mathbb{R}$. In other words, the components of f are scalar-valued functions.

Example 1

An example of such a function is the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \sin x_2 \\ x_1^2 + x_2^2 \end{pmatrix} \qquad \text{for } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2.$$

The components of f, f_1 and f_2, are given by

$$f_1\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \sin(x_2) \qquad \text{and} \qquad f_2\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2 \qquad \text{for } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2.$$

In single-variable calculus, a function is said to be differentiable if the limit of its difference quotients exists. You probably remember that this definition was equivalent to the given function being "well approximated" by an affine function near the point of interest. It is this idea that generalizes to multiple dimensions. In multivariate calculus, a function is said to be differentiable if it is "well approximated" by a linear map near the point of interest. The precise definition is as follows:

Definition B.0.1. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and let $\mathbf{x_0} \in \mathbb{R}^n$. We say that f is differentiable at $\mathbf{x_0}$ if there is an $m \times n$ matrix $Df(\mathbf{x_0})$ such that*

$$\lim_{\mathbf{h}\to 0} \frac{\|f(\mathbf{x_0} + \mathbf{h}) - f(\mathbf{x_0}) - Df(\mathbf{x_0})\mathbf{h}\|}{\|\mathbf{h}\|} = 0. \tag{B.1}$$

Remark B.0.2. In (B.1), the limit is taken as $\mathbf{h} \in \mathbb{R}^n$ goes to zero which, if you recall from M122, is a fairly stringent requirement. Moreover, the euclidean norm in the numerator is the $\mathbb{R}^m$ version whereas the euclidean norm in the denominator is the $\mathbb{R}^n$ version.

As mentioned above, the above definition should be interpreted thus: When f is differentiable at $\mathbf{x_0}$, f can be approximated near $\mathbf{x_0}$ by the affine map $\mathbf{h} \mapsto Df(\mathbf{x_0})\mathbf{h} + f(\mathbf{x}_0)$. As a consequence of the above definition, it can be shown that there is only one matrix for which (B.1) holds; we call this matrix $Df(\mathbf{x_0})$ the Jacobian[1] matrix and it is given by

$$Df(\mathbf{x_0}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

where, because f is differentiable at x_0, the partial derivatives of the components of f, $\partial f_k/\partial x_j$ for $k = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, necessarily exist at $\mathbf{x_0}$. It is important to note (though it won't be for us) that the mere existence of partial derivatives at a point does not guarantee that function be differentiable at that point.

Example 2

Let's consider f from Example B. I claim that f is differentiable at $\mathbf{x_0} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. To see this, we first compute the partial derivatives of f at $\mathbf{x_0}$. We have

$$\frac{\partial f_1}{\partial x_1}(\mathbf{x_0}) = \frac{\partial}{\partial x_1}(\sin x_2)\Big|_{\binom{x_1}{x_2} = \binom{1}{0}} = 0,$$

[1] The Jacobian matrix is named after Carl Jacobi (1804-1851).

$$\frac{\partial f_1}{\partial x_2}(\mathbf{x_0}) = \frac{\partial}{\partial x_2}(\sin x_2)\Big|_{\binom{x_1}{x_2}=\binom{1}{0}} = \cos x_2\Big|_{\binom{x_1}{x_2}=\binom{1}{0}} = 1,$$

$$\frac{\partial f_2}{\partial x_1}(\mathbf{x_0}) = \frac{\partial}{\partial x_1}\left(x_1^2 + x_2^2\right)\Big|_{\binom{x_1}{x_2}=\binom{1}{0}} = 2x_1\Big|_{\binom{x_1}{x_2}=\binom{1}{0}} = 2$$

and

$$\frac{\partial f_2}{\partial x_2}(\mathbf{x_0}) = \frac{\partial}{\partial x_2}\left(x_1^2 + x_2^2\right)\Big|_{\binom{x_1}{x_2}=\binom{1}{0}} = 2x_2\Big|_{\binom{x_1}{x_2}=\binom{1}{0}} = 0$$

Therefore

$$Df(\mathbf{x_0}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}.$$

Now, given $\mathbf{h} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \in \mathbb{R}^2$,

$$\begin{aligned} f(\mathbf{x_0} + \mathbf{h}) - f(\mathbf{x_0}) - Df(\mathbf{x_0})\mathbf{h} & \\ &= f\begin{pmatrix} 1 + h_1 \\ 0 + h_2 \end{pmatrix} - f\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \\ &= \begin{pmatrix} \sin h_2 \\ (1 + h_1)^2 + h_2^2 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} h_2 \\ 2h_1 \end{pmatrix} \\ &= \begin{pmatrix} \sin h_2 \\ 1 + 2h_1 + h_1^2 + h_2^2 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} h_2 \\ 2h_1 \end{pmatrix} \\ &= \begin{pmatrix} \sin h_2 - h_2 \\ h_1^2 + h_2^2 \end{pmatrix}. \end{aligned}$$

