Chapter 1

What is Calculus? The Six Pillars

Calculus was invented by Isaac Newton and Gottfried Leibniz in the mid-17th century to solve real problems: To understand the orbits of planets. To understand what happens when an apple falls from a tree. To design lenses and clocks and all sorts of machines. Even to make money betting at dice. Over time, people realized how much calculus had to say about almost *everything* in the world. If it's real and it's changing, we can understand it better with calculus.

But as a 21st-century student learning calculus, it's easy to get lost in a maze of formulas, computational techniques, obscure theorems, and contrived problems. You can lose sight of what it's all for.

So what is calculus, anyway? A good working definition is

Calculus is the study of things that change.

- If a plant absorbs solar energy at a rate of 1000 watts, and if the sun is up for 12 hours each day, then the plant absorbs

$$1000 \text{ watts} \times 12 \text{ hours} = 12,000 \text{ watt-hours} = 12 \text{ kilowatt-hours}$$

  of energy each day. That's simple arithmetic. But if the amount of solar radiation changes over the course of the day, depending on the angle of the sun, then figuring out the plant's total energy intake involves adding up the *variable* intakes for the different hours of the day. That's calculus.

- With every beat, your heart delivers oxygen-rich blood to your organs and muscles. You need more oxygen when you exercise, so your heart beats faster. However, your heart is also less efficient when it beats faster, so if it beats too fast it will deliver less oxygen. At what heart rate do your muscles get the most oxygen? That's calculus.

- If a lake has 1000 fish, and if the fish population is growing at 100 fish/year, then it is easy to predict how many fish there will be next year (1100), the year after (1200) or the year after that (1300). But what if population growth slows down as the lake becomes more crowded? Understanding how that changes the answer requires calculus.

Fortunately, calculus isn't black magic. Almost everything we do boils down to six simple ideas that we call the **Six Pillars of Calculus**:

(1) **Close is good enough.** In algebra, we look for exact answers to problems. If somebody asks you to solve $x^2-3x+2=0$, you might factor that as $(x-2)(x-1)=0$, or you might apply the quadratic formula, or you might just try different numbers and see that $x=1$ and $x=2$ work. What you probably don't say is "There must be a solution somewhere between 0.9 and 1.1 ..., make that 0.99 and 1.01 ..., no, make that 0.999 and 1.001." But that's exactly the sort of reasoning we use in calculus! When faced with problems that can't be solved right away, we simplify things and look for an approximate answer. Then we try to improve that answer, and improve it some more, until we get something that is fairly accurate. If the calculations are too grungy to do by hand (and they often are), we program a computer to do our dirty work for us. Eventually, we get an answer that's good enough for whatever real-world task we have in mind. Then we stop and congratulate ourselves on a job well done.

Once in a while, we need the exact answer. When that happens, we approach that exact answer as the **limit** of better and better approximations. Almost all of the important formulas of calculus come from this limiting process.

(2) **Track the changes.** We can often tell more about something by looking at the way that it's changing than by asking where it is right now.

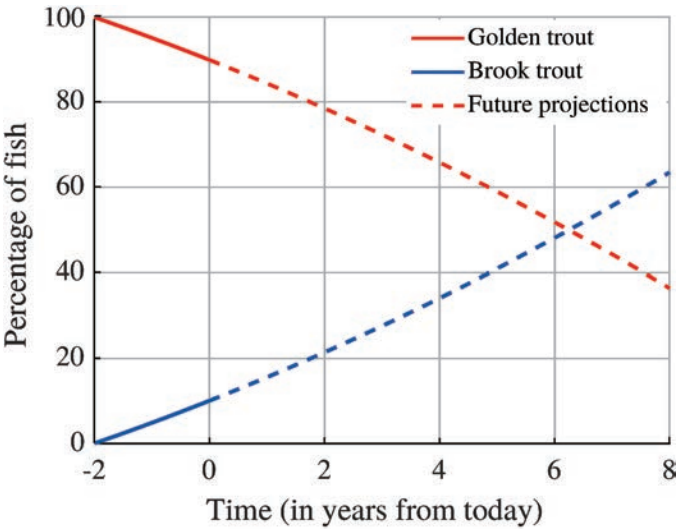


Figure 1.1. Which subspecies of trout is winning?

The Californian golden trout is the state freshwater fish of California. However, it is endangered due to competition from invasive species like brook trout that were introduced (either accidentally or on purpose) into its habitat. Suppose that a stream has 90% golden trout and 10% brook trout, but that 2 years ago it was almost all golden trout and that last year it had 95% golden trout and 5% brook trout, as shown in Figure 1.1. Which subspecies of trout is doing better?

If we only looked at current numbers, we might think that the golden trout were doing fine, as they outnumber the brook trout almost 10:1. However, if we look at the rate of change, we see that the brook trout population has doubled in the past year, while the golden trout population has declined. Within 10 or 20 years, the brook trout are likely to drive the local golden trout population to extinction.

Graphically, the rate of change of a quantity is closely related to the **slope** of the graph of that quantity. We're going to spend a lot of time studying slopes—both how to find them and what to do with them. In mathematical terms, this is called computing a **derivative**.

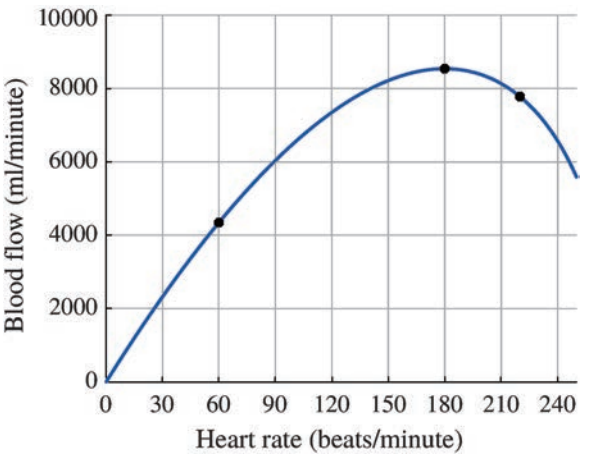


Figure 1.2. A schematic graph of blood flow vs. heart rate. The overall shape and scale of the graph are realistic, but the details are not medically accurate.

(3) **What goes up has to stop before it comes down.** The amount of blood that a typical 40-year-old person's heart pumps as a function of how fast it beats follows a curve similar to that shown in Figure 1.2. If it is currently beating at 60 beats per minute (bpm), then it can pump more blood by pumping faster, which is what happens when the person starts exercising. This part of the curve slopes upwards, so more beats mean more blood. However, if the heart is currently beating at 220 bpm, then it needs to slow down.[1] This part of the curve slopes downwards, so more beats means less blood. Peak blood flow is at 180 bpm, where the curve is flat and its slope is zero. By studying the rate of change of the curve and figuring

[1]Heart rates that high are very dangerous and can trigger fibrillation and cardiac arrest. The top of the curve is considered "maximum heart rate" (MHR) and athletes are generally advised to keep their heart rates below 85% of MHR.

out where that rate of change is zero, we can determine how much our hearts can (and can't) do.

(4) **The whole is the sum of the parts.** How much carbon dioxide (CO_2) will be in the atmosphere next year? That's a complicated question, because CO_2 enters and leaves the atmosphere in many ways, including plants converting CO_2 into sugar, that sugar decomposing back into CO_2 when the plant is eaten or rots, volcanic activity, and of course the burning of fossil fuels. By adding up all of these processes, we can find the net emission for each year: the amount of CO_2 put into the atmosphere minus the amount taken out.

Let's break things down year by year. The amount of CO_2 in the atmosphere next year is the amount there is now, plus this year's net emission. Similarly, this year's CO_2 is last year's CO_2 plus last year's net emission. Working backwards, year by year, we see that the CO_2 content of the atmosphere this year is the content a million years ago, plus the sum of all the net emissions in the last million years (roughly half of which have been added since 1950).

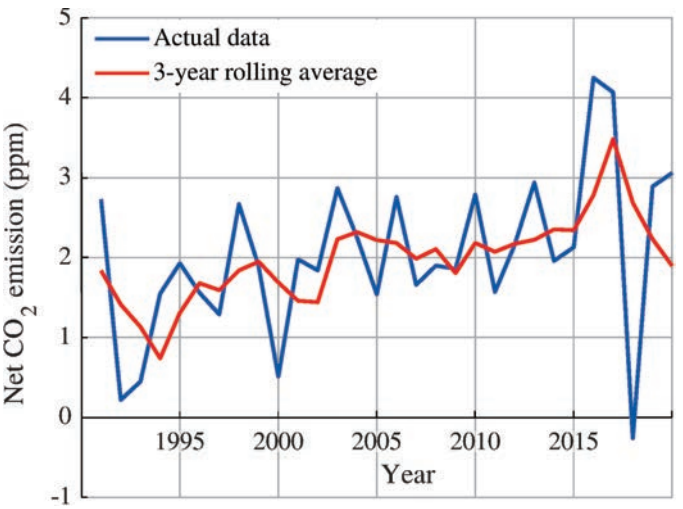


Figure 1.3. Global net CO_2 emissions by year, 1991–2020

If we plot the net emissions as a function of time, as in Figure 1.3, the sum of all those values is the same as the area under the curve. We're going to spend a lot of time talking about area, but it isn't because we're obsessed with geometry. It's because most bulk quantities, like the total power used by the city of Austin in 2007 or the total amount of water carried by the Colorado River in 2011, can be treated exactly like CO_2 concentration or area. Break the quantity you're studying into little pieces, estimate each piece, and add up the pieces. This process is called **integration**.

