Synthesis Lectures on
Mathematics & Statistics

SYNTHESIS
COLLECTION OF TECHNOLOGY

Eric Stade · Elisabeth Stade

Calculus: A Modeling and Computational Thinking Approach

Springer

Synthesis Lectures on Mathematics & Statistics

This series includes titles in applied mathematics and statistics for cross-disciplinary STEM professionals, educators, researchers, and students. The series focuses on new and traditional techniques to develop mathematical knowledge and skills, an understanding of core mathematical reasoning, and the ability to utilize data in specific applications.

Eric Stade · Elisabeth Stade

Calculus: A Modeling and Computational Thinking Approach

Springer

Eric Stade
University of Colorado Boulder
Boulder, CO, USA

Elisabeth Stade
University of Colorado Boulder
Boulder, CO, USA

ISSN 1938-1743 ISSN 1938-1751 (electronic)
Synthesis Lectures on Mathematics & Statistics
ISBN 978-3-031-24680-7 ISBN 978-3-031-24681-4 (eBook)
https://doi.org/10.1007/978-3-031-24681-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This first-semester calculus text represents a reimagining of (roughly the first half of) the revolutionary textbook *Calculus in Context* by James Callahan, David A. Cox, Kenneth R. Hoffman, Donal O'Shea, Harriet Pollatsek, and Lester Senechal, of the Five College Calculus Project. Many of the big ideas, and much of the content, of the present textbook derive, in modified form, from that one.

In both this book and *Calculus in Context*, the iconic "*SIR*" (for "Susceptible/Infected/Recovered") epidemic model is introduced at the very beginning. This model is developed from intuitive ideas, without requiring any of the foundational Calculus notions—limits, continuity, and so on—that often populate the first week or so of a beginning course in Calculus. Instead, those foundational notions are explored, here, only once the *SIR* equations have prompted a closer look at some of the mathematical underpinnings. In this way, the motivation—or at least a motivation—for studying calculus is situated front and center.

In the *SIR* scenario, it's not possible to quantify the populations of susceptible, infected, and recovered individuals with nice, compact formulas. Instead, numerical methods, using computers, are required to analyze these populations. Thus, in a traditional calculus course, where simple formulas are typically foregrounded, and numerical methods are generally eschewed, the *SIR* model cannot be studied in real depth. By contrast, here (as in *Calculus in Context*), we embrace numerical methods! Indeed, we see *SIR* as an excellent "excuse" to acquire some basic programming notions and skills. (All such skills and notions are developed along the way; none are required at the outset.)

And the programming, in turn, helps to elucidate the calculus. More specifically, the concept of a rate of change, which is perhaps *the* central idea of Calculus, is also at the heart of the "Euler's method" algorithm that we use to solve the *SIR* equations. Familiarity with Euler's method then promotes facility with rates of change.

Programming also requires certain linear, "computational" ways of thinking. Thus programming fulfills a role, in this book, that rigorous, formal proofs have sometimes played in earlier calculus texts. Of course, there's nothing wrong with rigorous, formal proofs. They can be powerful things of beauty. But programming is a skill that has perhaps wider, more direct utility in life beyond Calculus.

All of the above perhaps suggests that this text is well off of the beaten path. This may be true in some respects. But this book does cover most of the "standard" calculus notions: derivatives; differentiation rules and formulas, linear, polynomial, trigonometric, exponential, and logarithmic functions; inverse functions; optimization; Riemann sums; antiderivatives; integration; the Fundamental Theorem of Calculus; integration by substitution; separation of variables; and so on. The approach is different, though: it's contextual, and it's computational.

A number of the contexts considered in this book arise from the Life Sciences. In addition to the *SIR* model, we investigate (single and dual species) population growth, circadian rhythms, neural impulses, and several other biological phenomena. However, this text is not intended (solely) as a Calculus for Life Sciences textbook. Other contexts—projectile motion, work and force, power and energy, radioactive decay, solar energy, and so on—are explored as well. And again, the Calculus notions developed here are universal. Our intention is that this book be suitable for a wide variety of first-semester Calculus courses.

As noted above, this book owes a tremendous debt of gratitude to the Five College Calculus Project, and to James Callahan, David A. Cox, Kenneth R. Hoffman, Donal O'Shea, Harriet Pollatsek, and Lester Senechal, authors of *Calculus in Context*. All of the Five College folks have been incredibly generous and gracious in letting us use their material freely. We thank them not only for this material, but for their groundbreaking approach to Calculus.

Further, we thank them for allowing us the freedom to reimagine and update their text using our own voices, ideas, perspectives, and experiences in the classroom. Theirs is a difficult book to improve upon, but we did try to make the present text our own. Key new features of our textbook include the following:

- Additional topics related to epidemics, including familiar, COVID-age notions of reproduction number and herd immunity;
- Programming exercises and examples based on computers, and on the free, web-based, Sage (also known as Sagemath) computer mathematics package, instead of programmable calculators;
- Additional scaffolding, in the form of numerous new worked examples;
- A variety of new contexts, including an Ebola outbreak, circadian rhythms, the genetic toggle switch, diffusion across a membrane, neural impulses, the Colorado flood, solar energy, and so on;
- Selection and ordering of topics that is somewhat more in line with "traditional" Calculus I courses;
- A number of additional exercises, from "drill" to abstract.