Thus, for $\mathbf{h} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$,

$$\begin{aligned} 0 \leq \frac{\|f(\mathbf{x_0} + \mathbf{h}) - f(\mathbf{x_0}) - Df(\mathbf{x_0})\mathbf{h}\|}{\|\mathbf{h}\|} &= \frac{\left\|\begin{pmatrix} \sin h_2 - h_2 \\ h_1^2 + h_2^2 \end{pmatrix}\right\|}{\left\|\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}\right\|} \\ &= \frac{\sqrt{(\sin h_2 - h_2)^2 + (h_1^2 + h_2^2)^2}}{\sqrt{h_1^2 + h_2^2}} \\ &\leq \frac{|\sin h_2 - h_2|}{\sqrt{h_1^2 + h_2^2}} + \frac{|h_1^1 + h_2^2|}{\sqrt{h_1^2 + h_2^2}} \\ &\leq \frac{|\sin h_2 - h_2|}{|h_2|} + \sqrt{h_1^2 + h_2^2} \qquad \text{(B.2)} \end{aligned}$$

where we have used the triangle inequality, $\sqrt{a^2 + b^2} \leq |a| + |b|$, and the fact that dividing by $|h_2|$ instead of $\sqrt{h_1^2 + h_2^2}$ can only make the fraction larger. Upon noting that

$$\lim_{\theta \to 0} \frac{\sin\theta - \theta}{\theta} = 0$$

(this can be seen using elementary trigonometry or a careful study of the power series expansion of the sine function), the inequality (B.2) guarantees that

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|f(\mathbf{x_0}+\mathbf{h}) - f(\mathbf{x_0}) - Df(\mathbf{x_0})\mathbf{h}\|}{\|\mathbf{h}\|} = 0$$

by virtue of the squeeze theorem. Hence f is differentiable at $\mathbf{x_0} = \begin{pmatrix}1\\0\end{pmatrix}$ with Jacobian matrix

$$Df(\mathbf{x_0}) = Df\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0 & 1\\ 2 & 0\end{pmatrix}.$$

Example 3

Consider the function $f : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$f\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x^2\\ x+y\\ 1+y\end{pmatrix}$$

I claim that f is differentiable at $\mathbf{x}_0 = (1,0)^\top$. To this end, it's easy to see that the Jacobian matrix is given by

$$Df(\mathbf{x}_0) = \begin{pmatrix}2 & 0\\ 1 & 1\\ 0 & 1\end{pmatrix}. \tag{B.3}$$

We compute

$$\begin{aligned} f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\mathbf{h}) &= \begin{pmatrix}(1+h_1)^2\\ 1+h_1+h_2\\ 1+h_2\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}2h_1\\ h_1+h_2\\ h_2\end{pmatrix}\\ &= \begin{pmatrix}(1+h_1)^2 - 1 - 2h_1\\ 1+h_1+h_2-1-h_1-h_2\\ 1+h_2-1-h_2\end{pmatrix}\\ &= \begin{pmatrix}h_1^2\\0\\0\end{pmatrix}\end{aligned}$$

Correspondingly

$$\|f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\mathbf{h})\| = \sqrt{(h_1^2)^2+0^2+0^2} = h_1^2$$

and so

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x_0}) - D(\mathbf{h})\|}{\|\mathbf{h}\|} = \lim_{(h_1,h_2)^\top\to(0,0)^\top} \frac{h_1^2}{\sqrt{h_1^2+h_2^2}}.$$

Noting that

$$0 \le \frac{|h_1|}{\sqrt{h_1^2+h_2^2}} \le 1$$

we see that

$$0 \leq \frac{h_1^2}{\sqrt{h_1^2 + h_2^2}} = |h_1| \frac{|h_1|}{\sqrt{h_1^2 + h_2^2}} \leq |h_1|$$

for all $(h_1, h_2)^\top \neq (0,0)^\top$. In view of the squeeze theorem, we find that

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{\|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x_0}) - D(\mathbf{h})\|}{\|\mathbf{h}\|} = \lim_{(h_1,h_2)^\top \to (0,0)^\top} \frac{h_1^2}{\sqrt{h_1^2 + h_2^2}} = 0.$$

Thus f is differentiable at $\mathbf{x}_0 = (1,0)^\top$ and its derivative is given by (B.3)

Now, it's your turn.

Exercise 51

Consider the function $g : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$g(\mathbf{x}) = \begin{pmatrix} 2x_1 + x_2^2 \\ 1 + x_1 + \cos x_3 \end{pmatrix} \qquad \text{for } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3.$$

1. Compute the Jacobian matrix $Dg(\mathbf{x})$ at an arbitrary point $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.
2. As I have done in the example above, using the definition, show that g is differentiable at the point $\mathbf{x_0} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

We are now ready to state the multivariate chain rule.

Theorem B.0.3 (The chain rule)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^k$ and consider the composition $g \circ f : \mathbb{R}^n \to \mathbb{R}^k$ defined by*

$$(g \circ f)(\mathbf{x}) = g(f(\mathbf{x})) = g \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} \qquad \text{for } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

If f is differentiable at $\mathbf{x_0} \in \mathbb{R}^n$ and g is differentiable at $\mathbf{y_0} = f(\mathbf{x_0}) \in \mathbb{R}^m$, then $g \circ f$ is differentiable at $\mathbf{x_0}$ where the Jacobian matrix for $g \circ f$ is given by

$$D(g \circ f)(\mathbf{x_0}) = Dg(\mathbf{y_0}) Df(\mathbf{x_0}) \tag{B.4}$$

where $Dg(\mathbf{y_0})$ is the Jacobian matrix for g evaluated at $\mathbf{y_0} = f(\mathbf{x_0})$ and $Df(\mathbf{x_0})$ is the Jacobian matrix for f evaluated at $\mathbf{x_0}$.