Note that this is very different from how integration is often taught. In most calculus classes, students are taught that integration is in some sense the opposite of differentiation. That isn't exactly wrong, as we'll see with the 5th Pillar, but it

misses the point. Integration is about adding up the pieces, which is why it applies to a host of problems. Anti-derivatives are a great tool for actually computing integrals, but they don't explain why such-and-such quantity is represented by such-and-such integral. For that, we need the 4th Pillar.

(5) **One step at a time.** A famous Chinese proverb says, "A journey of a thousand miles begins with a single step". The journey then continues with about two million additional steps. By understanding what happens at each step, we can understand the entire journey.

Instead of asking what the concentration of CO_2 *is*, we can ask how much it *changed* in a short period of time, like a year. Geometrically, the change per year is the slope of the concentration curve, shown in Figure 1.4.[2] The steeper the curve is, the faster the concentration is changing.

In the 1990s, net emissions were typically between 1 and 2 parts per million (ppm) per year, while in the 2010s they were typically between 2 and 3 ppm per year. You can see that the concentration curve is steeper (has a higher slope) in the 2010s than in the 1990s.

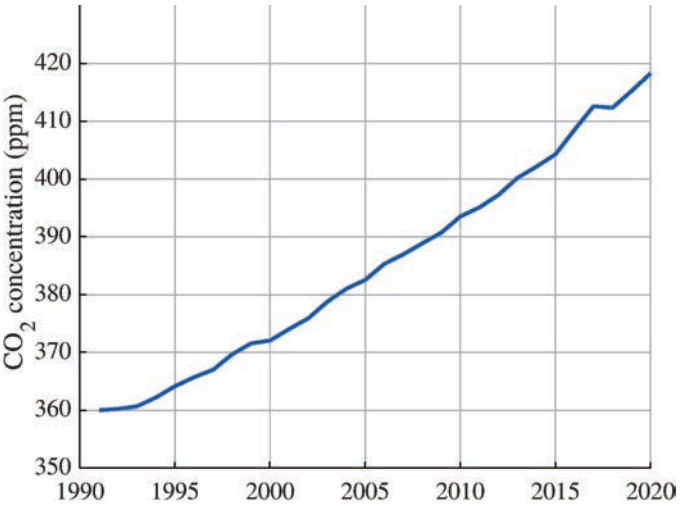


Figure 1.4. Peak concentration of CO_2 by year, 1990–2020

How are Figures 1.3 and 1.4 related? The new emissions in one year is the same as the change in the concentration from that year to the next. This means that the concentration is a running total of the net emissions and corresponds to the **area** under the net emission curve. The net emissions is the rate at which the concentration is changing and corresponds to the **slope** of the concentration curve. Mathematically, the concentration is the **integral** of the net emission and the net emission is the **derivative** of the concentration. The **Fundamental Theorem of Calculus** (FTC) relates derivatives and integrals in general. With it, we

[2]The concentration of CO_2 varies by season. The graph shows the peak value in May of each year. On May 8, 2018, the detectors measured a concentration of 418.74, more than 6 ppm higher than on any other day that year. That outlier is not included in the graph.

can use what we already know about derivatives to understand integrals. In fact, it is the single most powerful tool we have for evaluating integrals.

The FTC comes from thinking about changes one step at a time, but there is more to "one step at a time" than just the fundamental theorem. For instance, suppose that a herd of wild animals, currently containing 1000 individuals, is growing at 6% per year. How many animals will there be in 30 years? Using our initial population and the growth law, we can estimate the net births minus deaths in the first year, and from that we can compute the population a year from now. From that, we can estimate the net births minus deaths in the second year and get our population two years from now. Continuing the process, step by step, we can accurately project the population of the herd far into the future.

Finally, "one step at a time" is great slogan for how to approach calculus in general. Many problems are way too complicated to be done all at once. By breaking big problems into sequences of smaller problems, and by solving these smaller problems one at a time, we can accomplish wonders.

(6) **One variable at a time.** Many functions involve two or more input variables. The boiling point of water depends on our elevation and on how much salt we put in the water. In the summer, the heat index depends on temperature and humidity. In the winter, the wind chill factor depends on temperature and wind speed. The price of widgets depends on supply and demand. The value of an oil field depends on how much oil it produces and on the price of oil.

To understand functions of two (or more) variables, we always hold everything but one variable fixed and study just that variable. Asking how a change in temperature changes the heat index is a question about a function of just one variable that we already know how to answer. Similarly, we can figure out how changing the humidity affects the heat index. Putting the two answers together, we can understand how changing both temperature and humidity affects the heat index.

It's tempting to say "that's all there is!", but that isn't really true. Over the centuries, lots of really smart people have cooked up lots of really smart ways to solve lots of really hard problems. Now, with the help of computers, we can solve even more problems. In the next ten chapters, we're going to follow in the footsteps of these masters and learn some of their results. Yes, there will be formulas to memorize and techniques to practice and algorithms to implement on computers. And yes, it will take work to really absorb everything.

But hopefully this voyage of discovery won't be a mystery. Every new formula or technique or algorithm will lead straight back to the Six Pillars, so it will be connected to every other formula or technique or algorithm. If you remember the pillars, you'll have a framework for organizing all the little details. Over time, you'll forget many of those details, and that's OK. As long as you remember the pillars and are willing to look up the details as needed, you'll be able to use calculus for your whole life, not just in your classes and in your job, but in understanding the wild and complicated world we live in.

Chapter 2

Predicting the Future: The SIR Model

In December 2019, a strange form of pneumonia started affecting residents of Wuhan, China. The virus causing this disease was quickly identified and the disease was named COVID-19, for "**CO**rona**VI**rus **D**isease 20**19**".[1] By April 2020, Covid-19 had spread around the world and had mutated into a more transmissible form, leading to drastic lockdowns to control the spread. Schools closed or went online, most people were told to stay home except for essential errands like buying food, and everybody was told to practice "social distancing". Despite these measures, Covid-19 continued to spread, and by mid-2022 had killed at least 6 million people worldwide (with some estimates being twice that number) including over a million Americans. By the time you read this, the death toll will almost certainly be even higher. Throughout the pandemic, everybody wanted to know what would happen next, which kept mathematical modelers (including the author) very busy.

In this chapter we will use four of the Pillars of Calculus to get a handle on the spread of disease.

- It's impossible to keep track of the behavior of every person in the world, so we devise a simplified model to describe average behavior. *Close is good enough!*
- This involves understanding the rate at which people become sick and the rate at which they recover. *Track the changes!*
- Once we understand where things stand and how fast they are changing, we can make realistic predictions about what is likely to happen in the future, as well as what we can do to improve that trajectory. Our projections are only accurate for a short time, so we combine a lot of short-term projections to get a long-term projection. *One step at a time!*

[1] After a while, people dropped the "19", stopped capitalizing the whole word, and just wrote "Covid" or "covid".

- Even without these long-term projections, we can understand how the epidemic will peak by comparing the factors that increase infection with those that decrease infection. *What goes up has to stop before it comes down!*

We will begin with a fairly simple model, called the SIR model, that works very well with diseases ranging from ebola to measles to the common cold. We'll see how to set up the model, well see how to use the model to give quantitative predictions of how an epidemic will run, and we'll see how to analyze the model theoretically. Covid has features that make it slightly harder to model, but we'll return to Covid at the end of the chapter and see how modelers tweaked the SIR model to understand the pandemic.

2.1. Worried Sick

Suppose that you are the principal of a school where a few students have just come down with measles. You've heard that measles is very contagious and you're worried about the disease spreading. Do you need to shut the school down? Or is it enough just to tell parents to keep their sick kids at home? In order to make sense of the situation, you will need to run different scenarios and forecast the number of sick children for the next month or two. You hire a mathematical consultant (your Friendly Author), and together we attack the problem.

The first thing we do is define our quantities. Let t denote time, measured in days, and let $I(t)$ be the number of infected children at time t. We call t and I **variables** because they change. The time t is our **input variable**, while I is an **output variable**. The graph of I as a function of t will look something like the plot in Figure 2.1. There is an initial phase where more and more people get sick, a moment of peak infection, and a long tail as everybody recovers. We want the rise to be as slow as possible, the peak to be as low as possible, and the decline to be as fast as possible.

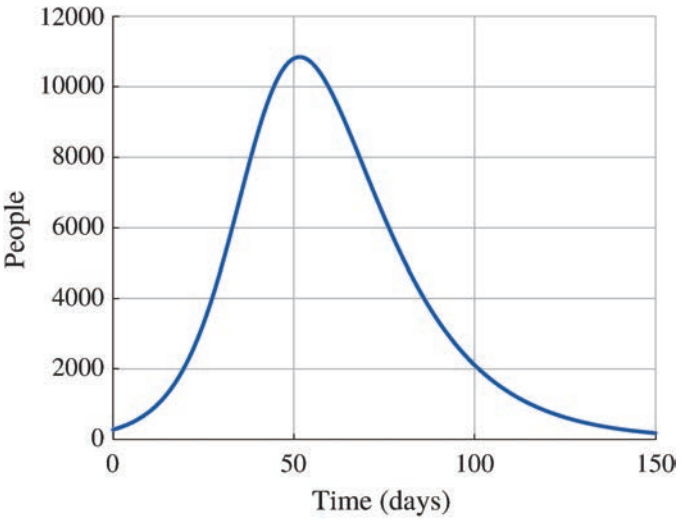


Figure 2.1. Infected children as a function of time

Before we try to control I, we need to understand what makes I change over time. That is, we need to make a mathematical **model** for what is going on. The model needs to take into account all the *important* features of what's happening in the world, while being simple enough to be solvable. We are going to ignore a *lot* of details, because *close is good enough!*

Once we have our model, we have to **analyze** it. In this step, we don't care where our equations came from. We just have to solve them. Maybe we can find a formula for the answer. More likely, we can't find a formula, but we can run the model on a computer to generate accurate predictions. Once again, *close is good enough!*

Finally, we need to **interpret** our results. Math can tell us that such-and-such variable will have such-and-such value at such-and-such time, but we need to understand biology to say what that means for the health of our school.

