Multivariable Calculus with Applications to the Sciences

Lecture Notes

Adolfo J. Rumbos

April 30, 2019

Contents

Chapter 1

Preface

All questions of interest in the sciences involve more than one variable and functions of more than one variable. Thus, the single variable Calculus that we have learned up to this point is very limited in its applicability to the analysis of problems arising in the sciences. Even in the case in which the functions of interest in some application can be assumed to be functions of a single variable (as illustrated in the example from epidemiology to be discussed in the next section), the fact that a problem requires more than one of those functions puts us in the realm of multiple variables. It is for that reason that we need to learn the concepts and methods of Multivariable Calculus.

In this course we will learn Multivariable Calculus in the context of problems in the life sciences. Throughout these notes, as well as in the lectures and homework assignments, we will present several examples from Epidemiology, Population Biology, Ecology and Genetics that require the methods of Calculus in several variables.

In addition to applications of Multivariable Calculus, we will also look at problems in the life sciences that require applications of probability. In particular, the use of probability distributions to study problems in which randomness, or chance, is involved, as is the case in the study of genetic mutations.

Chapter 2

Introductory Examples

In this chapter we present two examples that will help motivate the mathematical topics that will be covered in this course. The first example is a system of equations from Epidemiology that provides a simple model for the spread of a contagious disease. The second example is from Population Ecology and prescribes the interactions between predator and prey species in a simple model.

2.1 Modeling the Spread of a Disease

Example 2.1.1 (A simple SIR Model)**.** In a simple mathematical model for a disease that is spread through infections transmitted between individuals in a population, the population is divided into three compartments pictured in Figure 2.1.1. The first compartment, $S(t)$, denotes the set of individuals in the

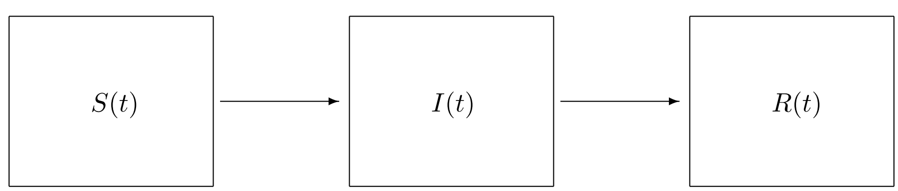

Figure 2.1.1: SIR Compartments

population that are susceptible to acquiring the disease at time t; the second compartment, $I(t)$, denotes the set of infected individual who can also infect others, also at time t; and the third compartment, $R(t)$, denotes the set of individuals who had the disease and who have recovered from the disease at time t.

We assume that the functions S, I and R are differentiable functions of time. Thus, the techniques that we learned in single variable Calculus can be applied to these functions. We also assume that the total number of individuals in the population,

$$N = S(t) + I(t) + R(t),$$

is constant.

Susceptible individuals can get infected through contact with infectious individuals and move to the infected class. This is indicated by the arrow going from the $S(t)$ compartment to the $I(t)$ compartment in Figure 2.1.1. In this simple model, we assume that the individuals in compartment $R(t)$ can no longer get infected.

In addition to the assumptions that we have made so far, we also assume the following:

- The rate at which susceptible individuals get infected is proportional to product of number of susceptible individuals and the number of infected individuals with constant of proportionality $\beta > 0$. We can write this in symbols as
$$\text{Rate of Infection} = \beta SI.$$
- The rate at which infected individuals recover is proportional to the number of infected individuals with constant of proportionality $\gamma > 0$. We can write this in symbols as
$$\text{Rate of Recovery} = \gamma I$$

We would like to understand the flow of individuals from one compartment to another according to the flow arrows pictured in Figure 2.1.1 and the assumptions that we have stated so far. One way to understand the flows is to look at the rates of change of the numbers of individuals in each compartment. For instance, the rate of change of the number of individuals in the infected compartment,

$$I'(t) \quad \text{or} \quad \frac{dI}{dt},$$

has to be accounted for by the rate at which individuals enter the compartment from the susceptible class by way of infections, and the number of individuals that leave the class by way of recovery. We can express this mathematically by means of the equation

$$\frac{dI}{dt} = \beta SI - \gamma I \tag{2.1}$$

The equation in (2.1) is an example of what is known as a *conservation principle*; it expresses the fact that, since the total number of individuals in the population is to remain constant, the rates of change of the number of individuals in a given compartment have to be accounted for by the rates at which individuals enter or leave a given class, or compartment. The expression in (2.1) is also an example of a *differential equation*.

Similar considerations lead to two additional differential equations

$$\frac{dS}{dt} = -\beta SI \tag{2.2}$$

and

$$\frac{dR}{dt} = \gamma I. \tag{2.3}$$

Putting the differential equations in (2.1), (2.2) and (2.3) together leads to the following system of differential equations:

$$\begin{cases} \dfrac{dS}{dt} &= -\beta SI; \\ \dfrac{dI}{dt} &= \beta SI - \gamma I; \\ \dfrac{dR}{dt} &= \gamma I. \end{cases} \tag{2.4}$$

The system in (2.4) is known in the literature as the Kermack–McKendrick SIR model. It first appeared in the scientific literature in 1927.

One of the goals of this course is to develop some of the concepts from Multivariable Calculus that will help us in the analysis of systems like the one in (2.4). An examination of the right–hand side of the equations in (2.4) reveals that the quantities $S(t)$, $I(t)$ and $R(t)$ have to be studied simultaneously, since their rates of change are intertwined. Thus, it makes sense to consider the triple

$$(S(t), I(t), R(t)), \quad \text{for } t \text{ in some interval of time.} \tag{2.5}$$

The expression in (2.5) defines a **vector–valued function** of a single variable, t. As t varies, the image of the function defined in (2.5) traces a curve in three dimensional space, as pictured in Figure 2.1.2. This curve is an example of a **parametrized curve**, an this is where we will begin our study of the topics from Multivariable Calculus in this course.

2.2 Preliminary Analysis of a Simple SIR Model

In many applications, analysis of two dimensional systems suffices to understand and solve the problems. We illustrate this in the following example in which we perform a preliminary analysis of the SIR model developed in Example 2.1.1.

Example 2.2.1 (Preliminary Analysis of a Simple SIR Model)**.** We begin with the observation that, in the system in (2.4), the total size, N, of the population is constant. Thus, from the equation

$$S(t) + I(t) + R(t) = N, \quad \text{for all } t,$$

we can solve for $R(t)$ in terms of $S(t)$ and $I(t)$ to get

$$R(t) = N - S(t) - I(t), \quad \text{for all } t.$$

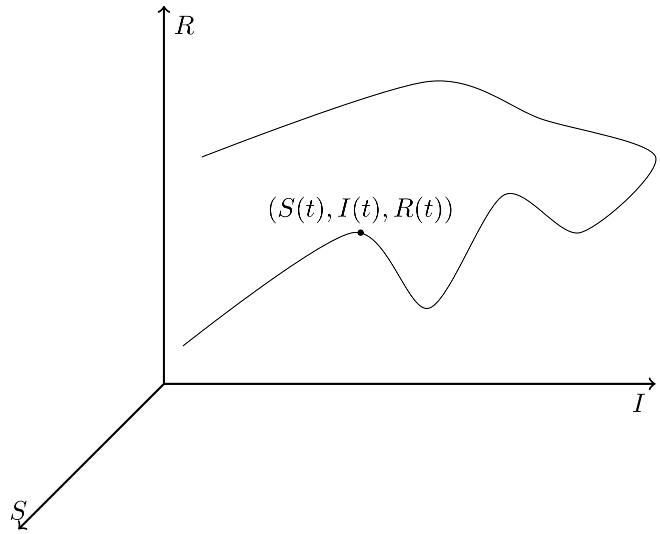

Figure 2.1.2: Curve in SIR–Space

Thus, if we can determine the number of susceptible and infectious individuals at any time t, we'll be able to determine the number of recovered individuals at any time t. Hence, it suffices to study the two–dimensional system

$$\begin{cases} \dfrac{dS}{dt} = -\beta SI; \\ \\ \dfrac{dI}{dt} = \beta SI - \gamma I. \end{cases} \tag{2.6}$$

We would like to determine the pairs $(S(t), I(t))$, which can be pictured as points in the SI–plane, whose components satisfy the equations in (2.6).

Suppose that initially (at time $t = 0$) there are I_o infectious individuals and S_o susceptible individuals. We would like to determine $S(t)$ and $I(t)$ for $t > 0$.

The initial point (S_o, I_o) is shown in Figure 2.2.3, as well as a possible solution curve. In the rest of this example we will see how to justify the shape of the curve drawn in Figure 2.2.3.

The system of equations in (2.6) gives information about the derivatives

$$S'(t) = -\beta S(t)I(t) \tag{2.7}$$

and

$$I'(t) = \beta S(t)I(t) - \gamma I(t), \tag{2.8}$$

of the quantities S and I, respectively. It follows from (2.7) that, in the case in which, S and I are positive, $S'(t) < 0$; so that, the number of susceptible

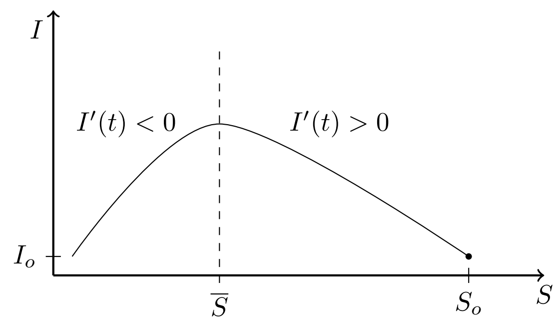

Figure 2.2.3: Curve in the SI–Plane

individuals in the population can only decrease. On the other, rewriting (2.8) as

$$I'(t) = \beta I(t)\left[S(t) - \frac{\gamma}{\beta}\right], \tag{2.9}$$

we see that the answer to the question of whether the number of infected individuals will increase or decrease will depend on whether or not S is bigger than the value

$$\overline{S} = \frac{\gamma}{\beta}. \tag{2.10}$$

The value of $\overline{S}$ is shown in Figure 2.2.3 for the case in which

$$S_o > \overline{S}. \tag{2.11}$$

In this case, for values of t for which $\overline{S} < S(t) \leqslant S_o$, $I'(t) > 0$, as indicated in Figure 2.2.3. Thus, the number of infected individuals will increase, and therefore the disease will spread in this case. Note also that, in the case in which (2.11), the number of infected individuals will increase to a largest value at a time $\bar{t}$ for which $I'(\bar{t}) = 0$ (see Figure 2.2.3). The number of infectious individuals reaching a maximum value indicates an epidemic. After reaching the maximum value, the number of infectious individuals begins to decrease because, according to (2.9) and (2.10), $S(t) < \overline{S}$ implies that $I'(t) < 0$, as shown in Figure 2.2.3.

On the other hand, in the case in which

$$S_o < \overline{S}, \tag{2.12}$$

$S(t) < \overline{S}$ for all $t \geqslant 0$; so that, according to (2.9) and (2.10), $I'(t) < 0$ for all $t \geqslant 0$ and, therefore, the number of infected individuals will decrease from I_o and the disease will not spread.

Finally, observe that, in view of (2.10), the inequality in (2.11) can be rewritten as

$$S_o > \frac{\gamma}{\beta},$$

from which we get that

$$\frac{\beta S_o}{\gamma} > 1. \tag{2.13}$$

The expression on the left–hand side of the inequality in (2.13) is usually denoted by R_o, and is called the **reproduction number**, or **reproductive number**. It is a very important number in epidemiology. When it can be computed, or estimated, R_o provides important information that can be used to determine whether a given disease will spread or not. In particular, since the inequality in (2.13) is equivalent to (2.11), we see that, if $R_o > 1$, the disease will spread. On the other hand, if $R_o < 1$, it follows from (2.12) that the disease will not spread.

2.3 A Predator–Prey System

Examples of applications that are amenable to the two–dimensional analysis illustrated in the previous section are provided by systems that model the interaction of two species that live in the same ecosystem. The simplest of those types of systems is the following predator–prey system known as the Lotka–Volterra system.

Example 2.3.1 (Lotka–Volterra System)**.** Let $x(t)$ and $y(t)$ denote the population densities of two species living in the same ecosystem at time t. We assume that the x and y are differentiable functions of t. Assume also that the population of density y depends solely on the density of the species of density x. We may quantify this by prescribing that, in the absence of the species of density x, the per–capita growth rate of species of density y is a negative constant:

$$\frac{y'(t)}{y(t)} = -\gamma, \quad \text{for all } t \text{ with } x(t) = 0, \tag{2.14}$$

for some positive constant γ. We will see later in this course that (2.14) implies that the population of density y will eventually go extinct in the absence of the species of density x.

On the other hand, in the absence of the species of density y, the species of density x will experience unlimited growth according to

$$\frac{x'(t)}{x(t)} = \alpha, \quad \text{for all } t \text{ with } y(t) = 0, \tag{2.15}$$

where α is a positive constant.

When both species are present, the per–capita growth rate of the species of population density x is given by

$$\frac{x'(t)}{x(t)} = \alpha - \beta y, \quad \text{for all } t, \tag{2.16}$$

where β is a positive constant, and the species of density y has a per–capita growth rate given by

$$\frac{y'(t)}{y(t)} = -\gamma + \delta x, \quad \text{for all } t, \tag{2.17}$$

for some positive constant δ.

The equations in (2.16) and (2.17) describe a predator–prey interaction in which predator species, y, relies solely on the prey species, x, for sustenance; while the only factor that can hinder the growth of the prey species, x, is the presence of the predator species y.

The equations in (2.16) and (2.17) form a system of differential equations,

$$\begin{cases} \dfrac{dx}{dt} &= \alpha x - \beta xy; \\[2ex] \dfrac{dy}{dt} &= \delta xy - \gamma y, \end{cases} \tag{2.18}$$

known as the Lotka–Volterra system. We will analyze this system in subsequent sections in these notes.

Chapter 3

Parametrized Curves

The curve pictured in Figure 2.1.2 is an example of a **parametrized curve** in three–dimensional space. It is the image of a vector–valued function of a single variable. In the example discussed in Section 2.1, this function is given by

$$(S(t), I(t), R(t)), \tag{3.1}$$

for t in some interval of time, J. In this case, we call t a parameter, and the curve traced by the points (3.1) is a parametrized curve in three dimensions.

In many applications, phenomena can be described by parametrized curves in two dimensions. We saw instances of this in Example 2.2.1 and in the Lotka–Volterra system derived in Section 2.3. For that reason, we begin this chapter by studying parametrized curves in the plane.

3.1 Differentiable Paths in the Plane

The set of points $(S(t), I(t))$ discussed in Example 2.2.1 trace out a curve in the SI–place, pictured in Figure 2.2.3, as the parameter t varies. The solutions $(x(t), y(t))$ of the Lotka–Volterra system in (2.18) trace out curves in the xy–plane as t varies. In both of these instances we obtain a parametrized curves in the plane. Before we give the definition of a parametrized curve, we will first define a **differentiable path**.

Definition 3.1.1 (Differentiable Paths in the xy–plane)**.** Let J denote an open interval of real numbers and $x\colon J \to \mathbb{R}$ and $y\colon J \to \mathbb{R}$ denote functions that are differentiable[1] in J. The function $\sigma\colon J \to \mathbb{R}^2$ defined by

$$\sigma(t) = (x(t), y(t)), \quad \text{for } t \in J,$$

[1]The function $x\colon J \to \mathbb{R}$ is differentiable at $t \in J$ means that $\lim_{h\to 0} \dfrac{x(t+h) - x(t)}{h}$ exists. This limit is usually denoted by $x'(t)$ and, geometrically, it gives the slope of the tangent line to the graph of $x = x(t)$ at the point $(t, x(t))$. We say that x is differentiable in J if x is differentiable at every t in J.

is called a differentiable path. Its derivative, $\sigma'(t)$, is given by

$$\sigma'(t) = (x'(t), y'(t)), \qquad \text{for } t \in J.$$

Example 3.1.2. Let $J = \mathbb{R}$ and $\sigma\colon \mathbb{R} \to \mathbb{R}^2$ be given by

$$\sigma(t) = (\cos t, \sin t), \qquad \text{for } t \in \mathbb{R}. \tag{3.2}$$

Since, $\cos\colon \mathbb{R} \to \mathbb{R}$ and $\sin\colon \mathbb{R} \to \mathbb{R}$ are differentiable functions, the path σ is differntiable and

$$\sigma'(t) = (-\sin t, \cos t), \qquad \text{for } t \in \mathbb{R}.$$

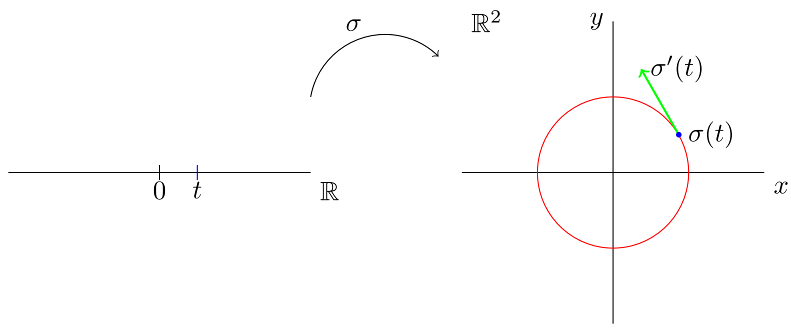

Figure 3.1.1: Sketch of circular path σ in Example 3.1.2

The path σ takes each point t on the real line to a point $\sigma(t) = (x(t), y(t))$ in the xy–plane, $\mathbb{R}^2$, where, according to (3.2),

$$x(t) = \cos t \quad \text{and} \quad y(t) = \sin t, \qquad \text{for all } t \in \mathbb{R}. \tag{3.3}$$

See the sketch in Figure 3.1.1.

It follows from (3.3) that the x and y coordinates of the point $\sigma(t)$ satisfy the equation

$$x^2 + y^2 = 1,$$

which is the equation of a circle of radius 1 centered at the origin in $\mathbb{R}^2$. This circle, which is pictured in Figure 3.1.1, is the image of the path σ.

The derivative of the map σ at the point $\sigma(t)$ is pictured in Figure 3.1.1 as a arrow at $\sigma(t)$. If we imagine the path σ as giving the coordinates of a point in the plane at time t, then $\sigma'(t)$ points in the direction of motion of the point along the circle as time increases.

Definition 3.1.3 (Image of a Path)**.** Given a path $\sigma\colon J \to \mathbb{R}^2$, the image of a the path is the set

$$\{\sigma(t) \mid t \in J\}.$$

Unless σ is the constant function $\sigma(t) = (c_1, c_2)$, where c_1 and c_2 are real numbers, the image of a differentiable path $\sigma\colon J \to \mathbb{R}^2$ is a curve in the plane. For instance, in Example 3.1.2, we saw that the image of $\sigma(t) = (\cos t, \sin t)$, for $t \in \mathbb{R}$, is the circle,

$$C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\},$$

of radius 1 centered at the origin $(0, 0)$.

If $\sigma(t) = (c_1, c_2)$, for all $t \in \mathbb{R}$, where c_1 and c_2 are constant, then the image of σ is the point (c_1, c_2).

A differentiable path $\sigma\colon J \to \mathbb{R}^2$ can be interpreted as modeling the motion of a particle in the xy–plane. At each time $t \in J$, $\sigma(t) = (x(t), y(t)$ gives the coordinates of the point at time t in the Cartesian plane $\mathbb{R}^2$. The derivative of the path, $\sigma'(t)$, in this interpretation, gives the velocity of the particle at time t. We will justify this interpretation in a subsequent section in these notes.

In Example 3.1.2, we saw that a particle moving in the plane according to the path $\sigma\colon \mathbb{R} \to \mathbb{R}^2$ given by $\sigma(t) = (\cos t, \sin t)$, for $t \in \mathbb{R}$, traces a circle of radius 1 centered at the origin,

$$C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}. \tag{3.4}$$

We notice that this circle is traced out once completely if we restrict the domain of σ from $\mathbb{R}$ to the interval $J = [0, 2\pi)$. We then get $\sigma\colon [0, 2\pi) \to \mathbb{R}^2$ given by

$$\sigma(t) = (\cos t, \sin t), \quad \text{for } 0 \leqslant t < 2\pi. \tag{3.5}$$

The path defined in (3.5) is an example of a parametrization of the circle C in (3.4).

3.2 Parametrized Curves in the Plane

Definition 3.2.1 (Parametrized Curves in the xy–plane)**.** Let J denote an interval of real numbers. A differentiable path $\sigma\colon J \to \mathbb{R}^2$ is said to be a parametrization of the curve $C = \{\sigma(t) \mid t \in J\}$ if and only if

(i) $\sigma\colon J \to \mathbb{R}^2$ is one–to–one, or injective, and

(ii) $\sigma'(t) \neq (0, 0)$ for all $t \in J$.

We then say that C is a parametrized curve.

A path $\sigma\colon J \to \mathbb{R}^2$ is one–to–one if and only if, for any two points t_1 and t_2 in J with $t_1 \neq t_2$,

$$\sigma(t_1) \neq \sigma(t_2).$$

Example 3.2.2. Let $\sigma\colon [-\pi, \pi) \to \mathbb{R}^2$ be given by

$$\sigma(t) = (\cos t, \sin t), \quad \text{for } -\pi \leqslant t < \pi. \tag{3.6}$$

We show that σ defined in (3.6) is one–to–one. To do this, we show that, if t_1 and t_2 are in the interval $[-\pi, \pi)$ and $\sigma(t_1) = \sigma(t_2)$, then it must be the case that $t_1 = t_2$.

Thus, suppose that t_1 and t_2 are in the interval in $[-\pi, \pi)$ and that

$$\sigma(t_1) = \sigma(t_2).$$

Then,

$$(\cos t_1, \sin t_1) = (\cos t_2, \sin t_2),$$

from which we get

$$\cos t_1 = \cos t_2, \tag{3.7}$$

and

$$\sin t_1 = \sin t_2. \tag{3.8}$$

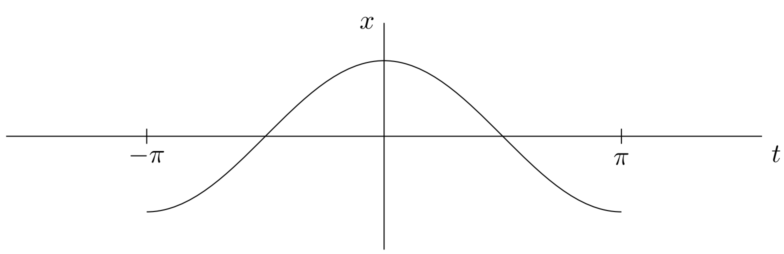

Figure 3.2.2: Sketch of graph of $x = \cos t$ over $[-\pi, \pi)$

It follows from (3.7) and the sketch of the function cos shown in Figure 3.2.2 that, if $t_1 \neq t_2$, then t_1 and t_2 must have opposite signs.

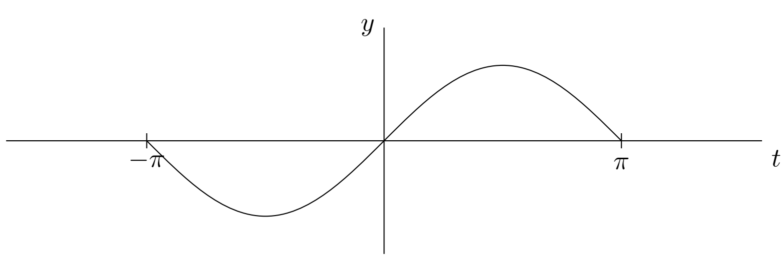

Figure 3.2.3: Sketch of graph of $y = \sin t$ over $[-\pi, \pi)$

However, if t_1 and t_2 have opposite signs, say $t_1 < 0 < t_2$, it follows the sketch of the sin function in Figure 3.2.3 that

$$\sin t_1 < \sin t_2,$$

which is in direct contradiction with the assertion in (3.8), unless $t_2 = \pi$, but we are excluding that case.

Hence, if $\sigma(t_1) = \sigma(t_2)$, and $t_1, t_2 \in [-\pi, \pi)$, it must be the case that $t_1 = t_2$. Therefore, $\sigma\colon [-\pi, \pi) \to \mathbb{R}^2$, where σ is given by (3.6), must be one–to–one.

It follows from the result of Example 3.2.2 that the path $\sigma\colon [-\pi, \pi) \to \mathbb{R}^2$ given by

$$\sigma(t) = (\cos t, \sin t), \quad \text{for } -\pi \leqslant t < \pi,$$

is a parametrization of

$$\{\sigma(t) \mid -\pi \leqslant t \leqslant \pi\} = C,$$

where C is the unit circle in $\mathbb{R}^2$.

In what follows we provide several examples of curves in the plane and their parametrizations.

Example 3.2.3. Let

$$x(t) = e^t, \quad \text{for all } t \in \mathbb{R},$$

and

$$y(t) = e^{-t}, \quad \text{for all } t \in \mathbb{R},$$

and

$$C = \{(x(t), y(t)) \mid t \in \mathbb{R}\}.$$

Then, C is a parametrized curve.

In this example we see how to sketch C.

Observe that $x(t) > 0$ and $y(t) > 0$ for all t because the exponential function is always positive (i.e. $e^a > 0$ for all $a \in \mathbb{R}$); thus, the curve C must lie in the first quadrant. Observe also that

$$x(t) \cdot y(t) = 1, \quad \text{for all } t,$$

so that C must be the portion of the hyperbola

$$xy = 1,$$

that lies in the first quadrant (see Figure 3.2.4).

Note the $(x(0), y(0)) = (1, 1)$; so, the point $(1, 1)$ is on the curve C; this point is shown in the picture in Figure 3.2.4.

Example 3.2.4 (Points)**.** Given a point, (x_o, y_o), in the plane, the constant functions

$$x(t) = x_o, \quad \text{for all } t \in \mathbb{R},$$

and

$$y(t) = y_o, \quad \text{for all } t \in \mathbb{R},$$

form a parametrization of the point (x_o, y_o).

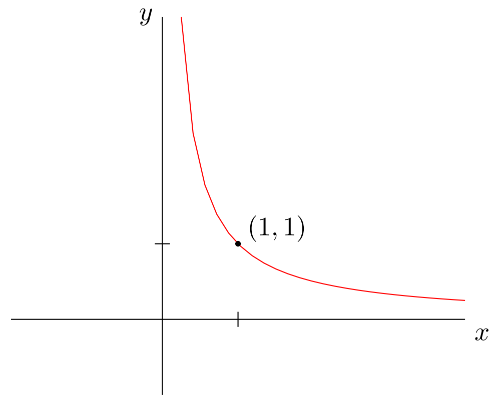

Figure 3.2.4: Sketch of C in Example 3.2.3

Example 3.2.5 (Straight Lines in the Plane)**.** Fix a point (x_o, y_o) in $\mathbb{R}^2$ and let a and b be real numbers such that $a \neq 0$. We consider a curve C given by the parametric equations

$$\begin{cases} x = at + x_o; \\ y = bt + y_o, \end{cases} \tag{3.9}$$

for $t \in \mathbb{R}$.

Since we are assuming that $a \neq 0$, we can solve for t in the first equation in (3.9) to get

$$t = \frac{1}{a}(x - x_o),$$

and then substitute for t in the second equation in (3.9) to get

$$y = \frac{b}{a}(x - x_o) + y_o, \tag{3.10}$$

which is the equation of a straight line through the point (x_o, y_o) with slope

$$m = \frac{b}{a}. \tag{3.11}$$

Figure 3.2.5 shows a sketch of the straight line given by the equation in (3.10) for the case in which the slope m in (3.11) is positive. This is the image of the parametric equations in (3.9) for the case in which $a \neq 0$ and the a and b have the same sign.