Before stating the next exercise, let's make a couple of notes about the theorem. First, notice how this parallels the chain rule from your introductory calculus course. In single-variable calculus, you learned that

$$(g \circ f)'(x) = g'(f(x)) f'(x)$$

where g and f were functions from $\mathbb{R}$ into $\mathbb{R}$. This is precisely the theorem above where the Jacobian matrices are simply 1×1 matrices consisting of ordinary derivatives. In this case, the matrix multiplication encoded in (B.4) is simply scalar multiplication. The second note about the theorem (and that which is the most satisfying observation for me) has to do with the respective dimensions of Jacobian matrices. First note that, because f is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ and g is a mapping from $\mathbb{R}^m$ to $\mathbb{R}^k$, the composition $g \circ f$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^k$. Correspondingly, the Jacobian matrices Df, Dg and $D(g \circ f)$ have dimensions $m \times n$, $k \times m$ and $k \times n$ respectively. The only way to multiply these matrices to get something sensible is to multiply the $k \times m$ matrix by (from the right) the $m \times n$ matrix which yields a $k \times n$ matrix. If you compare this with (B.4), this is precisely what the theorem says!

Exercise 52: Verifying the chain rule

Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ and $g : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \sin x_1 \\ \cos x_1 \\ x_1 x_2 \end{pmatrix} \qquad \text{for } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$$

and

$$g\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} y_1^2 + y_2^2 \\ y_3 \end{pmatrix} \qquad \text{for } \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3$$

respectively.

1. How many rows and columns does the Jacobian matrix of f have? Compute the Jacobian matrix $Df(\mathbf{x})$ at $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

2. How many rows and columns does the Jacobian matrix of g have? Compute the Jacobian matrix $Dg(\mathbf{y})$ at $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$.

3. Evaluate the Jacobian matrix Dg at the element $\mathbf{y} = f(\mathbf{x})$ where $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and simplify.

4. As it pertains to matrix dimension, confirm that it makes sense to multiply $Dg(f(\mathbf{x}))$ by (from the right) $Df(\mathbf{x})$? Using the results above, multiply and simplify the matrix product.

$$Dg(f(\mathbf{x}))Df(\mathbf{x}).$$

5. Explain why it makes sense to consider the composition $g \circ f$. For which natural numbers a and b does $g \circ f : \mathbb{R}^a \to \mathbb{R}^b$?

6. Compute and simplify

$$h(\mathbf{x}) = (g \circ f)(\mathbf{x}) = (g \circ f)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = g\begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ f_3(\mathbf{x}) \end{pmatrix} \qquad \text{for } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

7. How many rows and columns does the Jacobian matrix of h have? Compute the Jacobian matrix $Dh(\mathbf{x})$ at $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

8. By simplifying your results (if necessary), conclude that (B.4) holds.

Appendix C

Linear Algebra

As so much of our work on linear differential equations makes use of linear algebraic methods, this appendix is dedicated to give a brief account of some important concepts and tools from linear algebra. Let's first recall the definition of a vector space.

Definition C.0.1. *Let V be a non-empty set equipped with two operations $+$ and $\cdot$ described as follows. Given any $v, w \in V$, the vector sum of v and w, denoted by $v + w$, defines another element in V. For any $v \in V$ and $\alpha \in \mathbb{R}$, the multiplication of v by the scalar α, denoted by $\alpha \cdot v$ or αv, defines another element of V. With these two operations, V is said to be a vector space over $\mathbb{R}$ if the following properties are satisfied:*

For any $v, u, w \in V$, $v + (u + w) = (v + u) + w$	*(Associativity)*
For any $v, u \in V$, $v + u = u + v$	*(Communitivity)*
There is an element $\mathbf{0} \in V$, called the zero vector, such that $\mathbf{0} + v = v$ for all $v \in V$.	*(Existence of zero)*
For every $v \in V$, there exists an element $-v \in V$, called the additive inverse of v, such that $v + (-v) = \mathbf{0}$.	*(Existence of inverse)*
For any $\alpha, \beta \in \mathbb{R}$ and $v \in V$, $\alpha(\beta v) = (\alpha\beta)v$	*(Compatibility of multiplication)*
For any $\alpha \in \mathbb{R}$ and $v, u \in V$, $\alpha(u + v) = \alpha u + \alpha v$	*(Distributive Property 1)*
For any $\alpha, \beta \in \mathbb{R}$ and $v \in V$, $(\alpha + \beta)v = \alpha v + \beta v$.	*(Distributive Property 2)*

You should recall from linear algebra that d-dimensional euclidean space $\mathbb{R}^d$ is a vector space (with the usual component-wise addition and scalar multiplication) and, in fact, every finite-dimensional vector space V is isomorphic[1] to $\mathbb{R}^d$ for some natural number d. In this way, the euclidean spaces $\mathbb{R}^d$ represent a very large class of vector spaces. The vector spaces of interest for us live beyond this category.