In other words, predicting the future is a three-step process:

(1) **Model** our system mathematically. Define appropriate variables and write down some equations that describe how these variables change with time. This step requires real-world understanding as well as math.

(2) **Solve** the model to determine what each variable will be at some future time. This step is 100% math, using techniques that we are about to develop.

(3) **Interpret** the results. Take our mathematical results and make real-world sense of them.

We'll tackle these one at a time, which is a lesson in itself. *One step at a time.* We need to break hard problems into bite-sized tasks, and then do each task in turn.

2.2. Building the SIR Model

To understand how germs spread, we divide our population of students (and maybe teachers and staff) into three groups, as in Figure 2.2:

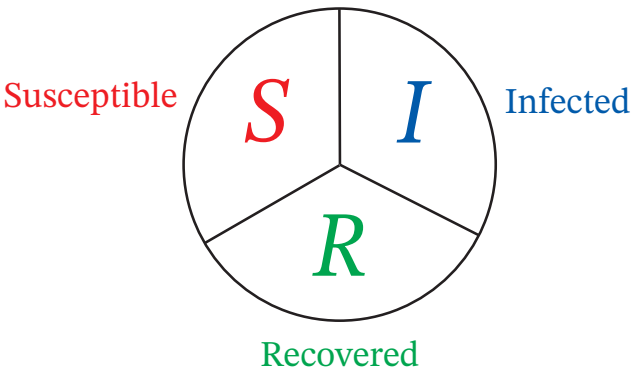


Figure 2.2. Three categories of people

- **Susceptible** people, or *Susceptibles*. These are people who might get sick if they're exposed to the disease.
- **Infected** people, or *Infecteds*: These are people who have already gotten sick. Not only are they sick, but they are spreading germs to everybody they meet.
- **Recovered** people, or *Recovereds*: These are people who can't get sick. Maybe they used to be sick and then recovered. Maybe they were vaccinated. Maybe they have some natural immunity that protects them. For whatever reason, they're no longer in the game.

Next we define our variables. Let $S(t)$ be the number of Susceptibles at time t, let $I(t)$ be the number of Infecteds at time t, and let $R(t)$ be the number of Recovereds at time t. Obviously, there are differences among the people in each category, but we're going to ignore those differences. This is a model, not a complete description of reality! Instead, we're going to talk about the behavior of the *average* Susceptible, the *average* Infected, and the *average* Recovered. We then ask two questions:

(1) At what rate do Susceptibles get exposed and become Infecteds?

(2) At what rate do Infecteds clear their bodies of germs (but not antibodies!) and become Recovereds?

From these rates, we can figure out the rates S', I', and R' at which the quantities S, I, and R are changing, and from those we can predict the future.

Losing Patients Through Recovery. Diseases typically don't last forever. Measles lasts about two weeks.[2] Of course the duration can vary from one person to the next, but because it takes an average of 14 days to recover from measles, it's a good approximation to say that about 1/14 of all the measles patients (Infecteds) will recover each day. That is,

$$\text{today's change in the Recovered population} \approx \frac{I(\text{today})}{14}. \tag{2.1}$$

Likewise,

$$\text{tomorrow's change in the Recovered population} \approx \frac{I(\text{tomorrow})}{14}, \tag{2.2}$$

and in general

$$\text{the change in the Recovered population on day } t \approx \frac{I(t)}{14}. \tag{2.3}$$

Note the units in this equation. $I(t)$ and $R(t)$ are numbers of *people*, and the number of new R's on any given day is also measured in people. But the *rate* R' at which R is changing is measured in people/day. That is,

$$R'(t) = \frac{I(t)}{14 \text{ days}}.$$

The numerator has units of people, the denominator has units of days, and the ratio has units of people/day.

[2] There are several stages to measles, including incubation, a period of high fever, and a period including the telltale rash, all spread over a total of two to three weeks. Because the period when you can infect others lasts about two weeks, we're going to simplify things and just say that the disease lasts two weeks.

This is an example of a **rate equation**. A rate equation describes the rate at which something is changing in terms of other data. In this case, it describes R' in terms of I.

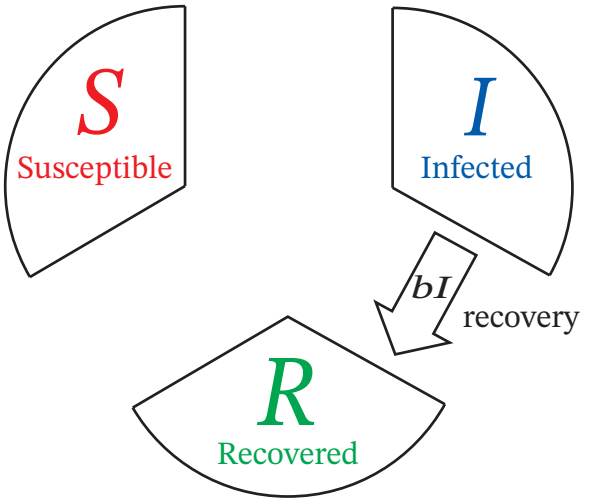


Figure 2.3. With the passage of time, Infecteds become Recovereds.

More generally, if T is the average length of time that people stay sick, then we expect

$$R' = \frac{I}{T}. \tag{2.4}$$

If we define $b = 1/T$, then we can write our rate equation without fractions;

$$R' = bI; \tag{2.5}$$

see Figure 2.3. R' has units of people/day, I has units of people, and b has units of (1/day)s. The parameter b is called the **recovery coefficient**, because it says how quickly people recover.

The number b is called a **parameter**. It isn't a variable, because it doesn't change during the epidemic. However, it may take on different values for different diseases or different populations.[3] By adjusting the value of b, we can use the same rate equation to model ebola, the common cold, seasonal flu, and many other diseases as well.

Gaining patients: Transmission. Recovery is good. If a city has 21,000 people sick with measles, then people recover at a rate of

$$\frac{\text{21,000 people}}{\text{14 days}} = \frac{\text{1500 people}}{\text{day}}. \tag{2.6}$$

As long as fewer than 1500 people get sick each day, the epidemic will subside.

To understand transmission, let's look at things from the perspective of a single Susceptible, who we'll call Joe. Joe will come in contact with a certain fraction p of

[3] For instance, measles is more dangerous in adults over 20 and in children under 5 than in teenagers or children over 5. The average duration depends on the age of the patient.

the Infecteds each day, for a total of pI contacts. The fraction p is typically very small, but the Infected population I can be very large. The more Infecteds there are, the more chances there are of Joe being exposed to the disease. Every time that Joe comes in contact with an Infected, there is a probability q that he will breathe in enough germs to become sick himself. That is, there is a probability pqI per day of Joe becoming infected.

We don't actually care about p and q separately, as all that matters is their product. We define $a = pq$, and call a the **transmission coefficient**. Like b, this is a parameter, not a variable. Different diseases in different communities will have different values of a, but we can use the same reasoning for all of them.

Finally, we look at the entire population of Susceptibles, and not just at Joe. If there are S Susceptibles, each of whom has a probability aI (per day) of becoming an Infected, then each day there will be approximately aSI Susceptibles who become Infecteds. That is,

$$S' = -aSI; \tag{2.7}$$

see Figure 2.4. S and I have units of people and S' has units of people/day, so a must have units of 1/(people $\times$ days). Also note the minus sign in our equation! The more Susceptibles become Infecteds, the fewer Susceptibles are left.

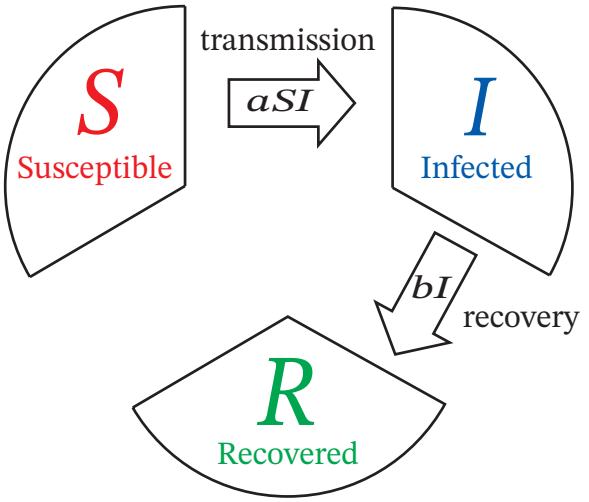


Figure 2.4. Through contact with Infecteds, Susceptibles become Infected.

What about I? Completing the model. So far we have figured out how the numbers of Recovereds and Susceptibles change, but what about the number of Infecteds? We have

$$\begin{aligned} I' &= +(\text{Rate at which people get sick}) \\ &\quad -(\text{Rate at which they recover}) \\ &= aSI - bI. \end{aligned} \tag{2.8}$$

Another way to see this is that $S + I + R$ is the total number of people out there, and that doesn't change. Because $S' + I' + R' = 0$, we must have $I' = -S' - R'$.

