

Chapter 2

Using Calculus to Model Epidemics

This chapter shows you how the description of changes in the number of sick people can be used to build an effective model of an epidemic. Calculus allows us to study change in significant ways.

In the United States, we have eradicated polio and smallpox, yet, despite vigorous vaccination campaigns, measles remains a persistent pest. Why were we able to eradicate polio but not measles? You can work out the answer to this question in the Herd Immunity project after you have studied this chapter. Calculus as the language of change really does give us deep insights. Plain English is not as powerful in understanding certain consequences of change. We want you to see an example immediately because the primary goal of our course is to show you that calculus has important things to contribute to many real problems.

The computer makes this study possible at such an early stage. The change model involves simple but tedious arithmetic. The formulas for this arithmetic look awfully messy, but you will be able to make computations easily in the short tables contained in the exercises. The computer will carry out much longer computations based on the same idea; the electronic exercises show you how to program the appropriate computations. The computer conveniently provides numerical and graphical results, so we can analyze our results and begin to develop insight into model epidemics.

Differential equations are ahead of our story only in a technical, not conceptual sense. You will see that the equations say what the various rates of change are, but you will not “integrate” them symbolically. You don’t need to know anything about symbolic derivatives or integrals (not even the ones in Chapter 1). Concentrate on the idea of how infectious individuals change the number of sick people over time, and let the computer do the arithmetic. The moral of the chapter is that these simple rates of change give us important information with the help of the computer. Seeing this will also raise new questions but at a higher level. Later, calculus and the computer together will answer these questions.

2.1 The First Model

An epidemic is a large short-term outbreak of a disease. This section develops a simple model of the spread of a disease.

Human epidemics are often spread by contact with infectious people, although sometimes there are “vectors,” such as mosquitos, rats and fleas, or mice and ticks involved in disease transmission. There are

many kinds of contagious diseases, such as smallpox, polio, measles, and rubella, that are easily spread through casual contact. Other diseases, such as gonorrhea, require more intimate contact.

Another important difference between the first group of diseases and gonorrhea is that they “confer immunity” to someone who recovers from it, whereas gonorrhea does not. In other words, once you recover from rubella, you cannot catch it again. This feature offers the possibility of control through vaccination. In this section we will formulate our first model of the spread of a disease such as rubella to which people are susceptible, then infectious, and finally, recovered and immune. (Rubella is commonly called “German measles.” It is usually mild, but may produce severe birth defects if contracted by a pregnant woman during the first trimester.)

A scientific model is necessarily a simplification. We cannot include contacts between particular individuals or even other diseases of the population, diet, the amount of travel, the weather and other factors that have “small” effects on the spread of the disease, because we would not be able to make computations and predictions with such a complicated model. By identifying the “main” effects, we can formulate a model that is simple enough to use to compute and predict the overall spread of a disease. That is the good news. The bad news is that our model cannot predict specifics, such as when you will contract rubella. Some exercises will ask you to use the model to make predictions, but others will ask you to think about and discuss limitations of the model. The usefulness of a model requires us to make correct scientific judgments about which effects are “small” and which are not. We need to test our models against known data both to measure important parameters and to verify that we have neglected only small effects. Once this is done, we can use the model to make predictions in new situations.

2.1.1 Basic Assumptions

We will make the following assumptions in formulating our model:

1. **SIR:** All individuals fit into one of the following categories:
 - Susceptible:** those who can catch the disease
 - Infectious:** those who can spread the disease
 - Removed:** those who are immune and cannot spread the disease
2. **The population is large** but fixed in size and confined to a well-defined region. You might imagine the population to be a large university during the semester, when relatively little outside travel takes place.
3. **The population is well mixed;** ideally, everyone comes in contact with the same fraction of people in each category every day. Again, imagine the multitude of contacts a student makes daily at a large university.

When formulating models, it is very important to make your assumptions clear. We have listed them explicitly to help make them clear. The exercises will ask you about the limitations of these assumptions.

The categories of people and the way we have assumed that they move between the categories can be summarized in the graphic “compartments” shown in Figure 2.1.

We want our mathematical model to be able to compute the number of people in each of the compartments at any given time. In mathematical models it is also important to keep a list of variables. Variables are letters that stand for quantities that can change. We must express the main properties of the epidemic in terms of these variables.

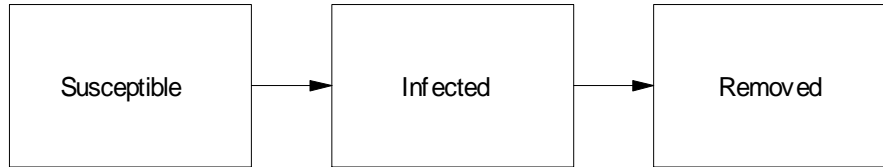


Figure 2.1: Disease Compartments

2.1.2 Variables for Model 1

t	=	the time in days with $t=0$ at the start of observation
S	=	the number of susceptible individuals
I	=	the number of infectious individuals
R	=	the number of removed individuals

These variables keep track of the number of individuals in each of the compartments as shown in Figure 2.1. Time t is the independent variable; and S , I , and R are dependent on time. In other words, they are functions of time, but they are not functions given by explicit formulas like the ones in Chapter 28 on high school review (such as $y = x^3$). Even though we will be able to compute $S = S[t]$, there is no explicit formula for S in terms of t .)