In what follows, we denote by I an open sub-interval of the real numbers, i.e., $I = (a, b)$

[1] We will discuss isomorphisms shortly.

where $-\infty \le a < b \le \infty$. We define the following spaces of functions:

$$\begin{aligned} C^0(I) &= \{f : I \to \mathbb{R} : f \text{ is continuous on } I\} \\ C^1(I) &= \{f : I \to \mathbb{R} : f \text{ is once differentiable on } I \text{ and } f' \text{ is continuous on } I\} \end{aligned}$$

and

$$C^2(I) = \{f : I \to \mathbb{R} : f \text{ is twice differentiable on } I \text{ and } f'' \text{ is continuous on } I\}.$$

Upon recalling that differentiable functions are necessarily continuous, we recognize that the above function spaces can be equivalently described by

$$C^n(I) = \left\{ f \in C^{n-1}(I) : f^{(n)} \in C^0(I) \right\}$$

for $n = 1, 2$ where the statement that $f^{(n)} \in C^0(I)$ should be taken to mean that the nth derivative of f, $f^{(n)} = \frac{d}{dx} f^{(n-1)}$, exists and is a member of $C^0(I)$. In fact, this definition generalizes for all (positive) natural numbers n, though we will only be interested in $C^0(I)$, $C^1(I)$ and $C^2(I)$. In light of the above definition, we see that

$$C^2(I) \subseteq C^1(I) \subseteq C^0(I) \tag{C.1}$$

where $A \subseteq B$ means that A is a subset of B, i.e., all the members of A are also members of B. On $C^0(I)$ we define addition and scalar multiplication as follows: For $f, g \in C^0(I)$, $f + g$ is the function defined by

$$(f + g)(x) = f(x) + g(x) \qquad \text{for } x \in I.$$

From Calculus 1, you recall that the sum of two continuous functions is a continuous function and therefore $f + g \in C^0(I)$. Similarly, for any $\alpha \in \mathbb{R}$ and $f \in C^0(I)$, αf is the function defined by

$$(\alpha f)(x) = \alpha f(x) \qquad \text{for } x \in I;$$

it is necessarily a member of $C^0(I)$. These notions of addition and scalar multiplication are exactly what you think they are: the addition and scalar multiplication of functions. For example, the functions f and g defined by

$$f(x) = x^2 \qquad \text{and} \qquad g(x) = \sin x \qquad \text{for } x \in \mathbb{R}$$

are members of $C^0(\mathbb{R})$. Their sum is the function $f + g \in C^0(\mathbb{R})$ whose rule is

$$(f + g)(x) = x^2 + \sin x \qquad \text{for } x \in \mathbb{R}.$$

As we discussed previously, $C^0(I)$ is a vector space over $\mathbb{R}$. Though monotonous, checking this fact is a relatively straightforward task; it relies primarily on the fact that $\mathbb{R}$ is itself a vector space over $\mathbb{R}$. Our next concept is the notion of a subspace.

Definition C.0.2. *Let V be a vector space over $\mathbb{R}$ and let S be a non-empty subset of V. The set S is said to be a subspace of V if the following two conditions hold:*

1. *For any $v, u \in S$, $v + u \in S$* *(S is closed under addition)*

2. *For any $\alpha \in \mathbb{R}$ and $v \in S$, $\alpha v \in S$* *(S is closed under scalar multiplication)*

One recalls that a vector space always contains the so-called trivial subspaces $S = V$ and $S = \{\mathbf{0}\}$. That is, the set consisting of only the zero vector is a subspace of V and the whole of V is always a subspace of itself. You will probably recall from linear algebra that, equipped with the addition and scalar multiplication from V, a subspace S of V

can be viewed as a vector space in its own right.

Undoubtedly, you are familiar with subspaces of euclidean space $\mathbb{R}^d$. For instance, the nontrivial subspaces of $\mathbb{R}^3$ are either planes through the origin or lines through the origin. Similarly in a previous homework, we showed that $C^2(I)$ and $C^1(I)$ are subspaces of $C^0(I)$. In fact, it can be checked easily that $C^n(I)$ is a subspace of $C^m(I)$ for all $n \geq m \geq 0$. Further, the set of infinitely differentiable functions on I,

$$C^\infty(I) = \bigcap_{n=1}^{\infty} C^n(I),$$

is a subspace of $C^m(I)$ for every $m \geq 0$.

We now turn to our focus to linear operators between vector spaces. We recall the following definition.

Definition C.0.3. *Let V and W be vector spaces over $\mathbb{R}$. A linear operator from V to W is a function $T : V \to W$ satisfying the property that, for any $\alpha, \beta \in \mathbb{R}$ and $v, w \in V$,*

$$T(\alpha v + \beta w) = \alpha T(v) + \beta T(w).$$

We note that the sum on the left-hand side is the vector sum in V while the sum on the right-hand side is the vector sum in W. A linear operator from V to W is said to be an isomorphism if it is also a bijection (one to one and onto). The vector spaces V and W are said to be isomorphic if there is an isomorphism from V to W.