Putting everything together, we have a system of three rate equations. Together, they're called the **SIR model**:

$$\begin{aligned} S'(t) &= -aS(t)I(t), \\ I'(t) &= aS(t)I(t) - bI(t), \\ R'(t) &= bI(t), \end{aligned} \tag{2.9}$$

where

- t is the time. Depending on the setting, you may want to measure t in days, weeks, months, or years.
- $S(t)$ is the number of Susceptibles at time t, measured in people.
- $I(t)$ is the number of Infecteds at time t, measured in people.
- $R(t)$ is the number of Recovereds at time t, measured in people.
- b is the recovery coefficient, which is the reciprocal of the average time T that the disease lasts. The units of b are the reciprocal of whatever units we are using for time. This parameter can vary from disease to disease.
- a is the transmission coefficient. This can depend both on the disease and on the community. If there are outbreaks in several different places, b will be more or less the same in each place, but a will typically be bigger in smaller communities (where each Susceptible knows a greater fraction of the Infecteds) and smaller in larger communities. The units of a are 1/(time × people).

2.3. Analyzing the Model Numerically

Now that we have our model, let's use it to predict the future. Suppose that a disease that lasts an average of 15 days is spreading across a college campus of 50,000 people. We measure t in days, we suppose that $a = 0.000004$/person-day and $b = \frac{1}{15}$/day, and we start with $S(0) = 40,000$, $I(0) = 2100$, and $R(0) = 7900$. What will S, I, and R be in two days? In five days? In ten days? For that matter, what were the values yesterday? A week ago? A month ago?

A naive approach is to use the SIR equations to compute S', I', and R' once and for all:

$$\begin{aligned} S' &= -0.000004(40,000)(2100) &&= -336 \text{ people/day}, \\ I' &= 0.000004(40,000)(2100) - 2100/15 &&= 196 \text{ people/day}, \\ R' &= 2100/15 &&= 140 \text{ people/day}. \end{aligned} \tag{2.10}$$

If we have 2100 Infecteds on day 0, and if I is growing at a rate of $I' = 196$ people/day, then we should expect $2100 + 196 = 2296$ Infecteds tomorrow. We should expect $2296 + 196 = 2100 + 2(196) = 2492$ Infecteds the day after tomorrow. After a week, we should expect $2100 + 7(196) = 3464$ Infecteds, after a month we should have

$2100 + 30(196) = 7980$ Infecteds, and after a year we should have $2100 + 365(196) = 73,640$ Infecteds. In general, after t days we should have

$$I(t) \approx I(0) + I'(0)t = 2100 + 196t. \tag{2.11}$$

This is called a **linear approximation**. Instead of finding the exact equation for $I(t)$, we found the equation of the line that has the right value and the right slope at $t = 0$. We then approximate the value of the true $I(t)$ function by the value of the linear function $2100 + 196t$.

Before we move on, let's recall some basic facts about things that change at a constant rate. If we are driving at a constant speed of 57 miles per hour and pass milepost 253 at 3:00, where will we be at 4:00? At 5:00? At time t?

From 3:00 to time t is $t - 3$ hours. As rate $\times$ time = distance, and as we are going at 57 mph, we will travel $57(t - 3)$ miles in that time. Adding that to our starting point at milepost 253, we will find ourselves at milepost

$$x(t) = 253 + 57(t - 3) \tag{2.12}$$

at time t. (At least if we're traveling in the direction where the mile markers are increasing. Otherwise we might find ourselves at milepost $253 - 57(t - 3)$.)

The same idea works for any quantity that is changing at a constant rate, not just for position. If a quantity Q is changing at rate Q', and if Q starts at a value Q_0 at time t_0, then what will Q be at time t? Because it grows at rate Q' for time $(t - t_0)$, it will increase by $Q' \times (t - t_0)$. Adding that to our starting value of Q gives

$$Q(t) = Q_0 + Q' \times (t - t_0). \tag{2.13}$$

This is the equation of a straight line in **point-slope form**. The linear approximation (2.11) is a special case of this, with I instead of Q, and with starting time $t_0 = 0$. We'll have a lot more to say about equations of lines and linear approximations in Chapter 3.

Returning to the SIR model, we can use the linear approximation (2.11) to study the past as well as the future. According to this approximation, yesterday we had around $2100 - 196 = 1904$ Infecteds, a week ago we had $2100 - 7(196) = 728$ Infecteds, and a month ago we had $2100 - 30(196) = -3780$ Infecteds.

How much do you trust those numbers? You should take them with a grain (or more) of salt. In particular, the estimate $I(-30) \approx -3780$ is absurd. You can't have a negative number of Infected people! You also can't have $I(365) = 73,640$, since that's more than the entire population.

In making those estimates, something went seriously wrong. You should get into the habit of always asking yourself whether answers make sense. Reality check! Before reading on, take a minute to think about what went wrong with the methods we used to get our negative answer and our overly positive answer.

The problem with our linear approximation is that it assumes that the rate of change of I is 196 people/day, always was 196 people/day, and always will be 196 people/day. In reality, we know that $I' = 196$ people/day *today*, and it's realistic to assume that I' won't change much in the next few days, but in a few weeks, or a few months,

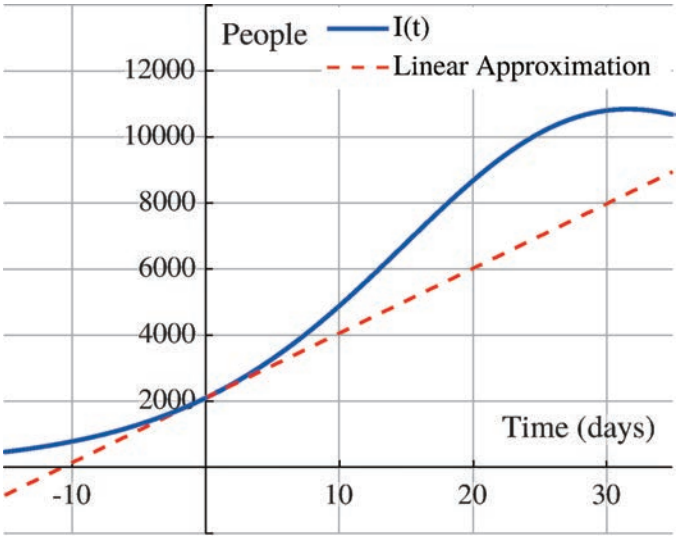


Figure 2.5. The linear approximation is negative when $t < -10.75$, but the actual $I(t)$ isn't.

it could change by a lot. This means that we can trust our linear approximation when $t = 1$ or $t = 2$ or $t = -1$ or $t = -2$, but we shouldn't trust it when $t = 30$ or 365 or -30. The actual situation is shown in Figure 2.5. The linear approximation tracks $I(t)$ very closely for $-2 < t < 2$, but it does not account for the curvature in the graph of $I(t)$. As t gets more and more negative, the graph of $I(t)$ flattens out and stays positive, while the linear approximation goes negative. Meanwhile, as t gets more and more positive, $I(t)$ goes up and then down, while the linear approximation just keeps on growing.

To do better than our naive approximation, we need a way of estimating S' and I' and R' not just right now, but also in the future and the past. Here's a way to estimate $S(10)$, $I(10)$, and $R(10)$.

First, we need to decide how far we can trust our linear approximation. Since I seems to be changing at about 10% each day, trusting it for a couple of days is reasonable. We're going to pick a time interval of two days, which we'll call Δt days, and predict the future $\Delta t = 2$ days at a time.

(1) Use the initial values $S(0)$, $I(0)$, and $R(0)$, together with the SIR equations, to estimate $S'(0)$, $I'(0)$, and $R'(0)$.

(2) Use a linear approximation to estimate $S(2) \approx S(0) + 2S'(0)$, as well as $I(2) \approx I(0) + 2I'(0)$ and $R(2) \approx R(0) + 2R'(0)$.

(3) Use the values of $S(2)$, $I(2)$, and $R(2)$, together with the SIR equations, to estimate $S'(2)$, $I'(2)$, and $R'(2)$.

(4) Use a linear approximation to estimate $S(4) \approx S(2) + 2S'(2)$, etc.

(5) Use those values and the SIR equations to approximate $S'(4)$, etc.

(6) Lather, rinse, repeat. At each time t, plug the estimated values of $S(t)$, $I(t)$, and $R(t)$ into the SIR equations to get estimated values of $S'(t)$, $I'(t)$, and $R'(t)$. Then

Table 2.1. Projecting forward with $\Delta t = 2$

t	$S(t)$	$I(t)$	$R(t)$	$S'(t)$	$I'(t)$	$R'(t)$
0	40,000	2,100	7,900	-336	196	140
2	39,328	2,492	8,180	-392	226	166
4	38,544	2,944	8,512	-454	257.5	196.5
6	37,636	3,459	8,905	-520.5	290	230.5
8	36,595	4,039	9,366	-591.5	322	269.5
10	35,412	4,683	9,905			

use a linear approximation to compute $S(t + \Delta t) \approx S(t) + \Delta t\, S'(t)$, etc. Table 2.1 shows the results for the first ten days.

Likewise, we can go backward in time, using a small negative time step. We can take $\Delta t = -2$ days and go from $t = 0$ to $t = -2$ to $t = -4$, etc., all the way back to to $t = -30$. The results are shown in Table 2.2, with some times skipped and the values of S', I', and R' omitted to save space. As you can see, $I(-30)$ is *not* negative.