We are particularly grateful to James Callahan, who was our primary liaison to the Five College Calculus Project, and who engaged in many conversations and email exchanges

with us. These communications with Jim made the writing of this text so much easier and made the end product so much better.

We also thank a number of University of Colorado faculty, including David Webb of the School of Education, Robin Dowell and Mike Klymkowsky of the Department of Molecular, Cellular, and Developmental Biology, and Richard Holley (Emeritus) of the Department of Mathematics, for the various ways in which they helped to shape this text.

Additionally, deep, deep gratitude goes out to all of the students who influenced this text. This includes all of the graduate student instructors and Teaching Assistants, as well as the undergraduate Learning Assistants, with whom we have taught a course based on this text. It also includes the hundreds of students who have taken that course. All of these students have had profound impacts on both the content and the pedagogy behind this text.

Finally, we thank our extraordinary sons Jack and Nick, for enduring everything that went into the writing of this text, but primarily just for being Jack and Nick.

Boulder, USA

Eric Stade
Elisabeth Stade

Contents

A Context for Calculus

1

1.1 Introduction: Calculus and Prediction

Calculus is often described as the mathematics of *change*. And one of the most significant things we can do about change is predict it.

If we can predict the way a situation will evolve, then we can prepare for it. Perhaps more critically: if we can predict how the evolution of a phenomenon will depend on present, or imposed, conditions, then we can endeavor to establish the conditions that will yield the most desirable outcome. So prediction is a good thing to be able to do.

For these reasons, we like to think of calculus as the mathematics of *prediction*. And all of calculus rests, at some level, on the following simple *prediction principle:*

If you know how fast you're going, then you know how far you'll get in a given amount of time.

The prediction principle

Now this principle, as stated, may read like a statement about moving objects—about position, displacement, velocity, and so on. And certainly this principle does apply to all of those things. Indeed, Calculus was invented (discovered?) in the mid-to-late 1600s, by Isaac Newton and Gottfried Leibniz (more or less independently), in part to address problems in planetary motion.

But the principle has much, much broader relevance. It applies in any situation where you know the rate at which things are changing (that is, you know "how fast you're going"). It applies, then, to many, many, many phenomena—populations, epidemics, temperature, climate, neural impulses, protein concentrations, chemical reactions, radioactive decay, etcetera, etcetera—that evolve with respect to time (or that vary spatially, or depend on some other dimension or dimensions).

E. Stade and E. Stade, *Calculus: A Modeling and Computational Thinking Approach*,
Synthesis Lectures on Mathematics & Statistics,
https://doi.org/10.1007/978-3-031-24681-4_1

Of course, since calculus is the mathematics of prediction, it *is* mathematics, among other things. Therefore, to apply calculus—and the above principle—to the study of an evolving phenomenon, we first need to describe that phenomenon in mathematical terms. Doing so is called **(mathematical) modeling**. The art and science of mathematical modeling will constitute major emphases of this book.

Models that describe, mathematically, how fast you're going are called **dynamical systems**. And once we have a dynamical system in hand, a major goal, if not *the* major goal, is to *solve* the system—that is, to use it to describe the evolution of the phenomenon. Solution of dynamical systems will therefore constitute another major emphasis of this book.

Building and solving dynamical systems are worthy ends in themselves, because of their powers of prediction. But dynamical systems also serve as engaging means of introducing, and developing, the fundamental constructs and concepts of calculus itself. In other words: the applications drive the mathematics just as much as the mathematics drives the applications. We think of this perspective as constituting a **contextual** view of calculus.

One thing about all of this that will, we hope, become clear later in this chapter is that *solution of dynamical systems can be computationally intensive*. In fact, it's completely impractical, in general, to try and solve these systems "by hand." For this reason we will, at many places in this book, apply computer mathematics software—we'll use the Sage package; packages like Mathematica and MATLAB work similarly—to the solution of these systems. We will discuss and develop the relevant computing techniques and notions as needed. (No previous computing experience is assumed.) Our approach to computing will be contextual too.

Certainly, not all of life, or even science, is about prediction. We often wish to study the past, or the present, as much as the future. Yet investigations of the past, and examinations of the present, are themselves often performed with an eye towards knowing and impacting the future. Moreover, analyses of the past—for example, approximating the age of an artifact using carbon dating—are often achieved through the application of predictive models that work just as well in either time direction.

1.2 The Spread of Disease: The SIR Model

Many human diseases are contagious: you "catch" them from someone who is already infected. Contagious diseases are of many kinds. Smallpox, polio, plague, and Ebola are severe and can be fatal, while the common cold and the childhood illnesses of measles, mumps, and rubella are usually relatively mild. Moreover, you can catch a cold over and over again, but you get measles only once. A disease like measles is said to "confer immunity" on someone who recovers from it.

Some diseases have the potential to affect large segments of a population; they are called *epidemics* (from the Greek words *epi*, upon + *demos*, the people.) *Epidemiology* is the scientific study of these diseases.

An epidemic is a complicated matter, but the dangers posed by contagion—and especially by the appearance of new and hard-to-control diseases, like COVID-19[1]—compel us to learn as much as we can about the nature of epidemics. Mathematics offers a very special kind of help.