We recall from linear algebra that linear operators between euclidean spaces are given by matrices. For example, if A is an $m \times n$ matrix, the function $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T_A \mathbf{x} = A\mathbf{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{for } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

defines a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

We recall that the *kernel* of a linear transformation $T : V \to W$ is defined to be the set

$$\ker(T) = \{v \in V : T(v) = \mathbf{0}\}.$$

In the above definition, $\mathbf{0}$ is the zero vector of W. Also, the *range* of T (or image of V under T) is defined to be the set

$$\operatorname{Ran}(T) = T(V) = \{w \in W : T(v) = w \text{ for some } v \in V\}.$$

As it turns out, these objects tell us quite a lot.

Exercise 53

Let V and W be vector spaces and let T be a linear transformation from V to W.

1. Show that $\ker(T)$ is a subspace of V.
2. Show that $\operatorname{Ran}(T)$ is a subspace of W.
3. Show that T is injective (one-to-one) if and only if $\ker(T) = \{\mathbf{0}\}$ (where $\mathbf{0}$ is the zero vector of V).

4. Show that T is onto if and only if $\text{Ran}(T) = W$.

The following exercise continues these ideas further when we have a specific subspace U of V lying around.

Exercise 54

Let $T : V \to W$ be a linear operator from V to W and let U be a subspace of V.

1. Show that $T(U) = \{w \in W : T(u) = w \text{ for some } u \in U\}$ is a subspace of W.
2. The restricted map $S = T|_U$ is the map $S : U \to W$ defined by $S(u) = T(u)$ whenever $u \in U$. Show that S is a linear operator from U to W.
3. If, additionally, T is one-to-one, show that the restricted map S is also one to one.
4. If T is one-to-one, show that the restricted map S is an isomorphism from U onto $T(U)$.

Exercise 55

For any natural number n, consider the space of polynomials $\mathcal{P}_n(I)$ consisting of polynomial functions $p : I \to \mathbb{R}$ of the form

$$p(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$$

where $a_n, a_{n-1}, \ldots, a_1, a_0$ are real numbers.

1. Show that $\mathcal{P}_n(I)$ is a subspace of $C^2(I)$.
2. What is the zero vector in $\mathcal{P}_n(I)$? What are its coefficients?
3. As in the case of euclidean space, polynomials in $\mathcal{P}_n(I)$ can be specified by simply specifying a finite collection of real numbers $a_0, a_1, \ldots, a_n$. Therefore, though it is a subspace of $C^2(I)$, $\mathcal{P}_n(I)$ "feels" a lot like euclidean space. To make this precise, lets define a function $E_n : \mathbb{R}^{n+1} \to \mathcal{P}_n(I)$ in the following way. Given a vector
$$\mathbf{a} = \begin{pmatrix} a_n \\ \vdots \\ a_1 \\ a_0 \end{pmatrix} \in \mathbb{R}^{n+1}$$
define the polynomial $E_n(\mathbf{a})$ by
$$(E_n(\mathbf{a}))(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots a_1 t + a_0.$$
Show that E_n is a linear operator from $\mathbb{R}^{n+1}$ to $\mathcal{P}_n(I)$.
4. Show that E is an isomorphism.
5. Conclude that $C^2(I)$ has an isomorphic copy of $\mathbb{R}^n$ (i.e. a subspace which is isomorphic to $\mathbb{R}^n$) for each n.

Theorem C.0.4. *A vector space V is of dimension d if and only if there exists an isomorphism $T : V \to \mathbb{R}^d$.*

The proof of the preceding theorem is straightforward and appears as standard material in linear algebra; I will therefore omit it. The same goes for the following theorem.

Theorem C.0.5. *Let V be a d-dimensional vector space. If the vectors $v_1, v_2, \dots, v_d \in V$ (of which there are V elements) are linearly independent, then they form a basis of V.*

In our next exercise (which will make use of your results above), we will discuss the "dimension" of $C^2(I)$. It is through this lens that our suspicion that $C^2(I)$ is much much larger than $\mathbb{R}^n$ is confirmed. It is a fact that a vector space V is of dimension n (contains a basis of n elements) if and only if it is isomorphic to $\mathbb{R}^n$, i.e., there is an isomorphism $T : V \to \mathbb{R}^n$. A vector space V is therefore finite dimensional if and only if it is isomorphic to $\mathbb{R}^n$ for some n. A vector space is said to be infinite dimensional if it is not finite dimensional and hence V is not isomorphic to $\mathbb{R}^n$ for any n.

Exercise 56

Let's assume, to reach a contradiction, that $C^2(I)$ is finite dimensional and is therefore isomorphic to $\mathbb{R}^n$ for some n. In this case, there is an isomorphism $T : C^2(I) \to \mathbb{R}^n$. By restricting T to the subspace $\mathcal{P}_n$ of $C^2(I)$, by the penultimate exercise, we obtain an isomorphism S from $\mathcal{P}_n$ onto a subspace $R = T(\mathcal{P}_n)$ of $\mathbb{R}^n$. That is, if $T : C^2(I) \to \mathbb{R}^n$ is an isomorphism, $S : \mathcal{P}_n \to R = T(\mathcal{P}_n)$ defined by

$$S(p) = T(p)$$

for each $p \in \mathcal{P}_n$ is an isomorphism from $\mathcal{P}_n$ onto R which is necessarily a subspace of $\mathbb{R}^n$. Using the results of the previous exercise, let $E_n : \mathbb{R}^{n+1} \to \mathcal{P}_n$ be the isomorphism from $\mathbb{R}^{n+1}$ to $\mathcal{P}_n$.