If $b = 0$ and $a \neq 0$ in

$$\begin{cases} x = at + x_o; \\ y = y_o, \end{cases}$$

for $t \in \mathbb{R}$. In this case, C is a horizontal line through (x_o, y_o). This line is pictured in Figure 3.2.6 has equation $y = y_o$.

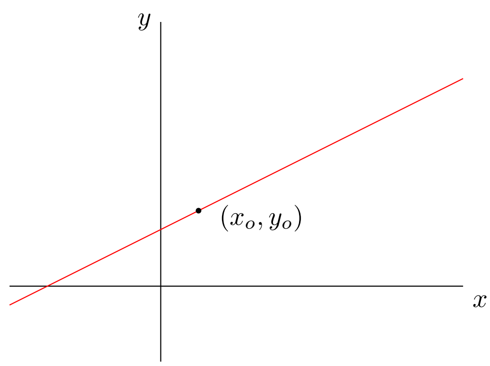

Figure 3.2.5: Sketch of straight line C in Example 3.2.5

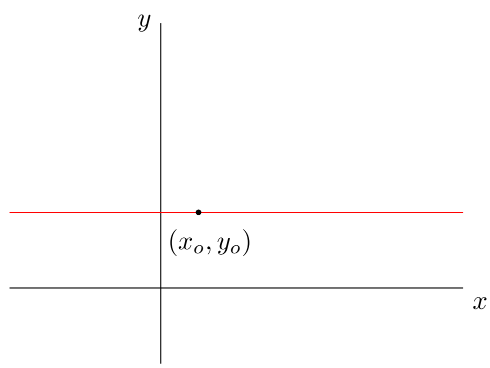

Figure 3.2.6: Sketch of straight line given by (3.10) with $b = 0$

Finally, if $a = 0$ but $b \neq 0$, the parametric equations in (3.9) become

$$\begin{cases} x = x_o; \\ y = bt + y_o, \end{cases}$$

for $t \in \mathbb{R}$.

In this case, the curve C is the vertical line given by the equation $x = x_o$ and pictured in Figure 3.2.7.

Example 3.2.6 (Line Segments)**.** Consider a pair of distinct points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ in the xy–plane. These are shown in the sketch in Figure 3.2.8 for the case $x_1 < x_2$ and $y_2 > y_1$.

We would like to construct a parametrization of the directed line segment that goes from the point P_1 to the point P_2. Assume that $x_1 < x_2$ and put

$$m = \frac{y_2 - y_1}{x_2 - x_1};$$

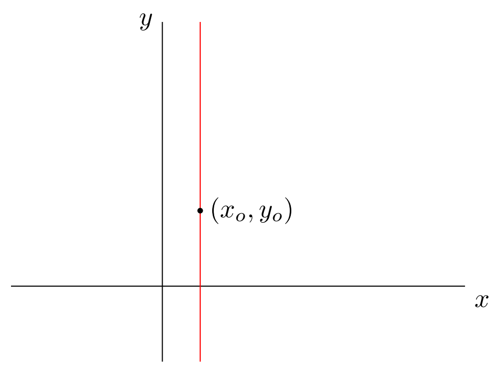

Figure 3.2.7: Sketch of straight line parametrized by (3.9) with $a = 0$

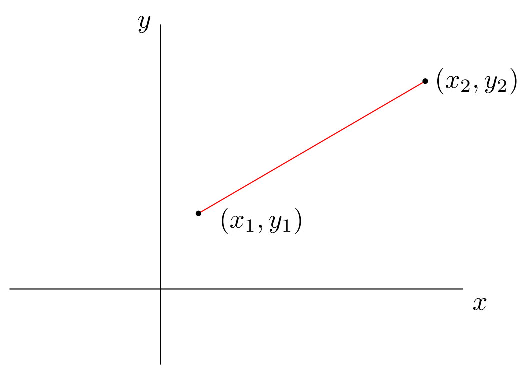

Figure 3.2.8: Sketch of straight line segment in Example 3.2.6

this is the slope of the line segment $\overline{P_1P_2}$. The equation of the straight line going through the points P_1 and P_2 is then given by

$$y = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1. \tag{3.12}$$

The line segment $\overline{P_1P_2}$ is then portion of the straight line given in (3.12) for values of x between x_1 and x_2.

Using the the parametrization

$$x = x_1 + (x_2 - x_1)t, \quad \text{for } 0 \leqslant t \leqslant 1, \tag{3.13}$$

we see that $x = x_1$ when $t = 0$, and $x = x_2$ when $t = 1$.

Substituting the values of x in (3.13) into (3.12) we then get that

$$y = (y_2 - y_1)t + y_1, \quad \text{for } 0 \leqslant t \leqslant 1. \tag{3.14}$$

Combining (3.13) and (3.14), we then get the parametrization

$$\begin{cases} x = (x_2 - x_1)t + x_1; \\ y = (y_2 - y_1)t + y_1, \end{cases} \quad \text{for } 0 \leqslant t \leqslant 1, \tag{3.15}$$

for the straight line segment connecting P_1 to P_2.

we can use the parametric equations in (3.15) to define the differentiable path $\sigma\colon [0,1] \to \mathbb{R}^2$ given by

$$\sigma(t) = ((x_2 - x_1)t + x_1, (y_2 - y_1)t + y_1), \quad \text{for } 0 \leqslant t \leqslant 1. \tag{3.16}$$

Note from (3.16) that $\sigma(0) = (x_1, y_1)$, the coordinates of the point P_1, and $\sigma(1) = (x_2, y_2)$, the coordinates of the point P_2.

Example 3.2.7. Find a parametrization for the directed line segment from $P_1(5,3)$ to $P_2(1,1)$.

Solution: Use the parametric equations in (3.15) in the previous example to compute

$$\begin{cases} x = 5 - 4t; \\ y = 3 - 2t, \end{cases} \quad \text{for } 0 \leqslant t \leqslant 1.$$

A sketch of the segment is shown in Figure 3.2.9. □

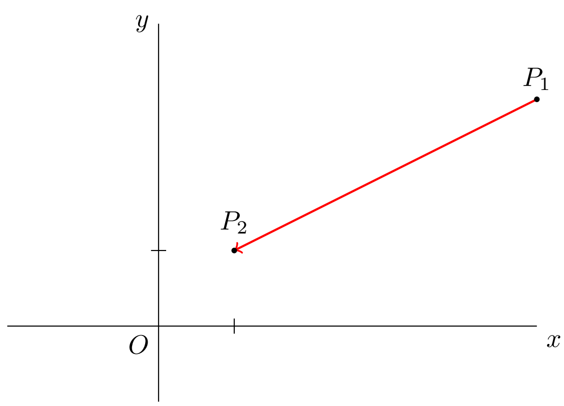

Figure 3.2.9: Sketch of straight line segment from $(5,3)$ to $(1,1)$

Example 3.2.8 (Circles)**.** Let C denote the circle in the xy–plane of radius $r > 0$ and centered at the point (x_o, y_o). Then, every point (x, y) in C is at a distance of r from the point (x_o, y_o). Thus,

$$(x - x_o)^2 + (y - y_o)^2 = r^2. \tag{3.17}$$

Divide both sides of the equation in (3.17) by r^2 to obtain

$$\frac{(x - x_o)^2}{r^2} + \frac{(y - y_o)^2}{r^2} = 1,$$

or

$$\left(\frac{x - x_o}{r}\right)^2 + \left(\frac{y - y_o}{r}\right)^2 = 1. \tag{3.18}$$

Recalling the trigonometric identity

$$\cos^2 t + \sin^2 t = 1, \qquad \text{for all } t \in \mathbb{R},$$

we can set

$$\frac{x - x_o}{r} = \cos t \quad \text{and} \quad \frac{y - y_o}{r} = \sin t,$$

or

$$x - x_o = r\cos t \quad \text{and} \quad y - y_o = r\sin t;$$

so that, the equations

$$\begin{cases} x &= x_o + r\cos t; \\ y &= y_o + r\sin t, \end{cases} \tag{3.19}$$

give a parametrization of C provided the values of the parameter t are confined to an interval of real numbers of length 2π; for example, $0 \leqslant t < 2\pi$, or $-\pi < t \leqslant \pi$. Observe that the direction given by the parametrization in (3.19) is in the counterclockwise sense (see Figure 3.2.10). To see why this assertion is true,

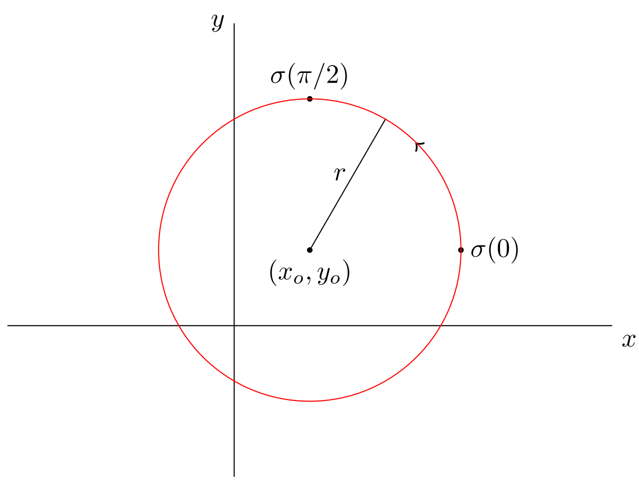

Figure 3.2.10: Sketch of Circle of Radius r and Center (x_o, y_o)

define the path $\sigma\colon [0, 2\pi) \to \mathbb{R}$ by

$$\sigma(t) = (x_o + r\cos t, y_o + r\sin t), \qquad \text{for } 0 \leqslant t < 2\pi,$$

and note that $\sigma(0) = (x_o + r, y_o)$ and, a quarter of the time later, $\sigma(\pi/2) = (x_o, y_o + r)$.

Observe that the choice

$$\frac{x - x_o}{r} = \sin t \quad \text{and} \quad \frac{y - y_o}{r} = \cos t,$$

will also satisfy the equation of the circle of radius r around (x_o, y_o) in (3.18). Thus, the set of equations

$$\begin{cases} x = x_o + r\sin t; \\ y = y_o + r\cos t, \end{cases} \tag{3.20}$$

for $0 \leqslant t < 2\pi$ is also a parametrization of the circle C given by the equation (3.17). However this parametrization is oriented in the clockwise sense (see Figure 3.2.11).

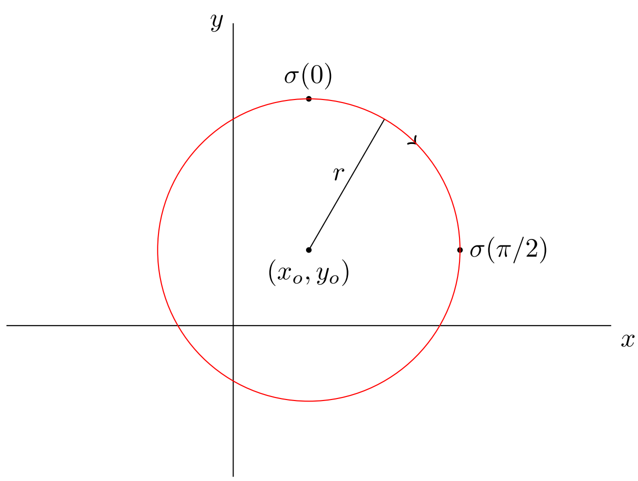

Figure 3.2.11: Sketch of Circle Parametrized by (3.20)

Example 3.2.9. Give a parametrization for the semicircle, C, from the point $P(0, 2)$ to the point $Q(0, 0)$.

Solution: Figure 3.2.12 shows a sketch of C. We use the parametrization in (3.20) with $x_o = 0$, $y_o = 1$ and $r = 1$, with t restricted to $0 \leqslant t \leqslant \pi$, to get

$$\begin{cases} x = \sin t; \\ y = 1 + \cos t, \end{cases} \quad \text{for } 0 \leqslant t \leqslant \pi.$$

□

Example 3.2.10 (Ellipses)**.** The graph of the equation

$$\frac{(x - x_o)^2}{a^2} + \frac{(y - y_o)^2}{b^2} = 1 \tag{3.21}$$

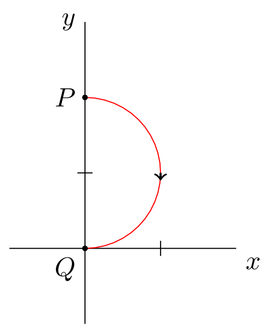

Figure 3.2.12: Sketch of Semicircle C in Example 3.2.9

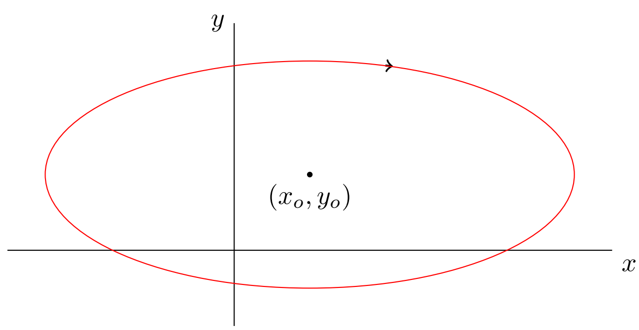

Figure 3.2.13: Sketch of Ellipse in Example 3.2.10

is an ellipse with center (x_o, y_o) and vertices at (x_o-a, y_o), (x_o+a, y_o), (x_o, y_o-b) and (x_o, y_o+b). A possible sketch is shown in Figure 3.2.13. As we did when we constructed parameterizations for circles, we can use the trigonometric identity

$$\cos^2 t + \sin^2 t = 1, \quad \text{for all } t \in \mathbb{R}$$

and set

$$\frac{x - x_o}{a} = \cos t \quad \text{and} \quad \frac{y - y_o}{b} = \sin t,$$

to get

$$x - x_o = a\cos t \quad \text{and} \quad y - y_o = b\sin t;$$

so that, the equations

$$\begin{cases} x = x_o + a\cos t; \\ y = y_o + b\sin t, \end{cases} \quad \text{for } 0 \leqslant t < 2\pi, \tag{3.22}$$

parametrize the ellipse given by (3.21). As was the case for the circle, the parametrization in (3.22) is oriented in the counterclockwise sense as shown in

Figure 3.2.13. Similarly, a parametrization in the clockwise sense is given by the equations

$$\begin{cases} x = x_o + a\sin t; \\ y = y_o + b\cos t, \end{cases} \quad \text{for } 0 \leqslant t < 2\pi, \tag{3.23}$$

has a clockwise orientation.

Example 3.2.11. Let C denote the portion of the ellipse given by the graph of the equation

$$4x^2 + y^2 = 4, \tag{3.24}$$

in the first quadrant of the xy–plane, from the point $P(0, 2)$ to the point $Q(1, 0)$. Give a parametrization for C.

Solution: Figure 3.2.14 shows a sketch of C.

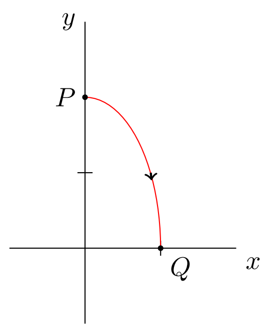

Figure 3.2.14: Sketch of Curve C in Example 3.2.11

Divide both sides of the equation in (3.24) by 4 to get

$$x^2 + \frac{y^2}{4} = 1. \tag{3.25}$$

We see from (3.25) that the ellipse is centered at the origin (so that, $x_o = y_o = 0$) with $a = 1$ and $b = 2$. Since the orientation is in the clockwise sense (see sketch in Figure 3.2.14), we use the parametric equations in (3.23) with t restricted to go from 0 to $\pi/2$:

$$\begin{cases} x = \sin t; \\ y = 2\cos t, \end{cases} \quad \text{for } 0 \leqslant t < \pi/2. \tag{3.26}$$

□

Example 3.2.12 (Graphs of Functions)**.** Let f denote a differentiable function of a single variable defined over some open interval containing a and b, where $a < b$. We let C denote the portion of the graph of $y = f(x)$ for $a \leqslant x \leqslant b$; that is,

$$C = \{(x, f(x)) \in \mathbb{R}^2 \mid a \leqslant x \leqslant b\}.$$

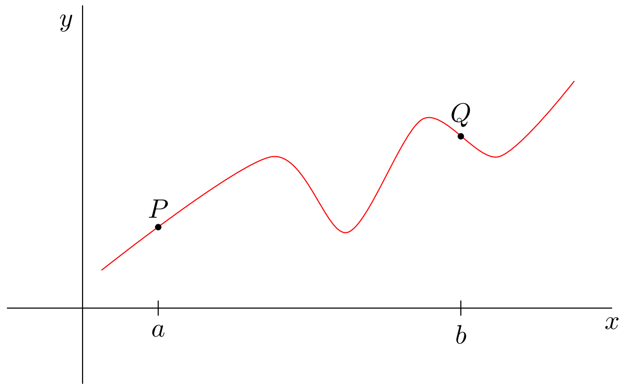

Figure 3.2.15: Sketch of Graph of $y = f(x)$

Figure 3.2.15 illustrates what may happen in a general situation. The curve C is the portion of the graph of f that lies between the points P and Q in Figure 3.2.15.

In order to paremetrize C, we can consider x as a parameter and set

$$x = t;$$

so that

$$y = f(t).$$

Hence, the equations

$$\left\{\begin{array}{rcl} x & = & t; \\ y & = & f(t), \end{array}\right. \qquad \text{for } a \leqslant t \leqslant b, \tag{3.27}$$

will parametrize C.

Example 3.2.13. Let C denote the portion of the parabola given by the equation

$$y = x^2$$

from the point $P(-1, 1)$ to the point $Q(2, 4)$.

Give a parametrization for C.

Solution: A sketch of C is shown in Figure 3.2.16.

Use the equations in (3.27) with $f(x) = x^2$, $a = -1$, and $b = 2$ to get

$$\left\{\begin{array}{rcl} x & = & t; \\ y & = & t^2, \end{array}\right. \qquad \text{for } -1 \leqslant t \leqslant 2.$$

□

In the next example with construct another parametrization of the curve in Example 3.2.11.

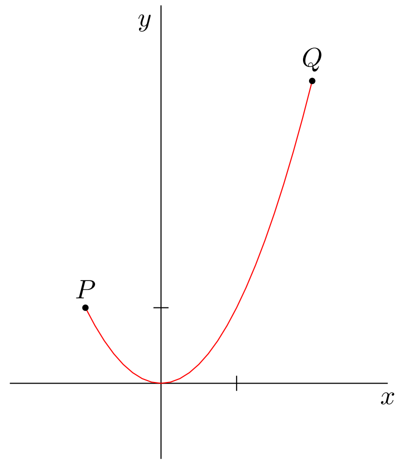

Figure 3.2.16: Sketch of Graph of $y = x^2$ from $x = -1$ to $x = 2$

Example 3.2.14. Let C denote the portion of the ellipse given by the graph of the equation

$$4x^2 + y^2 = 4, \tag{3.28}$$

in the first quadrant of the xy–plane, from the point $P(0, 2)$ to the point $Q(1, 0)$. Give a parametrization for C.

Solution: A sketch of the graph of C is shown in Figure 3.2.14.

Observe that C can also be realized as the graph of a function f that can be obtained by solving the equation in (3.28) for y. We obtain

$$f(x) = 2\sqrt{1 - x^2}, \quad \text{for } -1 \leqslant x \leqslant 1.$$

Thus, the equations

$$\begin{cases} x = t; \\ y = 2\sqrt{1 - t^2}, \end{cases} \quad \text{for } 0 \leqslant t \leqslant 1,$$

also constitute a parametrization of the curve C. □

Chapter 4

Vector Fields

In this course, we will focus on two–dimensional vector fields. These are function, F, from a domain in $\mathbb{R}^2$ to $\mathbb{R}^2$; we write,

$$F\colon D \to \mathbb{R}^2,$$

where D is the domain of the vector field.

In the second section of this chapter, we present a few examples of two–dimensional vector fields and their geometric representation and interpretations, and in the third section we present the concept of the flow of a field, which relates vector fields to that paths and parametrized curves that we studied in previous chapter.

We will first introduce the language of vectors, which will be used throughout the rest of these notes

4.1 Vectors in the Plane

In Example 3.2.6 of Section 3.2 we saw how to parametrize the directed line segment, $\overrightarrow{P_1P_2}$, from the point $P_1(x_1, y_1)$ to the point $P_2(x_2, y_2)$. Indeed, one parametrization is given by the equations in (3.15)

$$\left\{ \begin{array}{rcl} x & = & (x_2 - x_1)t + x_1; \\ y & = & (y_2 - y_1)t + y_1, \end{array} \right. \quad \text{for } 0 \leqslant t \leqslant 1. \tag{4.1}$$

Setting

$$a = x_2 - x_1 \tag{4.2}$$

and

$$b = y_2 - y_1, \tag{4.3}$$

the parametric equations in (4.1) can be rewritten as

$$\left\{ \begin{array}{rcl} x & = & at + x_1; \\ y & = & bt + y_1, \end{array} \right. \quad \text{for } 0 \leqslant t \leqslant 1. \tag{4.4}$$

We can use the parametric equations in (4.4) to define a differentiable path $\sigma\colon [0,1] \to \mathbb{R}^2$ given by

$$\sigma(t) = (at + x_1, bt + y_1), \quad \text{for } 0 \leqslant t \leqslant 1. \tag{4.5}$$

A sketch of the image of the path σ in (4.5) is shown in Figure 4.1.1, for the

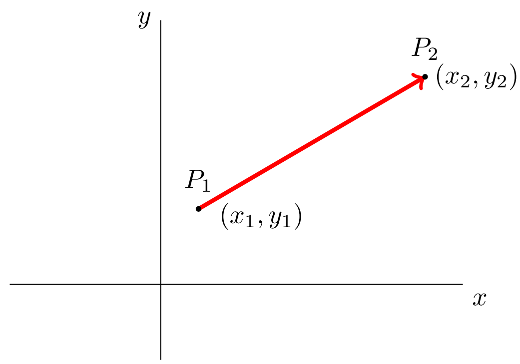

Figure 4.1.1: Sketch of directed line segment from P_1 to P_2

case $a > 0$ and $b > 0$, which, according to (4.2) and (4.3), correspond to $x_1 < x_2$ and $y_1 < y_2$, respectively.

The derivative of the path σ in (4.5) is

$$\sigma'(t) = (a, b), \quad \text{for } 0 \leqslant t \leqslant 1, \tag{4.6}$$

where a and b are given by (4.2) and (4.3), respectively.

The constant path, σ', in (4.6) is an example of a **row vector** in $\mathbb{R}^2$.

Definition 4.1.1 (Row Vectors in $\mathbb{R}^2$)**.** A row vector in $\mathbb{R}^2$ is a pair (a, b), where a and b are real numbers. We can also write (a, b) as

$$[\ a \quad b\ ]. \tag{4.7}$$

The expression in (4.7) is a 1×2 matrix (an arrow of numbers of 1 row and two columns).

In this course, we will also be dealing with column vectors.

Definition 4.1.2 (Column Vectors in $\mathbb{R}^2$)**.** A column vector in $\mathbb{R}^2$ is a the 2×1 matrix

$$\begin{pmatrix} a \\ b \end{pmatrix},$$

where a and b are real numbers.

Remark 4.1.3 (Notations and Conventions)**.** We will treat row vectors and column vectors as distinct objects; thus,

$$[\,a \quad b\,] \neq \begin{pmatrix} a \\ b \end{pmatrix},$$

even though, in some applications, they two vectors might be used to represent the same point (a, b) in the plane.

We say that two column vectors are equal if and only if their corresponding components are equal; that is,

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \quad \text{if and only if} \quad a_1 = a_2 \text{ and } b_1 = b_2.$$

A similar definition of equality applies to row vectors.

We will denote vectors by the symbols v, w, u, etc. Thus, we will write

$$v = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \quad w = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}, \quad \text{and} \quad u = \begin{pmatrix} a_3 \\ b_3 \end{pmatrix},$$

etc.

We note that in other texts you might find the notation $\overrightarrow{v}$, $\overrightarrow{w}$, $\overrightarrow{u}$, etc. for vectors. In these notes, we will avoid putting arrows on top of symbols, except in the case of $\overrightarrow{PQ}$ to indicate a directed line segment from a point P to a point Q. In other texts, you may also find the boldface notation v, w, u, etc. for vectors. We will also avoid this convention.

4.1.1 Interpretations of Vectors

Vectors in $\mathbb{R}^2$ can be used to locate points in the plane; they can also be used to indicated displacement from a point to another.

Vectors as displacements

The directed line segment $\overrightarrow{P_1P_2}$ pictured in Figure 4.1.1 can be used to represent the displacement of a particle moving in a straight line that starts at the point $P_1(x_1, y_1)$ and ends at the point $P_2(x_2, y_2)$. We will denote this displacement by v and express it as the column vector

$$v = \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}. \tag{4.8}$$

If we are only interested in the distance and direction of the displacement, and not in the beginning and ending points of the displacement, we can also draw the vector v as directed line segment that starts at the origin and that has the same length and direction as $\overrightarrow{P_1P_2}$. This is shown in Figure 4.1.2.

A peculiar property that vectors have is that we will consider the arrows shown in Figure 4.1.2 as representing the same vector v given in (4.8). Indeed,

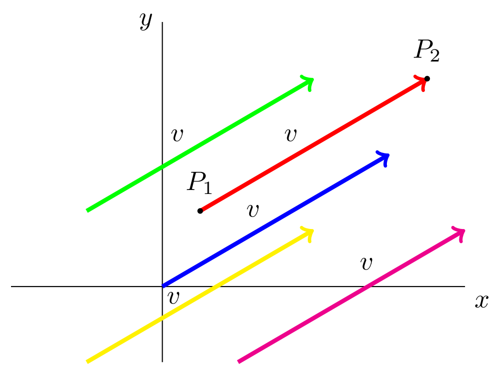

Figure 4.1.2: Vectors as displacements

any arrow that has the same length and same direction as $\overrightarrow{P_1P_2}$ can be used to represent the same vector v. In Figure 4.1.2 we picture a few more of those arrows representing the same vector v. We say that all the arrows representing the vector v are in the same class. This strange property of vectors tuns out to be very useful in many applications in which vectors are used to represent velocities, forces, and other physical entities in which the magnitude and the direction of the entity are what matter.

Vectors as points

When a directed line segment is drawn with its starting point at the origin $O(0,0)$ and its ending point at $P(a,b)$, as shown in Figure 4.1.3, it represents

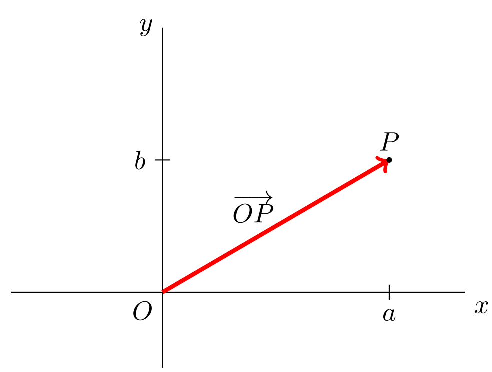

Figure 4.1.3: Vectors as points

the point in $\mathbb{R}^2$ with coordinates (a,b).