2.1.3 Derivation of the Equations of Change

Now we want to add a mathematical description of the way individuals move among the susceptible, infectious, and removed compartments. This will add two parameters to our model, a and b , that are specific to a particular S-I-R disease. Parameters are letters that have a fixed value during the entire problem but can change from one problem to the next. We begin with the simplest change - how one recovers.

A person with rubella is infectious for about 11 days. The infectious compartment will contain people who have had the disease for 1 day, 2 days, and so on up to 11 days. At 12 days an individual moves from the infectious to the removed compartment. Of course, not everyone gets the disease at exactly 8:00 am. Also, as the epidemic spreads, more people may be in the 1 day group than in the 11 day group. Nevertheless, we will assume that the different groups in the infectious compartment are roughly the same size so that we can make a simple statement about recovery: For rubella, the number of people added to the removed compartment tomorrow is $\frac{1}{11}$ th of the infectious compartment today,

$$\text{new people in } R\text{-compartment tomorrow} = \frac{1}{11}I$$

A general disease, from which one recovers and gains immunity, has the day's change in the recovered compartment given by,

$$\text{new people in } R\text{-compartment tomorrow} = b I$$

where the parameter b controls how quickly people recover, or $1/b$ is the number of days one stays in the infectious compartment. For rubella, $b = 1/11$.

Notice that the units of the right side of the above equation for the day's change in R are "number of people per day." We will keep careful account of units.

The next change we describe is how one spreads the disease. If you are infectious with rubella and you sit through a whole class next to a person who is susceptible, then you will probably spread the disease to that person. If you sit through class next to an immune person, you will not spread the disease. We are supposing that our population is well mixed, so people contact others from each compartment every day. The disease-causing contacts by a single infectious person is given by the number of "close contacts" made with susceptible people during the period of infectiousness. We will assume that we can find an average number of such contacts for everyone and denote this number by a parameter c

c = the average number of contacts per infective during the whole infectious period

The parameter c is the average total number of "close" contacts per infective. Some diseases are more contagious than others, so we want a parameter that describes the contagiousness. The parameter c is a combined measure of the level of mixing in the population and the contagiousness of the disease. The problem is: How do we measure c ? Later we will show that c can be measured indirectly with the help of calculus.

The number of contacts per infective each day is given by $a = c \cdot b$, because b equals $1/(\text{the number of days infectious})$. This means that if I people are infectious on a particular day,

$a \cdot I$ = the total number of adequate contacts per day

Only the contacts with susceptible people result in a new case of the disease, so we multiply this total times the fraction of susceptible people. Our final parameter, n , is the fixed total population. This makes the number of new cases of the disease,

$$\text{new cases tomorrow} = a \frac{S}{n} I$$

where

n = the fixed size of the population

2.1.4 Parameters of the S-I-R Model

b = $1/(\text{the average number of days infectious})$
 c = the average total number of contacts per infectious person
 $a = b \cdot c$ the daily rate of contacts per infective
 n = the total number of people in the population

Notice the units of a , b , and c :

c in average number of close contacts per infective
 b in $\frac{1}{\text{days}}$ or $\frac{1}{b}$ = number of days infective
 a in $\frac{\text{number (of adequate contacts)}}{\text{number (infectives)} \times \text{days (infectious)}}$

New cases cause the size of S to decrease and the size of I to increase by the same amount. Recoveries cause the size of I to decrease and the size of R to increase by that amount. It is best to write all our equations in terms of increases in the variables S , I , R . We simply take a "negative increase" to mean a decrease. The variable S has a negative increase $-a \frac{S}{n} I$, the variable R has a positive increase $b I$, and I has an input from the susceptible compartment and an output to the removed compartment, $a \frac{S}{n} I - b I$. The direction of change is shown in Figure 2.2.



Figure 2.2: Compartments with Parameters

2.1.5 The First Equations of Change

$$\begin{aligned} \text{tomorrow's increase in } S &= -a\frac{S}{n}I \\ \text{tomorrow's increase in } I &= a\frac{S}{n}I - bI \\ \text{tomorrow's increase in } R &= bI \end{aligned}$$

Mathematical models need to be expressed in terms of variables. The “equations” above are not written in terms of the variables we listed. In fact, the left-hand side is a phrase. The phrase is clear, and we want you to make some computations and think about formulating the equation in terms of the variables.

Exercise Set 2.1

- Use the daily changes above to compute 4 days of a rubella epidemic. You need the specific parameters, a , b , and c .
 - For rubella, we know that $\frac{1}{b} = 11$ and the average contact number is $c = 6.8$. Use these values to compute the parameter a .
 - Suppose a population initially (at $t = 0$) has

$$\begin{aligned} S &= 10,000 \\ I &= 1,000 \\ R &= 19,000 \end{aligned}$$

Use the equations of change above to complete a table showing the first four days of the epidemic.

t	0	1	2	3	4
S	10,000				
I	1,000				
R	19,000				

Table 2.1: Tabular epidemic

- Use your completed table to graph S versus t , I versus t , and R versus t .