1. Using only the fact that S and E_n are linear maps (you don't need any special properties of E_n), show that the composition $S \circ E_n$ is defined and is a linear map from $\mathbb{R}^{n+1} \to R$.

2. Using the fact that S and E_n are isomorphisms, conclude that $S \circ E_n$ is an isomorphism from $\mathbb{R}^{n+1}$ to R.

3. Using your result of the above item, show that such an isomorphism T cannot exist and conclude that $C^2(I)$ is infinite dimensional.

Appendix D

Refinements of the Picard-Lindelöff Theorems

Note here

We here prove the Picard-Lindelöff Theorem, a version which is more precise than that given in the body of this text, from which we can deduce the relevant existence and uniqueness theorems relevant for linear equations. Consider a $n \times n$ system given by

$$\dot{\mathbf{x}}(t) = F(t, \mathbf{x}) \tag{D.1}$$

where we shall assume that $F : I \times \mathcal{D} \to \mathbb{R}^n$ where I is a non-trivial open interval of $\mathbb{R}$ and $\mathcal{D}$ is a non-trivial open subset of $\mathbb{R}^n$. Further, we assume that F is continuous on $I \times \mathcal{D}$. Let's also consider the initial value problem

$$\begin{cases} \dot{\mathbf{x}} = F(t, \mathbf{x}) \\ \mathbf{x}(t_0) = \mathbf{x_0} \end{cases} \tag{D.2}$$

for $t_0 \in I$ and $\mathbf{x}_0 \in \mathcal{D}$. We recall: $\mathbf{x}$ is an analytic solution to (D.2) if there is an open interval J of I containing t_0, i.e., $t_0 \in J \subseteq I$, on which $\mathbf{x}$ is continuously differentiable (i.e. $\mathbf{x} \in C^1(J : \mathbb{R}^n)$) and for which

$$\dot{\mathbf{x}}(t) = F(t, \mathbf{x}(t))$$

for all $t \in J$ and $\mathbf{x}(t_0) = \mathbf{x}_0$.

Suppose that $\mathbf{x}(t)$ is an analytic solution to (D.2), then

$$\mathbf{x}(t) - \mathbf{x_0} = \mathbf{x}(t) - \mathbf{x}(t_0) = \int_{t_0}^{t} \dot{\mathbf{x}}(s)\, ds = \int_{t_0}^{t} F(s, \mathbf{x}(s))\, ds$$

for all $t \in J$ by virtue of the fundamental theorem of calculus (Part II). We write this as

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} F(s, \mathbf{x}(s))\, ds \tag{D.3}$$

for $t \in J$; this is called an integral equation for $\mathbf{x}$. Conversely, suppose it is known that $\mathbf{x} \in C^0(J : \mathbb{R}^n)$ solves (D.3) on J. Trivially, we see that

$$\mathbf{x}(t_0) = \mathbf{x}_0 + \int_{t_0}^{t_0} F(s, \mathbf{x}(s))\, ds = \mathbf{x}_0 + 0 = \mathbf{x_0}.$$

By virtue of the fact that F is continuous, we appeal to the fundamental theorem of calculus (Part I) to observe that $\mathbf{x}$ is differentiable and

$$\dot{\mathbf{x}}(t) = \frac{d}{dt}\mathbf{x}(t) = \frac{d}{dt}\mathbf{x}_0 + \frac{d}{dt}\int_{t_0}^{t} F(s,\mathbf{x}(s))\,ds = 0 + F(t,\mathbf{x}(t)) = F(t,\mathbf{x}(t))$$

for all $t \in J$. Furthermore, because we have assumed that F is continuous, we can conclude that $\dot{\mathbf{x}}(t)$ is also continuous. We have therefore proven the following proposition.

Proposition D.0.1. *Let J be an interval containing t_0. A function $\mathbf{x}(t)$ is an analytic solution to the initial value problem* (D.2) *on J (necessarily $\mathbf{x} \in C^1(J : \mathbb{R}^n)$) if and only if $\mathbf{x}(t) \in C^0(J : \mathbb{R}^n)$ and satisfies the integral equation* (D.3) *on J.*

Remark D.0.2. The integrals above are vector-valued Riemann integrals. Their construction is extremely similar to the construction given for real-valued functions and we shall take for granted all standard results. We note, for a function $G : [a,b] \to \mathbb{R}^n$ with components $G_j : [a,b] \to \mathbb{R}$ for $j = 1,2,\dots n$, all Riemann-integrable on $[a,b]$, the Riemann integral of G is the element of $\mathbb{R}^n$ given by

$$\int_a^b G(s)\,ds = \int_a^b \begin{pmatrix} G_1(s) \\ G_2(s) \\ \vdots \\ G_n(s) \end{pmatrix} ds = \begin{pmatrix} \int_a^b G_1(s)\,ds \\ \int_a^b G_2(s)\,ds \\ \vdots \\ \int_a^b G_n(s)\,ds \end{pmatrix}$$

Remark D.0.3. In the proposition above, it's interesting and important to note that the (single) integral equation (D.3) contains all of the information (both the differential equation and the initial condition) of the initial value problem (3.1)

Note here

In view of the preceding proposition, we shall focus on solving the integral equation (D.3) under various conditions on F. The following theorem, while not the strongest possible result, makes use of a general condition on F, called a local Lipschitz condition, under which we establish the existence of solutions.