We will denote the vector $\overrightarrow{OP}$ by v and will write it a column vector

$$v = \begin{pmatrix} a \\ b \end{pmatrix}. \tag{4.9}$$

where (a, b) are the coordinates of the point P. We say that the vector v is in **standard position** (the starting point of the directed line segment representing the vector is the origin $O(0,0)$). Thus, any point (a, b) in $\mathbb{R}^2$ corresponds to a vector in standard position given by (4.9).

4.1.2 Geometric Properties of Vectors in $\mathbb{R}^2$

Definition 4.1.4 (Norm, or Magnitude, of a Vector)**.** Given a vector $v = \overrightarrow{OP}$ given by (4.9), in standard position, its magnitude, or length, is the distance from the point P to the origin. We will denote this distance by $\|v\|$; so that,

$$\|v\| = \sqrt{a^2 + b^2}, \tag{4.10}$$

and call it the **Euclidean norm** of the vector v.

To see a justification of the definition of the norm of v in (4.10), refer to the sketch in Figure 4.1.4, and apply the Pythagorean Theorem.

For future use and reference here we state some properties of the norm of a vector, which can be derived from Definition 4.1.4. Before we state these properties, we define the scalar multiple cv, where $c \in \mathbb{R}$ and v is given by (4.9) to be

$$cv = \begin{pmatrix} ca \\ cb \end{pmatrix}. \tag{4.11}$$

Proposition 4.1.5 (Properties of the Euclidean Norm)**.** Let v be a vector in $\mathbb{R}^2$ given by (4.9), and $c \in \mathbb{R}$. Then, the following statements are true.

(i) $\|v\| \geqslant 0$ and $\|v\| = 0$ if and only if both components of v are 0; i.e., $a = 0$ and $b = 0$.

(ii) $\|cv\| = |c|\ \|v\|$, where $|c|$ denotes the absolute value of c.

Definition 4.1.6 (Direction of a Vector)**.** Given a vector $v = \overrightarrow{OP}$ given by (4.9), in standard position, denote by θ the angle that the directed line segment $\overrightarrow{OP}$ makes with the positive x–axis. See the sketch in Figure 4.1.4.

If $a > 0$ and $\|v\| > 0$ (as pictured in Figure 4.1.4), then

$$\cos\theta = \frac{a}{\|v\|},$$

from which we get that

$$a = \|v\| \cos\theta. \tag{4.12}$$

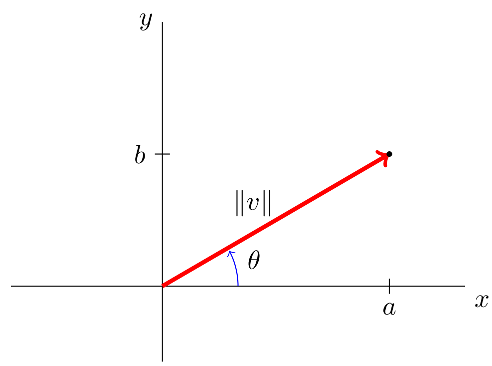

Figure 4.1.4: Norm and direction of a vector

Note that (4.12) gives the first component of the vector v in all cases (i.e., regardless of whether a is positive, negative, or zero.

Similar calculations to those leading to (4.12), with b and $\sin\theta$, instead of a and $\cos\theta$, yield

$$b = \|v\| \sin\theta. \tag{4.13}$$

Combining (4.12) and (4.13), and assuming that $a \neq 0$ and $\|v\| \neq 0$, we get that

$$\tan\theta = \frac{b}{a}. \tag{4.14}$$

The formula in (4.14) can be used to compute θ in terms of the components of the vector.

Example 4.1.7. Let P be a point in the xy–plane with Cartesian coordinates $(-1, 1)$, and put $v = \overrightarrow{OP}$, where O denotes the origin in $\mathbb{R}^2$ Then,

$$v = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

and v is in standard position as pictured in Figure 4.1.5. We compute the norm of v to get, using (4.10),

$$\|v\| = \sqrt{(-1)^2 + 1^2} = \sqrt{2}.$$

Then, using (4.12) and (4.13),

$$\cos\theta = -\frac{1}{\sqrt{2}} \text{ and } \sin\theta = \frac{1}{\sqrt{2}},$$

from which we get that $\theta = \dfrac{3\pi}{4}$.

Alternatively, We could have used the formula in (4.14) for $\tan\theta$ and the information that the point P is in the second quadrant to get that

$$\theta = \pi - \arctan(1) = \pi - \frac{\pi}{4} = \frac{3\pi}{4}.$$

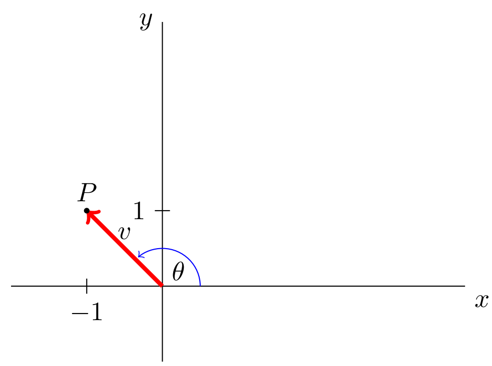

Figure 4.1.5: Vector v in Example 4.1.7

4.1.3 Algebra of Vectors in $\mathbb{R}^2$

There are algebraic operations that we can define in the set of vectors in $\mathbb{R}^2$. We have already mentioned scalar multiplication.

Scalar Multiplication

Given $c \in \mathbb{R}$ and $v = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$, the scalar multiple, cv, is defined to be

$$cv = \begin{pmatrix} ca \\ cb \end{pmatrix}.$$

Example 4.1.8 (Scalar multiples of a vector in standard position)**.** Let

$$L = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; \begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} a \\ b \end{pmatrix}, \ t \in \mathbb{R} \right\}; \tag{4.15}$$

That is, L is the set of scalar multiples of v, where

$$v = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Assume that $\|v\| \neq 0$.

Now, a vector $\begin{pmatrix} x \\ y \end{pmatrix}$ is in L, according to the definition of L in (4.15), if and only if

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} at \\ bt \end{pmatrix}, \quad \text{for some } t \in \mathbb{R},$$

from which we get the parametric equations

$$\left\{ \begin{array}{rcl} x &=& at; \\ y &=& bt, \end{array} \right. \quad \text{for } t \in \mathbb{R}. \tag{4.16}$$

The equations in (4.16) are a parametrization of a straight line from through the origin $(0,0)$ and the point (a,b) in $\mathbb{R}^2$. Thus, L is a straight line in the direction of the vector v. This is shown in Figure 4.1.6. Hence, all the multiples

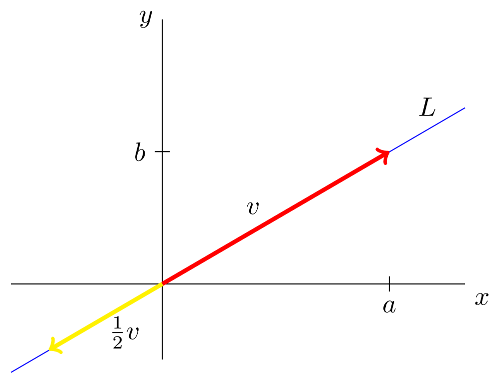

Figure 4.1.6: Line generated by v

of v lie in a line through the origin along the vector v; that is, the line through the points $(0,0)$ and (a,b). We note that, if $t > 0$, tv lies along the direction of v; and, if $t < 0$, tv points in the opposite direction to that of v. The sketch in Figure 4.1.6 shows the vector $-\dfrac{1}{2}v$, for the case in which both a and b are assumed to be positive.

Vector Addition

Given two vectors

$$v = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \quad \text{and } w = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \tag{4.17}$$

in $\mathbb{R}^2$, we define the **vector sum** of v and w, denote $v + w$, by

$$v + w = \begin{pmatrix} a_1 + a_2 \\ b_1 + b_2 \end{pmatrix}; \tag{4.18}$$

that is, $v + w$ is the vector in $\mathbb{R}^2$ whose components are the sum of the corresponding components of v and w.

Example 4.1.9. Let $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $w = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Then,

$$v + w = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+3 \\ 2+1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

Figure 4.1.7 shows a pictorial representation of the vector addition in Example 4.1.9.

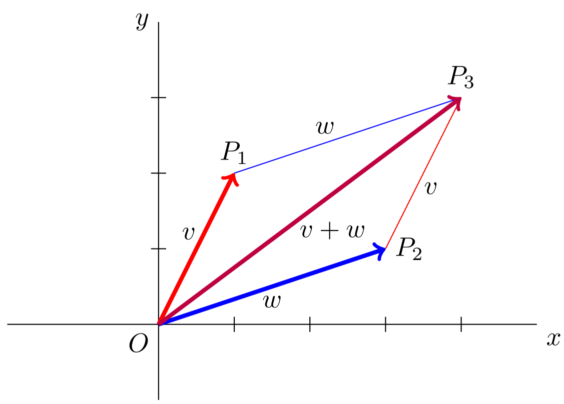

Figure 4.1.7: Parallelogram Rule

The vectors v, w and $v+w$ are drawn in standard position in Figure 4.1.7. The tip of v is then at $P_1(1,2)$, the tip of w is at $P_2(3,1)$, and the tip of $v+w$ is at $P_3(4,3)$. Figure 4.1.7 also shows the vector w translated so that its starting point is at P_1, and the vector v translated so that its starting point is at P_2. We then see that $w = \overrightarrow{P_1P_3}$ and $v = \overrightarrow{P_2P_3}$ in the figure. Consequently, O,P_1, P_2 and P_3 are the vertices of a parallelogram, since $\overrightarrow{OP_1}$ and $\overrightarrow{P_2P_3}$ are parallel, and $\overrightarrow{OP_2}$ and $\overrightarrow{P_1P_3}$ are parallel. Thus, vector addition is also called the **Parallelogram Rule**. This rule also shows that

$$v + w = w + v, \tag{4.19}$$

since, to go from O to P_3 along the parallelogram in Figure 4.1.7, one can go from O to P_1 and then from P_1 to P_3 (this is $v+w$), or from O to P_2 and then from P_2 from P_3 (this is $w+v$).

The parallelogram rule illustrated in Figure 4.1.7 provides a geometric interpretation of vector addition: To compute $v+w$, sketch v in standard position with its tip at point P_1. Then, translate the vector w and place it with starting point at the point P_1 and its tip at the point P_3. The vector sum is the vector from O to P_3. This is a vector along a diagonal of the parallelogram shown in Figure 4.1.7.

The vector expression in (4.19) is called the commutative property of vector addition. It can also be shown algebraically using the definition of vector addition in (4.17) and (4.18). There are other properties of vector addition and scalar multiplication that can be shown algebraically using the definitions of those operations. We state some of those properties in the following proposition.

Proposition 4.1.10 (Properties of Vector Addition and Sclar Multiplication)**.** Let u, v, w denote vectors in $\mathbb{R}^2$ and t and s be scalars.

(i) **Commutativity of Vector Addition**

$$v + w = w + v.$$

(ii) **Associativity of Vector Addition**

$$(u + v) + w = u + (v + w).$$

(iii) **Existence of an Additive Identity**

The vector $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ in $\mathbb{R}^2$ has the property that

$$v + \mathbf{0} = \mathbf{0} + v = v \quad \text{for all } v \text{ in } \mathbb{R}^2.$$

This follows from the fact that $x + 0 = x$ for all real numbers x.

(iv) **Existence of an Additive Inverse**

Given $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ in $\mathbb{R}^2$, the vector w defined by $v = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$ has the property that

$$v + w = \mathbf{0}.$$

The vector w is called an additive inverse of v.

(v) **Associativity of Scalar Multiplication**

Given scalars t and s and a vector v in $\mathbb{R}^2$,

$$t(sv) = (ts)v.$$

(vi) **Identity in Scalar Multiplication**

The scalar 1 has the property that

$$1v = v \quad \text{for all } v \in \mathbb{R}^2.$$

(vii) **Distributive Properties**

Given vectors v and w in $\mathbb{R}^2$, and scalars t and s,

(a) $t(v + w) = tv + tw$;

(b) $(t + s)v = tv + sv$.

Example 4.1.11 (Basis Vectors in $\mathbb{R}^2$)**.** Let $v = \begin{pmatrix} a \\ b \end{pmatrix}$ be any vector in $\mathbb{R}^2$. Using the definitions of vector addition and scalar multiplication we can write

$$v = \begin{pmatrix} a \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ b \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{4.20}$$

Setting

$$\hat{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \hat{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{4.21}$$

we see from (4.20) that the vector v can be written as

$$v = a\hat{i} + b\hat{j}. \tag{4.22}$$

The vectors $\hat{i}$ and $\hat{j}$ given in (4.21) are called the **standard basis** vectors in $\mathbb{R}^2$. The expression on the right–hand side of (4.22) is called a **linear combination** of the vectors $\hat{i}$ and $\hat{j}$. Thus, every vector v in $\mathbb{R}^2$ is a linear combination of the basis vectors $\hat{i}$ and $\hat{j}$.

We note that $\|\hat{i}\| = 1$ and $\|\hat{j}\| = 1$, for the standard basis vectors defined in (4.21). Any vector in $\mathbb{R}^2$ of norm 1 is said to be a **unit** vector.

Definition 4.1.12 (Unit Vectors)**.** A vector $u \in \mathbb{R}^2$ is said to be a unit vector if $\|u\| = 1$.

Given any vector v in $\mathbb{R}^2$, with $\|v\| \neq 0$, the formula

$$\hat{v} = \frac{1}{\|v\|} v$$

gives a unit vector in the direction of v.

As a convention in these notes, a hat ($\hat{}$) on top of a symbol denoting a vector indicates that the vector is a unit vector.

Example 4.1.13. We can use the commutative, associative, and distributive properties stated in Proposition 4.1.10 to express vector addition and scalar multiplication in terms of standard basis in Example 4.1.11.

Suppose v and w are given by

$$v = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad \text{and } w = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \tag{4.23}$$

respectively. Then, as in Example 4.1.11, we can write v and w in terms of the standard basis to get

$$v = x_1\hat{i} + y_1\hat{j} \quad \text{and} \quad w = x_2\hat{i} + y_2\hat{j}.$$

Then,

$$\begin{aligned} v + w &= (x_1\hat{i} + y_1\hat{j}) + (x_2\hat{i} + y_2\hat{j}) \\ &= x_1\hat{i} + y_1\hat{j} + x_2\hat{i} + y_2\hat{j} \\ &= x_1\hat{i} + x_2\hat{i} + y_1\hat{j} + y_2\hat{j} \\ &= (x_1 + x_2)\hat{i} + (y_1 + y_2)\hat{j}; \end{aligned}$$

so, we recover the definition of vector addition given in (4.17) and (4.18).

Similarly, for the case of scalar multiplication, let $c \in \mathbb{R}$ and compute

$$\begin{aligned} cv &= c(x_1\hat{i} + y_1\hat{j}) \\ &= cx_1\hat{i} + cy_1\hat{j}, \end{aligned}$$

where we have used the distributive property in Proposition 4.1.10.

Example 4.1.14 (Vector–Parametric Equation of a Straight line)**.** In this example we use vector addition and scalar multiplication to obtain a parametrization of the straight line through the point $P_o(x_o, y_o)$ in the direction of a nonzero vector v; see Figure 4.1.8.

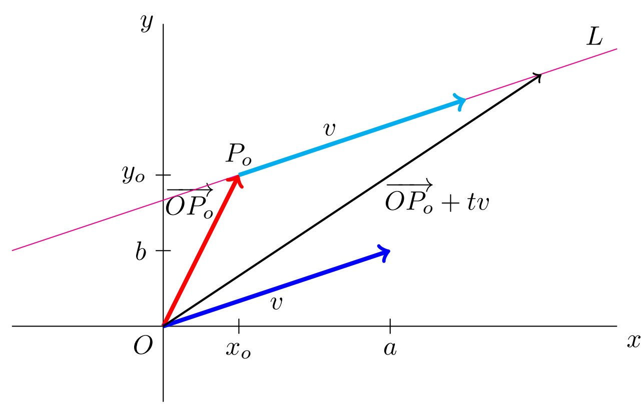

Figure 4.1.8: Vector Parametric Equation of a Straight Line

Suppose the vector v is given by

$$v = a\hat{i} + b\hat{j},$$

where $a^2 + b^2 \neq 0$.

Denote the line through $P_o(x_o, y_o)$ in the direction of v by L. We would like to find a vector–valued function $\sigma\colon \mathbb{R} \to \mathbb{R}^2$ whose image is L.

An arbitrary point

$$\sigma(t) = x(t)\hat{i} + y(t)\hat{j}, \quad \text{for } t \in \mathbb{R},$$

on the line L can be reached from the origin, O, by first going from the origin to the point P_o along the vector $\overrightarrow{OP_o}$ shown in Figure 4.1.8, and then going from the point P_o along the direction of v through a scalar multiple, tv, of the vector v. This is expressed as the vector equation

$$\sigma(t) = \overrightarrow{OP} + tv, \quad \text{for } t \in \mathbb{R}. \tag{4.24}$$

The expression in (4.24) is the **vector–parametric** equation of the line L.

As t varies over all real values, $\sigma(t)$ in (4.24) traces every point on L. For instance, when $t = 0$, $\sigma(0) = \overrightarrow{OP_o}$ determines the point $P(x_o, y_o)$; when $t = 1$, $\sigma(1) = \overrightarrow{OP_o} + v$ is the point at the tip of the vector v that lies on the line L when its starting point is at the point P_o.

The vector $\overrightarrow{OP_o}$ can be written as

$$x_o\hat{i} + y_o\hat{j}. \tag{4.25}$$

The scalar multiple, tv, of $\overrightarrow{v}$ is

$$tv = ta\hat{i} + tb\hat{j}, \tag{4.26}$$

Combining the expressions in (4.24), (4.25) and and (4.26) yields

$$x(t)\hat{i} + y(t)\hat{j} = x_o\hat{i} + y_o\hat{j} + at\hat{i} + bt\hat{j},$$

or

$$x(t)\hat{i} + y(t)\hat{j} = (x_o + at)\hat{i} + (y_o + bt)\hat{j}, \tag{4.27}$$

where we have used the properties of vector–addition and scalar multiplication in the right–hand side of (4.27).

Equating corresponding components of the vectors in (4.27) yields the parametric equations of the line L:

$$\begin{cases} x = x_o + at; \\ y = y_o + bt, \end{cases} \quad \text{for } t \in \mathbb{R},$$

which were the parametric equations of a straight line given in Example 3.2.5.

The Dot Product

Definition 4.1.15 (Transpose of a Vector in $\mathbb{R}^2$)**.** Given a column vector

$$v = \begin{pmatrix} a \\ b \end{pmatrix}$$

in $\mathbb{R}^2$, its **transpose**, denoted v^T, is the row–vector

$$v^T = \begin{bmatrix} a & b \end{bmatrix}.$$

Similarly, the transpose of a row–vector

$$R = \begin{bmatrix} x_1 & y_1 \end{bmatrix}$$

is the column vector

$$R^T = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}.$$

Remark 4.1.16. Column vectors and column vectors are different objects, even though they might have the same entries. As such, we can not add a column vector to a row vector. However, we can define a row–column product and a column–row product. These products will not yield vectors. In fact, we'll see shortly that a row–column product yields a scalar (a real number). On the other hand, a column–row product yields a 2×2 matrix. We will deal with 2×2 matrices in a subsequent chapter in these notes.

Definition 4.1.17 (Row–Column Product)**.** Given a row–vector

$$R = \begin{bmatrix} x_1 & y_1 \end{bmatrix}$$

and a column–vector

$$C = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix},$$

where x_1, x_2, y_1 and y_2 are real numbers. The row–column product, RC, is defined by

$$RC = \begin{bmatrix} x_1 & y_1 \end{bmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = x_1x_2 + y_1y_2. \tag{4.28}$$

We will use the row–column product given in Definition 4.1.17 to define the **dot product** of two vectors in $\mathbb{R}^2$.

Definition 4.1.18 (Dot Product of Vectors in $\mathbb{R}^2$)**.** Given vectors $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, the **dot product** of v and w is the real number (or scalar), denoted by $v \cdot w$, is the row–column of v^T and w. Thus, according to (4.28),

$$v \cdot w = v^T w = \begin{bmatrix} x_1 & y_1 \end{bmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = x_1x_2 + y_2y_2. \tag{4.29}$$

Remark 4.1.19. The dot product, $v \cdot w$, is also called the **inner product**, or scalar product, of v and w. In some texts, $v \cdot w$ is also denoted by $\langle v, w \rangle$.

Using the definition of the dot product given in (4.29), we can derive the following properties of the dot product.

Proposition 4.1.20 (Properties of the Dot Product)**.** Let v, w and u denote vectors in $\mathbb{R}^2$ and c_1 and c_2 denote real numbers.

(i) *(The dot product is positive definite).* $v \cdot v \geqslant 0$ for all $v \in \mathbb{R}^2$ and $v \cdot v = 0$ if and only if v is the zero vector.

(ii) *(The dot product is symmetric).* $v \cdot w = w \cdot v$.

(iii) *(The dot product is bi–linear).* $(c_1 v + c_2 w) \cdot u = c_1 v \cdot u + c_2 w \cdot u$, and $v \cdot (c_1 w + c_2 u) = c_1 v \cdot w + c_2 v \cdot u$.

Example 4.1.21 (The dot product and the Euclidean norm). Let $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ be a vector in $\mathbb{R}^2$ and use (4.29) to compute

$$v \cdot v = x_1^2 + x_2^2. \tag{4.30}$$

Compare (4.30) with the definition of the Euclidean norm in (4.10) to rewrite (4.30) as

$$v \cdot v = \|v\|^2. \tag{4.31}$$

The expression in (4.31) gives us an equivalent definition of the Euclidean norm in terms of the dot product:

$$\|v\| = \sqrt{v \cdot v}.$$

Example 4.1.22 (The dot product and the law of cosines). Let $v = x_1\hat{\imath} + y_1\hat{\jmath}$ and $w = x_2\hat{\imath} + y_2\hat{\jmath}$ denote vectors in $\mathbb{R}^2$ as pictured in Figure 4.1.9.

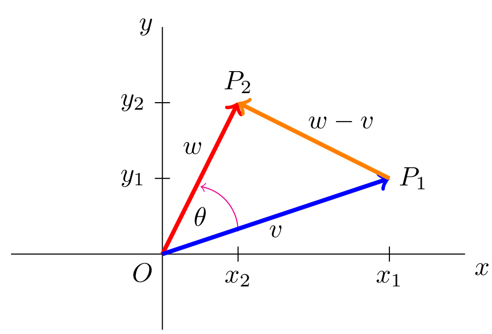

Figure 4.1.9: The dot product and the law of cosines

Assume that, with both v and w in standard position, the tip of v is at a point P_1 and the tip of w is at a point P_2. Then, the directed line segment $\overrightarrow{P_1P_2}$ has the property that

$$v + \overrightarrow{P_1P_2} = w,$$

by virtue of the parallelogram rule for vector addition; thus,

$$\overrightarrow{P_1P_2} = w - v.$$

This vector is sketched in Figure 4.1.9.

With v and w in standard position, we let θ denote the angle between v and w; this is the angle that the vector w makes with the half–line emanating from the origin in the direction of v. This angle is indicated in Figure 4.1.9.

The triangle with vertices O, P_1 and P_2 shown in Figure 4.1.9 has sides of lengths $\|v\|$, $\|w - v\|$ and $\|w\|$, with the angle between the sides of length $\|v\|$ and $\|w\|$ being $\|w\|$. Thus, the law of cosines implies that

$$\|w - v\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\| \cos\theta. \tag{4.32}$$

On the other hand, using the properties of the dot product and (4.31), we compute

$$\begin{aligned} \|w - v\|^2 &= (w - v) \cdot (w - v) \\ &= w \cdot w - w \cdot v - v \cdot w + v \cdot v \\ &= \|w\|^2 - v \cdot w - v \cdot w + \|v\|^2; \end{aligned}$$

so that,

$$\|w - v\|^2 = \|v\|^2 + \|w\|^2 - 2v \cdot w. \tag{4.33}$$

Comparing the equations in (4.32) and (4.33), we obtain that

$$v \cdot w = \|v\|\|w\| \cos\theta, \tag{4.34}$$

which serves as an alternate definition of the dot product of v and w in terms of the Euclidean norms of the vectors and the angle between them.

Example 4.1.23. Let $v = \hat{\imath}$ and $w = \hat{\imath} + \hat{\jmath}$. Then, $\|v\| = 1$, $\|w\| = \sqrt{1^2 + 1^2} = \sqrt{2}$, and

$$v \cdot w = \hat{\imath} \cdot (\hat{\imath} + \hat{\jmath}) = \hat{\imath} \cdot \hat{\imath} + \hat{\imath} \cdot \hat{\jmath} = 1.$$

Thus, if θ is the angle between v and w, according to (4.34),

$$\sqrt{2} \cos\theta = 1,$$

from which we get that

$$\cos\theta = \frac{1}{\sqrt{2}},$$

which implies that

$$\theta = \frac{\pi}{4},$$

or 45°, since v lies along the positive x–axis, and w lies in the first quadrant in the Cartesian plane

The calculations in Example 4.1.23 illustrate the fact that, if $\|v\|$ and $\|w\|$ are not zero, then we obtain from (4.34) that

$$\cos\theta = \frac{v \cdot w}{\|v\|\|w\|}. \tag{4.35}$$

It follows from (4.35) that, if v and w are non–zero vectors, the angle between them is $\frac{\pi}{2}$, or 90°, if and only if

$$v \cdot w = 0. \tag{4.36}$$

Vectors satisfying (4.36) are said to be **orthogonal**, or perpendicular.

Definition 4.1.24 (Orthogonality)**.** Vectors v and w in $\mathbb{R}^2$ are said to be orthogonal if and only if $v \cdot w = 0$.

Remark 4.1.25. Note that, according to Definition 4.1.24, the zero vector, $\mathbf{0} = 0\hat{\imath} + 0\hat{\jmath}$, is orthogonal to every vector in $\mathbb{R}^2$; that is,

$$\mathbf{0} \cdot w = 0, \qquad \text{for every } w \in \mathbb{R}^2.$$

In fact, $\mathbf{0}$ is the only vector in $\mathbb{R}^2$ that is orthogonal to every vector in $\mathbb{R}^2$. This follows from property that the dot product is positive definite.

4.1.4 An Application: Tangent lines and linear approximations

In this section we use the vector–parametric equation of a line to represent the tangent line to a curve C parametrized by a differentiable path $\sigma\colon J \to \mathbb{R}^2$, for some open interval J, at a given point $\sigma(t_o)$ on the curve, for some $t_o \in J$.

Definition 4.1.26 (Tangent line to a curve)**.** Let C be a curve parametrized by a differentiable path $\sigma\colon J \to \mathbb{R}^2$, where J is an open interval. For $t_o \in J$, the vector–parametric of the thangent line to C at the point $\sigma(t_o)$ is given by

$$\ell(t) = \sigma(t_o) + (t - t_o)\sigma'(t_o), \qquad \text{for } t \in \mathbb{R}. \tag{4.37}$$

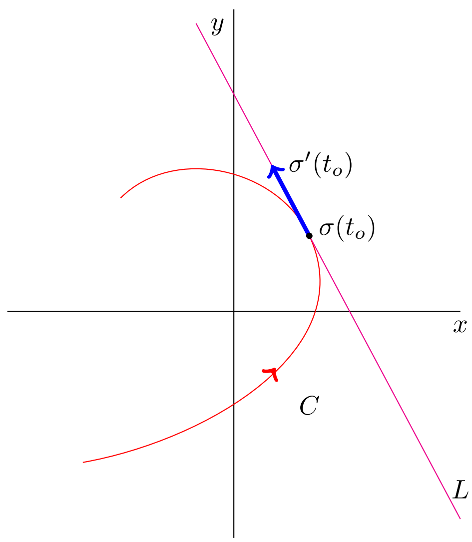

Figure 4.1.10: Sketch of curve C and tangent line at $\sigma(t_o)$

Figure 4.1.10 illustrates what can happen in a general situation. The line L in the figure is parametrized by (4.37).