Table 2.2. Projecting backward with $\Delta t = -2$

t	$S(t)$	$I(t)$	$R(t)$
0	40,000	2,100	7,900
-6	41,683	1,109	7,208
-12	42,591	562	6,848
-18	43,055	278	6,667
-24	43,286	136	6,577
-30	43,400	66	6,534

This method still doesn't give exact answers, but *close is good enough!* If we want to compute $I(10)$ with greater accuracy, we can use the same algorithm with $\Delta t = 1$ instead of $\Delta t = 2$, for an answer of $I(10) = 4783$ instead of 4683. That requires ten iterations instead of five. If we want even more accuracy, we can take $\Delta t = 0.1$ and do 100 iterations, getting $I(10) = 4876$, or take $\Delta t = 0.01$ and do 1000 iterations, getting $I(10) = 4886$. If we're willing to do the extra work, or if we program a computer to do the extra work for us, we can have as much accuracy as we want. (However, our model is only an approximation of the real world, so even if we can solve our model to great accuracy, that doesn't necessarily mean that we can predict the future with that much accuracy.)

This is the 5th Pillar of Calculus: *One step at a time.* Every big change is made up of many little changes. If we can understand each little change, we can put the pieces together to understand the big change.

2.4. Theoretical Analysis: What Goes Up Has to Stop Before It Comes Down

So far we have used the 1st, 2nd, and 5th Pillars of Calculus. By making approximations and tracking the changes in our variables S, I, and R and by putting a lot of short-term linear approximations together, we figured out how to obtain good projections of the future, and we were able to use those same projections to understand the past. Now we're going to tackle the question:

What is happening when $I(t)$ reaches its peak?

The key fact is that *the sign of I' tells you whether I is increasing or decreasing.* Whenever $I'(t) > 0$, $I(t)$ must be increasing. Whenever $I'(t) < 0$, I must be decreasing. At the very top of the curve, when $I(t)$ has stopped increasing and hasn't yet started decreasing, $I'(t)$ transitions from positive to negative. At that instant of time, we must have $I'(t) = 0$.

Let's figure out what is happening at that time. The SIR equations tell us that

$$I'(t) = aS(t)I(t) - bI(t) = I(t)\,(aS(t) - b)\,. \tag{2.14}$$

Since $I(t)$ is always positive, the sign of $I'(t)$ is the same as the sign of $aS(t) - b$. As long as $S(t) > b/a$, $I'(t)$ will be positive and $I(t)$ will increase. But eventually we will run out of Susceptibles. When $S(t)$ drops below b/a, $I'(t)$ will become negative and $I(t)$ will start to decrease.

The number b/a is called a **threshold**. In our example, the threshold was 16,667. This is *not* the value of $I(t)$ at the peak. Rather, it is the value of $S(t)$ when $I(t)$ hits its peak. In Figure 2.6, $I(t)$ hits its peak value of 10,870 when $t = 31.6$, which is when $S(t)$ passes through 16, 667.

If the threshold b/a is large and there aren't many Susceptibles to begin with, the outbreak will fizzle from the start. If b/a is small and $S(0)$ is large, there will be a

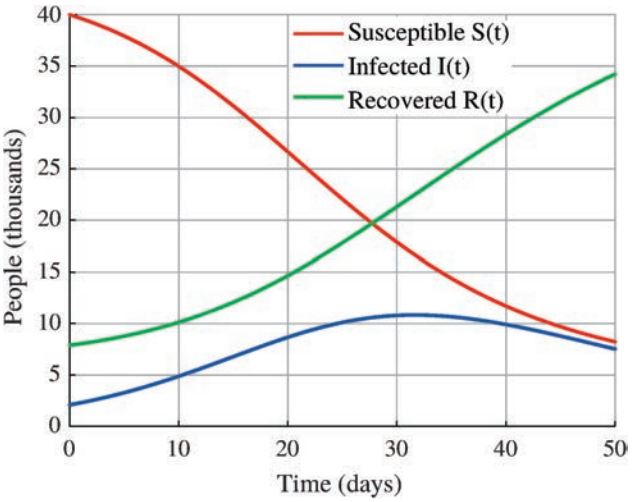


Figure 2.6. Over time, the number of Susceptibles drops, the number of Recovereds rises, and the number of Infecteds rises, reaches a peak, and then falls.

long period of growth before the epidemic peaks. As a public health consultant, our considerations for controlling the outbreak are the following:

(1) Make b as big as possible. Equivalently, make the duration $T = 1/b$ of the disease as small as possible. The less time that you're sick, the less time you have to infect others. Unfortunately, this is usually very difficult, especially with viral diseases. For instance, anti-viral drugs can sometimes reduce the length of the flu, but only by a small amount. A better approach is through testing and quarantining. Even if somebody hasn't *recovered* yet, we can try to *remove* them from the general population until they stop being infectious. That doesn't help the patient any, but when it comes to stopping an epidemic, removal is as good as recovery. For this reason, many people like to say that R stands for Removed rather than Recovered.

(2) Make a as small as possible. Do everything possible to minimize contact between Infecteds and Susceptibles. Stay home if you're sick. Wash your hands. Wear a mask. Use hand sanitizer. If the outbreak is severe, public health officials may need to take more drastic measures, like closing schools, canceling large public events, or issuing shelter-in-place orders.

(3) Reduce the number of Susceptibles through **vaccination,** which directly converts Susceptibles into Recovereds. That is, into people who aren't sick, won't get sick, and can't spread the disease. If we can push the starting value of S below the threshold, then we can prevent the epidemic from even starting. This is called **herd immunity**. People often think that it's important to get a shot to protect themselves, and it is. But it's even more important for them to get that shot to protect the rest of us!

2.5. Covid-19 and Modified SIR Models

The SIR model does a very good job of describing epidemics in well-mixed populations where the underlying conditions are changing slowly if at all. However, with just two parameters, it can't capture all of the details of a global crisis like the Covid-19 pandemic. In this section we'll go over several ways that modelers modified the SIR model to deal with Covid.

Scaled models and the replication number R_0. One problem with the SIR model is that the parameters depend on the size of the city where the outbreak is happening. Typically, the larger the population, the smaller a fraction of the population that each person knows or meets and the smaller the transmission coefficient a will be. However, the product of a and the total population tends to be the same in different cities. For this reason, it's useful to write a scaled version of the model.

Let $T = S + I + R$ be the total population. This number is of course constant, or at least approximately constant. We let $s = S/T$, $i = I/T$ and $r = R/T$ be the fractions of the population that are susceptible, infected, or recovered, so that $s + i + r = 1$. Since

$s' = S'/T$, $i' = I'/T$ and $r' = R'/T$, we can write rate equations for these quantities.

$$\begin{aligned} s' &= -aSI/T = -aTsi, \\ i' &= (aSI - bI)/T = aTsi - bi, \\ r' &= bI/T = bi. \end{aligned} \tag{2.15}$$

Modelers usually use the Greek letters β and γ in this scaled model, with

$$\beta = aT; \qquad \gamma = b. \tag{2.16}$$

This makes the scaled SIR equations:

$$\begin{aligned} s' &= -\beta si, \\ i' &= \beta si - \gamma i, \\ r' &= \gamma i. \end{aligned} \tag{2.17}$$

Conditions do vary from place to place, with β being larger in cities that are more crowded and where people are less careful, but β varies much less than a and T vary separately. As a result, it's sensible to talk about the values of β and γ for each disease.

In the early stages of an epidemic, when s is close to 1, people are getting sick at rate $\beta si \approx \beta i$ and recovering at rate γi. That is, there are β/γ people getting sick for every person who recovers. This ratio is called the **basic replication number**

$$R_0 = \frac{\beta}{\gamma}, \tag{2.18}$$

and represents the average number of new people that each sick person infects. If $R_0 > 1$, then the epidemic grows. If $R_0 < 1$, then the epidemic fizzles out.[4] For Covid, R_0 was originally estimated to be between 2 or 3. However, the speed at which it spread through the USA and Europe suggests that R_0 was actually higher. Later variants evolved to be even more infections, with values of R_0 around or even above 10.

Controlling an epidemic then amounts to getting s and R_0 as low as possible. In the short term, public health measures like wearing masks, closing schools, and staying home as much as possible can reduce β, and so can reduce R_0. Testing helps, too. Sick people don't have to infect others. If they can be identified and quarantined (or removed) from the general population, they can be infect**ed** without being infect**ious**.

However, such efforts can't last forever. Sooner or later people want to return to normal, they start behaving like they did pre-pandemic, and R_0 goes back up. The only long-term solution is reducing the fraction s of Susceptibles to below $1/R_0$, either through vaccination (best case) or natural infection (worst case).

The SEIR model. If Alice is sick and coughs on Bob, then Bob might get sick. If Bob then coughs on Carol, then Carol might get sick. However, if Bob coughs on Carol immediately after meeting Alice, then Carol is safe. It takes time for Alice's viruses to grow in Bob's body to the point that he can infect Carol.

[4] It's unfortunate that the basic replication number uses the same letter as the number of recovered individuals. However, since we're looking at r rather than R, and mostly care about s and i, this isn't really a problem.

To take this into account, modelers divide the population into four groups instead of three:

- **Susceptible** people (S, making a fraction $s = S/T$ of the population) are healthy but can become sick if they meet an infectious person.
- **Exposed** people (E, making a fraction $e = E/T$) are incubating the disease, but aren't sick enough to infect anybody else yet.
- **Infectious** people (I, making a fraction $i = I/T$) are a danger to everybody they meet.
- **Removed** people (R, making a fraction $r = R/T$) have recovered, have been quarantined, or have died. In any case, they aren't going to get sick again and they aren't going to infect anybody else.