First, we can try to draw out of the situation its essential features and describe them mathematically. Again, this process is called **(mathematical) modeling**. Second, we can use mathematical insights and methods to analyze the model. Any conclusion we reach about the model can then be interpreted to tell us something about the reality.

To give you an idea how this process works, we'll build a model—called the SIR model, for *susceptible, infected, recovered*—of an epidemic. This is a well-known, even iconic, model in epidemiology. Its basic purpose is to help us understand the way a contagious disease spreads through a population—to the point where we can even predict what fraction will fall ill, and when. Let's suppose the disease we want to model is like measles. Then it behaves (roughly) like the following.

1.2.1 Initial Setup

Our disease will entail the following three quantities, and their rates of change. These rates of change tell us "how fast we're going;" that is, they indicate the rates at which the indicated quantities are changing.

S: number of susceptible persons	S': rate of change of S
I: number of infected persons	I': rate of change of I
R: number of recovered persons	R': rate of change of R

It's common in mathematics to denote the rate of change of a quantity Q by Q'. We'll have quite a bit to say about rates of change as we proceed. In fact, they will constitute a (the?) major focus of this book. For now, we'll be content to understand a rate of change intuitively, as the "speed" at which a quantity changes, or evolves. Note that a rate of change Q' will, itself, typically be changing, since Q might change quickly at some times and slowly at others, may be increasing at times and decreasing at others, and so on.

We make the following assumptions about the nature of this disease:

- Everyone recovers eventually.
- The duration of infection is the same for everyone.
- Once recovered, you're immune, and can no longer infect anyone.
- Only a fraction of contacts with the disease cause infection.
- The units of S, I, and R are persons.

[1] COVID-19 is generally described as a *pandemic*, meaning an epidemic that has spread across multiple countries or continents.

- The units of time are days.
- The units of S', I', and R' are persons per day, written persons/day.
- The system is *closed;* this simply means that the total size of the population, which equals the sum $S + I + R$, does not change.

1.2.2 Thinking About S', I', and R'

In this section, we will develop a system of *rate equations*—mathematical formulas describing rates of change—for S, I, and R. Later, we'll apply a technique known as *Euler's method* to this system, to predict the evolution of our epidemic.

Let's begin by addressing R'. We deal with this rate of change first because its analysis is, in many ways, the easiest of the three.

Suppose infection lasts for k days. Also assume, in the absence of any definite information to the contrary, that the infected population is "uniform with respect to duration of infection," at any given point in time. That is, assume that there are, any any instant, just as many people in this population who have been infected for one day as there are who have been infected for two, or three, and so on, up to k days.

Then on any given day, one kth of the infected population will recover. In other words, the *rate of recovery*, in persons per day, is equal to $1/k$ times I. Recall that we are calling this rate R'. So, in symbols:

$$\boxed{R' = bI, \quad \text{where} \quad b = \frac{1}{k}.}$$

Rate equation for the recovered population

Here, b is constant, in that it doesn't change over the course of time. Of course, different diseases may entail different values of b. In mathematical modeling, a number that is constant within a given situation, but may vary from situation to situation, is called a *parameter*. Often, we will use uppercase letters for the variables (other than the time variable) in rate equations, and will use lowercase letters for the parameters (and the time variable).

When one variable quantity equals a constant times another, we say the two quantities are *proportional*. So in our present model, the *rate of change* of the recovered population is proportional to the *size* of the infected population.

Let's move on to examination of S', which is the next easiest rate of change to model. Let's suppose that:

(i) Each susceptible person comes into contact with a proportion, call it p, of the infected population each day. This implies that each susceptible person has contact with pI infected persons per day. This in turn implies that there are $pI \cdot S = pSI$ *total* contacts between susceptible and infected persons each day.

(ii) A certain proportion, call it q, of the above contacts cause infection.

The above tells us that there are $q \cdot pSI$ new infections occurring each day. Which in turn implies that the size of the susceptible population *decreases* by $qpSI$ persons each day. In other words, in persons/day, we have

$$\boxed{S' = -aSI, \quad \text{where} \quad a = qp.}$$

Rate equation for the susceptible population

The minus sign denotes a *negative* rate of change, meaning a *decrease* in the quantity in question.

Finally, we consider I'. Since $S + I + R$ is assumed to be constant, the sum $S' + I' + R'$ of the rates of change of the three subpopulations must be *zero*—any change in one of these quantities is offset by changes in the others. That is, by the above formulas for R' and S',

$$0 = S' + I' + R' = -aSI + I' + bI$$

or, solving for I',

$$\boxed{I' = aSI - bI.}$$

Rate equation for infected population

To summarize: under the conditions described above, we have

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

The *SIR* equations

Here:

- The number a is a positive parameter called the *transmission coefficient*. Recall from above that $a = qp$, where p is the proportion, or fraction, of the infected population with which each infected person comes into contact each day, and q is the proportion these of contacts that cause infection. Note that the *units* of a are 1/(person-day), or "inverse person-days." Why? Because the units on both sides of any equation must agree. So, considering for example the equation $S' = -aSI$, we see that inverse person-days are indeed the correct units for a, because they insure that the right-hand side of this equation has units

$$\frac{1}{\text{person-day}} \cdot \text{persons} \cdot \text{persons} = \frac{\text{persons}}{\text{day}},$$

which are the units of the left-hand side of this equation.