The table you have just filled out is one way to record S , I , and R as function of t . To find I at time 3, you simply look in the I row below the 3 in the t row. The mathematical notation for a function is

$$S[t], \quad I[t], \quad R[t]$$

The value of $I[3]$ is recorded in the I row below the 3 column of the t row. You computed $S[1]$, $I[1]$, and $R[1]$ from the change equations and the known values $S[0]$, $I[0]$, and $R[0]$. Next, you used $S[1]$, $I[1]$, and $R[1]$ to compute $S[2]$, $I[2]$, $R[2]$, and so on. This is a recursive calculation: step 3 requires steps 1 and 2.

2. Write formulas for (an unknown) $S[t + 1]$, $I[t + 1]$, and $R[t + 1]$ in terms of a (known) $S[t]$, $I[t]$, and $R[t]$. Write your formulas so that $S[t + 1]$, $I[t + 1]$, and $R[t + 1]$ are on the left-hand side of the equalities and expressions in $S[t]$, $I[t]$, and $R[t]$ are on the right.

The answer to Exercise 2 is our first formal mathematical epidemic model. It allows us to recursively compute the course of an S-I-R epidemic in steps of 1 day.

3. The equations of change are not needed to compute all three compartment sizes. If $S[t]$ and $I[t]$ are known, express $R[t]$ in terms of n , $S[t]$, and $I[t]$. (Hint: What is the total population in terms of these categories?)
4. What are the units of the expressions:
 - a) $a \cdot I$
 - b) $a \cdot I \cdot \frac{S}{n}$
 - c) $b \cdot I$

What do they each measure?

5. Log onto your computer account, and run the program *FirstSIR* in the Chapter 2 folder. It will help you check the values in your table and learn how to compute with the computer.
6. Conjectures

We want you to speculate on the course of different kinds of diseases. Suppose disease “A” is very virulent; that is, it is very easily spread. Sketch the graph of I versus t for such a disease. Now do the same for disease “B,” which is not so virulent. Suppose both diseases have an infectious period of 14 days. What is the value of b for both models of these diseases? How does the parameter a for disease A, a_A compare to the parameter a for disease B, a_B ? Later in this chapter, you should compare your speculations with computed results.
7. Which parts of our model are invalid for describing a flu epidemic on our campus that begins just before Thanksgiving break? Is our model valid for a flu epidemic that begins just after Thanksgiving break?

Problem 2.1.1

Build a model for a disease that confers a period of immunity for a limited time after recovery but for which immunity is lost after this period. What other basic assumptions do you make? Write a detailed explanation.

2.2 Shortening the Time Steps

A whole day of exposure may not be needed to transmit disease, so people may move from the susceptible compartment to the infectious compartment in a shorter period of time. This section adds that feature to our model.

Time units of “days” are natural for rubella because $b = 1/11$ days comes directly from the period of infection. We want to keep these units. Consider the new infections with daily rate

$$\text{number of new infectious per day} = a \frac{S}{n} I$$

In 12 hours, $\frac{1}{2}$ day, there will be only half of the day’s contacts, so

$$\text{new infections in } \frac{1}{2} \text{ day} = a \frac{S}{n} I \cdot \frac{1}{2}$$

In 4 hours, there will be $4/24 = \frac{1}{6}$ th of the day’s contacts, so

$$\text{new infections in 4 hours} = a \frac{S}{n} I \cdot \frac{1}{6}$$

In 1 hour, there will be $\frac{1}{24}$ th of the day’s contacts, so

$$\text{new infections in 1 hour} = a \frac{S}{n} I \cdot \frac{1}{24}$$

In 1 minute, there will be $1/(24 \times 60) = 1/1440$ th of the day’s contacts, so

$$\text{new infections in 1 minute} = a \frac{S}{n} I \cdot \frac{1}{1440}$$

In general in Δt days, the number of new infections will be

$$\text{new infections in } \Delta t \text{ days} = a \frac{S}{n} I \cdot \Delta t$$

Notice the units,

$$\begin{aligned} \left(a \frac{S}{n} I \right) \times \Delta t & \text{ in (successful contacts per day) times (fractional days)} \\ & = \text{successful contacts during the } \Delta t \text{ period} \end{aligned}$$

If we calculate cases in half-day steps, the new cases are as above and the number of susceptibles in a half day equals the old susceptibles less the new cases,

$$S \left[\frac{1}{2} \right] = S[0] - \left(a \frac{S[0]}{n} I[0] \right) \cdot \frac{1}{2}$$

This quantity could then be used to compute $S[1] = S[1/2]$ – new cases during the second half day. We would need to compute the infectious population at time $= 1/2$ using

$$I \left[\frac{1}{2} \right] = I[0] + \left(a \frac{S[0]}{n} I[0] - b I[0] \right) \cdot \frac{1}{2}$$

because $I[1/2]$ is needed to find the rate of change of S during the next half day. Once this is known,

$$S[1] = S\left[\frac{1}{2}\right] - \left(a \frac{S[\frac{1}{2}]}{n} I\left[\frac{1}{2}\right]\right) \cdot \frac{1}{2}$$

and

$$I[1] = I\left[\frac{1}{2}\right] + \left(a \frac{S[\frac{1}{2}]}{n} I\left[\frac{1}{2}\right] - bI\left[\frac{1}{2}\right]\right) \cdot \frac{1}{2}$$

This looks more complicated than it really is. All it amounts to is using the present values of S and I to compute the next value. To see why, just complete the table in the next exercise.