Theorem D.0.4. *Let $\mathcal{D}$ be a non-empty open subset of $\mathbb{R}^n$, I be an open interval and $F : I \times \mathcal{D} \to \mathbb{R}$. Suppose that F is continuous on $I \times \mathcal{D}$ and, for each $t_0 \in I$ and $x_0 \in \mathcal{D}$, there are positive constants δ_1 and δ_2 for and constant $L = L(t_0, x_0) \geq 0$ for which*

$$|F(s,\mathbf{x}) - F(s,\mathbf{y})| \leq L|\mathbf{x} - \mathbf{y}|$$

for all $s \in [t_0 - \delta_1, t_0 + \delta_1] \subseteq I$ and $\mathbf{x}, \mathbf{y} \in \mathcal{R}$ where

$$\mathcal{R} := \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{x}_0| \leq \delta_2\} \subseteq \mathcal{D}.$$

Then, given $(t_0, \mathbf{x}_0) \in I \times \mathcal{D}$, there exists $\delta > 0$ and $\mathbf{x}(t) \in C^1((t_0 - \delta, t_0 + \delta)$ which satisfies (D.3) *for all $t_0 - \delta < t < t_0 + \delta$.*

Proof. Our proof used the famous iterative method known as *Picard Iteraction*. Given $(t_0, \mathbf{x}_0) \in I \times \mathcal{D}$, let δ_1, δ_2 and L be as given in the statement of the theorem and define

$$\delta = \min\{\delta_1, \delta_2/M\}$$

where

$$M = \sup\{|F(t,\mathbf{x})| : |t - t_0| \leq \delta_1, |\mathbf{x} - \mathbf{x}_0| \leq \delta_2\};$$

this M is finite because F is continuous on $I \times \mathcal{D}$ and $[t_0 - \delta_1, t_0 + \delta_1] \times \mathcal{R}$ is a compact (closed and bounded) subset of $I \times \mathcal{D}$. For $t_0 - \delta \leq t \leq t_0 + \delta$, define

$$\mathbf{x}_1(t) = \mathbf{x}_0 + \int_{t_0}^{t} F(s, \mathbf{x}_0)\, ds$$

and, for each $n \geq 1$,

$$\mathbf{x}_{n+1}(t) = \mathbf{x}_0 + \int_{t_0}^{t} F(s, \mathbf{x}_n(s))\, ds.$$

for $n \geq 1$. Observe that, for each $t_0 - \delta \leq t \leq t_0 + \delta$,

$$|\mathbf{x}_1(t) - \mathbf{x}_0(t)| = \left| \int_{t_0}^{t} F(s, \mathbf{x_0})\, ds \right| \leq \int_{t_0}^{t} |F(s, \mathbf{x}_0)|\; ds \leq M|t - t_0| \leq M\delta \leq \delta_2 \tag{D.4}$$

and therefore $\mathbf{x}_1(t) \in \mathcal{R}$ for all $t_0 - \delta \leq t \leq t_0 + \delta$. Upon noting that, for each $t_0 - \delta \leq t \leq t + 0 + \delta$ and $n \geq 1$,

$$|\mathbf{x}_n(t) - \mathbf{x}_0| \leq \int_{t_0}^{t} |F(s, \mathbf{x}_{n-1}(s)|\; ds$$

an inductive argument (based one exactly reasoning above) ensures that $\mathbf{x}_n(t) \in \mathcal{R}$ for ever $t_0 - \delta \leq t \leq t_0 + \delta$ and $n \in \mathbb{N}$. We claim that, for each $n \geq 1$ and $t_0 - \delta \leq t \leq t_0 + \delta$,

$$|\mathbf{x}_n(t) - \mathbf{x}_{n-1}(t)| \leq M \frac{L^{n-1}|t - t_0|^n}{n!} \leq M \frac{L^{(n-1)}\delta^n}{n!}. \tag{D.5}$$

We have already shown the base case in (D.4). We therefore assume the induction hypothesis for $n \geq 1$. If $t_0 \leq t \leq t_0 + \delta$,

$$\begin{aligned}
|\mathbf{x}_{n+1}(t) - \mathbf{x}_n(t)| &= \int_{t_0}^{t} [F(s, \mathbf{x}_n(s)) - F(s, \mathbf{x}_{n-1}(s))]\; ds \\
&\leq L \int_{t_0}^{t} |\mathbf{x}_n(s) - \mathbf{x}_{n-1}(s)|\; ds \\
&\leq LML^{n-1} \int_{t_0}^{t} \frac{(s - t_0)^n}{n!}\, ds = ML^n \frac{|t - t_0|^{n+1}}{(n+1)!} \leq M \frac{L^n \delta^{n+1}}{(n+1)!}
\end{aligned}$$

where we have made use of the fact that $(s, \mathbf{x}_n(s)), (s, \mathbf{x}_{n-1}(s)) \in [t_0 - \delta, t_0 + \delta] \times \mathcal{R}$ for all $t_0 \leq s \leq t_0 \leq t \leq t_0 + \delta$. If, instead $t_0 - \delta \leq t \leq t_0$, a similar argument (while paying attention to the sign of $(s - t_0)$ yields precisely the same inequality and we have therefore completed the induction argument and thus proved our claim.