- The number b is a positive parameter called the *recovery coefficient*. Recall from above that $b = 1/k$, where k is the number of days that infection lasts. The units of b are 1/day, or inverse days.

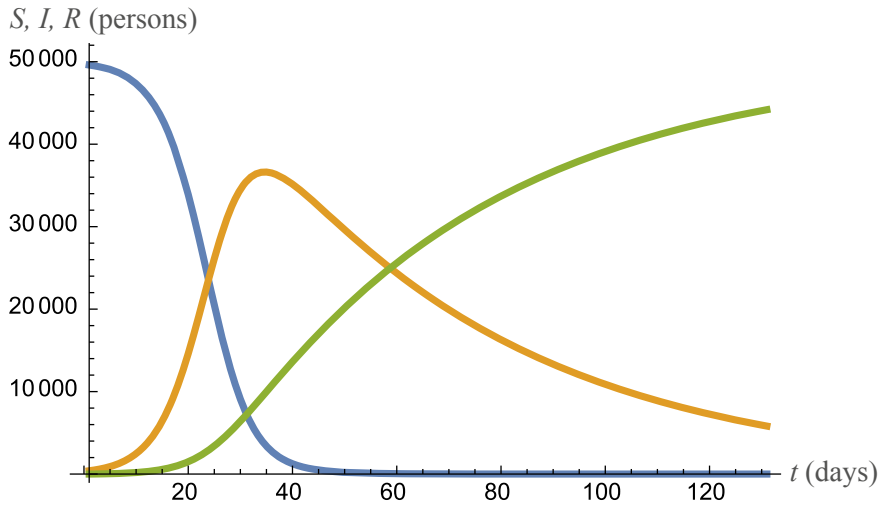
As noted above, S is a *decreasing* quantity. This is reflected in the above equation for S': the quantities a, R, and I are all positive, so $S' = -aSI$ is negative. And again, a negative rate of change corresponds to a decreasing quantity. Actually, it's conceivable that I or R might be zero at some point(s), in which case S' would be zero there too. So technically, it might be more precise to say that S is *nonincreasing*: its rate of change is never positive. To keep the terminology simple, though, we'll use the term "decreasing" even for quantities that are, strictly speaking, nonincreasing.

Similarly, R is an increasing quantity. The size I of the infected polulation can increase *or* decrease, since the sign of $I' = aSI - bI$ can be positive or negative, depending on the relative sizes of aSI and bI. In the next section, we'll look at the rise and fall of I more closely.

Also in the next section, we'll use our SIR equations, together with a generalized version of the *Euler's method* discussed in the previous section, to study the progress of our epidemic.

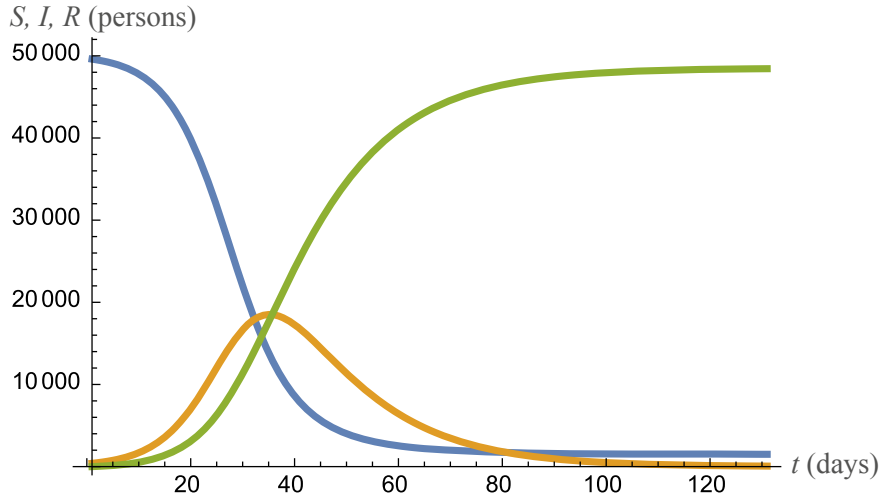
1.2.3 Exercises

1. The graph below depicts a disease evolving according to the above SIR model.



(a) Which curve is S, which is I, and which is R? Explain how you know which is which.
(b) About how long does it take before the susceptible population has decreased to half of its original size?
(c) About how large is the overall population (comprising all susceptible, infected, and recovered individuals)?
(d) About how long does it take before the infection peaks?
(e) About what is the size of the infected population at the time when infection peaks?
(f) About what is the size of the *susceptible* population at the time when infection peaks? (This value of S is called the *threshold value*, to be discussed in Sect. 1.4 below.)
(g) About how many people became infected over the course of the first 20 days? (Note: this is *not* the same as the number who are infected on day 20. Hint: look at the S curve.)

2. The graph below again depicts a disease evolving according to the SIR model.



In this graph, the initial values $S(0)$, $I(0)$ and $R(0)$, and the transmission coefficient a, are the same as in the graph in Exercise 1 above. But the two graphs correspond to different recovery coefficients b.
Which of these two graphs corresponds to the *larger* value of b? Please explain.

3. A town of population 100,000 is hit with a measles epidemic, which evolves according to the above SIR equations. This unique strain of the measles is known to last for twelve days.