Exercise Set 2.2

1. (a) Complete the table below where $t = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, 3, \frac{7}{2}, 4$:

t	0	$\frac{1}{2}$	1	$\frac{3}{2}$	2	$\frac{5}{2}$	3	$\frac{7}{2}$	4
S	10,000								
I	1,000								
R	19,000								

Table 2.2: Half-day table

- (b) Graph S , I , and R vs. t .
2. (a) What does the following quantity measure and in what units?

$$\left(a \frac{S[0]}{n} I[0] - bI[0]\right) \cdot \frac{1}{2}$$

- (b) Explain why the computation you did to complete the third column of the table in the previous exercise is represented functionally by

$$S\left[\frac{3}{2}\right] = S[1] - \left(a \frac{S[1]}{n} I[1]\right) \cdot \frac{1}{2}$$

$$I\left[\frac{3}{2}\right] = I[1] + \left(a \frac{S[1]}{n} I[1] - bI[1]\right) \cdot \frac{1}{2}$$

- (c) How many new infections will there be in 1 hour? Write formulas for $S[t + \frac{1}{24}]$ and $I[t + \frac{1}{24}]$ in terms of $S[t]$ and $I[t]$, which we assume are known, and where t has one of the values $t = \frac{k}{24}$, $k = 0, 1, 2, 3, \dots, 95$.
- (d) How many total computations would you need to do to compute $S[4]$, $I[4]$, and $R[4]$ in one hour steps? (Count all additions and multiplications. $t = 4$ means 4 days.)
- (e) Explain why the general recursive procedure for computation of $S[t]$ and $I[t]$ is

$$\text{Start with } \begin{array}{l} S[0] \\ I[0] \end{array}$$

For each successive $t = \Delta t, 2\Delta t, 3\Delta t, \dots$, compute the new function values,

$$S[t + \Delta t] = S[t] - \left(a \frac{S[t]}{n} I[t]\right) \cdot \Delta t$$

$$I[t + \Delta t] = I[t] + \left(a \frac{S[t]}{n} I[t] - bI[t]\right) \cdot \Delta t$$

The computer does not mind doing all the computations required even to find $S[400]$, $I[400]$, and $R[400]$ in one hour steps. You will use a “Do loop” to write a computer program for this computation.

3. Programming the Recursion

Run the program `SecondSIR` from the courseware. The program will not only do all the calculations necessary to solve the S - I - R model for small time steps but also will produce graphs of the results. The details are explained (within the program itself) first step by step, and then as a complete computation. First, use `SecondSIR` to verify your hand computations from Exercise 2.2 above. Second, use `SecondSIR` to calculate and graph a 3-month rubella epidemic.

2.3 The Continuous Variable Model

This section extends our disease model so that time varies continuously.

We will summarize what you should have learned from the text and exercises so far. The summary involves the awfully messy formulas mentioned in the introduction, but these lead to the simpler differential equations and, as you already know, they correspond to much simpler computer statements in **SecondSIR**.

First, it is fairly easy to give formulas for the average daily changes in S , I , and R , such as

$$\text{tomorrow's increase in } R = bI$$

Second, it is easy to use these equations to recursively tabulate the course of an epidemic as you did in the first exercise. Third, the *change* in a function of time such as $S[t]$ as time changes from t to $t + \Delta t$ is given by the difference

$$S[t + \Delta t] - S[t]$$

A general form of the solution to Exercise 2

$$\begin{aligned} S[t + \Delta t] &= S[t] - \left(a \frac{S[t]}{n} I[t] \right) \Delta t \\ I[t + \Delta t] &= I[t] + \left(a \frac{S[t]}{n} I[t] - b I[t] \right) \Delta t \end{aligned}$$

for $t = 0, \Delta t, 2\Delta t, 3\Delta t, \dots, k\Delta t, \dots$ may be written in the form

$$\begin{aligned} \frac{S[t + \Delta t] - S[t]}{\Delta t} &= -a \frac{S[t]}{n} I[t] \\ \frac{I[t + \Delta t] - I[t]}{\Delta t} &= a \frac{S[t]}{n} I[t] - b I[t] \end{aligned}$$

so both sides of the equations are rates of change. (Compute $R[t] = n - S[t] - I[t]$.) The first form is most useful in computing and looks simpler in **SecondSIR**. The second form is written as a rate of change in units of numbers of people per day,

$$\frac{S[t + \Delta t] - S[t]}{\Delta t} \quad \text{in numbers/day}$$

Recall that we verified that the units of the right-hand sides are numbers per day.