Example 4.1.27. A curve C is parametrized by the path $\sigma\colon \mathbb{R} \to \mathbb{R}^2$ given by

$$\sigma(t) = t\ \hat{\imath} + t^2\ \hat{\jmath}, \qquad \text{for } t \in \mathbb{R}. \tag{4.38}$$

Sketch the graph of C and give the vector–parametric equation of tangent line to C at the point $(1, 1)$.

Solution: To determine the image of the path σ given in (4.38), set $x = t$ and $y = t^2$, for $t \in \mathbb{R}$, so that $= x^2$. Thus, C is a parabola with vertex at the origin sketched in Figure 4.1.11.

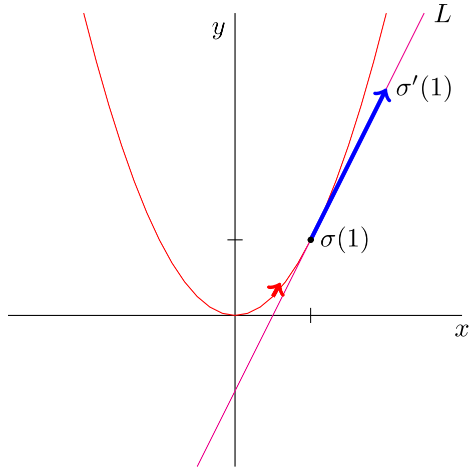

Figure 4.1.11: Sketch for Example 4.1.27

The point $(1, 1)$ correspond to $t_o = 1$. Thus, according to (4.37), the vector–parametric equation of the tangent line to C at the point $\sigma(1)$ is

$$\ell(t) = \sigma(1) + (t-1)\sigma'(1), \qquad \text{for } t \in \mathbb{R}, \tag{4.39}$$

where, differentiating the expression for σ in (4.38) with respect to t,

$$\sigma'(t) = \hat{i} + 2t\hat{j}, \qquad \text{for } t \in \mathbb{R};$$

so that,

$$\sigma'(1) = \hat{i} + 2\hat{j}.$$

This vector is sketched in Figure 4.1.11, and the line, L, parametrized by (4.39) is also shown in the sketch. □

An examination sketch in Figure 4.1.10, and the sketch in Figure 4.1.10 in Example 4.1.27, suggests that the tangent line can be used to approximate the path σ by the tangent line ℓ when t is close to t_o. We write

$$\sigma(t) \approx \ell(t) \qquad \text{for } t \text{ close to } t_o,$$

or,

$$\sigma(t) \approx \sigma(t_o) + (t - t_o)\sigma'(t_o), \qquad \text{for } t \text{ close to } t_o. \tag{4.40}$$

The right–hand side of (4.40) is called the **linear approximation** to the path σ for t near t_o.

Example 4.1.28. For the path $\sigma\colon \mathbb{R} \to \mathbb{R}^2$ given in Example 4.1.27, use the linear approximation to estimate $\sigma(1.1)$.

Solution: Use the linear approximation

$$\sigma(t) \approx \sigma(1) + (t-1)\sigma'(1), \quad \text{for } t \text{ near } 1,$$

to estimate

$$\sigma(1.1) \approx \sigma(1) + (1.1-1)\sigma'(1),$$

or

$$\sigma(1.1) \approx \hat{\imath} + \hat{\jmath} + (0.1)(\hat{\imath} + 2\hat{\jmath}),$$

or

$$\sigma(1.1) \approx 1.1\hat{\imath} + 1.2\hat{\jmath}.$$

□

4.2 Vector Fields in the Plane

A path, $\sigma\colon J \to \mathbb{R}^2$, is a vector–valued function that maps a real value $t \in J$ to a vector $\sigma(t)$ in the Cartesian plane. The vector $\sigma(t)$, drawn in standard position, can be viewed as the location of a particle in motion at time t. In this section we discuss another type of vector–valued function. This time the domain of the function is a subset, D, of the plane, instead of an open interval. Denoting this function by F, we have that

$$F\colon D \to \mathbb{R}^2. \tag{4.41}$$

A function F in (4.41), where D is a region of the Cartesian plane is called a **vector field**.

4.2.1 Examples of Vector Fields

We begin with examples in which the domain, D, of a vector field, F, is the entire xy–plane. We then have that

$$F\colon \mathbb{R}^2 \to \mathbb{R}^2.$$

We have already encountered examples of vector fields in these notes. In Example 2.3.1, we derived the system of differential equations

$$\begin{cases} \dfrac{dx}{dt} &= \alpha x - \beta xy; \\ \\ \dfrac{dy}{dt} &= \delta xy - \gamma y, \end{cases} \tag{4.42}$$

which models the interaction of two species in an ecosystem. The variable y, which is a differentiable function of time t, represents the population size, or density, of species (the predator species) that depends solely on a prey species of population density x, which is also a differentiable function of t, survival. The parameters α, β, γ and δ in (4.42) are assumed to be positive.

Consider the case in which

$$\alpha = \beta = \gamma = \delta = 1$$

in (4.42). We then obtain the system

$$\begin{cases} \dfrac{dx}{dt} & = & x - xy; \\ \dfrac{dy}{dt} & = & xy - y. \end{cases} \tag{4.43}$$

The system in (4.43) can be written in vector form as

$$\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} x - xy \\ xy - y \end{pmatrix}. \tag{4.44}$$

The left–hand side of the vector equation in (4.44) is the derivative of a differentiable path $\sigma\colon J \to \mathbb{R}^2$, for some open interval J, given by

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad \text{for } t \in J.$$

Denoting the right–hand side of the equation in (4.44) by $F(x, y)$, we can rewrite the equation in (4.44) as

$$\sigma'(t) = F(x, y), \tag{4.45}$$

where

$$F(x, y) = \begin{pmatrix} x - xy \\ xy - y \end{pmatrix}, \quad \text{for } (x, y) \in \mathbb{R}^2. \tag{4.46}$$

The function $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ in (4.46) is an example of a vector field. Each point $(x, y) \in \mathbb{R}^2$ gets assigned a vector, $F(x, y)$. Pictorially, the vector $F(x, y)$ gets drawn in the xy–plane as an arrow (directed line segment) with its starting point at (x, y). For instance, at $(1, 0)$ the vector field F in (4.46) yields the vector

$$F(1, 0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \hat{\imath}.$$

This vector is shown in Figure (4.2.12) with its starting point at $(1, 0)$.

Figure 4.2.12) also shows the vectors

$$F(1, 1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad F(0, 0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

the zero vector. These are shown as dots at $(1, 1)$ and the origin, respectively, in Figure 4.2.12.

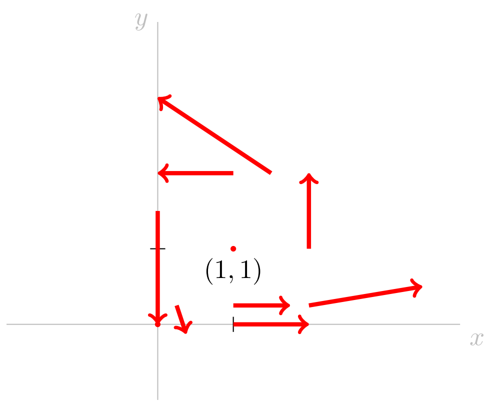

Figure 4.2.12: Sketch of vector field in (4.46)

Figure 4.2.12 also shows the vectors

$$F(0.25, 0.25) = \begin{pmatrix} 0.125 \\ -0.125 \end{pmatrix}, \ F(1, 0.25) = \begin{pmatrix} 0.75 \\ 0 \end{pmatrix}, \ F(2, 0.25) = \begin{pmatrix} 1.5 \\ 0.25 \end{pmatrix},$$

$$F(2, 1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \ F(1.5, 2) = \begin{pmatrix} -1.5 \\ 1 \end{pmatrix}, \ F(1, 2) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

and

$$F(0, 1.5) = \begin{pmatrix} 0 \\ -1.5 \end{pmatrix}.$$

We imagine that we plot the vectors $F(x, y)$ with their starting points at (x, y), for all $(x, y) \in \mathbb{R}^2$, to visualize the vector field in (4.46). This will yield a picture of the vector field that is more complete than the one shown in Figure 4.2.12. We can also use a mathematical software package to get a better picture of the vector field in (4.46). For instance, we can use WolframAlpha® to get a plot of the vector field in (4.46) in the first quadrant by typing

```
plot {x-xy, -y+xy}, {x,0,4},{y,0,4}
```

We obtain a graphics output shown in Figure 4.2.13. We note that the scale of the magnitudes of the vectors plotted by WolframAlpha® is different from the one that we used in the plot in Figure 4.2.12.

4.2.2 The Flow of a Vector Field

The WolframAlpha® command

```
plot {x-xy, -y+xy}, {x,0,4},{y,0,4}
```

yields a second graphics output shown in Figure 4.2.14.

We shall refer to the curves shown in Figure 4.2.14 as the **flow lines** of the vector field F given in (4.46). The WolframAlpha® graphics output also refers

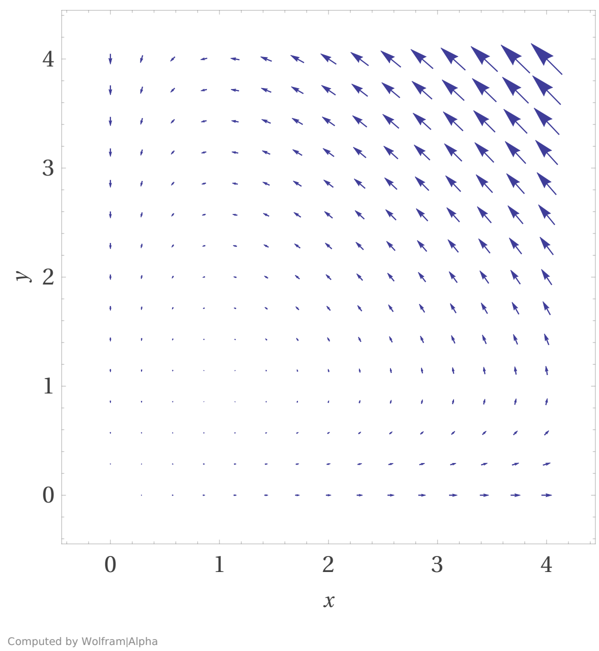

Figure 4.2.13: Sketch of vector field in (4.46)

to them as **integral curves**. They are also called **trajectories**, **orbits**, or **solution curves**.

To understand how the flow lines of a vector field $F(x, y)$ come about, refer to the vector equation in (4.45), which we restate here in slightly different form:

$$\sigma'(t) = F(\sigma(t)), \qquad \text{for } t \in J, \tag{4.47}$$

for some open interval J. We note that the vector–differential equation in (4.47) is another representation of the Lotka–Volterra system in (4.43). A difference between the two representations is that the equation in (4.47) makes the time dependence explicit, while time, t, is implicit in the system in (4.43). The vector equation in (4.47) also makes explicit the fact that the tangent vector $\sigma'(t)$ of the path $\sigma\colon J \to \mathbb{R}^2$ is prescribed by the vector field F. This observation will form the core of the definition of the flow of a vector field to be presented shortly.

We consider the general situation of a vector field $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$F(x, y) = f(x, y)\hat{\imath} + g(x, y)\hat{\jmath}, \qquad \text{for } (x, y) \in \mathbb{R}^2, \tag{4.48}$$

where f and g are real–valued functions of two variables. In the Lotka–Volterra example,

$$f(x, y) = x - xy, \qquad \text{for } (x, y) \in \mathbb{R}^2,$$

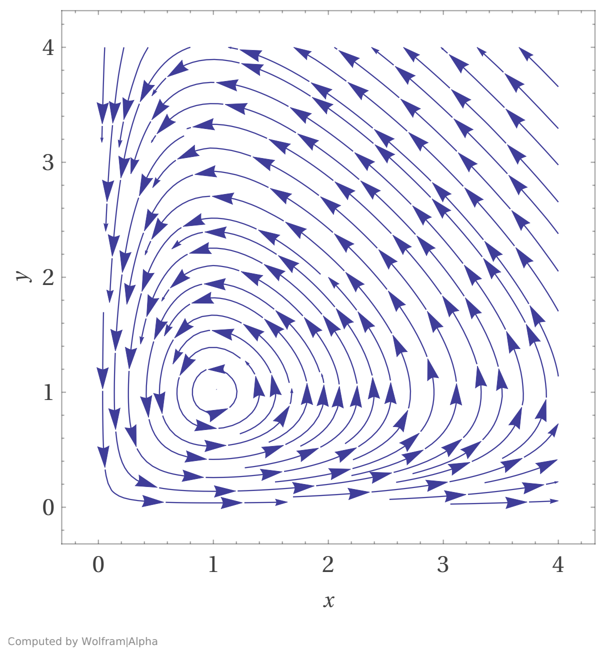

Figure 4.2.14: Flow of vector field in (4.46)

and

$$g(x,y) = -y + xy, \qquad \text{for } (x,y) \in \mathbb{R}^2.$$

The flow of the vector field F given in (4.48) consists of all curves generated by differentiable paths $\sigma\colon J \to \mathbb{R}^2$ satisfying the vector–differential equation in (4.47). This corresponds to the set of solution curves of the system of differential equations

$$\begin{cases} \dfrac{dx}{dt} & = & f(x,y); \\ \\ \dfrac{dy}{dt} & = & g(x,y), \end{cases} \tag{4.49}$$

An instance of the general system in (4.49) is the Lotka–Volterra system in (4.42). There are many problems in scientific applications that lead to two–dimensional systems like the one in (4.49). We will get to see several examples of these applications in these notes.

In the remainder of this section, we present a few simple examples of the system in (4.49) for which we will be able to compute the flow of the vector field in (4.48).

Example 4.2.1. Compute and sketch the flow of the vector field

$$F(x,y) = x\hat{\imath} - y\hat{\jmath}, \quad \text{for } (x,y) \in \mathbb{R}^2. \tag{4.50}$$

Solution: The flow of the vector field in (4.50) are the solutions curves of the system

$$\begin{cases} \dfrac{dx}{dt} & = & x; \\ \\ \dfrac{dy}{dt} & = & -y. \end{cases} \tag{4.51}$$

We note that each of the equations in the system in (4.51) can be solved independently. Indeed solutions of the first equation are of the form

$$x(t) = c_1 e^t, \quad \text{for } t \in \mathbb{R}, \tag{4.52}$$

and for some constant c_1. To see that the functions in (4.52) indeed solve the first equation in (4.52), we can differentiate with respect to t to get

$$x'(t) = c_1 e^t = x(t), \quad \text{for all } t \in \mathbb{R}.$$

We can also derive (4.52) by using **separation of variables**. We illustrate this technique by showing how to solve the second equation in (4.51); namely,

$$\frac{dy}{dt} = -y. \tag{4.53}$$

We fist rewrite the equation in (4.53) using the differential notation

$$\frac{1}{y}\,dy = -dt. \tag{4.54}$$

The equation in (4.54) displays the variables y and t separated on each side of the equal sign (hence the name of the technique). Next, integrate on both sides of (4.54) to get

$$\int \frac{1}{y}\,dy = \int -dt,$$

or

$$\int \frac{1}{y}\,dy = -\int dt. \tag{4.55}$$

Evaluating the indefinite integrals on both sides (4.55) then yields

$$\ln|y| = -t + k_1, \tag{4.56}$$

where k_1 is a constant of integration.

It remains to solve the equation in (4.56) for y in terms of t. To do this, we first apply the exponential function on both sides of (4.56) to get

$$e^{\ln|y|} = e^{-t+k_1},$$

or, using the properties of the exponential function,

$$|y| = e^{-t} \cdot e^{k_1},$$

or

$$|y(t)| = k_2 e^{-t}, \quad \text{for all } t \in \mathbb{R}, \tag{4.57}$$

where the constant e^{k_1} has been renamed k_2. Note that, since e^t is positive for all t, the expression in (4.57) can be rewritten as

$$|y(t)e^t| = k_2, \quad \text{for all } t \in \mathbb{R}. \tag{4.58}$$

Since we are looking for a differentiable function, y, that solves the second equation in (4.51), we may assume that y is continuous. Hence, the expression in the absolute value on the left–hand side of (4.58) is a continuous function of t. It then follows, by continuity, that $y(t)e^t$ must be constant for all values of t. Calling that constant c_2 we get that

$$y(t)e^t = c_2, \quad \text{for all } t \in \mathbb{R},$$

from which we get that

$$y(t) = c_2 e^{-t}, \quad \text{for all } t \in \mathbb{R}. \tag{4.59}$$

Combining the results (4.52) and (4.59) we see that the parametric equations for the solution curves of the system in (4.51) are given by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} c_1 e^t \\ c_2 e^{-t} \end{pmatrix}, \quad \text{for } t \in \mathbb{R}, \tag{4.60}$$

where c_1 and c_2 are arbitrary constants.

We will now proceed to sketch all types of solution curves determined by (4.60). These are determined by values of the constants c_1 and c_2. For instance, when $c_1 = c_2 = 0$, (4.60) yields the constant function

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \text{for } t \in \mathbb{R}. \tag{4.61}$$

The constant function in (4.61) is known as an **equilibrium solution** of the system in (4.51). It corresponds to origin, $(0, 0)$, of the Cartesian plane, and is sketched as a dot at the origin in Figure 4.2.15. Since $(0, 0)$ corresponds an equilibrium solution of the system in (4.51), $(0, 0)$ is called and **equilibrium point** of the system.

Next, if $c_1 \neq 0$ and $c_2 = 0$, then the solution curve

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} c_1 e^t \\ 0 \end{pmatrix}, \quad \text{for } t \in \mathbb{R},$$

will lie in the positive x–axis, if $c_1 > 0$, or in the negative x–axis if $c_1 < 0$. These two possible trajectories are shown in Figure 4.2.15. The figure also shows the

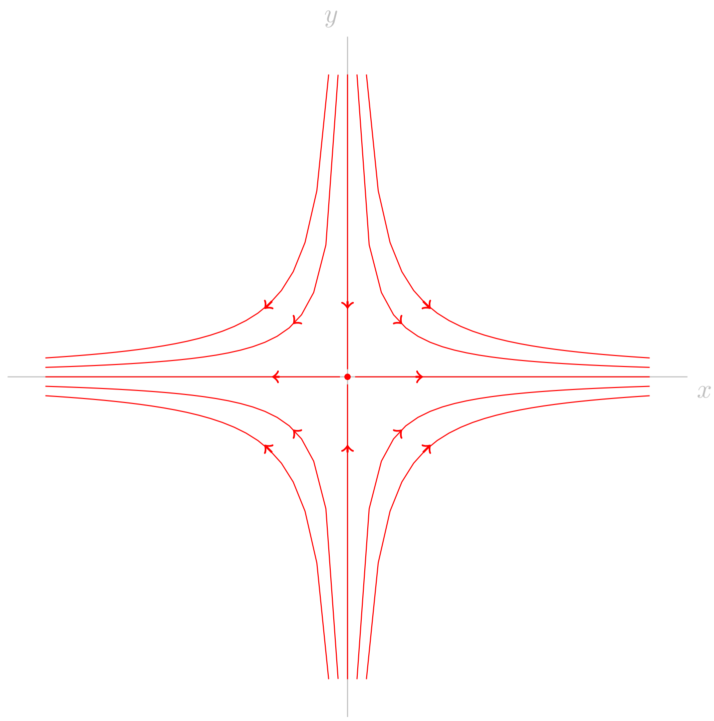

Figure 4.2.15: Sketch of Flow of Vector Field F in (4.50)

trajectories going away from the origin, as indicated by the arrows pointing away from the origin. The reason for this is that, as t increases, the exponential e^t increases.

Similarly, for the case $c_1 = 0$ and $c_2 \neq 0$, the solution curve

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 \\ c_2 e^{-t} \end{pmatrix}, \qquad \text{for } t \in \mathbb{R},$$

will lie in the positive y–axis, if $c_2 > 0$, or in the negative y–axis if $c_2 < 0$. In this case, the trajectories point towards the origin because the exponential e^{-t} decreases as t increases.

The other trajectories in the flow of the system in (4.51)correspond to the case in which $c_1 \cdot c_2 \neq 0$. To see what these trajectories look like, we combine the two parametric equations of the curves,

$$\left\{ \begin{array}{rcl} x & = & c_1 e^t; \\ y & = & c_2 e^{-t}, \end{array} \right. \tag{4.62}$$

into a single equation involving x and y by eliminating the parameter t. This can be done by multiplying the equations in (4.62) to get

$$xy = c_1 c_2,$$

or

$$xy = c, \tag{4.63}$$

where we have written c for the product c_1c_2. The graphs of the equations in (4.63) are hyperbolas for $c \neq 0$. A few of these hyperbolas are sketched in Figure 4.2.15. Observe that all the hyperbolas in the figure have directions associate with them indicated by the arrows. The directions can be obtained from the formula for the solution curves in (4.60), or from the differential equations in the system in (4.51). For instance, in the first quadrant ($x > 0$ and $y > 0$), we get from the differential equations in (4.51) that $x'(t) > 0$ and $y'(t) < 0$ for all t; so that, the values of x along the trajectories increase, while the y–values decrease. Thus, the arrows point down and to the right as shown in Figure 4.2.15. □

Example 4.2.2. Sketch the flow of the vector field

$$F(x,y) = x\hat{i} + 2y\hat{j}, \quad \text{for all } (x,y) \in \mathbb{R}^2. \tag{4.64}$$

Solution: We compute the solutions curves to the system of differential equations

$$\begin{cases} \dfrac{dx}{dt} &= x; \\ \\ \dfrac{dy}{dt} &= 2y. \end{cases} \tag{4.65}$$

We proceed as in Example 4.2.1 and compute solutions of each equation in (4.65) separately. We may use separation of variables, or simply write down the solutions.

Solutions of the first equation in (4.65) are of the form

$$x(t) = c_1 e^t, \quad \text{for } t \in \mathbb{R},$$

for some constant c_1; while those of the second equation are of the form

$$y(t) = c_2 y^{2t}, \quad \text{for } t \in \mathbb{R},$$

for some constant c_2.

Thus, the parametric equations for the solution curves of the system in (4.65) are given by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} c_1 e^t \\ c_2 e^{2t}) \end{pmatrix}, \quad \text{for } t \in \mathbb{R}, \tag{4.66}$$

where c_1 and c_2 are arbitrary constants.

Next, we sketch all types of solution curves for the system (4.65).

As in Example 4.2.1, if $c_1 = c_2 = 0$, the equation in (4.66) yields the origin; if $c_1 \neq 0$ and $c_2 = 0$, the trajectories are along the x–axis and point away from the origin; while, if $c_1 = 0$ and $c_2 \neq 0$, the trajectories lie on the y–axis and point away from the origin. These are all sketeched in Figure 4.2.16.

Trajectories when $c_1 \neq 0$ and $c_2 \neq 0$ can be obtained by eliminating the parameter t from the parametric equations

$$x = c_1 e^t \tag{4.67}$$

and

$$y = c_2 e^{2t}, \tag{4.68}$$

This can be done by squaring on both sides of the equation in (4.67) to get

$$x^2 = c_1^2 e^{2t}$$

and combining this with (4.68) to get

$$y = cx^2, \tag{4.69}$$

where we have written c for $c_2/{c_1}^2$. The graphs of the equations in (4.69) are parabolas. Thus, the trajectories for the system in (4.65) lie on parabolas given by (4.69) and they all emanate from the origin. A few of these are sketched in Figure 4.2.16. □

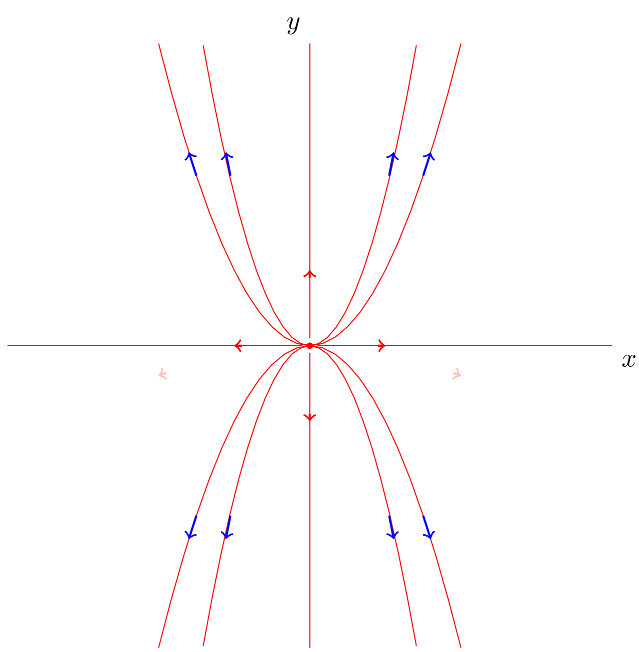

Figure 4.2.16: Sketch of Flow of F in (4.64)

Example 4.2.3. Sketch the flow of the vector field

$$F(x, y) = -y\hat{i} + x\hat{j}, \quad \text{for all } (x, y) \in \mathbb{R}^2. \tag{4.70}$$

Solution: We compute the solutions curves of the system of differential equations

$$\begin{cases} \dfrac{dx}{dt} &= & -y; \\ \\ \dfrac{dy}{dt} &= & x. \end{cases} \tag{4.71}$$

We note that, in this case, the differential equations in (4.71) cannot be solved separately. We have to combine them somehow. To do this, we think of y as an implicit function of x, and apply the Chain Rule.

Imagine that, in some portion of the curve y is an explicit function of x; so that,

$$y = y(x).$$

At the same time, x is a function of t; so that,

$$x = x(t);$$

consequently,

$$y = y(x(t)), \qquad \text{for } t \in J, \tag{4.72}$$

where J is some open interval.

The equation in (4.72) displays y as the composition of a function x and a function of t. Since we are assuming that x and y are differentiable functions, the Chain Rule applies, and we obtain from (4.72) that

$$\frac{dy}{dt} = y'(x)\frac{dx}{dt}, \tag{4.73}$$

or

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt}. \tag{4.74}$$

For points on the solution curves at which $\dfrac{dx}{dt} \neq 0$, we obtain from (4.74) that

$$\frac{dy}{dx} = \frac{\dfrac{dy}{dt}}{\dfrac{dx}{dt}}. \tag{4.75}$$

We can use the expression in (4.75) to combine the two differential equations in (4.71) into one differential equation. Indeed, substituting the right–hand sides of the equations in (4.71) into (4.75) yields

$$\frac{dy}{dx} = -\frac{x}{y}. \tag{4.76}$$

Note that the differential equation in (4.76) involves only the variables x and y. Note also that the variables in equation (4.76) can be separated to yield

$$y\ dy = -x\ dx. \tag{4.77}$$

Integrating on both sides of (4.77),

$$\int y\ dy = -\int x\ dx,$$

yields

$$\frac{1}{2}y^2 = -\frac{1}{2}x^2 + c_1, \tag{4.78}$$

for some constant of integration c_1.