Susceptible individuals become Exposed at a rate proportional to SI, just as in the SIR model. Exposed individuals become Infectious, and Infectious individuals become Removed, through the passage of time. After rescaling to express everything in terms of fractions of the population, our rate equations become

$$\begin{aligned} s' &= -\beta si, \\ e' &= \beta si - \alpha e, \\ i' &= \alpha e - \gamma i, \\ r' &= \gamma i. \end{aligned} \tag{2.19}$$

The new parameter α describes how long the incubation period is. On average, a person who gets infected is Exposed for time $1/\alpha$ and then is Infectious for time $1/\gamma$ before finally becoming Removed.

SEIR models behave a lot like SIR models with the same values of β and γ, except that the time delay causes the onset of the epidemic to grow a bit more slowly. In particular, we still talk about $R_0 = \beta/\gamma$ and the "herd immunity" threshold for starting to recover from the epidemic is still when $s = 1/R_0$.

The SIRS model. Immunity doesn't last forever. Over time, you lose the antibodies you had to a virus. Worse, germs can mutate into new forms that your initial antibodies don't recognize.[5] Either way, being Removed isn't really permanent.

We take this into account by adding a new parameter δ to our model to indicate the rate at which Removed individuals rejoin the ranks of Susceptibles.

$$\begin{aligned} s' &= \delta r - \beta si, \\ i' &= \beta si - \gamma i, \\ r' &= \gamma i - \delta r. \end{aligned} \tag{2.20}$$

The average duration of a person's immunity is $1/\delta$. This new model is called the SIRS model, for Susceptible-Infected-Removed-Susceptible. We can similarly turn the SEIR model into an SEIRS model.

[5] That's why you need a new flu shot each year. You may still be immune to last year's flu strain, but not to this year's.

Asymptomatic transmission and other compartments. One of the most dangerous things about Covid is its ability to be transmitted by people who don't even know that they are sick, either because they don't yet have symptoms (**presymptomatic transmission**) or because they never got symptoms (**asymptomatic transmission**). At the beginning of the pandemic, roughly a third of Covid cases were asymptomatic. This fraction has increased over time as Covid has evolved. While asymptomatic carriers tended to shed fewer viruses than symptomatic or presymptomatic carriers, they tended to go out a lot more. This made disease control very difficult.

To take asymptomatic transmission into account, modelers divided the Infected category (or the Infectious category for the SEIR model) into two groups, I_s and I_a for "symptomatic" and "asymptomatic". They had different rates β_s and β_a of infecting others and different recovery rates γ_s and γ_a. (They might take the same time to actually recover from the disease, but symptomatic individuals get removed from circulation a lot faster than asymptomatic carriers.) The SIR version of the model would then look like

$$\begin{aligned} s' &= -\beta_s s i_s - \beta_a s i_a, \\ i_s' &= \mu(\beta_s s i_s + \beta_a s i_a) - \gamma_s i_s, \\ i_a' &= (1-\mu)(\beta_s s i_s + \beta_a s i_a) - \gamma_a i_a, \\ r' &= \gamma_s i_s + \gamma_a i_a, \end{aligned} \tag{2.21}$$

where μ is the fraction of cases that develop symptoms.

You can also write down versions of the SEIR and SIRS models that take asymptomatic transmission into account. These models are more complicated than the basic SIR model, but they can be solved numerically in exactly the same way that we solved the SIR model. It's too complicated to do by hand, but programming a computer to solve equations (2.21) really isn't any harder than programming one to solve the basic model.

In practice, most of the models that were used to understand the spread of Covid had many compartments. Besides sorting people according to whether they were susceptible, exposed, infectious, or removed, there were compartments for being hospitalized, for being in intensive care, and for dying. People were sorted by age and sex, since Covid hit older people and men much harder than it hit younger people and women. To understand patterns of disease spread within a single city, people were also sorted by occupation, race, and zip code, as it turned out that all three factors were important. Once vaccines were developed, there were separate compartments for people who were vaccinated or who had recovered from Covid. Some of the models were very complicated. However, they were still just souped-up versions of SIR, with more details but the same underlying reasoning.

2.6. Same Song, Different Singer: SIR and Product Marketing

Throughout history, scientists have developed mathematical explanations for one purpose, only to find the same mathematical structure over and over again in completely

different settings. The physicist Eugene Wigner called this the "unreasonable effectiveness of mathematics". In this section we'll explore how the SIR model can be used in business.

Imagine that you work for a company that has just introduced a hot new phone app. People are learning about it by word of mouth, and more and more people are using your product. Of course, that growth can't go on forever. Eventually people will get tired of your app and will move on to the next great thing. In order to make the most of your product's popularity, you need to forecast usage for the next year or two. Having learned about the SIR model of how biological viruses spread, you decide to use the model to study how other things "go viral".

Instead of having Susceptibles, Infecteds, and Recovereds, you divide the population of potential customers into three groups:

(1) **Potential** users, or *Potentials*: These are people who might use the product in the future if they hear enough good things about it from their friends.

(2) **Active** users, or *Actives*: These are people who have already adopted the product. Not only are they using it, but they are spreading the word about it.

(3) **Rejected** users, or *Rejecteds*: These are people who won't ever use the product. Maybe they just aren't interested. Maybe they used to use the product but got tired of it. Maybe our app doesn't even work on their brand of phone. For whatever reason, they're no longer reachable.

As with diseases, people go from the first category (Susceptible or Potential) into the second (Infected or Active) through contact with people in the second category. The only difference is that transmission happens by word of mouth, either literally talking face to face or via texting or social media. Meanwhile, people go from the second category to the third through the passage of time. Every day, there's a chance that an Active user will get bored with the product or will find out about something better.

We define a parameter a that says how well Active users talk their friends into using the app, and a parameter b that says how quickly people stop using the app. We call a the **transmission coefficient**, exactly as before, and call b the **attrition coefficient**. If we let $S(t)$, $I(t)$ and $R(t)$ denote the numbers of Potential, Active and Rejected users,[6] then the equations are exactly the same as before:

$$\begin{aligned} S'(t) &= -aS(t)I(t), \\ I'(t) &= aS(t)I(t) - bI(t), \\ R'(t) &= \ bI(t), \end{aligned} \tag{2.22}$$

Since the equations are the same, the way we solve them is the same, the thresholds are the same, and the trajectories are the same.

If we want, we can make a scaled version of this model, with parameters β and γ and a replication number $R_0 = \beta/\gamma$ that describes how many new customers each Active user attracts before abandoning the product. We can introduce time delays with an SEIR model. We call allow for Rejected users to become interested in the product

[6] We could use different letters and call this the PAR model, but we won't.

again with an SIRS model. To a mathematician, the epidemic and the product launch are exactly the same.

There is a big difference, of course, in what we want to see happen. If an epidemic has an initial value of S that is below the threshold, causing the outbreak to fizzle, then we congratulate the health authorities who kept us safe. If a product launch has an initial value of S that is below the threshold, causing the product to fizzle (something we call a **product bust**), then the marketing team is likely to get fired. The math is the same, but the interpretation is completely different.

In marketing, we generally want the opposite of what the health authorities want. We want $I(t)$ to grow as quickly as possible, to reach a peak that is as high as possible, and to die out as slowly as possible. This means that we want to

(1) Make b as *small* as possible. Equivalently, make the usage time $T = 1/b$ as big as possible. This has to do with the quality of our product. The more exciting our product is, and the more we provide updates and other continuing benefits to the users, the lower the attrition coefficient will be.

(2) Make a as *big* as possible. Do everything possible to encourage word-of-mouth communication between Active and Potential users.

(3) Pick a market where $S(0)$ is as *big* as possible. Even the best products can fizzle if they're directed at the wrong market. A lemonade stand will do a lot better in Texas in August than in Minnesota in February.

2.7. Chapter Summary

The Main Ideas.

- *Close is good enough.* Most of the time, we don't need an exact answer. A good simple approximation is often more useful than a complicated formula.
- Mathematical models help us to understand real-world problems.
  - Take what we know about our problem.
  - Throw out the unimportant details and keep the main features.
  - Express those features mathematically.
  - Use math to **analyze** the model.
  - **Interpret** the results. Turn numbers and equations into real-world conclusions. Without this last step, the rest is useless.
- *Track the changes.* Many useful models involve **rate equations** that describe how fast something is changing in terms of its current state. Many of the same equations show up in different settings. The SIR equations model the spread of disease. The same equations also model market penetration.
- Once you know how fast something is changing, you can use a **linear approximation** to predict its future or explore its past. This is accurate for a short time, but can't be trusted over long time intervals.
- *One step at a time.* If you don't trust a linear approximation to predict what will happen next year, just use it to predict what will happen tomorrow. Then use

those results to predict what will happen the day after tomorrow, then the day after that, and so on.

- *What goes up has to stop before it comes down.* You can learn a lot by studying the point when something stops moving: that is, when the rate of change equals zero.

Expectations. You should be able to:

- Model epidemics with a particular set of rate equations (the SIR model).
- Explain how the parameters in the SIR model relate to properties of the disease and the community where it's spreading.
- Use the SIR equations, together with a linear approximation, to predict future infection rates.
- Iterate this process to generate a table of values for several times.
- Relate the sign of I' to whether an epidemic is getting better or worse, and relate the size of I' to how fast this is happening.
- Use the same tools to analyze a product launch.