The time step Δt (or difference in time, abbreviated by Greek capital delta- t) is arbitrary, so it is not difficult to imagine t and Δt taking arbitrary continuous real values. We also want to let our basic variables take continuous real values so that we can formulate a calculus model.

2.3.1 The Continuous S-I-R Variables

For our continuous epidemic model, we choose

$$\begin{aligned} t &= \text{time measured in days continuously from } t = 0 \text{ at the start of the epidemic} \\ s &= \text{the fraction of the population that is susceptible} = \frac{S}{n} \\ i &= \text{the fraction of the population that is infectious} = \frac{I}{n} \\ r &= \text{the fraction of the population that is removed} = \frac{R}{n} \end{aligned}$$

where n is the (fixed) size of the total population. Dividing both sides of our previous equations by n , we obtain

$$\left(\frac{S[t + \Delta t]}{n} - \frac{S[t]}{n} \right) \frac{1}{\Delta t} = -a \frac{S[t]}{n} \frac{I[t]}{n}, \text{ etc.},$$

so that, in terms of the fraction variables, we obtain rate equations

$$\begin{aligned} \frac{\Delta s}{\Delta t} &= \frac{s[t + \Delta t] - s[t]}{\Delta t} = -a s[t] i[t] \\ \frac{\Delta i}{\Delta t} &= \frac{i[t + \Delta t] - i[t]}{\Delta t} = a s[t] i[t] - b i[t] \end{aligned}$$

where $\Delta s = s[t + \Delta t] - s[t]$ and $\Delta i = i[t + \Delta t] - i[t]$ represent the differences in the susceptible and infectious fractions corresponding to the change in time Δt . (We can compute $r[t] = 1 - s[t] - i[t]$.) Briefly, this is the expression

$$\begin{aligned} \frac{\Delta s}{\Delta t} &= -a s i \\ \frac{\Delta i}{\Delta t} &= a s i - b i \end{aligned}$$

Both sides of the above equations are average rates of change from time t to time $t + \Delta t$.

In calculus, the limit of the rate of change of a continuous function as the time step tends to zero is called the derivative,

$$\begin{aligned} \frac{ds}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{\Delta s}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{s[t + \Delta t] - s[t]}{\Delta t} \\ \frac{di}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{\Delta i}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{i[t + \Delta t] - i[t]}{\Delta t} \end{aligned}$$

Since the right-hand sides of the above rate equations do not depend on Δt , we can interpret the limiting equations as the system of S-I-R differential equations.

2.3.2 S-I-R Differential Equations

$$\begin{aligned} \frac{ds}{dt} &= -a s i \\ \frac{di}{dt} &= a s i - b i \end{aligned} \quad \text{(S-I-R DE's)}$$

Methods of calculus will let us compute interesting things about this continuous model, but for now we know that if we let Δt be “small,” then we can find $s[t]$ and $i[t]$ in recursive steps of Δt by specifying $a, b, s[0], i[0]$ and using a computer program like **SecondSIR** corresponding to the formulas

$$\begin{aligned} s[t + \Delta t] &= s[t] - (a s[t] i[t]) \Delta t \\ i[t + \Delta t] &= i[t] + (a s[t] i[t] - b i[t]) \Delta t \end{aligned}$$

The discrete recursive form is closer to the “differential” form of the equations

$$\begin{aligned} ds &= -a s i dt \\ di &= (a s i - b i) dt \end{aligned}$$

where

$$\begin{aligned} ds &\approx s[t + \Delta t] - s[t] = -a s[t] i[t] \Delta t \\ di &\approx i[t + \Delta t] - i[t] = (a s[t] i[t] - b i[t]) \Delta t \end{aligned}$$

The interpretation of the approximation or “limit” is more difficult for the differential form of the equations, because both sides tend to zero as Δt tends to zero. Actually, the functions $s[t]$ and $i[t]$ computed with the recursive formulas depend on the step size, Δt . You already may have seen that this process “converges” when you worked with **SecondSIR**. This convergence is called “Euler’s Method” of approximate solution of differential equations. It is an important “limit” based on a very simple idea and will be a recurring theme in this course.

Exercise Set 2.3

1. *What is the meaning of fractional values of $S, I,$ and R in Exercise 2.1.5 of Section 2.1? Are the fractional variables $s, i,$ and r meaningful in a small population?*
2. *The total population is part of a model with $S, I,$ and $R,$ but not with $s, i,$ and $r.$ Explain how we might use the history of a rubella epidemic with one size population to make predictions about a new and different size one.*

3. *Convergence and a Fast Solver*

*Are the recursive formulas for $s[t + \Delta t]$ exact solutions to the S - I - R differential equations? Run the **ThirdSIR** program and compare solutions with step sizes of $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$ over the time interval $0 \leq t \leq 4.$*

The point of Euler’s method is that you can start at a value and move a little in the direction of change given by a differential equation to find a nearby value. Once a new value is known, you can repeat the process with the known value. If the steps are small enough, this approximates the solution. Try this in the next exercise, and remember that you need only a rough sketch, rather than high numerical accuracy.