Multiplying on both sides of the equations in (4.78) by 2, and setting $C = 2c_1$, yields

$$y^2 = -x^2 + C,$$

or

$$x^2 + y^2 = C, \tag{4.79}$$

for some constant C.

We see from equation (4.79) that the constant C is non–negative. We also see that solution curves of the system in (4.71) lie on concentric circles around the origin of radius $\sqrt{C}$, is $C > 0$; or at the origin if $C = 0$. In the latter case, if $C = 0$ in (4.79), then

$$x^2 + y^2 = 0,$$

which implies that $x = y = 0$; so that,

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0, \end{pmatrix}$$

for all $t \in \mathbb{R}$ in this case. This yields the equilibrium point $(0, 0)$ for the system in (4.71).

Figure 4.2.17 shows the equilibrium point of the system in (4.71) and few of the other orbits of the system. There is a direction associated with each of the trajectories in the flow of the vector field in (4.70) sketched in Figure 4.2.17. These directions are indicated by arrows on the concentric circles in the figure. To see how those directions are obtained, look at the information given by the system of differential equations in (4.71). In the first quadrant, $x > 0$ and $y > 0$; so that, according to the differential equations in (4.71), $x'(t) < 0$ and $y'(t) > 0$, for all $t \in \mathbb{R}$. Hence, x decreases in the first quadrant as t increases, while y increases. This is indicated by the two arrows sketched in the first quadrant in Figure 4.2.17. Thus, the concentric circles in the flow of the vector in (4.70) sketched in Figure 4.2.17 are oriented in the counterclockwise sense. □

Remark 4.2.4 (Time-Derivative Notation Convention)**.** In the equation in (4.73) we see the appearance of two derivatives of y. There is the derivative with respect to x, which we denote with the usual prime notation y'. There is also the derivative of y with respect to t, $\dfrac{dy}{dt}$. To distinguish between the two kinds of derivatives, we introduce the notation $\dot{y}$ to denote the derivative of y with respect to t; so that,

$$\dot{y} = \frac{dy}{dt}.$$

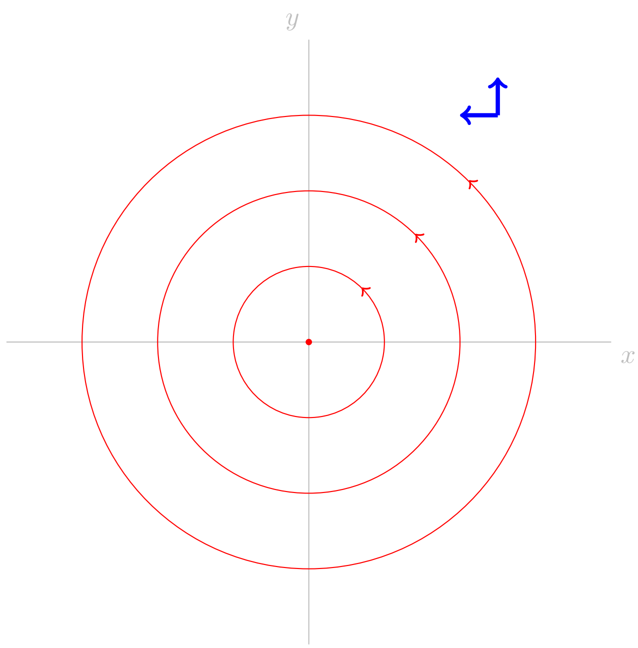

Figure 4.2.17: Sketch of Flow of Vector Field in (4.70)

With this notation we have

$$\dot{x} = \frac{dx}{dt}.$$

We shall reserve the prime notation to denote the derivative with respect to x; so that,

$$y' = \frac{dy}{dx} \quad \text{and} \quad x' = 1.$$

With this notation convention, the equation in (4.73) can be rewritten as

$$\dot{y} = y' \cdot \dot{x}.$$

By the same token, the equation in (4.75) can be rewritten as

$$\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}},$$

provided that $\dot{x} \neq 0$.

With the notation introduced in Remark 4.2.4, the flow of a vector field $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$F(x,y) = f(x,y)\hat{i} + g(x,y)\hat{j}, \quad \text{for } (x,y) \in \mathbb{R}^2,$$

are the solution curves of the system

$$\begin{cases} \dot{x} = f(x,y); \\ \dot{y} = g(x,y). \end{cases}$$

We can parametrize the orbits sketched in Figure 4.2.17, and given by the equation in (4.79), by the parametric equations

$$\begin{cases} x(t) = a\cos(t+\phi); \\ y(t) = a\sin(t+\phi), \end{cases} \quad \text{for } t \in \mathbb{R}, \tag{4.80}$$

where a is a non–negative constant, and $\phi \in \mathbb{R}$.

We can check that the coordinates, $(x(t), y(t))$, for $t \in \mathbb{R}$, of the points on the curve parametrized by the equations in (4.80), satisfy the equation in (4.79) with $C = a^2$. We can also check that x and y satisfy the system of differential equations in (4.71).

In the following example, we show another way to obtain the solutions in (4.80) of the system of differential equations in (4.71).

Example 4.2.5. Let F be as in Example 4.2.3 and let

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad \text{for } t \in J, \tag{4.81}$$

for some open interval J, be a parametrization of curve in the flow of the field F. Then, x and y are solutions of the system in (4.71); so that,

$$\begin{cases} \dot{x} = -y; \\ \dot{y} = x. \end{cases} \tag{4.82}$$

For each $t \in J$, define $r(t)$ to be the distance from from the point $(x(t), y(t))$ to the origin $O(0,0)$. Consequently, $r(t)$ is the Euclidean norm of the vector $\sigma(t)$ given in (4.81); so that,

$$r(t) = \|\sigma(t)\|, \quad \text{for } t \in J. \tag{4.83}$$

Next, define $\theta(t)$ to be the angle (in radians) that the vector $\sigma(t)$ in (4.81) makes with the positive x–axis (see sketch in Figure 4.2.18); so that,

$$\tan(\theta(t)) = \frac{y(t)}{x(t)}, \tag{4.84}$$

provided that $x(t) \neq 0$.

It also follows from the sketch in Figure 4.2.18 that

$$x(t) = r(t)\cos\theta(t), \quad \text{for all } t \in J, \tag{4.85}$$

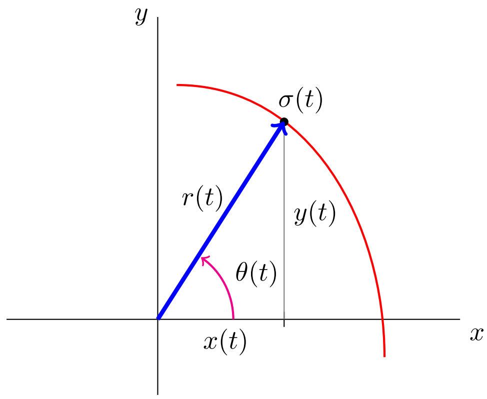

Figure 4.2.18: Path $\sigma(t)$

and

$$y(t) = r(t)\sin\theta(t), \quad \text{for all } t \in J. \tag{4.86}$$

We first consider the case in which

$$\sigma(t) \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \text{for all } t \in J. \tag{4.87}$$

It follows from (4.83) that

$$r(t) = \sqrt{(x(t))^2 + (y(t))^2}, \quad \text{for all } t \in J. \tag{4.88}$$

Hence, in view of (4.87), the Chain Rule implies that r as given in (4.88) is a differentiable function of t, since it is a composition of differentiable functions.

Similarly, writing (4.84) as

$$\theta(t) = \arctan\left(\frac{y(t)}{x(t)}\right), \quad \text{provided } x(t) \neq 0, \tag{4.89}$$

we see that θ is a differentiabld function of t, by virtue of the Chain Rule.

We get from (4.88) that

$$r^2 = x^2 + y^2, \tag{4.90}$$

and from (4.84), or (4.89), that

$$\tan\theta = \frac{y}{x}. \tag{4.91}$$

We would like to obtain expressions for the derivatives of r and t with respect to t in terms of $\dot{x}$ and $\dot{y}$.

Taking the derivative with respect to t on both sides of the expression in (4.90) and using the Chain Rule, we obtain

$$2r\frac{dr}{dt} = 2x\dot{x} + 2y\dot{y},$$

from which we get

$$\frac{dr}{dt} = \frac{1}{r}(x\dot{x} + y\dot{y}), \qquad \text{for } r > 0. \tag{4.92}$$

Similarly, taking the derivative with respect to t on both sides of (4.91) and applying the Chain Rule and the Quotient Rule,

$$\sec^2\theta \, \frac{d\theta}{dt} = \frac{x\dot{y} - y\dot{x}}{x^2}, \qquad \text{for } x \neq 0;$$

so that, using the trigonometric identity

$$1 + \tan^2\theta = \sec^2\theta$$

and (4.91),

$$\left(1 + \frac{y^2}{x^2}\right)\frac{d\theta}{dt} = \frac{x\dot{y} - y\dot{x}}{x^2}, \qquad \text{for } x \neq 0. \tag{4.93}$$

Next, multiply both sides of the equation in (4.93) by x^2, for $x \neq 0$, to get

$$\left(x^2 + y^2\right)\frac{d\theta}{dt} = x\dot{y} - y\dot{x};$$

so that, in view of (4.90),

$$\frac{d\theta}{dt} = \frac{1}{r^2}(x\dot{y} - y\dot{x}), \qquad \text{for } r > 0. \tag{4.94}$$

Combine the equations in (4.92) and (4.94) to obtain the change of variables equations

$$\left\{\begin{array}{rcl} \dot{r} & = & \dfrac{1}{r}(x\dot{x} + y\dot{y}); \\ \\ \dot{\theta} & = & \dfrac{1}{r^2}(x\dot{y} - y\dot{x}). \end{array}\right. \tag{4.95}$$

The equations in (4.95) will allow us to change the system in (4.82), that is written in terms of the Cartesian coordinates x and y, to a new system written in terms of the variables r and θ defined by the equations (4.90) and (4.91). Indeed, substituting the expressions for $\dot{x}$ and $\dot{y}$ given by the right–hand sides of the equations in (4.82) into the right–hand side of the first equation in (4.95) yields

$$\dot{r} = \frac{1}{r}(x(-y) + yx) = 0,$$

or

$$\dot{r} = 0. \tag{4.96}$$

Similarly, substituting the expressions for $\dot{x}$ and $\dot{y}$ given by the right–hand sides of the equations in (4.82) into the right–hand side of the second equation in (4.95) yields

$$\dot{\theta} = \frac{1}{r^2}(xx - y(-y)) = \frac{1}{r^2}(x^2 + y^2);$$

so that, in view of (4.90),

$$\dot{\theta} = 1. \tag{4.97}$$

Putting together the equations in (4.96) and (4.97) yields the system

$$\begin{cases} \dot{r} &= 0; \\ \\ \dot{\theta} &= 1. \end{cases} \tag{4.98}$$

The system in (4.98) is the system in (4.82) in terms of the variables r and θ. Note that the system in (4.98) can be integrated to yield

$$\begin{cases} r(t) &= a; \\ \\ \theta(t) &= t + \phi, \end{cases} \qquad \text{for } t \in \mathbb{R}, \tag{4.99}$$

where a and ϕ are constants of integration.

The expressions for r and θ in (4.99) can now be used, in conjunction with (4.85) and (4.86), to yield the solutions of the the system in (4.82):

$$\begin{cases} x(t) &= a\cos(t + \phi); \\ \\ y(t) &= a\sin(t + \phi), \end{cases} \qquad \text{for } t \in \mathbb{R},$$

which are the functions given in (4.80).

4.3 An Application: Conservation of Momentum

In this section we present an application of the concepts and ideas discussed in Chapter 3 and in the previous two section to study of motion of a particle in the plane.

Consider a particle of mass m (measured in kilograms, for instance) that moves in the Cartesian plane along a path $\sigma\colon J \to \mathbb{R}^2$ given by

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \qquad \text{for } t \in J, \tag{4.100}$$

where J denotes an interval of time.

The variable t in (4.100) is measured in units of time (for instance, seconds). The variables x ad y in (4.100) are measured in units of distance (say, meters).

The function σ in (4.100) locates the particle at time t as it moves along a curve in the plane traced by the path σ.

We can also write the path σ in (4.100) as

$$\sigma(t) = x(t)\hat{\imath} + y(t)\hat{\jmath}, \qquad \text{for } t \in J. \tag{4.101}$$

If we assume that the component functions of σ in (4.101) are differentiable functions of time, then the time derivative of σ,

$$\dot{\sigma}(t) = \dot{x}(t)\hat{i} + \dot{y}(t)\hat{j}, \quad \text{for } t \in J, \tag{4.102}$$

gives the velocity of the particle at time t.

Thus, at every time t, the velocity vector $\dot{\sigma}(t)$ in (4.102) is tangent to the path at the point $\sigma(t)$; see the sketch in Figure to see an illustration of this situation.

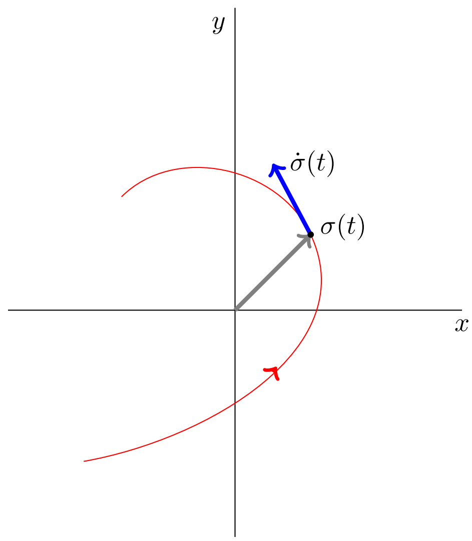

Figure 4.3.19: Path of a moving particle

If, in addition, we assume that the functions x and y are twice differentiable, we can compute the second derivative of the path σ,

$$\ddot{\sigma}(t) = \ddot{x}(t)\hat{i} + \ddot{y}(t)\hat{j}, \quad \text{for } t \in J. \tag{4.103}$$

The vector $\ddot{\sigma}$ defined in (4.103) is called the **acceleration** of the particle. The acceleration vector $\ddot{\sigma}$ gives the rate of change of the velocity vector and has units of length over squared time (say, meters per second squared).

Law of Conservation of Momentum. The **momentum** of a particle of mass m moving with velocity $\dot{\sigma}(t)$ along a path $\sigma(t)$ is given by

$$p(t) = m\dot{\sigma}(t), \quad \text{for } t \in J., \tag{4.104}$$

The law of conservation of momentum states that the rate of change of the momentum of a particle has to be accounted for by the vector sum of the forces acting on the particle. In symbols,

$$\dot{p} = F, \tag{4.105}$$

where the symbol F on the right–hand side of (4.105) denotes the vector sum of all the forces acting on the particle of mass m.

Using the definition of momentum in (4.104), and assuming that the mass of the particle is constant, the law of conservation of momentum in (4.105) reads

$$m\ddot{\sigma} = F. \tag{4.106}$$

The expression in (4.106) is also known as **Newton's Second Law of Motion**.

In this section, we present two applications of the principle of conservation of momentum in (4.106).

4.3.1 Trajectory of a Baseball

Assume that a baseball, after it leaves the batter box, travels in a plane in space, which we identify with the xy–plane. The variable x will denote the horizontal displacement of the ball, while y denotes the vertical displacement of the ball.

Assume that, at time $t = 0$ the location of the ball is at (x_o, y_o), the velocity of the ball is

$$\dot{\sigma}(0) = v_{ox}\hat{\imath} + v_{oy}\hat{\jmath}. \tag{4.107}$$

An any time $t \geqslant 0$, the location of the center of mass of the baseball is given by the vector–valued function

$$\sigma(t) = x(t)\hat{\imath} + y(t)\hat{\jmath}, \quad \text{for } t \geqslant 0. \tag{4.108}$$

See Figure 4.3.20 for an illustration of a possible path of the baseball.

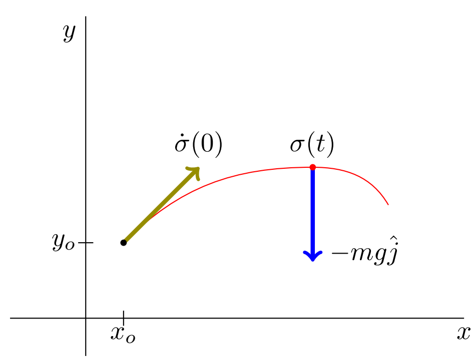

Figure 4.3.20: Possible path of a baseball

We would like to determine the values of the vector–valued function in (4.108) for all $t \geqslant 0$.

We assume that the components of the path σ defined in (4.108) are twice differentiable functions of time t. We can therefore apply the law of conservation of momentum in (4.106).

In this example we assume that the effects of wind speed and air drag are negligible in comparison with the gravitational force acting in the baseball. Consequently, we may assume that the vector sum of the forces acting on the baseball at any point $\sigma(t)$ in the path of the baseball is

$$F = -mg\hat{j},$$

as illustrated in Figure 4.3.20. Consequently, we obtain from (4.106) that

$$m(\ddot{x}(t)\hat{i} + \ddot{y}(t)\hat{j}) = -mg\hat{j}, \quad \text{for all } t \geqslant 0,$$

or

$$\ddot{x}(t)\hat{i} + \ddot{y}(t)\hat{j} = -g\hat{j}, \quad \text{for all } t \geqslant 0, \tag{4.109}$$

after canceling m from both sides of the equation.

Equating corresponding components in the vector equation in (4.109), we obtain the system of differential equations

$$\left\{ \begin{array}{rcl} \ddot{x} & = & 0; \\ \ddot{y} & = & -g. \end{array} \right. \tag{4.110}$$

We can integrate the equations in (4.110) to obtain

$$\left\{ \begin{array}{rcl} \dot{x} & = & c_1; \\ \dot{y} & = & -gt + c_2, \end{array} \right. \tag{4.111}$$

where c_1 and c_2 are constants of integration.

Next, substitute the initial condition in (4.107) into the equations in (4.111) to get

$$\left\{ \begin{array}{rcl} v_{ox} & = & c_1; \\ v_{oy} & = & c_2, \end{array} \right.$$

which yield the values of the integration constants in (4.111); so that

$$\left\{ \begin{array}{rcl} \dot{x} & = & v_{ox}; \\ \dot{y} & = & -gt + v_{oy}. \end{array} \right. \tag{4.112}$$

The equations in (4.112) can in turn be integrated to yield

$$\left\{ \begin{array}{rcl} x(t) & = & v_{ox}t + c_1; \\ y(t) & = & -\frac{1}{2}gt^2 + v_{oy}t + c_2, \end{array} \right. \tag{4.113}$$

where c_1 and c_2 are constants of integration.

Since we are given that the baseball is at (x_o, y_o) at time $t = 0$, it follows from (4.113) that $c_1 = x_o$ and $c_2 = y_o$. We therefore obtain from (4.113) that

$$\begin{cases} x(t) &= v_{ox}t + x_o; \\ y(t) &= -\dfrac{1}{2}gt^2 + v_{oy}t + y_o, \end{cases} \quad \text{for } t \geqslant 0. \tag{4.114}$$

The equations in (4.114) give a parametrization of the trajectory of the baseball from the time it leaves the batter box at the point (x_o, y_o) with initial velocity given in (4.107).

We can determine the shape of the curve parametrized by the equations in (4.114)by solving for the parameter t in the first equation,

$$t = \frac{x - x_o}{v_{ox}},$$

and substituting into the second equation to get

$$y = -\frac{1}{2}g\left(\frac{x - x_o}{v_{ox}}\right)^2 + v_{oy}\frac{x - x_o}{v_{ox}} + y_o, \tag{4.115}$$

which is the equation of a parabola in the xy–plane with vertex at the point with x–coordinate

$$x = x_o + \frac{v_{ox}v_{oy}}{g}$$

that opens downward. A sketch of that parabola is shown in Figure 4.3.21.

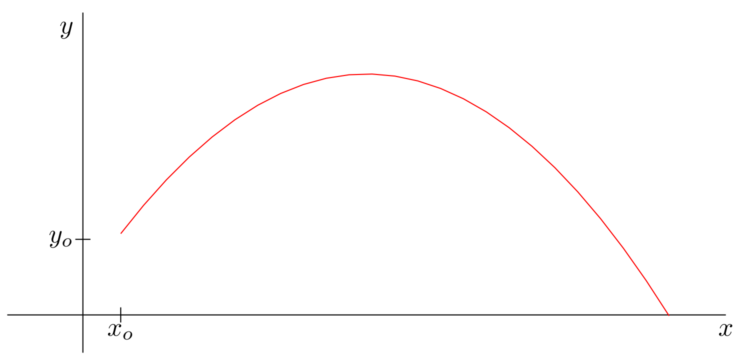

Figure 4.3.21: Sketch of trajectory of baseball

4.3.2 Mass–Spring System

An object of mass m is attached to a spring. The object is lying on a horizontal board that is assumed to be friction–less of the purpose of this example. See Figure 4.3.22 for a schematic of this set–up.

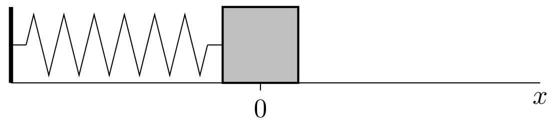

Figure 4.3.22: Mass–spring system at equilibrium

Let $x(t)$ denote the x–coordinate of the object. We assume that x is a twice–differentiable function of t. We assume that when the spring is not stretched or compressed, the center of mass of the object has coordinate $\bar{x} = 0$. We refer to this as the equilibrium position of the system. This is the situation depicted in Figure 4.3.22.

Assume that at time $t = 0$ the spring is stretched so that the center of mass of the object has x–coordinate x_o, where $x_o > 0$. Assume also that, at that time the object is at rest; so that, at time $t = 0$, we have the following initial conditions

$$\left\{\begin{array}{rcl} x(0) & = & x_o; \\ \\ \dot{x}(0) & = & 0. \end{array}\right. \tag{4.116}$$

This situation is depicted in Figure 4.3.23.

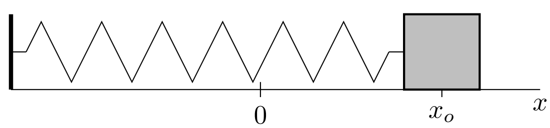

Figure 4.3.23: Mass–spring system at time $t = 0$

At time $t = 0$ the object is released. We would like to predict the location of the center of mass of the object at any time $t \geqslant 0$. To do this, we use the law of conservation of momentum in (4.106).

In this case,

$$\sigma(t) = x(t)\hat{\imath}, \quad \text{for } t \in \mathbb{R},$$

because the motion is occurring in the direction of the x–axis. Then,

$$\ddot{\sigma}(t) = \ddot{x}(t)\hat{\imath}, \quad \text{for } t \in \mathbb{R}. \tag{4.117}$$

To model the forces acting on the object, first we use Hooke's Law, which states that, for small stretching or compression, the magnitude of the force exerted by the spring on the object is proportional to the stretching (or compression) of the spring. In symbols, if F_s denotes the force the spring exerts on object,

$$\|F_s\| = k|x|,$$

where k is a positive constant of proportionality. Furthermore, the force F_s opposes the direction of the displacement. Consequently,

$$F_s = -kx\hat{\imath}. \tag{4.118}$$

Next, if we assume that other forces acting on the object (e.g., the force of friction between the object and the board on it the object slides) are negligible by comparison with the force of the spring, we can conclude that the total vector sum on the forces acting on the object is given by F_s in (4.118); so that,

$$F = -kx\hat{\imath}. \tag{4.119}$$

Hence, combining (4.117), (4.119), and the Law of Conservation of Momentum in (4.106),

$$m\ddot{x}\hat{\imath} = -kx\hat{\imath},$$

from which we get that

$$m\ddot{x} = -kx. \tag{4.120}$$

Like the equations in (4.110) that we obtained in Example 4.3.1, the equation in (4.120) is a second–order differential equations because the highest order derivative of the unknown function in the equation is the second derivative with respect to time t. However, unlike the differential equations in (4.110), which we were able to integrate to obtain their solutions, we are not able to integrate the equation in (4.120) directly. In the remainder of this example, we will show how to obtain solutions of the equation in (4.120).

Introduce a new parameter ω by means of the equation

$$\omega^2 = \frac{k}{m}. \tag{4.121}$$

Then, the equation in (4.120) can be rewritten as

$$\ddot{x} = -\omega^2 x. \tag{4.122}$$

We will show how to find solutions of the second–order differential equation in (4.122) by turning it into a system of first–order differential equations like the ones we studied in Section 4.2.2.

Suppose that $x\colon J \to \mathbb{R}$ is a solution of the differential equation in (4.122), for some open interval J. Define y to be the time derivative of x; so that,

$$y(t) = \dot{x}(t), \qquad \text{for } t \in J. \tag{4.123}$$

It then follows that y is a differentiable function of t with

$$\dot{y}(t) = \ddot{x}(t), \qquad \text{for } t \in J.$$

Hence, in view of (4.122) the function y solves the differential equation

$$\dot{y} = -\omega^2 x. \tag{4.124}$$

Thus, in view of (4.123) and (4.124), we see that the path $\sigma\colon J \to \mathbb{R}^2$ given by

$$\sigma(t) = x(t)\hat{\imath} + y(t)\hat{\jmath}, \qquad \text{for } t \in J,$$

solves the system of differential equations

$$\begin{cases} \dot{x} &= \quad y; \\ \\ \dot{y} &= -\omega^2 x. \end{cases} \tag{4.125}$$

The system of differential equations in (4.125) is reminiscent of the system in (4.71) that we solved in Example 4.2.3. We may therefore proceed to determine the solution curves, or trajectories, of the system in (4.125) as we did in Example 4.2.3 for the system in (4.71).

Proceeding as in Example 4.2.3, we can determine the trajectories of the system in (4.125) by first computing

$$\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}} = \frac{-\omega^2 x}{y};$$

so that,

$$\frac{dy}{dx} = -\frac{\omega^2 x}{y}. \tag{4.126}$$

We can solve the differential equation in (4.126) by separating variables:

$$y\ dy = -\omega^2 x\ dx. \tag{4.127}$$

Integrating on both sides of (4.127) yields

$$\frac{1}{2}y^2 = -\frac{1}{2}\omega^2 x^2 + c_1, \tag{4.128}$$

for some constant of integration c_1.

Multiply both sides of the equation in (4.128) by 2 and rearrange terms to get

$$\omega^2 x^2 + y^2 = c_2, \tag{4.129}$$

where we have set $c_2 = 2c_1$.

Next, divide both sides of the equation in (4.129) by ω^2 to get

$$x^2 + \frac{y^2}{\omega^2} = c_3, \tag{4.130}$$

where we have set $c_3 = \dfrac{c_2}{\omega^2}$.

We note that the constant c_3 in (4.130) is non–negative. To emphasize this fact, we set

$$c_3 = a^2,$$

for some non–negative real number a. We can then rewrite the equation in (4.130) as

$$x^2 + \frac{y^2}{\omega^2} = a^2. \tag{4.131}$$

If $a = 0$ in (4.131) we get that

$$x = y = 0;$$

so that, the origin is a possible trajectory of the system in (4.125). This is the equilibrium solution, $(0, 0)$, sketched in Figure 4.3.24.