2.8. Exercises

Rate of Change

2.1. A farmer is monitoring the number of fire ant mounds on his property. Suppose that there are 1200 mounds in March and 1250 mounds in May of the same year. Let $F(t)$ be the number of fire ant mounds at time t.

(a) What is the change in the number of fire ant mounds between March and May? Denote your answer by ΔF. What are the units for ΔF?

(b) What is the change in time between these two observations? Denote your answer by Δt. What are the units for Δt?

(c) At what rate did the number of mounds change from March to May? Call your answer m_1. What are the units for m_1?

(d) Assuming that the number of mounds is a linear function of time, how many mounds should we expect in August? What value of Δt are you using to find your answer? Use the notation and answers from parts (a), (b), and (c) to express your answer.

(e) Again assuming a linear relationship, how many mounds were there in January of the same year? What value of Δt are you using to find your answer?

(f) Find a linear formula for the number of mounds t months after March (that is, $t = 0$ means March).

2.2. In Exercise 2.1 you developed a formula for the number of fire ant mounds t months after March. Suppose that you wish to update your model using additional data. You observe that model you developed in Exercise 2.1 worked well through August, but that the *change* in the number of mounds between the months of August and September was 100. (That is, there were 100 more mounds in September than in August.)

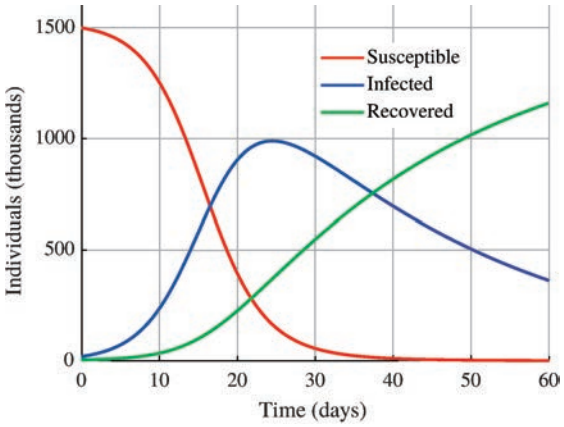
(a) What is the per month rate of change between August and September? Denote your answer by m_2.
(b) How does m_2 compare to the value of m_1 you found in Exercise 2.1(c)? What does this say about the seriousness of the fire ant infestation?
(c) Use m_2 to find an updated linear formula for predicting the number of mounds for October through December. (This is similar to the formula you found in Exercise 2.1(f).) State the correct domain for this expression.
(d) Use the information given in Exercise 2.1 and the result of Exercise 2.2(c) to write a piecewise function that expresses the number of ant mounds for the entire year (January through December), letting $t = 0$ represent the month of March.
(e) Use MATLAB to graph the piecewise function you found in part (d). Label your axes and give your graph an appropriate title. Use different colors for the two pieces of your function.
(f) Carefully explain why using the piecewise function to express monthly ant mounds for the entire year is better than simply using the linear function you found in Exercise 2.1(f) for the entire year.

2.3. Explain the term "rate equation". How is a rate *equation* different from the *rate of change m* associated with a linear function?

2.4. Suppose we have a rate equation $F' = \cdots$ for a quantity $F(t)$. At a starting time t_0, we have $F(t_0) = F_0$, and the rate equation tells us that $F'(t_0) = R_0$.
(a) Explain, using proper notation, how we can approximate the change in $F(t)$ between times $t = t_0$ and $t = t_0 + a$ using a linear approximation.
(b) Explain, using proper notation, how we can approximate $F(t_0 + a)$.
(c) If $F(t)$ is a linear function, would your answers in parts (a) and (b) still be approximations? Explain.

Interpreting SIR Models

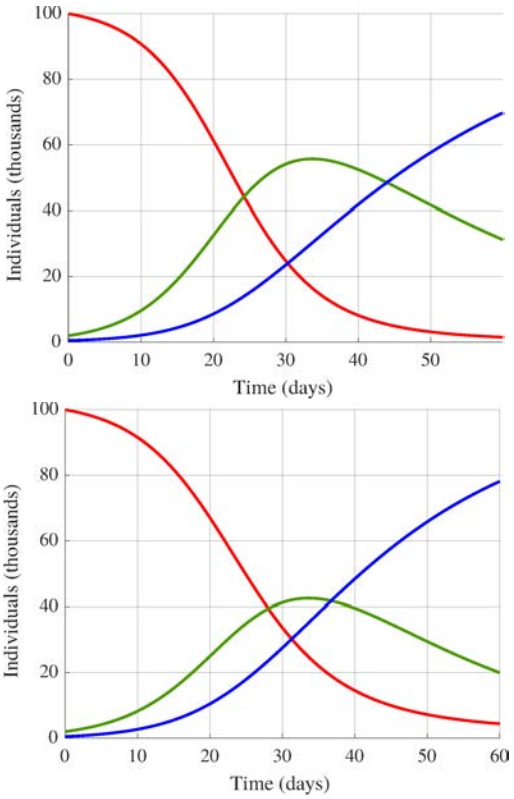
2.5. Give approximate answers to these questions about the following graph of SIR model behavior:



(a) When does the number of Infecteds reach its peak? How many people are infected at that time?

(b) Initially, how many people are Susceptible? How many days does it take for the Susceptible population to be cut in half?
(c) How many days does it take the Recovered population to reach 500,000? How many people recover by day 60?
(d) On what day is the size of the Infected population increasing most rapidly? When is it decreasing most rapidly? How can you tell?
(e) How many people became Infected at some time during the first 30 days? (Note that this is not the same as the number of people who are infected on day 30.) Explain how you found this information.

2.6. Below are two graphs depicting the spread of disease according to the usual SIR model. The initial values $S(0)$, $I(0)$, and $R(0)$ and the transmission coefficient a are the same for both graphs. However, the two graphs correspond to different recovery coefficients b.



(a) For each graph, indicate which curve is S, which is I, and which is R.
(b) Which graph corresponds to the larger value of b? Explain.

2.7. Consider an epidemic with the simplified rate equation $S' = -35{,}000$ persons per day. (This is not the usual SIR model.) Suppose the initial Susceptible population is 2,000,000 on December 14.
(a) How many Susceptibles will be left on Christmas Day (December 25)?
(b) When will the Susceptible population vanish entirely?

(c) How many Susceptibles were there on December 10?
(d) When were there 2,200,000 Susceptibles?

2.8. Suppose that a new disease is circulating in Smalltown, population 10,100, and that it evolves according to the usual SIR equations

$$\begin{aligned} S' &= -aSI, \\ I' &= aSI - bI, \\ R' &= bI. \end{aligned}$$

It appears that this disease lasts for an average of 20 days. It is also thought that there were initially 100 Infected people. That is, we had

$$\begin{aligned} S(0) &= 10{,}000, \\ I(0) &= 100, \\ R(0) &= 0. \end{aligned}$$

One day later, it was observed that there were eight *newly infected* people.
(a) Use the given information to complete the model. That is, find the transmission and recovery coefficients a and b.
(b) Using $\Delta t = 2$, find estimates for $S(4)$, $I(4)$, and $R(4)$.
(c) Find the threshold and explain its significance.
(d) Suppose that upon becoming Infected, patients are quarantined. It still takes time to identify Infected people, but the average length of time that they can infect others drops from 20 days to 10. What effect will this have on the threshold?

2.9. (a) Construct the appropriate SIR model for a disease with the following characteristics:
- Infected people remain sick for an average of 20 days.
- A typical Susceptible meets about 0.01% of the Infecteds each day (because he meets 0.01% of the entire population each day).
- It takes an average of eight contacts with an Infected before a Susceptible becomes Infected.

(b) How large does the Susceptible population need to be for an epidemic to grow?

2.10. A town's public health department has created an SIR model to predict the effects of this year's new strain of flu. The SIR equations are given by

$$\begin{aligned} S' &= -.0000002SI, \\ I' &= .0000002SI - I/15, \\ R' &= I/15. \end{aligned}$$

The initial values at $t = 0$ are thought to be

$$S(0) = 2{,}000{,}000, \quad I(0) = 21{,}000, \quad R(0) = 15{,}000.$$

(a) Using the initial rates, use the SIR model to find $S(1)$, $I(1)$, and $R(1)$.
(b) Using the SIR equations, calculate the rates of change $S'(1)$, $I'(1)$, and $R'(1)$. Then use these values to determine $S(2)$, $I(2)$, and $R(2)$.

(c) Using the values of $S(2)$, $I(2)$, and $R(2)$ that you computed in part (b), calculate the rates of change $S'(2)$, $I'(2)$, and $R'(2)$. Then estimate $S(3)$, $I(3)$, and $R(3)$.

(d) Go back to the starting time $t = 0$ and to the initial values

$$S(0) = 2{,}000{,}000, \qquad I(0) = 21{,}000, \qquad R(0) = 15{,}000.$$

Recalculate the values of $S(2)$, $I(2)$, and $R(2)$ by using a time step of $\Delta t = 2$. This only requires a single round of calculations, using $S'(0)$, $I'(0)$, and $R'(0)$. How do your answers compare to those computed in part (b)? Which estimates do you think are most accurate, those in part (b) or those in part (d)? Why?

2.11. **Herd Immunity.** Suppose there is a measles outbreak in a small community of 40,000 people, almost all of them Susceptible, with $b = 1/(14 \text{ days})$ and $a = 0.000003$.

(a) What is the threshold in this situation? How many people will get sick before the epidemic reaches its peak?

(b) Now suppose that 10,000 people (out of 40,000 total) are immune to measles, either because they already had the disease or because they were vaccinated. How many will get sick before the epidemic reaches its peak?