4. *The Geometry of Euler’s Method*

Let x and y be real variables. You are given that when $x = 0,$ then $y = 1,$

$$y[0] = 1 \quad \text{and that } y \text{ changes with respect to } x \text{ by the rule } \frac{dy}{dx} = -y$$

Sketch the graph of $y = y[x]$ by graphically using Euler’s Method.

HINTS: What is the slope of the graph at $(x, y) = (0, 1)$? What value of y do you get if you move along a graph of slope -1 for 0.1 x -units? What is the slope of the graph if you are at the point $(x, y) = (0.1, 0.9)$? What is the slope of the graph at a point $(x, y) = (?, 0.2)$? What is the slope of the graph at a point $(x, y) = (?, 0.0)$?

Now we are ready to start using the model we have developed.

Problem 2.3.1 DIFFERENCES IN PARAMETERS

Run the program **SIRsolver** from the courseware for the following cases of the parameters b and c .

1. A rubella epidemic, where $1/b = 11$, $c = 6.8$ and $a = b \cdot c$.
2. Measles, with a very high contact number $c = 15$ and $b = 1/8$.
3. Influenza, with a low contact number of $c = 1.4$ and $b = 1/3$.

The most important part of this problem is to analyze your results. What are the main differences among these three diseases? Which one is highly contagious, which is moderate and which is least contagious? Which one allows a larger portion of the population to escape infection? How many escape in each case? Write your summary analysis of the comparison among these diseases in a few brief paragraphs. (You may use a text cell in your copy of the program.)

If you made some conjectures in Exercise 6, test them by running this program with the appropriate parameters.

2.4 Analysis of Change

Now that we have a mathematically simple way to describe change in an epidemic and an effective way to experiment with our model using the computer, we want to return to some scientific questions about epidemics.

The point of our work so far in this chapter is this:

If we know **initial values** and **instantaneous rates of change**,

$$s[0] = s_0 \quad \frac{ds}{dt} = -a s i$$

$$i[0] = i_0 \quad \frac{di}{dt} = a s i - b i$$

then the computer program **SIRsolver** can give us the functions $(s[t], i[t])$ satisfying these two properties.

The computer solution can be used either numerically or graphically, but the functions $s[t]$ and $i[t]$ cannot be expressed as formulas (like high school functions such as $y = x^3$); we can only compute them by hand recursively (as in Exercise 2.2). (Explicit formulas are also important, and we will study their calculus later.) We want you to use the graphed computer solution together with the description of change given by the differential equations to analyse when an epidemic is increasing or declining (see Problem 2.4.1.)

Recall from Chapter 1 that $\frac{di}{dt} = \lim_{\Delta t \rightarrow 0} \frac{i[t+\Delta t] - i[t]}{\Delta t}$ represents the slope of the graph of i vs. t (although we used x and y there). This is because $i[t + \Delta t] - i[t]$ is the change in i during the time interval from t to $t + \Delta t$. If the slope is positive, the graph is rising, or i is increasing as t increases. Since the change law is $\frac{di}{dt} = a s i - b i$, this means that i is increasing when $a s i - b i$ is positive. The combination of graphics and this formula is a powerful tool.

Problem 2.4.1 EXPANDING AND DECLINING EPIDEMICS

We say an epidemic is expanding as long as the fraction of infectious people increases. It is declining when i decreases. See Figure 2.3 for the typical appearance of the graph of i vs. t in an epidemic. Where is i increasing on the graph? Where is it decreasing? When are the most people sick at one time? How many are sick at the maximum? At the point where i changes from increasing to decreasing, what can you say about the time rate of change of i , that is, the quantity $\frac{di}{dt}$? How is this related to the size of s , a , and b ? (Hint: What is the formula for $\frac{di}{dt}$?)

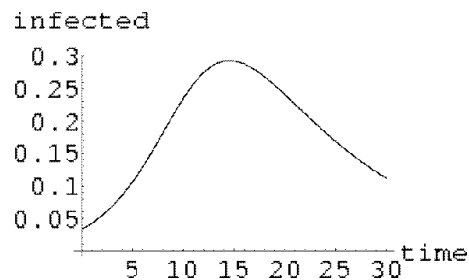


Figure 2.3: Infected vs. Time

2.5 Long-Term Change

Will anyone escape infection in our epidemic?

If some people never get sick, then as t grows larger and larger, $s[t]$ will NOT get closer and closer to zero, but rather will stay above the fraction of the population that does not get sick. This is because we are studying an S-I-R disease from which you can never become susceptible once you are infected.

Since $s[t]$ can only decrease, mathematically, we are asking if

$$\lim_{t \rightarrow \infty} s[t] = 0 \quad \text{or} \quad \lim_{t \rightarrow \infty} s[t] > 0$$

The limit as t tends toward infinity simply means the $s[t]$ approaches a value as t gets larger and larger. We denote this limiting value by

$$s[\infty] = \lim_{t \rightarrow \infty} s[t]$$

This is a difficult limit to compute analytically (see Exercise 3), but it turns out to be a key to a complete understanding of the S-I-R model.