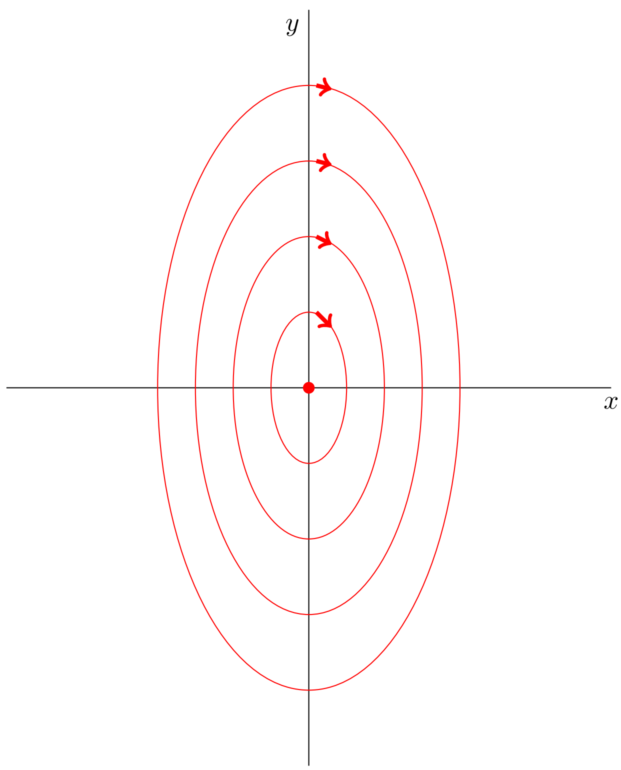

Figure 4.3.24: Sketch of phase portrait of the system in (4.125)

For the case $a > 0$, divide both sides of the equation in (4.131) by a^2 to get

$$\frac{x^2}{a^2} + \frac{y^2}{\omega^2 a^2} = 1. \tag{4.132}$$

The graph of the equation in (4.132) is a circle of radius a centered at the origin in the case in which $\omega = 1$. If $\omega \neq 1$, the graph is an ellipse with vertices $(-a, 0)$ and $(a, 0)$ on the x–axis and vertices $(0, -\omega a)$ and $(0, \omega a)$ on the y–axis. The sketch in Figure 4.3.24 shows a few of those ellipses for varies values of $a > 0$ in the case $\omega > 1$. The sketch also shows the direction along the orbits dictated by the system of differential equations in (4.125).

We next use a parametrization of the orbits sketched in Figure 4.3.24 to determine solutions of the system of differential equations in (4.125).

Rewrite the equation of the trajectories of the system in (4.125) given in (4.132) as

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{\omega a}\right)^2 = 1. \tag{4.133}$$

The equation in (4.133) suggests that we define

$$\frac{x}{a} = \sin(\theta(t)), \qquad \text{for } t \in \mathbb{R},$$

and

$$\frac{y}{\omega a} = \cos(\theta(t)), \qquad \text{for } t \in \mathbb{R},$$

where θ is a differentiable function of t; so that,

$$\begin{cases} x(t) = a\sin(\theta(t)); \\ y(t) = a\omega\cos(\theta(t)), \end{cases} \qquad \text{for } t \in \mathbb{R}, \tag{4.134}$$

are the parametric equations of the solution curves of the system in (4.125), some of which are sketched in Figure 4.3.24.

We have encountered the function θ in the parametric equations in (4.134) previously in these notes. In Example 4.2.5 we defined θ to the the angle that the path $\sigma(t)$ makes with the positive x–axis. We also saw in that example that θ is a differentiable function as a consequence of the Chain Rule.

Using the Chain Rule, we compute the derivatives of the functions in (4.134) to get

$$\begin{cases} \dot{x} = a\dot{\theta}\cos(\theta); \\ \dot{y} = -a\dot{\theta}\omega\sin(\theta), \end{cases}$$

which, in view of the definitions of x and y in (4.134) yields

$$\begin{cases} \dot{x} = \dfrac{\dot{\theta}}{\omega} y; \\ \dot{y} = -\dot{\theta}\omega x. \end{cases} \tag{4.135}$$

Comparing the equations in (4.135) with those in (4.125) we see that

$$\dot{\theta} = \omega. \tag{4.136}$$

The differential equation in (4.136) can be integrated to yield

$$\theta(t) = \omega t + \phi, \qquad \text{for all } t \in \mathbb{R}, \tag{4.137}$$

since ω is constant, where ϕ is a constant of integration.

Substituting the values of $\theta(t)$ in (4.137) into the equations in (4.134) yields

$$\left\{\begin{array}{rcl} x(t) & = & a\sin(\omega t+\phi); \\ & & \\ y(t) & = & a\omega\cos(\omega t+\phi), \end{array}\right. \quad \text{for } t\in\mathbb{R}, \tag{4.138}$$

The parametric equations in (4.138) are the solutions of the system of differential equations in (4.125). We note that a and ϕ in (4.138) are two constants of integration. The symbol ω is the parameter that appears in the second equation in (4.125).

The first equation in (4.138)

$$x(t) = a\sin(\omega t+\phi), \quad \text{for } t\in\mathbb{R}, \tag{4.139}$$

yields the general solution of the second order equation in (4.122), which we set out to solve at the outset of this example.

With the general form of the solution of the second–order equation in (4.122) given in (4.139), we can solve the problem that we stated at the start of this example: Determine the location, $x(t)$, of the object in the mass–spring system at any time $t\geqslant 0$, given that the object is at location $x(0)=x_o>0$ at time $t=0$, and is released from rest at that time. This problem can be stated as the following initial value problem (IVP):

$$\left\{\begin{array}{l} \ddot{x} = -\omega^2 x; \\ x(0) = x_o; \\ \dot{x}(0) = 0. \end{array}\right. \tag{4.140}$$

We have already seen that the general solution of the second–order differential equation in (4.140) is given in (4.139), where a and ϕ are constants of integration. We next determine values of a and ϕ so that the initial conditions are satisfied.

From (4.139) we obtain that

$$\dot{x}(t) = a\omega\cos(\omega t+\phi), \quad \text{for } t\in\mathbb{R}, \tag{4.141}$$

where we have used the Chain Rule.

Substitute 0 for t in (4.139) and (4.141), and use the initial conditions in (4.140) to get

$$\left\{\begin{array}{rcl} a\sin(\phi) & = & x_o; \\ a\omega\cos(\phi) & = & 0. \end{array}\right. \tag{4.142}$$

We first note that a cannot be 0; otherwise, $x(t)=0$, for all t, according to (4.139), and this is incompatible with the initial condition $x_o>0$. Hence, since we are also assuming that $\omega>0$, we get from the second equation in (4.142) that

$$\cos(\phi)=0;$$

thus, we can take

$$\phi = \frac{\pi}{2}. \tag{4.143}$$

Substituting the value of ϕ in (4.143) into the first equation in (4.142) then yields

$$a = x_o. \tag{4.144}$$

Substitute the values for a and ϕ in (4.144) and (4.143), respectively, into the formula for $x(t)$ in (4.139) to get

$$x(t) = x_o \sin\left(\omega t + \frac{\pi}{2}\right), \quad \text{for } t \in \mathbb{R},$$

which, using the trigonometric identity

$$\sin(A + B) = \sin A \cos B + \cos A \sin B,$$

can be rewritten as

$$x(t) = x_o \cos(\omega t), \quad \text{for } t \in \mathbb{R}. \tag{4.145}$$

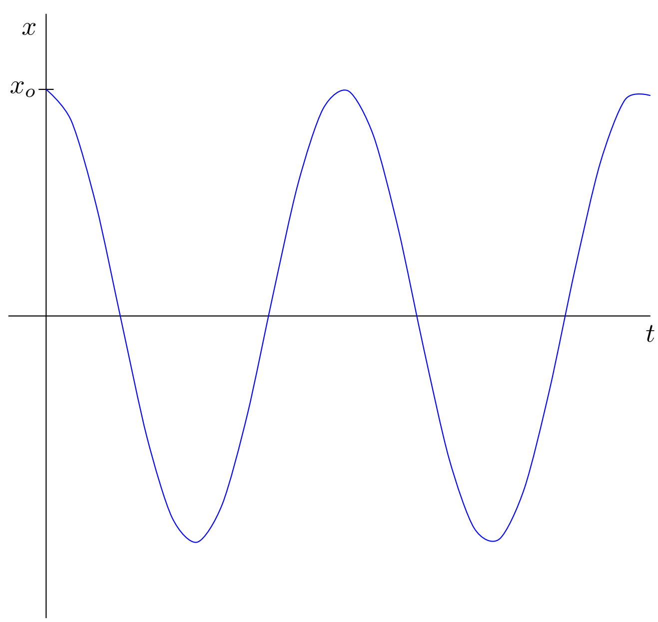

Figure 4.3.25: Sketch of x as a function of t

Figure 4.3.25 shows a sketch of the graph of x given in (4.145) as a function of t. We see from the sketch that the model in (4.122) predicts oscillations of the object in the mass–spring system about the equilibrium position $\overline{x} = 0$. The period, T, of the oscillations is given by the expression

$$\omega T = 2\pi,$$

from which we get that

$$T = \frac{2\pi}{\omega}. \tag{4.146}$$

The frequency, f, of the oscillations of the mass–spring system is given by the reciprocal of the period, T, in (4.146):

$$f = \frac{\omega}{2\pi}. \tag{4.147}$$

Thus, the parameter ω in the equation (4.122) is related to the frequency of oscillations of the mass–spring system. Indeed, ω, defined by the equation in (4.121) as

$$\omega = \sqrt{\frac{k}{m}}, \tag{4.148}$$

is called the angular frequency and is measured in units of radians per second.

Combining (4.147) and (4.148) we get the following formula for the frequency,

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}},$$

in terms of the original parameters of the mass–spring system in (4.120).

Chapter 5

Linear Vector Fields in Two Dimensions

5.1 Definition of a Linear Vector Fields

A function $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ is said to be a linear vector field if it is given by an expression of the form

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cd + dy \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \tag{5.1}$$

where a, b, c and d are real numbers (constants).

We note that, starting in this chapter, we adopt the convention of using columns to denote vectors in in $\mathbb{R}^2$. We are also dropping the arrow above the symbol to denote names of vectors. The context will make it clear when we are talking about vectors and not numbers. We shall also refer to numbers as "scalars" to distinguish them from vectors.

Examples of linear vector fields are

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ -y \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2,$$

and

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ x - y \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

The vector field associated with the Lotka–Volterra system, namley,

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha x - \beta xy \\ \beta xy - \gamma y \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$$

where α, β and γ are real constants, is not linear (Why?).

Another example of a two–dimensional vector field that is not linear is provided by

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

5.2 Matrices and Matrix Algebra

The general form of linear field given in (5.1) can be written in a more compact way by using matrix notation. In this section we discuss definitions of matrices and matrix products and present some of their properties.

A matrix is an array of numbers organized in rows of and columns. An $m \times n$ matrix consists m rows and n columns. In this course we will deal only with the cases in which m and n are 1 or 2.

A 2×1 matrix is a column vector

$$\begin{pmatrix} x \\ y \end{pmatrix}, \tag{5.2}$$

2 rows and 1 column. We will use column vectors of the form in (5.2) to represent vectors in $\mathbb{R}^2$.

A 1×2 matrix is a row vector of the form

$$\begin{pmatrix} a & b \end{pmatrix}. \tag{5.3}$$

Denote the column vector in (5.2) by v and the row vector in (5.3) by R.

Definition 5.2.1 (Row–Column Product)**.** The row-column product, Rv, is the scalar obtained by

$$R\mathrm{v} = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax + by. \tag{5.4}$$

We note two things about the product in (5.4):

(i) it is the first entry in the definition of the linear vector field in (5.1);

(ii) It is the dot product of the vectors $w = a\hat{\imath} + b\hat{\jmath}$ and $v = x\hat{\imath} + y\hat{\jmath}$.

Writing

$$\mathrm{w} = \begin{pmatrix} a \\ b \end{pmatrix}, \tag{5.5}$$

the **transpose** of the column vector w, denoted by w^T, is the column vector obtained from w in (5.5) as follows

$$\mathrm{w}^T = \begin{pmatrix} a \\ b \end{pmatrix}^T = \begin{pmatrix} a & b \end{pmatrix}$$

Thus, given two column vectors

$$\mathrm{w} = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{and} \quad \mathrm{v} = \begin{pmatrix} x \\ y \end{pmatrix},$$

we can from the row–column product of w^T and v to get

$$\mathrm{w}^T\mathrm{v} = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax + by,$$

which is the dot product of the vectors w and v in $\mathbb{R}^2$. We then have that

$$\mathrm{w}\cdot\mathrm{v} = \mathrm{w}^T\mathrm{v}.$$

In general, a 2×2 matrix is an array, A, of the form

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{5.6}$$

where a, b, c and d are real numbers.

The matrix A in (5.6) is made up of two rows

$$R_1 = \begin{pmatrix} a & b \end{pmatrix} \quad \text{and} \quad R_2 = \begin{pmatrix} c & d \end{pmatrix},$$

or two columns

$$\mathrm{v}_1 = \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad \mathrm{v}_2 = \begin{pmatrix} b \\ d \end{pmatrix}.$$

We can therefore write

$$A = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix}, \tag{5.7}$$

or

$$A = \begin{bmatrix} \mathrm{v}_1 & \mathrm{v}_2 \end{bmatrix}.$$

We can use the row–column product defined in Definition 5.2.1 to define the product of a matrix and a vector.

Definition 5.2.2 (Product of Matrix and a Vector)**.** Given a 2×2 matrix A and a (column) vector v, the product, Av, is the column vector obtained as follows: Write the matrix A in terms of its rows as in (5.7); then, compute

$$A\mathrm{v} = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix}\mathrm{v} = \begin{pmatrix} R_1\mathrm{v} \\ R_2\mathrm{v} \end{pmatrix} \tag{5.8}$$

Thus, if A is as given in (5.6) and v is the column vector given in (5.2), then

$$A\mathrm{v} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cd+dy \end{pmatrix}. \tag{5.9}$$

Comparing (5.1) and (5.9) we see that the linear vector field in (5.1) can be written as multiplication by the 2×2 matrix A given in (5.6). We then have that

$$F\begin{pmatrix} x \\ y \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2. \tag{5.10}$$

According to (5.10), every linear vector field $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ has a 2×2 matrix, A, associated with it.

Example 5.2.3. The matrix associated with the linear field $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -4x \end{pmatrix}, \quad \text{for } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2,$$

is

$$A = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix}.$$

Definition 5.2.4 (Matrix Multiplication)**.** Given 2×2 matrices A and B, write the matrix A in terms of its rows as in (5.7), and write B in terms of its columns,

$$B = \begin{bmatrix} \mathrm{v}_1 & \mathrm{v}_2 \end{bmatrix}.$$

The matrix product AB is the 2×2 matrix obtained as follows

$$AB = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \begin{bmatrix} \mathrm{v}_1 & \mathrm{v}_2 \end{bmatrix} = \begin{pmatrix} R_1\mathrm{v}_1 & R_1\mathrm{v}_2 \\ R_2\mathrm{v}_1 & R_2\mathrm{v}_2 \end{pmatrix}. \tag{5.11}$$

Example 5.2.5. Let A denote the 2×2 matrix

$$A = \begin{pmatrix} 3 & -1 \\ 5 & -3 \end{pmatrix}, \tag{5.12}$$

and B the 2×2 matrix

$$B = \begin{pmatrix} 1 & 1 \\ 5 & 1 \end{pmatrix}. \tag{5.13}$$

Then, using the formula in (5.11), we compute

$$AB = \begin{pmatrix} 3 & -1 \\ 5 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ -10 & 2 \end{pmatrix}. \tag{5.14}$$

Example 5.2.6. Let A and B be as given in (5.12) and (5.13), respectively. We can also form the matrix product BA:

$$BA = \begin{pmatrix} 1 & 1 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 5 & -3 \end{pmatrix} = \begin{pmatrix} 8 & -4 \\ 20 & -8 \end{pmatrix}. \tag{5.15}$$

Comparing the results in (5.14) and (5.15) we see that

$$AB \neq BA; \tag{5.16}$$

thus, matrix multiplication is not commutative.

The statement in (5.16) is justified by the notion of matrix equality.

Definition 5.2.7 (Matrix Equality)**.** Two matrices are said to be equal if and only if corresponding entries in the matrix are the same.

Example 5.2.8. Let A and B denote the 2×2 matrices given by

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

and

$$B = \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix}.$$

Then,

$$AB = \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{5.17}$$

Definition 5.2.9 (Zero Matrix)**.** The 2×2 matrix whose entries all all 0 is called the zero 2×2 matrix. We will denote it by the symbol O; so that,

$$O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Example 5.2.10. Let A and B denote the 2×2 matrices given by

$$A = \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix}$$

and

$$B = \begin{pmatrix} -3/2 & 1 \\ -1/2 & 0 \end{pmatrix}.$$

We compute

$$AB = \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} -3/2 & 1 \\ -1/2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We also compute

$$BA = \begin{pmatrix} -3/2 & 1 \\ -1/2 & 0 \end{pmatrix} \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Definition 5.2.11 (Identity Matrix)**.** The 2×2 matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is called the 2×2 identity matrix. We will denote it by the symbol I; so that,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For the matrices A and B in Example 5.2.10, we saw that

$$AB = BA = I.$$

When this happens we say that the matrix A is invertible.

Definition 5.2.12 (Invertible Matrix)**.** A 2×2 matrix A is said to be invertible if there exists a 2×2 matrix B such that

$$AB = BA = I. \tag{5.18}$$

If (5.18) holds true, we also say that B is the inverse of A and denote it by A^{-1}; so that,

$$AA^{-1} = A^{-1}A = I.$$

Definition 5.2.13 (Matrix Addition)**.** Given two matrices, A and B, of the same size, the matrix $A + B$ is obtained by adding corresponding entries. We have three cases two consider.

(i) Adding two 2×2 matrices.

Let A and B be 2×2 matrices given by

$$A = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix},$$

respectively. The sum $A + B$ is defined by

$$A + B = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{pmatrix}.$$

(ii) Adding two column–vectors in $\mathbb{R}^2$.

Let v_1 and v_2 be vectors in $\mathbb{R}^2$ given by

$$\mathrm{v}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad \text{and} \quad \mathrm{v}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix},$$

respectively. The vector sum $\mathrm{v}_1 + \mathrm{v}_2$ is defined by

$$\mathrm{v}_1 + \mathrm{v}_2 = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}.$$

(iii) Adding two row–vectors in $\mathbb{R}^2$.

Let R_1 and R_2 be row–vectors in $\mathbb{R}^2$ given by

$$R_1 = \begin{pmatrix} a_1 & b_1 \end{pmatrix} \quad \text{and} \quad R_2 = \begin{pmatrix} a_2 & b_2 \end{pmatrix},$$

respectively. The sum $R_1 + R_2$ is the row–vector defined by

$$R_1 + R_2 = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \end{pmatrix}.$$

Remark 5.2.14. The addition of a column–vector and a row vector is not defined, neither is the addition of a 2×2 matrix and a vector.

Example 5.2.15. Let A and B denote 2×2 matrices given by

$$A = \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 2 \\ -1 & 3 \end{pmatrix},$$

respectively. Then

$$A + B = O,$$

where O is the 2×2 zero matrix. We say that B is the additive inverse of A and write

$$B = -A.$$

Definition 5.2.16 (Scalar Multiplication)**.** Given a matrix A and a real number t, the matrix tA is obtained by multiplying every entry in the matrix by t. We have three cases two consider.

(i) Scalar multiple of a 2×2 matrix.

Let A be the 2×2 matrix given by

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The matrix tA is defined by

$$tA = \begin{pmatrix} ta & tb \\ tc & td \end{pmatrix}.$$

(ii) Scalar multiple of a column–vector in $\mathbb{R}^2$.

Let v be a vector in $\mathbb{R}^2$ given by

$$\mathrm{v} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

The vector $t\mathrm{v}$ is defined by

$$t\mathrm{v} = \begin{pmatrix} tx \\ ty \end{pmatrix}.$$

(iii) Scalar multiple of a row–vector in $\mathbb{R}^2$.

Let R be a row–vector in $\mathbb{R}^2$ given by

$$R = \begin{pmatrix} a & b \end{pmatrix}.$$

The row–vector tR is defined by

$$tR = \begin{pmatrix} ta & tb \end{pmatrix}.$$

Example 5.2.17. Let A denote the 2×2 matrix given by

$$A = \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix}.$$

Compute $A^2 + 3A + 2I$, where I is the 2×2 identity matrix.

Solution: First, we compute

$$A^2 = \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix} = \begin{pmatrix} -2 & 6 \\ -3 & 7 \end{pmatrix}.$$

Then

$$\begin{aligned} A^2 + 3A + 2I &= \begin{pmatrix} -2 & 6 \\ -3 & 7 \end{pmatrix} + 3 \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix} + 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 6 \\ -3 & 7 \end{pmatrix} + \begin{pmatrix} 0 & -6 \\ 3 & -9 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \end{aligned}$$

so that

$$A^2 + 3A + 2I = O,$$

the 2×2 zero matrix. □

5.2.1 Properties of Matrix Products

In this section we list a few of the properties of the matrix and vector operations that have been defined so far. These operations will be used in various matrix calculations in these notes.

Proposition 5.2.18 (Distributive Properties)**.**

(i) For 2×2 matrices A, B and C,

$$A(B + C) = AB + AC.$$

(ii) For 2×2 matrices A, B and C,

$$(A + B)C = AC + BC.$$

(iii) For a 2×2 matrix A and column–vectors v_1 and v_2 in $\mathbb{R}^2$,

$$A(\mathrm{v}_1 + \mathrm{v}_2) = A\mathrm{v}_1 + A\mathrm{v}_2.$$

(iv) For 2×2 matrices A and B, and a column–vector v in $\mathbb{R}^2$,

$$(A + B)\mathrm{v} = A\mathrm{v} + B\mathrm{v}.$$

(v) For a scalar t and column–vectors v_1 and v_2 in $\mathbb{R}^2$,

$$t(\mathrm{v}_1 + \mathrm{v}_2) = t\mathrm{v}_1 + t\mathrm{v}_2.$$

(vi) For a scalars t and r, and a column–vector v in $\mathbb{R}^2$,

$$(t + r)\mathrm{v} = t\mathrm{v} + r\mathrm{v}.$$

Proposition 5.2.19 (Associative Properties)**.**

(i) For 2×2 matrices A, B and C,

$$A(BC) = (AB)C.$$

(ii) For a 2×2 matrix A, a column–vector v, and a scalar t,

$$A(t\mathrm{v}) = tA\mathrm{v}$$

Remark 5.2.20. The properties in Proposition 5.2.18 and 5.2.19 can be derived by using the definition of the operations in Definitions 5.2.4, 5.2.13 and 5.2.16.

Example 5.2.21. In Example 5.2.17 we saw that the matrix

$$A = \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix}. \tag{5.19}$$

satisfies the equation

$$A^2 + 3A + 2I = O. \tag{5.20}$$

We can rewrite (5.20) as

$$A^2 + 3A = -2I. \tag{5.21}$$

We can then use the distributive properties to rewrite the left–hand side in (5.21) to get

$$A(A + 3I) = -2I, \tag{5.22}$$

where we have also used the fact that $A = AI$.

Next, multiply on both sides of (5.22) by the scalar $-\frac{1}{2}$, and use the distributive and associative properties, to get

$$A\left[-\frac{1}{2}(A + 3I)\right] = I. \tag{5.23}$$

It follows from (5.23) and Definition 5.2.12 that the matrix A given in (5.19) is invertible, and its inverse is given by

$$A^{-1} = -\frac{1}{2}(A + 3I),$$

or

$$\begin{aligned} A^{-1} &= -\frac{1}{2}A - \frac{3}{2}I \\ &= -\frac{1}{2}\begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix} - \frac{3}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1/2 & 3/2 \end{pmatrix} + \begin{pmatrix} -3/2 & 0 \\ 0 & -3/2 \end{pmatrix}; \end{aligned}$$

so that

$$A^{-1} = \begin{pmatrix} -3/2 & 1 \\ -1/2 & 0 \end{pmatrix}. \tag{5.24}$$

In the next section we will see another way to obtain the result in (5.24).

5.2.2 Invertible Matrices

In this section we will see that the 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{5.25}$$

has an inverse provided that $ad - bc \neq 0$.

The expression $ad - bc$ is called the determinant of the matrix A in (5.25) and will be denoted by $\det(A)$. We then have that

$$\det(A) = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc. \tag{5.26}$$

Assume that $\det(A) \neq 0$ and look for a 2×2 matrix B, given by

$$B = \begin{pmatrix} x & z \\ y & w \end{pmatrix}, \tag{5.27}$$

where x, y, z and w are unknowns to be determine shortly, such that

$$AB = I, \tag{5.28}$$

the 2×2 identity matrix, or

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & z \\ y & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

or

$$\begin{pmatrix} ax + by & az + bw \\ cx + dy & cz + dw \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5.29}$$

It follows from (5.29) that (5.28) is equivalent to the system of equations

$$\begin{cases} ax + by = 1 \\ cx + dy = 0 \\ az + bw = 0 \\ cz + dw = 1 \end{cases} \tag{5.30}$$

We first consider the case in which

$$a \neq 0 \quad \text{and} \quad c \neq 0. \tag{5.31}$$

In this case, we can solve the second equation in (5.30) for x to get

$$x = -\frac{d}{c}y, \tag{5.32}$$

and substitute into the first equation in (5.30) to get

$$-\frac{ad}{c}y + by = 1,$$

which can be solved for y to yield

$$y = -\frac{c}{ad - bc},$$

or

$$y = -\frac{c}{\det(A)}. \tag{5.33}$$

Combining (5.33) and (5.32) we get that

$$x = \frac{d}{\det(A)}. \tag{5.34}$$

Similarly, if (5.31) is true, then we can solve the third equation in (5.30) for z to get

$$z = -\frac{b}{a}w. \tag{5.35}$$

Substituting (5.35) into the last equation in (5.30) then yields

$$-\frac{bc}{a}w + dw = 1,$$

which can be solved for w to yield

$$w = \frac{a}{\det(A)}. \tag{5.36}$$

Combining (5.35) and (5.36) then yields

$$z = -\frac{b}{\det(A)}. \tag{5.37}$$

It follows from (5.27), (5.28), (5.33), (5.34), (5.37) and (5.36) that the matrix

$$B = \frac{1}{\det(A)}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \tag{5.38}$$

for the case in which $\det(A) \neq 0$ and (5.31) holds true, is such that

$$AB = I.$$

It can also be verified that $BA = I$ (see Problem 1 in Assignment #13). Thus the matrix in (5.38) is the inverse of A for the case in which $\det(A) \neq 0$ and (5.31) holds true.

Next, assume that $\det(A) \neq 0$ and

$$a = 0 \quad \text{or} \quad c = 0. \tag{5.39}$$

Suppose that $a = 0$; then,

$$\det(A) = -bc \neq 0; \tag{5.40}$$

so that

$$b \neq 0 \quad \text{and} \quad c \neq 0. \tag{5.41}$$

It then follows from the first equation in (5.30) that

$$by = 1,$$

from which we get that

$$y = \frac{1}{b}, \tag{5.42}$$

since $b \neq 0$ by the first condition in (5.41).