(a) What is the recovery coefficient b, and what are the units for b? Please explain.

Now suppose that, on day 15, 14,893 people are susceptible (that is, $S(15) = 14{,}893$) and 69,613 people are infected ($I(15) = 69{,}613$). Also suppose that, one tenth of a day later, the number of susceptibles has decreased to 13,856.

(b) What, at least approximately, is $S'(15)$ (the rate of change of S at $t = 15$)? What are the units of $S'(15)$? Hint: net change in a quantity (roughly) equals rate of change times elapsed time. (See Eq. (1.2) below.)

(c) What is the transmission coefficient a? (Hint: use information determined in part (b) of this exercise, together with the equation for S' appearing in the SIR model.) What are the units for a?

4. Consider an epidemic that progresses according to the usual SIR model, except that, now, recovered people become susceptible again (and can infect again) after m days. (The model for such a disease is sometimes called the *SIRS model*, to reflect the fact that the recovered population feeds back into the susceptible population.)
Modify the usual SIR equations to reflect this new feature (wherein recovered can become susceptible again). Hints:

(i) Your new equations will look a *lot* like the old ones, but with some *new terms* added on. These terms should account for the facts that, now, on average, $1/m$ of the recovered population gets *added to* susceptible population, and *subtracted from* the recovered population, on any given day.

(ii) Your new equations should involve unspecified parameters a, b, and c, where a and b are as above, and $c = 1/m$.

1.3 Prediction Using SIR

Euler's method amounts to a **big idea** that allows us to use "rate equations," like the above SIR equations, to predict. That big idea is this: if Q is any quantity, varying with time, then between any two instants—a "new" one and an "old" one, say—we have

$$\text{New}\,Q = \text{Old}\,Q + \Delta Q \tag{1.1}$$

where ΔQ denotes the *change* in Q, from the "old" instant to the "new" instant. Moreover, we have

$$\Delta Q \approx Q' \Delta t \tag{1.2}$$

where Δt is the elapsed time, and Q' is the rate of change of Q with respect to time. Equation (1.2) simply says: **net change in a quantity (roughly) equals the rate of change of that quantity, times elapsed time.** And Eq. (1.1) simply says: **the new value of a quantity is the old value plus any change in going from old to new.**

Remark 1.3.1 The "$\approx$" in Eq. (1.2) above means "is approximately equal to." Why are the two sides of Eq. (1.2) only approximately equal to each other? Because Q' itself is (typically) changing. Distance traveled, for example, only equals velocity times time if the velocity is constant over the interval of time in question. One can think of things this way: if the velocity is changing, then reading the speedometer at a given instant will not allow us to predict the distance traveled over the next hour *exactly*—or even very well, most likely. On the other hand, it will probably give a pretty good idea of distance traveled over the next second, say.

The upshot is that Eq. (1.2) should be "pretty accurate" if Δt is small. How small, and how accurate? These questions will be explored as we proceed. For now, we'll understand Eq. (1.2) in the sense—admittedly, a somewhat vague sense—just discussed.

Remark 1.3.2 We pause a for moment to reflect on the (approximate) formula (1.2). In particular, we observe that this formula captures the above prediction principle—"If you know how fast you're going, then you know how far you'll get in a given amount of time"—mathematically, and quite succinctly. Indeed, (1.2) tells us: if you know Q' (that is, "how fast you're going"—remember that Q' is the rate of change of Q), then you know, at least approximately, what the change ΔQ in Q will be (that is, "how far you'll get") after Δt units have transpired (that is, "in a given amount of time").

For these reasons, we give Eq. (1.2) a name: we call it **the prediction equation**. This equation will be central to a large proportion of our studies.

1.3.1 An Example

Now, let's *use* Eqs. (1.1) and (1.2), together with the above SIR equations, to predict, as follows.

Example 1.3.1 Consider a disease that behaves according to the above SIR model. Suppose the initial values $S(0)$, $I(0)$, and $R(0)$ of S, I, and R, at time $t = 0$, are given by

$$S(0) = 500, \quad I(0) = 10, \quad R(0) = 0.$$

As before, we'll take the units of S, I, and R to be persons, and the units of time t to be days. Let's suppose we also know that the transmission and recovery coefficients a and b are given by

$$a = 0.001 \text{ (person-day)}^{-1}, \quad b = 0.2 \text{ day}^{-1}.$$

Use this information to predict $S(4)$, $I(4)$, and $R(4)$, using

(i) stepsize $\Delta t = 2$;
(ii) stepsize $\Delta t = 4$.

Solution.
(i) As we're starting at $t = 0$, and using stepsize $\Delta t = 2$, the first values of S, I and R to be predicted are $S(2)$, $I(2)$, and $R(2)$. Let's begin with $S(2)$. We have

$$\begin{aligned} S(2) &= S(0) + \Delta S && \text{(by (1.1))} \\ &\approx S(0) + S'(0)\Delta t && \text{(by (1.2))} \\ &= S(0) + (-aS(0)I(0))\Delta t && \text{(by the } SIR \text{ equations)} \\ &= 500 + (-0.001 \cdot 500 \cdot 10) \cdot 2 && \text{(plug in numerical values)} \\ &= 500 - 10 = 490. \end{aligned}$$

Note: Because an approximation occurs *somewhere* (anywhere!) in the computation of $S(2)$, the final result of that computation is itself an approximation. So, in spite of the "=" appearing in the last step (and in various other steps) of the above computation, what we have actually found is that $S(2) \approx 490$, and not that $S(2) = 490$.