We observe that, for each $n \geq 1$ and $t_0 - \delta \leq t \leq t_0 + \delta$,

$$\mathbf{x}_n(t) = \sum_{k=1}^{n} (\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t))$$

given the telescoping nature of the sum. Consequently, provided we can show that the series

$$\sum_{k=1}^{\infty} (\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t))$$

converges, the sequence of functions $\{\mathbf{x}_n\}$ converges and

$$\mathbf{x}(t) = \lim_{n \to \infty} \mathbf{x}_n(t) = \sum_{k=1}^{\infty} (\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t))\,.$$

In fact, we will show that this series converges uniformly and absolutely on the interval $[t_0-\delta, t_0+\delta]$ and, because the summands are all continuous (being defined as integrals of continuous function), the limit $\mathbf{x}(t) \in C^0([t_0-\delta, t_0+\delta])$ by virtue of Theorem 7.12 of [14]. In view of (D.5), the summands $\mathbf{x}_k(t)-\mathbf{x}_{k-1}(t)$ are uniformly bounded in absolute value by $ML^{k-1}\delta^k/k!$ and since

$$\sum_{k=1}^{\infty} ML^{k-1}\frac{\delta^k}{k!} = \begin{cases} Me^{L\alpha}/L & L>0 \\ M & L=0 \end{cases}$$

(in particular, the series converges), the comparison test guarantees that the series $\sum_{k=1}^{\infty}(\mathbf{x}_k(t)-\mathbf{x}_{k-1}(t))$ converges uniformly on $[t_0-\delta, t_0+\delta]$. In other words, the uniform limit

$$\mathbf{x}(t) = \lim_{n\to\infty} \mathbf{x}_n(t)$$

exists for $t \in [t_0-\delta, t_0+\delta]$ and is necessarily a continuous function having the property that $\mathbf{x}(t) \in \mathcal{R}$ for ever $t_0-\delta \le t \le t_0+\delta$. Given that F is jointly continuous and bounded on $[t_0-\delta, t_0+\delta]\times\mathcal{R}$, it follows that

$$\begin{aligned}
\mathbf{x}(t) &= \lim_{n\to\infty} \mathbf{x}_n(t) \\
&= \mathbf{x}_0(t) + \lim_{n\to\infty}\int_{t_0}^{t} F(s, \mathbf{x}_{n-1}(s))\,ds \\
&= \mathbf{x}_0 + \int_{t_0}^{t} F\left(s, \lim_{n\to\infty}\mathbf{x}_{n-1}(s)\right)\,ds \\
&= \mathbf{x}_0 + \int_{t_0}^{t} F(s, \mathbf{x}(s))\,ds.
\end{aligned}$$

for all $t_0-\delta \le t \le t_0+\delta$. □

Note here:

Bibliography

[1] Spring Paint Oscillator, MIT Department of Physics, Technical Service Group. Video available at https://techtv.mit.edu/videos/d4c120a0db764a6b996b90b0de05869f/

[2] Mark S. Ashbaugh, Carmen C. Chicone and Richard H. Cushman. *The Twisting Tennis Racket.* Journal of Dynamics and Differential Equations. **3**(1): 67–85 (1991). doi:10.1007 BF01049489.

[3] Boyce, William E. and Richard C. DiPrima. Elementary Differential Equations and Boundary Value Problems. John Wiley & Sons, Inc. New York, 2nd ed. (1969)

[4] Braun, Martin. Differential Equations and Their Applications. Springer-Verlag, New York, 4th ed. (1991)

[5] L. C. Evans Partial Differential Equations. Graduate Studies in Mathematics, Volume 14. American Mathematical Society, Providence, Rhode Island, 2nd ed. (2010)

[6] Simmons, George F. Differential Equations with Applications and Historical Notes. McGraw-Hill Inc., New York. (1972)

[7] H. Goldstein, C. Poole and J. Safko. Classical Mechanics. Addison Wesley Publishing, San Francisco, 3rd ed. (2002)

[8] E. Isaacson and H. B. Keller. Analysis of Numerical Methods. Dover Publications, New York, 1994.

[9] Sears, Francis Weston. Mechanics, Wave Motion, and Heat. Addison-Wesley Publishing, Reading, MA. (1958)

[10] Otto Bretscher. Linear Algebra with Applications. Pearson Prentice Hall, Upper Saddle River, NJ. 4th ed. (2009)

[11] Jim Hefferon. Linear Algebra VCU mathematics textbook series. Online text available at http://joshua.smcvt.edu/linearalgebra

[12] J. D. Jackson. Classical Electrodynamics. John Wiley & Sons, New York, 3rd ed. (1999)

[13] Peano, Giuseppe (1889), "Sur le déterminant wronskien.", Mathesis (in French), IX: 75–76, 110–112, JFM 21.0153.01

[14] Rudin, Walter. Principles of Mathematical Analysis. McGraw-Hill Inc. New York. (1976)