Next, use the second condition in (5.41) and (5.40) to rewrite (5.42) as

$$y = -\frac{c}{\det(A)}. \tag{5.43}$$

Next, use the second condition in (5.41) to solve the second equation in (5.30) to get

$$x = -\frac{d}{c}y. \tag{5.44}$$

Combining (5.44) and (5.43) then yields that

$$x = \frac{d}{\det(A)}. \tag{5.45}$$

Continuing with the assumption that $a = 0$, and using the first condition in (5.41), we obtain from the third equation in (5.30) that

$$w = 0, \tag{5.46}$$

we can rewrite as

$$w = \frac{a}{\det(A)}, \tag{5.47}$$

since $a = 0$.

Finally, substituting the result in (5.46) into the last equation in (5.30) we get

$$cz = 1,$$

which can be solved for z to yield

$$z = \frac{1}{c}, \tag{5.48}$$

in view of the second condition in (5.41). We can then use (5.39) to rewrite (5.48) as

$$z = -\frac{b}{\det(A)}. \tag{5.49}$$

We note that the results in (5.45), (5.43), (5.49) and (5.47) are precisely the results in (5.33), (5.34), (5.37) and (5.36), respectively, which is the result that we obtained in the previous case. Consequently, in this case as well we obtain that the inverse of A in (5.25) is given by B in (5.38).

The second option in (5.39) (namely, $c = 0$) yields the same result. We therefore conclude that, if A given in (5.25) is such that $\det(A) \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{5.50}$$

Example 5.2.22. Let A be the matrix in Example 5.2.21; namely,

$$A = \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix}. \tag{5.51}$$

Then, $\det(A) = 2$; so that, $\det(A) \neq 0$. Thus, we can use the formula in (5.50) to compute the inverse of A in (5.51) to get

$$A^{-1} = \frac{1}{2} \begin{pmatrix} -3 & 2 \\ -1 & 0 \end{pmatrix},$$

which yields the same matrix in (5.24) obtained in Example 5.2.21.

5.3 Geometry of Linear Functions

In this section we provide a geometric interpretation for the determinant of a 2×2 matrix.

Let A denote the 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{5.52}$$

where a, b, c and d are real numbers.

We have seen that this matrix can be used to define a linear function $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2. \tag{5.53}$$

We would like to understand geometrically how the function T defined in (5.53) acts on $\mathbb{R}^2$. To do this, consider the sketch in Figure 5.3.1. The Figure

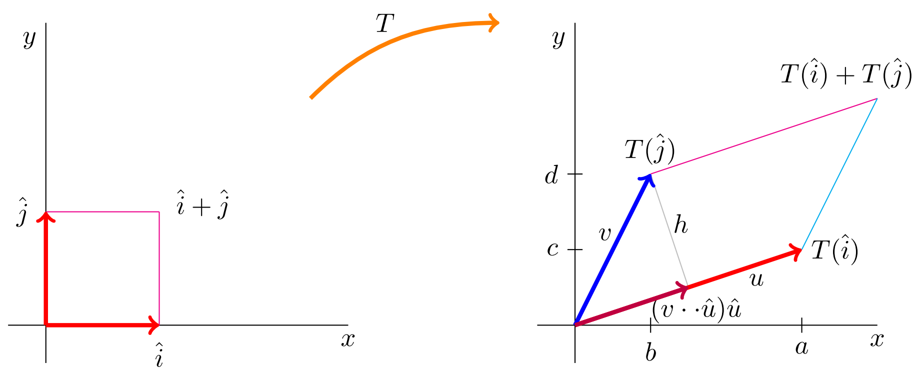

Figure 5.3.1: Action of a linear function

shows the unit vectors $\hat{i}$ and $\hat{j}$ and the square of vertices O, $\hat{i}$, $\hat{j}$ and $\hat{i}+\hat{j}$ on the left–side of the sketch, and the effect that the function T has on that square on the right–hand side of the sketch, for the case in which all the entries of the matrix A in (5.52) are assumed to be positive.

We compute

$$\begin{aligned} T(\hat{i}) &= A\hat{i} \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} a \\ c \end{pmatrix}. \end{aligned}$$

Thus, $T(\hat{i})$ is the first column of the matrix A given in (5.52).

Similarly,

$$\begin{aligned} T(\hat{j}) &= A\hat{j} \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} b \\ d \end{pmatrix}; \end{aligned}$$

so that, $T(\hat{j})$ is the second column of the matrix A given in (5.52).

The vectors $T(\hat{\imath})$ and $T(\hat{\jmath})$ are shown on the right–hand side of the sketch in Figure 5.3.1 for the case in which a, b, c and d are positive, $b < a$ and $c < b$. The sketch in the figure shows that T maps the square pictured on the left of the figure to the parallelogram shown on the right–hand side of the sketch. To see why this is the case, compute

$$\begin{aligned} T(\hat{\imath}+\hat{\jmath}) &= A(\hat{\imath}+\hat{\jmath}) \\ &= A\hat{\imath}+A\hat{\jmath} \\ &= T(\hat{\imath})+T(\hat{\jmath}); \end{aligned}$$

so that, T maps the corner of the square on the left of the sketch in Figure 5.3.1 at $\hat{\imath}+\hat{\jmath}$ to the vector sum, $T(\hat{\imath})+T(\hat{\jmath})$, of the images of $\hat{\imath}$ and $\hat{\jmath}$ under the function T. Thus, according to the parallelogram rule of vector addition, the image of $\hat{\imath}+\hat{\jmath}$ under T lies on the vertex of a parallelogram that is diagonally opposed to the origin. This parallelogram is shown on the right–hand side of the sketch in Figure 5.3.1.

We would like to compute the area of the parallelogram with vertices at O, $T(\hat{\imath})$, $T(\hat{\jmath})$ and $T(\hat{\imath})+T(\hat{\jmath})$ pictured on the right–hand side of the sketch in Figure 5.3.1 in terms of the entries of the matrix A in (5.52).

Denote $T(\hat{\imath})$ by u and $T(\hat{\jmath})$ by v; so that,

$$u = \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad v = \begin{pmatrix} b \\ d \end{pmatrix}. \tag{5.54}$$

We will denote the parallelogram with vertices at O, $T(\hat{\imath})$, $T(\hat{\jmath})$ and $T(\hat{\imath}) + T(\hat{\jmath})$ by $\mathcal{P}(u,v)$ and the dote the area of the parallelogram by area$(\mathcal{P}(u,v))$. Then,

$$\text{area}(\mathcal{P}(u,v)) = \|u\|h, \tag{5.55}$$

where $\|u\|$ is the length of the vector u and h is the height of the parallelogram; that is, h is the distance from the vector v to the line thorough the origin in the direction of u.

Using the result of Problem 5 in Assignment #5, we see that the point through u in the direction of u that is the closest to v is

$$(v\cdot\hat{u})\hat{u},$$

where

$$\hat{u} = \frac{1}{\|u\|}u, \tag{5.56}$$

a unit vector in the direction u. Consequently,

$$h = \|v-(v\cdot\hat{u})\hat{u}\|. \tag{5.57}$$

Squaring on both sides of the equation in (5.55), we get

$$(\text{area}(\mathcal{P}(u,v)))^2 = \|u\|^2h^2; \tag{5.58}$$

so that, substituting the value for h in (5.57) in (5.58) yields

$$(\text{area}(\mathcal{P}(u,v)))^2 = \|u\|^2\|v-(v\cdot\hat{u})\hat{u}\|^2. \tag{5.59}$$

Next, use the properties the dot product in (5.59) to compute

$$\begin{aligned}(\text{area}(\mathcal{P}(u,v)))^2 &= \|u\|^2(v-(v\cdot\hat{u})\hat{u})\cdot(v-(v\cdot\hat{u})\hat{u}) \\ &= \|u\|^2(v\cdot v-(v\cdot\hat{u})v\cdot\hat{u}-(v\cdot\hat{u})\hat{u}\cdot v+(v\cdot\hat{u})^2\hat{u}\cdot\hat{u});\end{aligned}$$

thus, using $v\cdot v=\|v\|^2$ and the fact that $\hat{u}$ is a unit vector,

$$(\text{area}(\mathcal{P}(u,v)))^2 = \|u\|^2(\|v\|^2-(v\cdot\hat{u})^2), \tag{5.60}$$

where we have also used the symmetry of the dot product.

Next, use the definition of $\hat{u}$ in (5.56) to get from (5.60) that

$$(\text{area}(\mathcal{P}(u,v)))^2 = \|u\|^2\left(\|v\|^2-\frac{1}{\|u\|^2}(v\cdot u)^2\right),$$

from which we get that

$$(\text{area}(\mathcal{P}(u,v)))^2 = \|u\|^2\|v\|^2-(v\cdot u)^2. \tag{5.61}$$

We can express the right–hand side of (5.61) in terms of the components of u and v in (5.54), namely

$$\|u\|^2=a^2+c^2, \quad \|v\|^2=b^2+d^2, \quad \text{and} \quad v\cdot u=ab+cd,$$

to get that

$$\begin{aligned}(\text{area}(\mathcal{P}(u,v)))^2 &= (a^2+c^2)(b^2+d^2)-(ab+cd)^2 \\ &= a^2b^2+a^2d^2+c^2b^2+c^2d^2-(a^2b^2+2abcd+c^2d^2) \\ &= a^2d^2+c^2b^2-2abcd \\ &= (ad)^2-2(ad)(bc)+(bc)^2,\end{aligned}$$

from which we get that

$$(\text{area}(\mathcal{P}(u,v)))^2 = (ad-bc)^2. \tag{5.62}$$

Taking the square root on both sides of (5.62) then yields

$$\text{area}(\mathcal{P}(u,v)) = |ad-bc|. \tag{5.63}$$

We recognize on the right–hand side of (5.63) the expression $ad-bc$ for the determinant of the matrix A in (5.52). We can therefore rewrite (5.63) as

$$\text{area}(\mathcal{P}(u,v)) = |\det(A)|, \tag{5.64}$$

where u and v are the columns of the matrix A. Hence, according to (5.64), the absolute value of the determinant of A gives the area of the parallelogram determined by the columns of A.

5.4 The Flow of Two–Dimensional Linear Fields

The goal of this section is to compute the flow of two–dimensional linear fields $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}, \quad \text{for } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2. \tag{5.65}$$

The flow of F in (5.65) is made up of curves parametrized by paths

$$\begin{pmatrix} x \\ y \end{pmatrix} : \mathbb{R} \to \mathbb{R}^2$$

satisfying the differential equations

$$\begin{cases} \dfrac{dx}{dt} & = & ax+by; \\ \\ \dfrac{dy}{dt} & = & cx+dy. \end{cases} \tag{5.66}$$

We can rewrite the system in (5.66) in vector form as

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}, \tag{5.67}$$

where A is the 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{5.68}$$

and where we have used the notation convention for time–derivatives introduced in Remark 4.2.4.

Example 5.4.1. Consider system

$$\begin{cases} \dot{x} & = & -9y; \\ \\ \dot{y} & = & x-6y. \end{cases} \tag{5.69}$$

The matrix corresponding to the system in (5.69) is

$$A = \begin{pmatrix} 0 & -9 \\ 1 & -6 \end{pmatrix}. \tag{5.70}$$

Let v be the vector

$$\mathrm{v} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \tag{5.71}$$

We note that

$$A\mathrm{v} = \begin{pmatrix} 0 & -9 \\ 1 & -6 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} -9 \\ -3 \end{pmatrix};$$

so that,

$$Av = -3v \tag{5.72}$$

We will show that the path

$$\begin{pmatrix} x \\ y \end{pmatrix} : \mathbb{R} \to \mathbb{R}^2$$

defined by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c\ e^{-3t} v, \quad \text{for } t \in \mathbb{R}, \tag{5.73}$$

where c is a constant and v is the vector in (5.71), is a solution of the equation

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}, \tag{5.74}$$

where A is the 2×2 matrix given in (5.70). Indeed, taking the derivative with respect to t on both sides of (5.73) yields

$$\begin{aligned} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} &= \frac{d}{dt}\left[c\ e^{-3t} v\right] \\ &= c\frac{d}{dt}\left[\ e^{-3t}\right] v \\ &= c(-3)e^{-3t} v \\ &= ce^{-3t}(-3v); \end{aligned}$$

so that, by virtue of (5.72)

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = ce^{-3t} Av.$$

Consequently, using the properties of the matrix product,

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A\left[ce^{-3t} v\right]. \tag{5.75}$$

Comparing (5.73) and (5.75) we see that

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix},$$

which is (5.74).

We have therefore obtained solutions of the system in (5.69) which lie on a line determined by the vector v in (5.70). These solutions are sketched in Figure 5.4.2. The sketch shows the origin, corresponding to $c = 0$ in (5.73), and two half lines pointing towards the origin correspond to $c > 0$ (in the first quadrant), and to $c < 0$ (in the third quadrant). The lines point towards the origin because of the decreasing exponential in the definition of the solutions in (5.73).

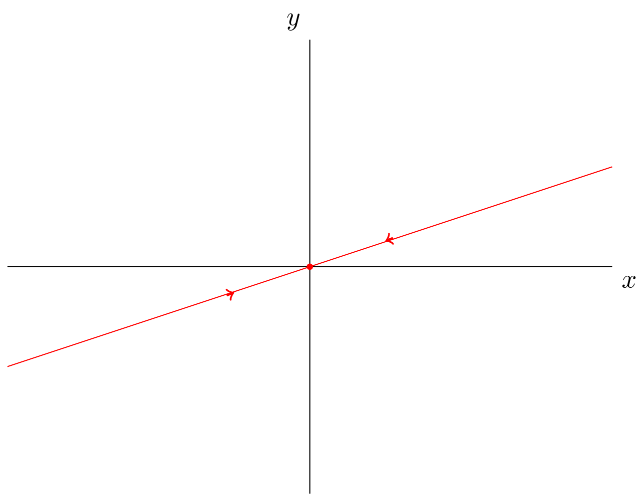

Figure 5.4.2: Sketch of (5.73)

Example 5.4.1 illustrates a special situation in the flow of linear fields. In some cases, trajectories along the flow will lie on lines through the origin. We shall refer to these special solutions as "line solutions."

Line solutions for a system of the form

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}, \tag{5.76}$$

where A is a 2×2 matrix, occur when there exists a nonzero vector v in $\mathbb{R}^2$ such that

$$A\mathrm{v} = \lambda \mathrm{v}, \tag{5.77}$$

for some scalar λ. When this is the case, the paths

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c\, e^{\lambda t}\mathrm{v}, \quad \text{for } t \in \mathbb{R}, \tag{5.78}$$

for arbitrary constant c, yields solutions of the system in (5.76). The solutions in (5.78) yield the origin for the case $c = 0$, and two half-lines in the direction of the vector v pointing towards the origin if $\lambda < 0$, or away from the origin if $\lambda > 0$.

We are able to find line solutions of the system in (5.76) as long as we are able to find scalars λ for which the equation in (5.77) holds true. This is a very special situation; when it occurs, we call the scalar λ an *eigenvalue* of the matrix A; a corresponding nonzero vector v for which (5.77) holds true is called an eigenvector for λ.

5.4.1 Eigenvalues and Eigenvectors

Given a 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{5.79}$$

we say that a scalar λ is an eigenvalue of A if there exists a nonzero vector

$$\mathrm{v} = \begin{pmatrix} x \\ y \end{pmatrix}$$

such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}. \tag{5.80}$$

Thus, to find eigenvalues for the matrix A in (5.79), we first need to find conditions on the entries of A that will guarantee that the equation in (5.80) has nonzero solutions. The equation in (5.80) can be written as system of two linear equations

$$\begin{cases} ax + by &= \lambda x; \\ cx + dy &= \lambda y, \end{cases}$$

or

$$\begin{cases} (a-\lambda)x + by &= 0; \\ cx + (d-\lambda)y &= 0. \end{cases} \tag{5.81}$$

By inspection we see that the system in (5.81) has the zero solution.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

For this reason, this solution is usually referred to as the *trivial solution*. However, in some cases, the system in (5.81) might have infinitely many solutions (this would be the case in which the two equations in (5.81) represent the same line). This occurs, according to the result in Problems 3 and 4 in Assignment #12, when

$$(a-\lambda)(d-\lambda) - bc = 0,$$

or

$$\lambda^2 - (a+d)\lambda + ad - bc = 0. \tag{5.82}$$

The equation in (5.82) is called the characteristic equation of the matrix A in (5.79). Its solutions will be eigenvalues of A. The polynomial

$$p_A(\lambda) = \lambda^2 - (a+d)\lambda + ad - bc \tag{5.83}$$

is called the *characteristic polynomial* of A. The roots, or zeros, of $p_A(\lambda)$ are the eigenvalues of A. We have encountered the expression $ad - bc$ in the characteristic polynomial in (5.83); it is the determinant of A, denoted by $\det(A)$. The coefficient $a + d$ of λ expression for $p(\lambda)$ in (5.83) is the sum of the entries along the main diagonal of A, and it is called the *trace* of A; we write,

$$\mathrm{trace}(A) = a + d.$$

We can therefore write the characteristic polynomial of A as

$$p_{_A}(\lambda) = \lambda^2 - \text{trace}(A)\lambda + \det(A). \tag{5.84}$$

Once we find an eigenvalue, λ, we can find a corresponding eigenvector by solving the system of equations in (5.81) for the specific value of λ. We illustrate this procedure in the following example.

Example 5.4.2. Find the eigenvalues and corresponding eigenvectors of the matrix

$$A = \begin{pmatrix} 3 & -1 \\ 5 & -3 \end{pmatrix}. \tag{5.85}$$

Solution: The trace of the matrix A in (5.85) is $\text{trace}(A) = 0$, its determinant is $\det(A) = -4$. Then, according to (5.84), the characteristic polynomial of A is

$$p_{_A}(\lambda) = \lambda^2 - 4,$$

which factors into

$$p_{_A}(\lambda) = (\lambda + 2)(\lambda - 2);$$

so that, the eigenvalues of A in (5.85) are

$$\lambda_1 = -2 \quad \text{and} \quad \lambda_2 = 2. \tag{5.86}$$

Next, we compute eigenvectors corresponding to λ_1 and λ_2 in (5.86). We do this by solving the system of equations in (5.81) for $\lambda = \lambda_1$ and for $\lambda = \lambda_2$.

For $\lambda = -1$, the system in (5.81) yields

$$\begin{cases} 5x - y &= 0; \\ 5x - y &= 0, \end{cases}$$

which is the single equation

$$y = 5x. \tag{5.87}$$

Since we are looking for a nontrivial solution os the system, we can set $x = 1$ in (5.87) to get $y = 5$; so that,

$$\text{v}_1 = \begin{pmatrix} 1 \\ 5 \end{pmatrix} \tag{5.88}$$

is an eigenvector corresponding to $\lambda_1 = -2$.

Similarly, for $\lambda = \lambda_2 = 2$, we obtain the system

$$\begin{cases} x - y &= 0; \\ x - y &= 0, \end{cases}$$

which is equivalent to the equations

$$y = x,$$

from which we get that

$$\text{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{5.89}$$

is an eigenvector corresponding to the eigenvalue $\lambda_2 = 2$. □

5.4.2 Line Solutions

An advantage of knowing eigenvalues and eigenvectors of a 2×2 matrix, A, is that they provide special kind of solutions of the system

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}. \tag{5.90}$$

For instance, if A has a real eigenvalue, λ, with a corresponding eigenvector, v, then the path $\begin{pmatrix} x \\ y \end{pmatrix} : \mathbb{R} \to \mathbb{R}^2$ defined by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c\, e^{\lambda t} \mathrm{v}, \quad \text{for } t \in \mathbb{R},$$

where c is a constant, solves the system in (5.90). An instance of this fact was seen in Example 5.4.1.

For the case in which the matrix A has two real eigenvalues, λ_1 and λ_2, with $\lambda_1 \neq \lambda_2$, corresponding eigenvectors v_1 and v_2, respectively, do not lie on the same line (see Problem 3 in Assignment #15). Consequently, we obtain two distinct line solutions of the system in (5.90),

$$\begin{pmatrix} x_1(t) \\ y_1(t) \end{pmatrix} = c_1\, e^{\lambda_1 t} \mathrm{v}_1 \quad \text{and} \quad \begin{pmatrix} x_2(t) \\ y_2(t) \end{pmatrix} = c_2\, e^{\lambda_2 t} \mathrm{v}_2 \quad \text{for } t \in \mathbb{R},$$

where c_1 and c_2 are constants. Furthermore, it is shown in courses in differential equations, that the expression

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1\, e^{\lambda_1 t} \mathrm{v}_1 + c_2\, e^{\lambda_2 t} \mathrm{v}_2 \quad \text{for } t \in \mathbb{R}, \tag{5.91}$$

where c_1 and c_2 are arbitrary constants, yields all solutions of the system in (5.90) (see also Problem 4 in Assignment #14).

The expression in (5.91) can be used to aid in sketching the flow the vector field

$$F\begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

We illustrate this in the next example.

Example 5.4.3. Sketch the flow of the vector field $F \colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x - y \\ 5x - 3y \end{pmatrix}, \quad \text{for } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2. \tag{5.92}$$

Solution: The flow of the field (5.92) is obtained by solving the pair of differential equations

$$\begin{cases} \dot{x} = 3x - y; \\ \dot{y} = 5x - 3y. \end{cases} \tag{5.93}$$

The system (5.93) can in turn be written in vector form as

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}, \tag{5.94}$$

where A is the matrix

$$A = \begin{pmatrix} 3 & -1 \\ 5 & -3 \end{pmatrix}. \tag{5.95}$$

We saw in Example 5.4.2 that the matrix A in (5.95) has eigenvalues $\lambda_1 = -2$ and $\lambda_2 = 2$, with corresponding eigenvectors given in (5.88) and (5.89), respectively; that is,

$$\mathrm{v}_1 = \begin{pmatrix} 1 \\ 5 \end{pmatrix} \quad \text{and} \quad \mathrm{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{5.96}$$

By the discussion preceding this example, all solutions of the system in (5.93) are given by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1\, e^{-2t}\mathrm{v}_1 + c_2\, e^{2t}\mathrm{v}_2 \quad \text{for } t \in \mathbb{R}, \tag{5.97}$$

where c_1 and c_2 are arbitrary constants, and v_1 and v_2 are as given in (5.96).

The general form of the solutions given in (5.97) is very helpful in sketching the the flow of the field in (5.92). We first note that, if $c_1 = c_2 = 0$, (5.97) yields the origin, $(0, 0)$, as a solution. If $c_1 \neq 0$ and $c_2 = 0$, the flow curves will lie on the line through the origin parallel to the vector v_1; both solution curves will point towards the origin because the exponential e^{-2t} decreases with increasing t. On the other hand, if $c_1 = 0$ and $c_2 \neq 0$, the trajectories lie on the in the direction of the vector v_2 and point away from the origin because the exponential e^{2t} increases with increasing t. We have therefore obtained the origin and four line solutions. All of these are shown in Figure 5.4.3. Figure 5.4.3 shows other possible flow curves of the field in (5.92). In the next section we will see how to sketch those curves. □

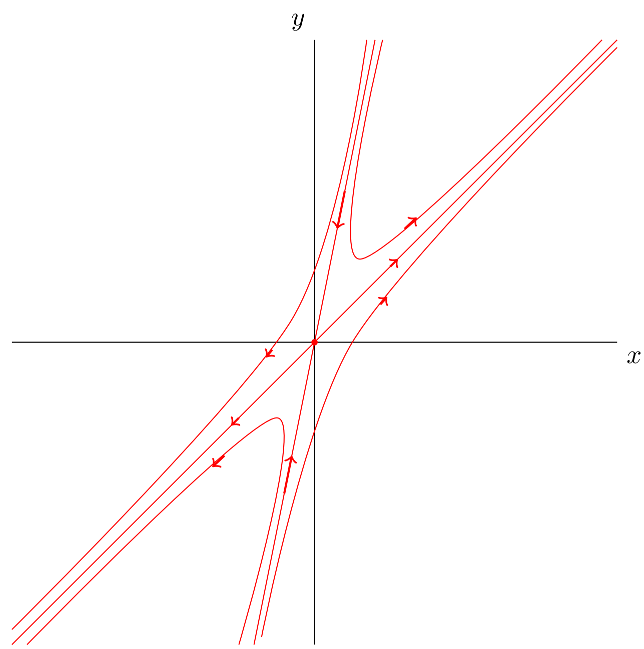

Figure 5.4.3: Sketch of Flow of the Field in (5.92)

Chapter 6

Linear Functions and Linear Approximations

6.1 Definition of linear functions

We have seen that a linear function $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ is of the form

$$T\begin{pmatrix} x \\ y \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{for } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \tag{6.1}$$

where A is a 2×2 matrix with real entries

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; \tag{6.2}$$

so that,

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}, \quad \text{for } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2. \tag{6.3}$$

Let $v_1 \in \mathbb{R}^2$ and $v_2 \in \mathbb{R}^2$ and use the definition of $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ in (6.1) to compute

$$T(v_1 + v_2) = A(v_1 + v_2),$$

where A is the 2×2 matrix given in (6.2); so that, using the distributive property of matrix algebra,

$$T(v_1 + v_2) = Av_1 + Av_2,$$

or

$$T(v_1 + v_2) = T(v_1) + T(v_2), \tag{6.4}$$

by the definition of T in (6.1).

Similarly, for any real number c and any vector $v \in \mathbb{R}^2$,

$$T(cv) = A(cv) = cAv,$$

from which we get that

$$T(cv) = cT(v). \tag{6.5}$$

All linear functions from $\mathbb{R}^2$ to $\mathbb{R}^2$ satisfy the two properties in (6.4) and (6.5). Indeed, the expressions in (6.4) and (6.5) are the properties that define a linear function.

We will see other examples of functions satisfying properties (6.4) and (6.5) in the case of functions from $\mathbb{R}^2$ to $\mathbb{R}$. In fact, the components of linear vector field $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ in (6.3), namely the functions $\ell_1\colon \mathbb{R}^2 \to \mathbb{R}$ and $\ell_2\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$\ell_1(v) = ax + by, \qquad \text{for } v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \tag{6.6}$$

and

$$\ell_2(v) = cx + dy, \qquad \text{for } v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2,$$

respectively, are linear functions.

We verify that the function $\ell_1\colon \mathbb{R}^2 \to \mathbb{R}$ give in (6.6) satisfies the properties in (6.4) and (6.5), namely

$$\ell_1(v_1 + v_2) = \ell_1(v_1) + \ell_1(v_2), \qquad \text{for } v_1, v_2 \in \mathbb{R}^2, \tag{6.7}$$

and

$$\ell_1(cv) = c\ell_1(v), \qquad \text{for } v \in \mathbb{R}^2 \text{ and } c \in \mathbb{R}, \tag{6.8}$$

respectively.

To verify (6.7), set $v_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$, and use the definition of ℓ_1 in (6.6) to compute

$$\ell_1(v_1 + v_2) = \ell_1 \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} = a(x_1 + x_2) + b(y_1 + y_2);$$

so that, using the distributive, associative and commutative properties,

$$\begin{aligned} \ell_1(v_1 + v_2) &= ax_1 + ax_2 + by_1 + by_2 \\ &= ax_1 + by_1 + ax_2 + by_2 \\ &= (ax_1 + by_1) + (ax_2 + by_2) \\ &= \ell_1(v_1) + \ell_1(v_2), \end{aligned}$$

which is (6.7).

Similarly, to establish (6.8), let $v = \begin{pmatrix} x \\ y \end{pmatrix}$ and $c \in \mathbb{R}$, and compute, using the definition of ℓ_1 in (6.6),

$$\ell_1(cv) = \ell_1 \begin{pmatrix} cx \\ cy \end{pmatrix} = a(cx) + b(cy);$$

which, by virtue of the associative and distributive properties can be written as

$$\ell_1(cv) = c(ax + by) = c\ell_1(v),$$

which is (6.8).