Next, we compute $R(2)$ (we'll save $I(2)$ for last, because the equation for I' is less simple than the one for R'):

$$\begin{aligned} R(2) &= R(0) + \Delta R && \text{(by (1.1))} \\ &\approx R(0) + R'(0)\Delta t && \text{(by (1.2))} \\ &= R(0) + (bI(0))\Delta t && \text{(by the } SIR \text{ equations)} \\ &= 0 + (0.2 \cdot 10) \cdot 2 && \text{(plug in numerical values)} \\ &= 0 + 4 = 4. \end{aligned}$$

To find $I(2)$, we use the fact that, by assumption, $S + I + R$ is constant. Since, initially (at $t = 0$), this sum equals $500 + 10 + 0 = 510$, we have

$$I(2) = 510 - S(2) - R(2) \approx 510 - 490 - 4 = 16.$$

To summarize our first "step" of part (i) of this example: we've found that

$$S(2) \approx 490, \quad I(2) \approx 16, \quad R(2) \approx 4 \quad \text{(persons)}. \tag{1.3}$$

For the next step—estimating $S(4)$, $I(4)$, and $R(4)$—we imagine now that $t = 2$ is our "old," or starting, value of t, and that $t = 4$ is our "new," or final, value of t. We then proceed as above, using the (approximate) values of $S(2)$, $I(2)$, and $R(2)$ just computed, and summarized in Eq. (1.3). So, by the same kind of reasoning as we used in the first step,

$$\begin{aligned} S(4) &= S(2) + \Delta S \\ &\approx S(2) + S'(2)\Delta t \\ &= S(2) + (-aS(2)I(2))\Delta t \\ &\approx 490 + (-0.001 \cdot 490 \cdot 16) \cdot 2 \\ &= 490 - 15.68 = 474.32. \end{aligned}$$

Note that, this time, we use the symbol "≈" in two different instances—the first time because, as before, net change is only approximately equal to rate of change times elapsed time; the second time because the "old" values of S, I, and R that we're using (that is, the values at $t = 2$) are, themselves, approximations. (And never mind the 32 hundredths of a person who is presumably part of this susceptible population at $t = 4$. While math may be used to model real life, the two aren't the same, which is probably a good thing for that 0.32 of a person.)

Similarly,

$$\begin{aligned} R(4) &= R(2) + \Delta R \\ &\approx R(2) + R'(2)\Delta t \\ &= R(2) + (bI(2))\Delta t \\ &\approx 4 + (0.2 \cdot 16) \cdot 2 \\ &= 4 + 6.4 = 10.4, \end{aligned}$$

and

$$I(2) = 510 - S(4) - R(4) \approx 510 - 474.32 - 10.4 = 25.28.$$

In sum, then: using $\Delta t = 2$, we have found that

$$S(4) \approx 474.32, \quad I(4) \approx 25.28, \quad R(4) \approx 10.4 \quad \text{(persons)}. \tag{1.4}$$

(ii) In much the same manner as above, we find that

$$\begin{aligned} S(4) &= S(0) + \Delta S \\ &\approx S(0) + S'(0)\Delta t \\ &= S(0) + (-aS(0)I(0)) \cdot \Delta t \\ &\approx 500 + (-0.001 \cdot 500 \cdot 10) \cdot 4 \\ &= 500 - 20 = 480, \end{aligned} \qquad \begin{aligned} R(4) &= R(0) + \Delta R \\ &\approx R(0) + R'(0)\Delta t \\ &= R(0) + (bI(0))\Delta t \\ &\approx 0 + (0.2 \cdot 10) \cdot 4 \\ &= 0 + 8 = 8, \end{aligned}$$

and

$$I(4) = (S(0) + I(0) + R(0)) - S(4) - R(4) \approx (500 + 10 + 0) - 480 - 8 = 22.$$

In sum: using $\Delta t = 4$, we find that

$$S(4) \approx 480, \quad I(4) \approx 22, \quad R(4) \approx 8 \quad \text{(persons)}. \tag{1.5}$$

Compare (1.5) with (1.4): not surprisingly, these estimates are different. Again, the smaller stepsize $\Delta t = 2$ yields better results than $\Delta t = 4$, because the rates of change S, I, and R

are themselves continuously changing. And the smaller Δt is, the more often we recalibrate, to adjust for this change.

Using essentially the "Euler's method" algorithm implemented above, but with stepsize $\Delta t = 0.001$, we would find that

$$S(4) \approx 463.57, \quad I(4) \approx 31.30, \quad R(4) \approx 15.13 \quad \text{(persons)}.$$

These numbers are still approximations, but they are closer to the truth.

Note that, to approximate $S(4)$, $I(4)$, and $R(4)$ using stepsize $\Delta t = 0.001 = 1/1000$, we need to compute $S(t)$, $I(t)$, and $R(t)$ at $4 \cdot 1,000 = 4,000$ different values of t. Needless to say, we would not, and did not, do these computations by hand. We used a computer, together with the open-source Sage mathematical software package, which is very similar to MATLAB, Mathematica, and other mathematical software that you may have seen.