(c) If enough people are immunized, the epidemic will fizzle from the start. How many people would have to be immunized for that to happen?

(d) When enough people are immunized to keep outbreaks from spreading, the population is said to have herd immunity. Explain how the number of vaccinations needed for herd immunity is related to the threshold b/a.

(e) If you were a public health official, what would you say to somebody who doesn't want to get vaccinated and who insists that their (not) getting vaccinated is none of your business?

2.12. **Replication Number.** The ratio of how many people get sick to how many recover is $aSI/(bI) = aS/b$. At the beginning of an epidemic, when S is almost the entire population, this is the same as the basic replication number $R_0 = \beta/\gamma$ that we saw in our scaled SIR model. If $R_0 > 1$, then the epidemic will grow, at least at first. If $R_0 < 1$, it will fizzle out.

(a) Suppose that the total population is 40,000. If $R_0 = 1.2$, what is the threshold? How many people will get sick before the epidemic reaches its peak? What if $R_0 = 10$? In general, how are R_0, the total population, and the threshold related?

(b) In a town with total population T and a disease with basic replication number R_0, how many people would need to become immune to achieve herd immunity?

Food for thought: The number of people needed to reach herd immunity is the same as the number of people who have to get sick *before* the epidemic reaches its peak. However, a substantial number of people can still get sick *after* the peak. For instance, suppose that $R_0 = 2$. The peak will be reached when half the population has gotten sick, but nearly 30% of the population will get sick later. Only about 20% will avoid the

disease altogether. This is one reason why it is much better to immunize people before an epidemic than to wait for herd immunity during an epidemic.

2.13. **Superspreader Events.** One way for a disease to spread is through big public gatherings such as sporting events, music festivals and political protests. By giving sick individuals the opportunity to interact with a lot of other people, these gatherings effectively increase the transmission coefficient a.[7]

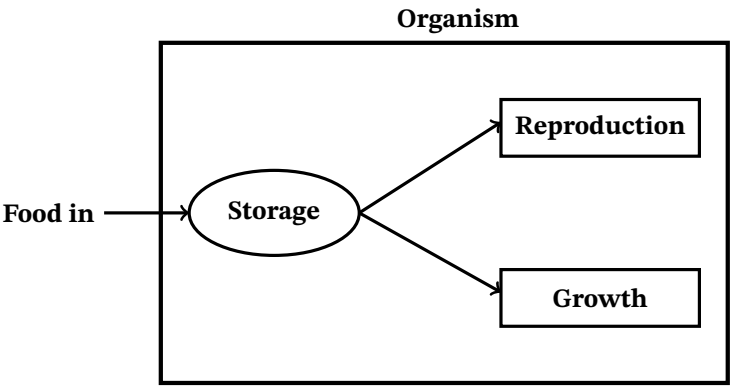
(a) Suppose that we are in the setting of Exercise 2.10, except that a music festival doubles the chance that a Susceptible will become Infected. What is the new transmission coefficient?

(b) Compute the new threshold if the festival of part (a) takes place. Compare your answer to the threshold in the original setting of Exercise 2.10.

(c) Suppose you have a model for a small community where $S(0) = 40,000$. It is estimated that, if the threshold is 30,000, then the hospitals will be able to keep up with the flow of sick patients. However, if the threshold is less than 30,000, then the epidemic will progress so fast that the hospitals won't be able to keep up. If $b = 1/14$ and $a = 0.000002$, will the health care system be able to handle the strain?

(d) Now suppose that the town hosts an athletic competition that increases a to 0.000003. What will the new threshold be? What will happen to the town?

2.14. **Dynamic Energy Budget**. The reasoning we used to build the SIR model, namely tracking the inputs and outputs for each variable to come up with a rate of change, also helps us build models that have nothing to do with disease spread. For instance, plants and animals, in all of their wondrous variety, can be viewed as consumers. Plants must consume water, nutrients, and sunlight to survive. Animals must consume plants or other animals. The energy in the resources they consume is utilized in growth, reproduction, and basal maintenance. (To simplify the model, we have ignored several other energy requirements.) The following is a simple schematic that represents the basic components of a dynamic energy budget.



[7] An August 2020 motorcycle rally drew almost half a million people to the small town of Sturgis, SD. Within weeks, South Dakota and its neighboring states were leading the country in Covid infections per capita.

(a) Suppose energy (measured in calories) enters storage (S) via a food source at rate p_E (measured in calories per day). Energy leaves storage at rate p_L. A fraction κ of that energy is used for growth G, while a fraction $1-\kappa$ is used for reproduction R. Write down a system of rate equation for S, G, and R.

(b) Suppose that a fraction m of the energy designated for growth is actually used for organism maintenance. (These are the processes essential to keeping an organism alive, such as respiration, cell repair, etc.) Revise your model to show the energy used for organism maintenance separately from the energy used for growth.[8]

Exponential Growth. In many cases, the rate of change of a quantity is proportional to the quantity itself. We will see many different examples of that in Chapter 6, but the two simplest are money earning interest and population growth. We will examine these in Exercises 2.15–2.17.

If a bank account containing \$10,000 earns \$150/year in interest, then an account at the same bank containing \$20,000 should earn \$300/year in interest. Either way, the interest is 1.5%/year of the bank balance.

If a city of 100,000 is growing at a rate of 3000 people per year, then we would expect a similar city of 200,000 (or even the same city a few years later) to grow by 6000 people per year. That is, if $P(t)$ is the population at time t, then the **net growth rate** P' is proportional to P:

$$P' = kP,$$

where k is a constant. In this example, $k = .03$/year, or 3%/year.

2.15. In the equation $P' = kP$ for the population of a city, the number k is called the **per capita growth rate**, with the Latin phrase "per capita" literally meaning "per head". Explain why the units for k are 1/year.

2.16. In 2020, the population of Australia was 25 million and was growing at 1.36%/year, while the population of the United States was 330 million and was growing at 0.71%/year. Assume that both countries' populations keep growing at these per capita rates.

(a) Let $A(t)$ and $U(t)$ be the populations of Australia and the United States, respectively. Write down the rate equations that govern the growth of these functions.

(b) What were the net growth rates $A'(2020)$ and $U'(2020)$ in 2020?

(c) In general, does a country with a higher per capita growth rate necessarily have a higher net growth rate?

(d) On average, how many seconds did it take for the population of Australia to increase by 1 in 2020? For the population of the United States? (A year is about 31.6 million seconds.)

Food for Thought: At these per capita growth rates, the populations of Australia and the United States would become equal in 2417, at which time each country would have

[8] Plants don't need much energy to keep themselves going, so m is typically small for plants. m is larger for animals, especially warm-blooded animals (like humans) that spend a lot of energy just keeping their body temperatures up.

about 4.3 *billion* people. Exponential growth is great when it comes to money, but it can be scary in other cases.

2.17. A bank account is earning interest at a constant percentage rate. When the account had \$3200, it was growing at a rate of \$79/year.
   (a) Write an equation that links the net growth rate (meaning P', not the interest rate) to the bank balance.
   (b) At a later time, the bank balance is growing at \$100/year. What is the balance at that time?

Scaling the Time Variable. In a rate equation, the quantity Y' has units, namely whatever units we use for Y divided by whatever units we use for time. If Y is people and t is measured in days, then Y' is measured in people/day. In Section 2.5, we generated a scaled model, where S, I, and R were replaced with dimensionless quantities s, i, and r that are proportional to S, I, and R. In Exercises 2.18–2.20, we will see what happens when we similarly replace the time t with a dimensionless variable.

2.18. Suppose that $i' = 0.014$/week. That is, i is increasing at a rate of 0.014 per week.
   (a) How fast is i increasing per *day*? How fast is i increasing per *year*? Which is bigger, the growth per year or the growth per day?
   (b) Let t_d, t_w, and t_y be the number of days, weeks, and years since a fixed starting time. How are the variables related? Which is bigger, t_d or t_y?

To make a model dimensionless, we pick a fixed unit of time T. T might be a second, an hour, a day, a year, or 3723.849 years. We then measure time in units of T. That is, we define a new dimensionless variable $\tau = t/T$ that counts how many T's have elapsed since $t = 0$. If T is small, then τ will be big, while if T is big, then τ is small. A very *large* number of nanoseconds equals a very *small* number of centuries. For variables Y and P that change in time, we let $\dot{Y}$ and $\dot{P}$ denote the changes in Y and P *per change in* τ.

2.19. How is $\dot{Y}$ related to Y'? (*Hint*: if Y' is constant, how much will Y change by the time that $\tau = 1$?)

2.20. Suppose that $P' = rP$. Let $T = r^{-1}$.
   (a) Compute $\dot{P}$ in terms of P.
   (b) Compute T when $r = 1\%$/year, when $r = 3\%$/month, and when $r = 0.06$/minute.

Food for Thought: You should discover that the trajectory of $P' = rP$ looks essentially the same for all of these values of r, only we trace out that trajectory at different speeds. If we can understand the solution to $P' = rP$ for one value of the parameter r (say, for $r = 1$), then we can understand the solution to $P' = rP$ for all values of r.

2.21. Consider the scaled SIR equations (2.17). Let $T = \gamma^{-1}$. Write down a system of equations for $\dot{s}$, $\dot{i}$, and $\dot{r}$.

Food for Thought: The resulting equations still have one dimensionless parameter, namely the basic replication number R_0. The details of an epidemic depend on a lot of different parameters and initial conditions, but the general shape of the curve only depends on R_0.