We could have also established (6.7) and (6.8) by realizing that $\ell_1(v)$ can be written as the dot product of the vector $w = \begin{pmatrix} a \\ b \end{pmatrix}$ with the vector $v = \begin{pmatrix} x \\ y \end{pmatrix}$; that is,

$$\ell_1(v) = w \cdot v, \quad \text{for all } v \in \mathbb{R}^2. \tag{6.9}$$

The properties in (6.7) and (6.8) can then be derived from (6.9) and the properties of the dot product.

In fact, any linear function $\ell\colon \mathbb{R}^2 \to \mathbb{R}$ is of the form on (6.9); that is, $\ell(v)$, for any $v \in \mathbb{R}^2$, is the dot product of v with a vector $w \in \mathbb{R}^2$:

$$\ell(v) = w \cdot v, \quad \text{for all } v \in \mathbb{R}^2. \tag{6.10}$$

If $w = \begin{pmatrix} a \\ b \end{pmatrix}$, then the expression for ℓ in (6.10) can be written in matrix form as

$$\ell\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2. \tag{6.11}$$

6.2 Linear Approximations

Given a function $f\colon D \to \mathbb{R}$, defined in a domain $D \subseteq \mathbb{R}^2$, and $(x_o, y_o) \in D$, we would like to approximate $f(x, y)$, for (x, y) near (x_o, y_o), by an expression of the form

$$f(x_o, y_o) + \ell\begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix}, \tag{6.12}$$

where $\ell\colon \mathbb{R}^2 \to \mathbb{R}$ is a linear function.

The expression in (6.12) is called a *linear approximation* to $f(x_o, y_o)$ for (x, y) near (x_o, y_o). We write,

$$f(x, y) \approx f(x_o, y_o) + \ell\begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix}, \quad \text{for } (x, y) \text{ near } (x_o, y_o). \tag{6.13}$$

Out of all the linear approximations to $f(x, y)$ for (x, y) near (x_o, y_o), given in (6.13), we would like to find the best one.

Example 6.2.1. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x, y) = x^2 + y^2, \quad \text{for all } (x, y) \in \mathbb{R}^2.$$

We would like to find a linear function $\ell\colon \mathbb{R}^2 \to \mathbb{R}$ such that

$$f(x, y) \approx 2 + \ell\begin{pmatrix} x - 1 \\ y - 1 \end{pmatrix}, \quad \text{for } (x, y) \text{ near } (1, 1),$$

and the approximation is the best possible.

To understand the concept of the best linear approximation to a real–valued function near a point, we define the *error* in the approximation

$$E(x,y) = f(x,y) - f(x_o,y_o) - \ell\begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix}, \qquad \text{for } (x,y) \text{ near } (x_o,y_o). \quad (6.14)$$

Note that $E(x_o,y_o) = 0$, since $\ell\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0$, for any linear function $\ell\colon \mathbb{R}^2 \to \mathbb{R}$. We would like the error term in (6.14) to approach 0 as (x,y) approaches (x_o,y_o); in symbols,

$$\lim_{\|(x,y)-(x_o,y_o)\|\to 0} E(x,y) = 0, \quad (6.15)$$

where

$$\|(x,y)-(x_o,y_o)\| = \sqrt{(x-x_o)^2+(y-y_o)^2}. \quad (6.16)$$

We will say that a good linear approximation is one for which the error term approaches 0 at a quadratic rate; that is,

$$|E(x,y)| \leqslant C\|(x,y)-(x_o,y_o)\|^2, \quad (6.17)$$

for some positive constant C.

Note that (6.17) implies (6.15) by virtue of the Squeeze Lemma.

Assume that $(x,y) \neq (x_o,y_o)$ and that (x,y) is near (x_o,y_o). Then, we can divide both sides of the inequality in (6.17) by $\|(x,y)-(x_o,y_o)\|$ to get

$$\frac{|E(x,y)|}{\|(x,y)-(x_o,y_o)\|} \leqslant C\|(x,y)-(x_o,y_o)\|.$$

Thus, by the Squeeze Lemma,

$$\lim_{\|(x,y)-(x_o,y_o)\|\to 0} \frac{|E(x,y)|}{\|(x,y)-(x_o,y_o)\|} = 0. \quad (6.18)$$

We will use (6.18) in the definition of the best linear approximation to $f(x,y)$ for (x,y) near (x_o,y_o).

To simplify notation, we rewrite the limit in (6.18) as

$$\lim_{(x,y)\to(x_o,y_o)} \frac{|E(x,y)|}{\|(x,y)-(x_o,y_o)\|} = 0, \quad (6.19)$$

Where $(x,y) \to (x_o,y_o)$ is understood as $\|(x,y)-(x_o,y_o)\| \to 0$ or, according to (6.16),

$$\sqrt{(x-x_o)^2+(y-y_o)^2} \to 0.$$

Definition 6.2.2 (Linear Approximation to a Real–Value Function)**.** Let D denote a domain in $\mathbb{R}^2$ and $f\colon D \to \mathbb{R}$ be a real–valued function defined in D. Let $(x_o,y_o) \in D$. Suppose that there exists a linear function $\ell\colon \mathbb{R}^2 \to \mathbb{R}$ such that

$$f(x,y) = f(x_o,y_o) + \ell\begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix} + E(x,y), \qquad \text{for } (x,y) \text{ near } (x_o,y_o), \quad (6.20)$$

where the error term E satisfies (6.19). We then say that

$$f(x_o, y_o) + \ell \begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix}$$

is the linear approximation, or *first–order approximation*, to $f(x, y)$ for (x, y) near (x_o, y_o).

Example 6.2.3. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x, y) = x^2 + y^2, \quad \text{for all } (x, y) \in \mathbb{R}^2. \tag{6.21}$$

We would like to find a the linear approximation to f near $(1, 1)$.

Compute

$$\begin{aligned} (x-1)^2 + (y-1)^2 &= x^2 - 2x + 1 + y^2 - 2y + 1 \\ &= x^2 + y^2 - 2x - 2y + 2 \\ &= f(x, y) - 2(x-1) - 2(y-1) - 2. \end{aligned}$$

Consequently,

$$f(x, y) = f(1, 1) + 2(x-1) + 2(y-1) + (x-1)^2 + (y-1)^2. \tag{6.22}$$

Thus, defining $\ell\colon \mathbb{R}^2 \to \mathbb{R}$ by

$$\ell \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{for all } (x, y) \in \mathbb{R}^2, \tag{6.23}$$

and

$$E(x, y) = (x-1)^2 + (y-1)^2, \quad \text{for all } (x, y) \in \mathbb{R}^2, \tag{6.24}$$

we see from (6.22) that

$$f(x, y) = f(1, 1) + \ell \begin{pmatrix} x - 1 \\ y - 1 \end{pmatrix} + E(x, y), \tag{6.25}$$

where $\ell\colon \mathbb{R}^2 \to \mathbb{R}$ is the linear function defined in (6.23), and the error term E is given in (6.24).

Next, suppose that $(x, y) \neq (1, 1)$ and divide both sides of the equation in (6.24) by $\|(x, y) - (x_o, y_o)\|$ to get

$$\frac{|E(x, y)|}{\|(x, y) - (1, 1)\|} = \sqrt{(x-1)^2 + (y-1)^2}, \quad \text{for } (x, y) \neq (1, 1), \tag{6.26}$$

where we have used (6.16).

It follows from (6.26) that

$$\lim_{(x,y)\to(1,1)} \frac{|E(x,y)|}{\|(x,y)-(1,1)\|} = 0.$$

Consequently,

$$f(1,1) + \ell\begin{pmatrix} x-1 \\ y-1 \end{pmatrix} = 2 + 2(x-1) + 2(y-1), \qquad \text{for } (x,y) \in \mathbb{R}^2,$$

is the linear approximation to the function f defined in (6.21) near $(1,1)$, according to Definition 6.2.2.

6.3 Linear Approximations and Partial Derivatives

In this section we describe a general procedure for computing linear approximations for a large class of real–valued functions, $f\colon D \to \mathbb{R}$, defined on a subsets, D, of $\mathbb{R}^2$.

Let $(x_o, y_o) \in D$ and suppose that f has a linear approximation at (x_o, y_o) according to Definition 6.2.2; so that,

$$f(x,y) = f(x_o,y_o) + \ell\begin{pmatrix} x-x_o \\ y-y_o \end{pmatrix} + E(x,y), \qquad \text{for } (x,y) \text{ near } (x_o,y_o), \tag{6.27}$$

where the error term E satisfies

$$\lim_{(x,y)\to(x_o,y_o)} \frac{|E(x,y|}{\|(x,y)-(x_o,y_o)\|} = 0, \tag{6.28}$$

and $\ell\colon \mathbb{R}^2 \to \mathbb{R}$ is a linear function given by

$$\ell\begin{pmatrix} x \\ y \end{pmatrix} = ax + by, \qquad \text{for } (x,y) \in \mathbb{R}^2, \tag{6.29}$$

for some real constants a and b.

We consider the case in which

$$(x,y) = (x_o + h, y_o), \tag{6.30}$$

where $|h|$ is sufficiently small.

It follows from (6.30) that

$$\lim_{h\to 0}(x,y) = (x_o, y_o);$$

consequently, by virtue of (6.30),

$$(x,y) \to (x_o,y_o) \text{ is equivalent to } h \to 0. \tag{6.31}$$

Therefore, the limit in (6.28) is equivalent to

$$\lim_{h\to 0} \frac{|E(x_o+h,y_o)|}{|h|} = 0. \tag{6.32}$$

We then get from (6.20) that

$$f(x_o+h,y_o) = f(x_o,y_o) + \ell\begin{pmatrix} h \\ 0 \end{pmatrix} + E(x_o+h,y_o), \quad \text{for small } |h|, \tag{6.33}$$

where the error term $E(x_o+h,y_o)$ satisfies (6.32).

Using the definition of ℓ in (6.29), we obtain from (6.33) that

$$f(x_o+h,y_o) = f(x_o,y_o) + ah + E(x_o+h,y_o), \quad \text{for small } |h|, \tag{6.34}$$

where the error term $E(x_o+h,y_o)$ satisfies (6.32).

It follows from (6.34) that, if $h\neq 0$ and $|h|$ is small,

$$\frac{f(x_o+h,y_o)-f(x_o,y_o)}{h} = a + \frac{E(x_o+h,y_o)}{h}, \quad \text{for } h\neq 0 \text{ and } |h| \text{ small.} \tag{6.35}$$

Now, it follows from (6.32) that

$$\lim_{h\to 0} \frac{E(x_o+h,y_o)}{h} = 0.$$

Consequently, the limit as $h\to 0$ of the expression on the left–hand side of (6.35) exists, and is given by

$$\lim_{h\to 0} \frac{f(x_o+h,y_o)-f(x_o,y_o)}{h} = a. \tag{6.36}$$

The limit on the left–hand side of (6.36), if it exists, is called the **partial derivative** of f with respect to x at the point (x_o,y_o), and is denoted by

$$\frac{\partial f}{\partial x}(x_o,y_o).$$

Similar calculations to those leading to (6.36) can be used to show that

$$\lim_{k\to 0} \frac{f(x_o,y_o+k)-f(x_o,y_o)}{k} = b, \tag{6.37}$$

where b is given in the definition of the linear function $\ell\colon \mathbb{R}^2\to\mathbb{R}$ in (6.29).

The limit on the left–hand side of (6.37), if it exists, is called the partial derivative of f with respect to y at the point (x_o,y_o), and is denoted by

$$\frac{\partial f}{\partial y}(x_o,y_o).$$

Definition 6.3.1 (Partial Derivatives)**.** Let $f\colon D \to \mathbb{R}$ denote a function of two variables, x and y, defined in a domain, D, in $\mathbb{R}^2$. Let (x, y) be a point in D. If

$$\lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h} \tag{6.38}$$

exists, we call the limit the partial derivative of f with respect to x at (x, y), and denote it by

$$\frac{\partial f}{\partial x}(x,y).$$

Thus, if the limit in (6.38) exists, we write

$$\frac{\partial f}{\partial x}(x,y) = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}. \tag{6.39}$$

Similarly, if

$$\lim_{k\to 0} \frac{f(x,y+k) - f(x,y)}{k}$$

exists, we call the limit the partial derivative of f with respect to y at (x, y), and denote it by

$$\frac{\partial f}{\partial y}(x,y),$$

and we write

$$\frac{\partial f}{\partial y}(x,y) = \lim_{k\to 0} \frac{f(x,y+k) - f(x,y)}{k}. \tag{6.40}$$

Remark 6.3.2 (Computing partial derivatives)**.** For the case of a real–values function of a single variable, $g\colon I \to \mathbb{R}$, where I is an open interval of real numbers, and $x \in I$, if

$$\lim_{h\to 0} \frac{g(x+h) - g(x)}{h}$$

exists, we call it the derivative of g at x and denote it by $g'(x)$; so that,

$$g'(x) = \lim_{h\to 0} \frac{g(x+h) - g(x)}{h}. \tag{6.41}$$

Thus, comparing the right–side in (6.39) with the right–hand side in (6.41) suggests that, to compute the partial derivative of f with respect, we may think of the value of y as fixed (constant), and proceed by computing an ordinary derivative with respect to x while holding y constant. The following examples illustrate this procedure.

Example 6.3.3. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = xy, \quad \text{for all } (x,y) \in \mathbb{R}^2.$$

Compute the partial derivatives of f with respect to x and with respect to y.

Solution: Thinking of y as constant, we take the derivative with respect to x to get

$$\frac{\partial f}{\partial x}(x,y) = y, \quad \text{for all } (x,y) \in \mathbb{R}^2.$$

Similarly, thinking of x as fixed and taking the derivative with respect to y, we get

$$\frac{\partial f}{\partial y}(x,y) = x, \quad \text{for all } (x,y) \in \mathbb{R}^2.$$

□

Example 6.3.4. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = e^{-x^2-y^2}, \quad \text{for all } (x,y) \in \mathbb{R}^2.$$

Compute the partial derivatives of f with respect to x and with respect to y.

Solution: Thinking of y as constant, we take the derivative with respect to x to get

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= \frac{\partial}{\partial x}\left[e^{-x^2-y^2}\right] \\ &= e^{-x^2-y^2} \cdot \frac{\partial}{\partial x}\left[-x^2-y^2\right], \end{aligned}$$

where we have used the Chain–Rule.

Consequently,

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= e^{-x^2-y^2} \cdot (-2x) \\ &= -2xe^{-x^2-y^2}, \end{aligned}$$

for $(x,y) \in \mathbb{R}^2$.

Similarly, thinking of x as fixed and taking the derivative with respect to y, we get

$$\frac{\partial f}{\partial y}(x,y) = -2ye^{-x^2-y^2}, \quad \text{for all } (x,y) \in \mathbb{R}^2.$$

□

We have shown in this section that if $f\colon D \to \mathbb{R}$ satisfies

$$f(x,y) = f(x_o,y_o) + a(x-x_o) + b(y-y_o) + E(x,y), \tag{6.42}$$

for (x,y) near (x_o,y_o) in D, where the error term satisfies

$$\lim_{(x,y)\to(x_o,y_o)} \frac{|E(x,y|}{\|(x,y)-(x_o,y_o)\|} = 0,$$

then a and b in (6.42) are necessarily the partial derivatives of f at (x_o,y_o) with respect to x and with respect to y, respectively. Thus, we define the linear approximation to f at (x_o,y_o) to be the affine function $L\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$L(x,y) = f(x_o,y_o) + \frac{\partial f}{\partial x}(x_o,y_o)\cdot(x-x_o) + \frac{\partial f}{\partial y}(x_o,y_o)\cdot(y-y_o), \tag{6.43}$$

for$(x, y) \in \mathbb{R}^2$.

$L(x, y)$ approximates $f(x, y)$ when (x, y) is very close to (x_o, y_o). We write

$$f(x, y) \approx f(x_o, y_o) + \frac{\partial f}{\partial x}(x_o, y_o) \cdot (x - x_o) + \frac{\partial f}{\partial y}(x_o, y_o) \cdot (y - y_o), \qquad (6.44)$$

for (x, y) in D sufficiently close to (x_o, y_o).

Example 6.3.5. Let $D = \{(x, y) \in \mathbb{R}^2 \mid x > 0 \text{ and } y > 0\}$ and $f\colon D \to \mathbb{R}$ be given by

$$f(x, y) = \sqrt{x^2 + y^3}, \qquad \text{for } (x, y) \in D.$$

Note that $f(1, 2) = 3$. Then, according to (6.44), the linear approximation to f at $(1, 2)$ is

$$L(x, y) = 3 + \frac{\partial f}{\partial x}(1, 2) \cdot (x - 1) + \frac{\partial f}{\partial y}(1, 2) \cdot (y - 2), \qquad \text{for } (x, y) \in \mathbb{R}^2, \quad (6.45)$$

where

$$\frac{\partial f}{\partial x}(x, y) = \frac{x}{\sqrt{x^2 + y^3}}, \qquad \text{for } (x, y) \in D,$$

and

$$\frac{\partial f}{\partial y}(x, y) = \frac{3y^2}{2\sqrt{x^2 + y^3}}, \qquad \text{for } (x, y) \in D.$$

Consequently,

$$\frac{\partial f}{\partial x}(1, 2) = \frac{1}{3},$$

and

$$\frac{\partial f}{\partial y}(1, 2) = 2.$$

Hence, in view of (6.45),

$$L(x, y) = 3 + \frac{1}{3}(x - 1) + 2(y - 2), \qquad \text{for } (x, y) \in \mathbb{R}^2. \qquad (6.46)$$

We can use the expression in (6.46) to approximate $f(x, y)$ for (x, y) close to $(1, 2)$. For example,

$$\begin{aligned} f(0.99, 2.03) &\approx L(0.99, 2.03) \\ &= 3 + \frac{1}{3}(0.99 - 1) + 2(2.03 - 2) \\ &= 3 + \frac{1}{3}(-0.01) + 2(0.03) \\ &\approx 3 - 0.003 + 0.06; \end{aligned}$$

so that,

$$f(0.99, 2.03) \approx 3.057.$$

6.4 Partial Derivatives and the Gradient

The linear approximation to f at $(x_o, y_o) \in D$ given in (6.43) can be written as $f(x_o, y_o)$ plus the dot product of the vector

$$\nabla f(x_o, y_o) = \frac{\partial f}{\partial x}(x_o, y_o)\ \hat{i} + \frac{\partial f}{\partial y}(x_o, y_o)\ \hat{j} \tag{6.47}$$

with the vector

$$(x - x_o)\ \hat{i} + (y - y_o)\ \hat{j} = x\ \hat{i} + y\ \hat{j} - (x_o\ \hat{i} + y_o\ \hat{j});$$

so that,

$$L(x, y) = f(x_o, y_o) + \nabla f(x_o, y_o) \cdot [x\ \hat{i} + y\ \hat{j} - (x_o\ \hat{i} + y_o\ \hat{j})], \tag{6.48}$$

for $(x, y) \in \mathbb{R}^2$.

The expression in (6.47) is called the **gradient** of f at (x_o, y_o).

Definition 6.4.1 (Gradient of a function)**.** Given a function real–valued function $f\colon D \to \mathbb{R}$ defined in some domain D in $\mathbb{R}^2$, suppose that the partial derivatives of f exist at all points $(x, y) \in D$. The gradient of f is the vector field $\nabla f\colon D \to \mathbb{R}^2$ given by

$$\nabla f(x, y) = \frac{\partial f}{\partial x}(x, y)\ \hat{i} + \frac{\partial f}{\partial y}(x, y)\ \hat{j}, \quad \text{for all } (x, y) \in D.$$

Example 6.4.2. Let $D = \{(x, y) \in \mathbb{R}^2 \mid (x, y) \neq (0, 0)\}$ and define $f\colon D \to \mathbb{R}$ by

$$f(x, y) = \ln\left(\sqrt{x^2 + y^2}\right), \quad \text{for } (x, y) \in D.$$

Compute $\nabla f(x, y)$ for all $(x, y) \in D$.

***Solution**:* Write the function f as

$$f(x, y) = \frac{1}{2}\ln(x^2 + y^2), \quad \text{for all } (x, y) \in D,$$

and use the Chain–Rule to compute

$$\frac{\partial f}{\partial x}(x, y) = \frac{x}{x^2 + y^2}, \quad \text{for all } (x, y) \in D.$$

Similarly,

$$\frac{\partial f}{\partial y}(x, y) = \frac{y}{x^2 + y^2}, \quad \text{for all } (x, y) \in D.$$

Thus,

$$\nabla f(x, y) = \frac{x}{x^2 + y^2}\ \hat{i} + \frac{y}{x^2 + y^2}\ \hat{j}, \quad \text{for all } (x, y) \in D,$$

or

$$\nabla f(x, y) = \frac{1}{x^2 + y^2}\ (x\ \hat{i} + y\ \hat{j}), \quad \text{for all } (x, y) \in D.$$

□

6.5 The Gradient and the Chain Rule

Let D denote a domain in the xy–plane and $f\colon D \to \mathbb{R}$ be a real–valued function defined in D that has continuous partial derivatives in D. For $(x_o, y_o) \in D$, we write the expression for the linear approximation of f at (x_o, y_o) in (6.48) as

$$L(x, y) = f(x_o, y_o) + \nabla f(x_o, y_o) \cdot [(x, y) - (x_o, y_o)], \tag{6.49}$$

for $(x, y) \in \mathbb{R}^2$, where we have written (x, y) for the vector $x\,\hat{\imath} + y\,\hat{\jmath}$ and (x_o, y_o) for the vector $x_o\,\hat{\imath} + y_o\,\hat{\jmath}$.

The function $L\colon \mathbb{R}^2 \to \mathbb{R}$ defined in (6.49) can be used to approximate $f(x, y)$ for values of (x, y) that are very close to (x_o, y_o); so that,

$$f(x, y) \approx f(x_o, y_o) + \nabla f(x_o, y_o) \cdot [(x, y) - (x_o, y_o)], \tag{6.50}$$

for (x, y) very close to (x_o, y_o).

Let $\sigma\colon I \to \mathbb{R}^2$ be a differentiable path, for an open interval of real numbers I. Assume that $\sigma(t) \in D$ for all $t \in I$. We can then consider the values of the function f on the path σ: $f(\sigma(t))$ for all $t \in I$. This defines a real valued function $g\colon I \to \mathbb{R}$ given by

$$g(t) = f(\sigma(t)), \quad \text{for all } t \in I. \tag{6.51}$$

Thus, the function g defined in (6.51) is the composition of the real valued function f with the vector–valued function σ.

Assume that $\sigma(t_o) = (x_o, y_o)$ for some $t_o \in I$. We would like to approximate $g(t)$ for t near t_o using the linear approximation L to f given in (6.49) and (6.50), and the linear approximation, $\ell\colon \mathbb{R} \to \mathbb{R}^2$, to the path σ near t_o given by

$$\ell(t) = \sigma(t_o) + (t - t_o)\sigma'(t_o), \quad \text{for all } t \in \mathbb{R}, \tag{6.52}$$

where σ' is the derivative of the path σ.

The line in (6.52) approximates the curve parametrized by σ for t very close to t_o; so that,

$$\sigma(t) \approx \sigma(t_o) + (t - t_o)\sigma'(t_o), \quad \text{for } t \in I \text{ very close to } t_o. \tag{6.53}$$

Thus, we can approximate the function g defined in (6.51) by

$$g(t) \approx f(\sigma(t_o) + (t - t_o)\sigma'(t_o)), \quad \text{for } t \in I \text{ close to } t_o. \tag{6.54}$$

We can then use the approximation to f given in (6.50) to get from (6.54) that

$$g(t) \approx f(\sigma(t_o)) + \nabla f(\sigma(t_o)) \cdot [(t - t_o)\sigma'(t_o)],$$

for $t \in I$ close to t_o, or

$$g(t) \approx g(t_o) + (t - t_o)\nabla f(\sigma(t_o)) \cdot \sigma'(t_o), \tag{6.55}$$

for $t \in I$ close to t_o, where we have used the definition of g in (6.51).

The right–hand side of the expression in (6.55) displays the linear approximation to g for $t \in I$ very close to t_o. For the case in which $t \neq t_o$ is very close to t_o, we obtain from (6.55) the approximation

$$\frac{g(t) - g(t_o)}{t - t_o} \approx \nabla f(\sigma(t_o)) \cdot \sigma'(t_o), \quad \text{for } t \neq t_o, \tag{6.56}$$

and $t \in I$ very close to t_o.

As t approaches t_o becomes exact and therefore

$$\lim_{t \to t_o} \frac{g(t) - g(t_o)}{t - t_o} = \nabla f(\sigma(t_o)) \cdot \sigma'(t_o). \tag{6.57}$$

The expression in (6.57) shows that, if f has continuous partial derivatives in D, and the path $\sigma \colon I \to \mathbb{R}$ is differentiable, then composition f and σ is differentiable, and its derivative is the dot product of the gradient of f with the velocity vector σ'. This is the Chain–Rule.

Proposition 6.5.1 (The Chain–Rule)**.** Let $f \colon D \to \mathbb{R}$ be a real–valued function defined on some domain, D, in the xy–plane, and let $\sigma \colon I \to \mathbb{R}^2$, for some open interval I, denote a differentiable path with $\sigma(t) \in D$ for all $t \in I$. Suppose that the partial derivatives of f exist and are continuous in D. Then, the composition $f \circ \sigma \colon I \to \mathbb{R}$ is differentiable in I, and

$$\frac{d}{dt}[f(\sigma(t))] = \nabla f(\sigma(t)) \cdot \sigma'(t), \quad \text{for all } t \in I.$$

Example 6.5.2 (Directional Derivatives)**.** Let $f \colon D \to \mathbb{R}$ have continuous partial derivatives in D and let $(x_o, y_o) \in D$. Let $\hat{u}$ denote a unit vector in $\mathbb{R}^2$ and define $\sigma \colon \mathbb{R} \to \mathbb{R}^2$ by

$$\sigma(t) = (x_o, y_o) + t\hat{u}, \quad \text{for all } t \in \mathbb{R}. \tag{6.58}$$

Then, $\sigma(0) = (x_o, y_o)$ and, for $|t|$ sufficiently small, $\sigma(t) \in D$, and we can apply the Chain-Rule to conclude that $f \circ \sigma$ is differentiable and

$$\frac{d}{dt}[f(\sigma(t))] = \nabla f(\sigma(t)) \cdot \sigma'(t), \quad \text{for } |t| \text{ sufficeintly small},$$

or, in view of (6.58),

$$\frac{d}{dt}[f((x_o, y_o) + t\hat{u})] = \nabla f((x_o, y_o) + t\hat{u}) \cdot \hat{u}, \tag{6.59}$$

for $|t|$ very close to 0.

Setting $t = 0$ in (6.59) we obtain

$$\frac{d}{dt}[f((x_o, y_o) + t\hat{u})]\Big|_{t=0} = \nabla f(x_o, y_o) \cdot \hat{u}. \tag{6.60}$$

The expression on the left–hand side of (6.60) is called the **directional derivative** of f at (x_o, y_o) in the direction of the unit vector $\hat{u}$, and will be denoted by $D_{\hat{u}} f(x_o, y_o)$. Thus, (6.60) can be rewritten as

$$D_{\hat{u}} f(x_o, y_o) = \nabla f(x_o, y_o) \cdot \hat{u}.$$