Problem 2.5.1 HOW MANY ESCAPE?

1) As an epidemic fades out, the fraction of infectious people tends to zero,

$$\lim_{t \rightarrow \infty} i[t] = 0$$

Why is this correct? Can you give scientific reasons, if not mathematical ones? Verify that our model of rubella predicts this by running the **SIRsolver** program to large values of t .

2) Does the susceptible fraction tend to zero as an epidemic runs its course,

$$s[\infty] = \lim_{t \rightarrow \infty} s[t] = 0?$$

What sorts of intuitive arguments can you make pro or con? Run the **SIRsolver** program for large values of time and estimate $s[\infty]$ from the computer solution. (The graph of $s[t]$ for a 4 month rubella epidemic is shown in the Figure 2.4. Does $s[t] \rightarrow 0$ on that graph?)

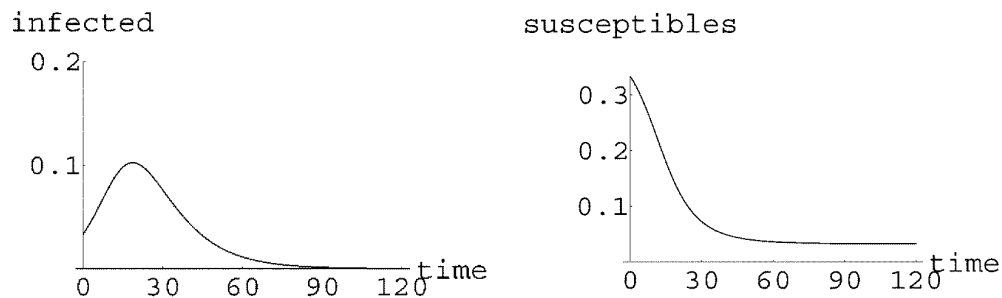


Figure 2.4: Four months of rubella

Problem 2.5.2 LIMITING SUSCEPTIBLES BY DISEASE

Experiment with $\lim_{t \rightarrow \infty} s[t]$ for rubella, measles, and influenza using the **SIRsolver** program. The necessary parameters, c and b , are given in Problem 2.3.1 above. The long-term graphs raise many questions such as, “How long does the epidemic last?” “Is there a formula for $s[\infty]$?” These are very difficult questions, but the computer will give you practical answers, if not formulas. Write a summary of your experiments on how $s[\infty]$ depends on c , b , and $s[0]$?

2.6 Calculus and the S-I-R Invariant

Calculus gives us a key insight into the epidemic model by finding an “invariant formula.” This formula gives us a way to measure the contact parameter c .

Up to this point, you should have followed all the mathematical derivations by working the exercises. With the help of the computer, you could even answer questions like

$$\text{Does } \lim_{t \rightarrow \infty} s[t] \text{ equal zero?}$$

At least approximately.

You might wonder, “Why bother with calculus if you have a computer?” The answer is that you can do more powerful calculations with both calculus and the computer than you can do with either one separately. We want to show you one example of calculus combined with computing, but we do NOT expect you to follow the details of the integral derivation at the end of this section. The derivation shows how explicit of high school functions ARE helpful, despite the fact that $s[t]$ cannot be given by a high school formula.

The derivation at the end of the section uses calculus to show that the S-I-R differential equations

$$\begin{aligned}\frac{ds}{dt} &= -a s i \\ \frac{di}{dt} &= a s i - b i\end{aligned}$$

imply that

$$i + s - \frac{1}{c} \text{Log}[s] = k, \quad \text{for the same constant } k \text{ and all } t$$

In other words, the combined quantity $i[t] + s[t] - \frac{1}{c} \text{Log}[s[t]]$ always has the same constant value, no matter what time t we take. (This is an example of a “first order invariant,” something like the energy in physics.) This is the key to measuring the parameter c in the model. (If we could not measure c by some experiment, the model would only be “descriptive” or have a “fudge factor.”)

Example 2.6.1 *Calculus Measures the Contact Number*

The invariant (in time) equation for s and i can be used to measure c from epidemic data. We have

$$s[\infty] - \frac{1}{c} \text{Log}[s[\infty]] = i[0] + s[0] - \frac{1}{c} \text{Log}[s[0]]$$

since $i[\infty] = 0$ and both sides equal the same constant k . At the start of an epidemic, $i[0]$ is small, so

$$s[\infty] - \frac{1}{c} \text{Log}[s[\infty]] \approx s[0] - \frac{1}{c} \text{Log}[s[0]]$$

and therefore approximately,

$$s[\infty] - s[0] = \frac{1}{c} (\text{Log}[s[\infty]] - \text{Log}[s[0]])$$

so

$$c \approx \frac{\text{Log}[s[\infty]] - \text{Log}[s[0]]}{s[\infty] - s[0]}$$

By measuring the susceptible fraction at the beginning and the end of an epidemic, we can compute c and then we can find a from the computation $a = b c$. Calculus gave us the exact invariant equation that s and i satisfy at all times. No substitute in recursive computer computations provides this equation. However, once we have this information from calculus, the computer again adds to our knowledge.