In Sect. 1.5, we'll discuss the use of Sage to these ends.

If we could somehow make sense of the above algorithm for the case $\Delta t = 0$, we would, in theory, have an *exact* solution to the SIR equations. As it turns out, the SIR system of rate equations does *not*, in fact, admit an exact, "closed-form" solution, meaning one where $S(t)$, $I(t)$, and $R(t)$ can be written as mathematical expressions in the variable t. Many other interesting "real-life" phenomena do. We'll discuss this further in the course of this text.

1.3.2 Summary: Euler's Method and SIR

Schematically, Euler's method, as applied to the SIR system of rate equations, looks like this (Fig. 1.1).

The starting point of this loop corresponds to using $S(0)$, $I(0)$, and $R(0)$ as the "current" values of S, I, and R. And what about the ending point? If one wants to predict all the way out to time T, say, and if one chooses a stepsize Δt, then one will need cycle through the above loop $T/\Delta t$ times. (Including the initial values $S(0)$, $I(0)$, and $R(0)$, one will then end up with $(T/\Delta t) + 1$ different values of S, I, and R.)

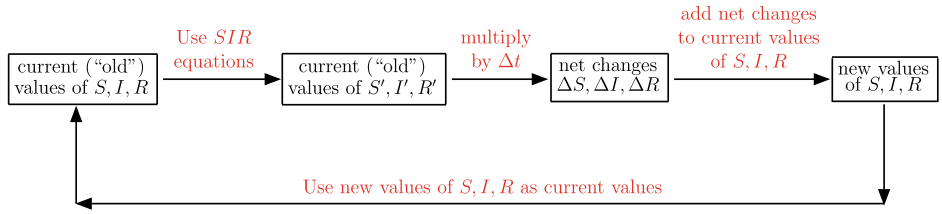


Fig. 1.1 Schematic diagram for Euler's method applied to the SIR equations

1.3.3 Exercises

1. We consider a measles epidemic with transmission coefficient $a = 0.000005$ (person-day)$^{-1}$, and recovery coefficient $b = 1/14$ day^{-1}. This epidemic is then modeled by the equations

$$\begin{aligned} S' &= -0.000005\, SI, \\ I' &= \quad 0.000005\, SI - I/14, \\ R' &= \qquad\qquad\qquad I/14. \end{aligned}$$

   We assume that the initial values of S, I, and R are:

$$S(0) = 45,400, \quad I(0) = 2,100, \quad R(0) = 2,500.$$

   (a) Calculate the "current" rates of change $S'(0)$, $I'(0)$, and $R'(0)$, and use these rates of change to estimate $S(1)$, $I(1)$, and $R(1)$.
   (b) Using the values of $S(1)$, $I(1)$, and $R(1)$ found in the previous exercise, calculate $S'(1)$, $I'(1)$, and $R'(1)$, and use these rates of change to estimate $S(2)$, $I(2)$, and $R(2)$.
   (c) Using the values of $S(2)$, $I(2)$, and $R(2)$ found in the previous exercise, calculate $S'(2)$, $I'(2)$, and $R'(2)$, and use these rates of change to estimate $S(3)$, $I(3)$, and $R(3)$.

2. Go back to the starting time $t = 0$, and to the initial values and parameter values specified in Exercise 1 above. Recalculate the values of S, I, and R at time $t = 2$, this time using stepsize $\Delta t = 2$. You should perform only a single round of calculations, using the rates $S'(0)$, $I'(0)$, and $R'(0)$. Which estimates of $S(2)$, $I(2)$, and $R(2)$ do you think are "better:" those of this exercise, or those of Exercise 1(b) above?
3. Repeat Exercise 1 above, with the same initial values of S, I, and R, and the same transmission coefficient, but this time with recovery coefficient $b = 1/5$ day^{-1}.
   How do your values of $S(3)$, $I(3)$, and $R(3)$ obtained here compare with those of Exercise 1(c) above? Explain these changes from a modeling perspective. That is: by considering the "real life" meaning of b, explain why it makes sense that changing b from $1/14$ to $1/5$ would result in the kinds of changes you see in $S(3)$, $I(3)$, and $R(3)$.
4. Repeat Exercise 1 above, with the same initial values of S, I, and R, and the same recovery coefficient, but this time with transmission coefficient $a = 0.00001$ (person-day)$^{-1}$.
   How do your values of $S(3)$, $I(3)$, and $R(3)$ obtained here compare with those of Exercise 1(c) above? Explain these changes from a modeling perspective. That is: by considering the "real life" meaning of b, explain why it makes sense that changing a from 0.000005 to 0.00001 would result in the kinds of changes you see in $S(3)$, $I(3)$, and $R(3)$.
5. Repeat Exercise 1 above for a disease that evolves according to the $SIRS$ model of Exercise 4 of Sect. 1.2.3. Use the same initial values of S, I, and R, and the same transmission and recovery coefficients, as in Exercise 1 above, but this time, also assume

a time interval of 10 days between recovery and becoming susceptible again. (That is, in the language of Exercise 4 of Sect. 1.2.3, assume $m = 10$.)
How do your values of $S(3)$, $I(3)$, and $R(3)$ obtained here compare with those of Exercise 1(c) above? Explain these changes from a modeling perspective. That is: by considering the "real life" differences between the SIR and $SIRS$ models, explain why it makes sense that changing from one to the other would result in the kinds of changes you see in $S(3)$, $I(3)$, and $R(3)$.

1.4 More on the SIR Model

We next consider some basic implications of the above SIR model—implications that do not require Euler's method or any similar iterative process.

1.4.1 Threshold Value S_T of S

As noted earlier, the susceptible population only decreases in size. Eventually, this population will become small enough that it can no longer sustain growth in the infected population (assuming the latter population is, initially, growing). At this point, I will *peak*, and thereafter will dwindle.

The question that we wish to address is: how small is small enough? How small *does* the susceptible population need to become before I peaks, and begins to decline? We'll answer in a moment, but first, let's give a *name* to this particularly important value of S.

Definition 1.4.1 In the context of the above SIR model, the *threshold value of S*, denoted S_T, is the value of S at which I peaks.

Figure 1.2 gives a graphical interpretation of S_T. We can also use our above SIR equations—in particular, the equation for I'—to deduce a *formula* for S_T, as follows.

The condition "I peaks" in the above definition of S_T indicates that I changes from *increasing* to *decreasing*. In terms of rates of change, this means I' changes from *positive* to *negative*. Now it stands to reason that, at a point where a quantity changes from being positive to being negative, it must equal *zero*. (This reasoning fails if the quantity in question has some kind of sudden "jump" from a positive to a negative value, but let's assume this is not the case. In our above model, there's no reason to think I' would jump so abruptly.)

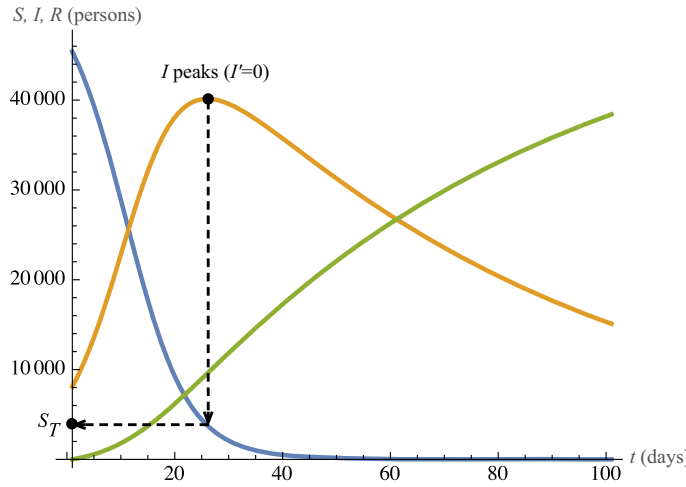


Fig. 1.2 Meaning of the threshold value S_T of S

So: S_T is the value of S where $I' = 0$. But by the above SIR equations, $I' = aSI - bI$. So we see that S_T satisfies the equation

$$aS_T I - bI = 0.$$

Factoring out the I gives

$$I(aS_T - b) = 0.$$

Assuming $I \neq 0$, so we can divide both sides of this equation by I to get

$$aS_T - b = 0.$$

Solving for S_T gives our final formula for the threshold value of S:

$$\boxed{S_T = \frac{b}{a}}$$

Threshold value S_T of S

For instance, in Example 1.3.1 above we have

$$S_T = \frac{0.2}{0.001} = 200.$$

In other words, as soon as the susceptible population has decreased from its initial value $S(0) = 500$ to just 200 remaining susceptible persons, the disease (specifically, the infected population) reaches its peak, and thereafter starts to wane.

Remark 1.4.1 At any point in time t where $S(t)$ is *smaller* than the threshold value S_T, I will be decreasing (as long as it is nonzero to begin with). To see this, suppose $S(t) < S_T$; that is, $S(t) < b/a$. Then, multiplying both sides by a, we get $aS(t) < b$, so $aS(t) - b < 0$, so

$$I'(t) = aS(t)I(t) - bI(t) = I(t)(aS(t) - b) < 0,$$

meaning $I'(t)$ is negative, and therefore I is decreasing at time t (and thereafter, since I can peak at most once).

Similarly, if $S(t)$ is *larger* than S_T, then I will be increasing at time t (but will start to decrease as soon as S becomes smaller than S_T).

1.4.2 Herd Immunity

One very nice application of the threshold value is to the understanding of *herd immunity*. This is the phenomenon where enough people have achieved immunity (through having been infected, through vaccination, etc.) that infection tapers off.

Herd immunity is often understood in terms of the following question: what *fraction*, or *proportion*, of the original susceptible population must become immune in order to ensure that the infected population decreases with time?

To answer note that, if a fraction f of the original $S(0)$ susceptible persons are immune, then the remaining number of susceptible people is $S(0) - fS(0)$. By Remark 1.4.1, infection will decrease if this number is less than $S_T = b/a$. This gives us an inequality that we can solve for f, as follows:

$$\begin{aligned}
S(0) - fS(0) &< \frac{b}{a} \\
S(0) &< \frac{b}{a} + fS(0) \\
S(0) - \frac{b}{a} &< fS(0) \\
\frac{S(0) - b/a}{S(0)} &< f \\
1 - \frac{b}{aS(0)} &< f.
\end{aligned}$$