Example 2.6.2 *How Calculus Finds the Time-Invariant Equation (Optional)*

Without justifying why this works, we compute the ratio

$$\left(\frac{di}{dt}\right) / \left(\frac{ds}{dt}\right) = \frac{a s i - b i}{-a s i}$$

cancel dt , and do some algebra

$$\begin{aligned} \frac{di}{ds} &= \frac{b}{a} \frac{1}{s} - 1 = \frac{1}{c} - 1 \\ di &= \frac{1}{c} \frac{1}{s} ds - 1 ds \end{aligned}$$

We will learn later that if $y = s + k$ for any constant, then $\frac{dy}{ds} = 1$ and that if $\frac{dy}{ds} = \frac{1}{c} \cdot \frac{1}{s}$ for a constant c , then $y = \frac{1}{c} \text{Log}[s]$, the natural logarithm. These facts of calculus say that the indefinite integrals are as follows:

$$\begin{aligned} \int di &= \frac{1}{c} \int \frac{1}{s} ds - \int ds \\ i &= \frac{1}{c} \text{Log}[s] - s + k \end{aligned}$$

for some constant k . This is the computation of the important time invariant equation for the epidemic model in five short steps.

Exercise Set 2.61. *Math Nerd Fever*

At the beginning of the semester last year, 96% of the students were not interested in calculus, but word of this chapter spread around campus infecting the population with a mythical disease Math Nerd Fever. Victims of the disease find that they become interested in calculus - and stay that way, though they only are successful in spreading the infection for a brief period when they are first excited about the chapter. At the end of the last academic year, 44% of the population was interested in calculus. What is the contact number for Math Nerd Fever? On the average, how many new math nerds does each infected student create?

2. *Suppose we measure a and b in a college population for a new strain of flu. What limitations would there be in using these parameters in predicting an epidemic in a nursing home?*3. *Computation of $\lim_{t \rightarrow \infty} s[t]$*

(a) *In the rubella epidemic with $c = 6.8$, $s[0] = 1/3$, $i[0] = 1/30$, find k .*

(b) *We know from the previous section that $i[t]$ tends to zero as t tends to infinity, so the limiting susceptible fraction also satisfies*

$$0 + s[\infty] - \frac{1}{c} \text{Log}[s[\infty]] = k$$

with the same value of k as you computed in part (1) of the exercise. We could view this as the following question: Find all solutions in s of the equation

$$s - \frac{1}{c} \text{Log}[s] - k = 0$$

Use the computer program **EpidemicRoots** to do this. Notice from the graph that there are two roots. The smaller one is $s[\infty]$.

Later in the course, we will see analytical and graphical ways to understand the invariant equation more clearly. (See the phase diagrams of Chapter 22.)

2.7 Chapter Summary

Problem 2.7.1 THE BIG PICTURE

Write a paragraph explaining how the language of calculus is used to describe the changes in the susceptible, infectious, and removed fractions of a population. In other words, explain what the S-I-R differential equations “say” in English. Notice how cryptic the calculus is compared to English.

Write a paragraph summarizing the way the computer recursively finds the course of an epidemic by using the change rules given by the differential equations and an initial value of s , i , and r . (In other words, describe Euler’s Method and the way you completed the tables in Exercises 2.1.5 and 2.2.) Be brief and give only a sample calculation and the main idea, rather than a lot of details. How does the computer interact with calculus to find the whole graph of the epidemic from the starting values and the formulas describing change due to a particular disease?

Finally, explain why the same model applies to a number of diseases by using different parameters a , b , and c . Do these parameters vary as time changes?

2.8 Projects

Once or twice a semester we want you to explore a topic in detail. There is a separate book of detailed Scientific Projects that go with this text. The following subsections give a brief description of the projects most closely related to this chapter. You can work those projects now or come back to them later.

2.8.1 The New York Flu Epidemic

This project compares the S-I-R model with actual data of the Hong Kong Flu epidemic in 1968-69. The data we have to work with are “observed excess pneumonia-influenza deaths.” In this case, our “removed” class includes people who have died. This is a little gruesome, but it is difficult to find data on actual epidemics unless they are extreme.

2.8.2 Vaccination Strategies for Herd Immunity

This project uses the mathematical model of an S-I-R disease to find a prediction of how many people in a population must be vaccinated in order to prevent the spread of an epidemic. In this project, you use data from around the world to make predictions on successful vaccination strategies. Dreaded diseases like polio and smallpox have virtually been eliminated in the last two generations, yet measles is a persistent pest right here on campus. The mathematical model can shed important light on the differences among these diseases. The mathematics is actually rather easy once you understand the basic concept of decreasing infectives.

2.8.3 Endemic S-I-S Diseases

Some infectious diseases do not confer immunity, such as strep throat, meningitis, or gonorrhea. In these diseases, there is no removed class, only susceptibles and infectives. As in the S-I-R diseases, we can make a mathematical model for S-I-S diseases.

An S-I-S disease has the potential to become endemic, that is, to approach a non-zero limit in the fractions of susceptible and infectious people. Once you have done some experiments, you will be able to show what happens mathematically and explore health policy implications.