9

Euler's Numerical Method

In the last chapter, we saw that a computer can easily generate a slope field for a given first-order differential equation. Using that slope field we can sketch a fair approximation to the graph of the solution y to a given initial-value problem, and then, from that graph, we find find an approximation to $y(x)$ for any desired x in the region of the sketched slope field. The obvious question now arises: Why not let the computer do all the work and just tell us the approximate value of $y(x)$ for the desired x ?

Well, why not?

In this chapter, we will develop, use, and analyze one method for generating a "numerical solution" to a first-order differential equation. This type of "solution" is not a formula or equation for the actual solution $y(x)$, but two lists of numbers,

$$\{ x_0 , x_1 , x_2 , x_3 , \ldots , x_N \} \qquad \text{and} \qquad \{ y_0 , y_1 , y_2 , y_3 , \ldots , y_N \}$$

with each y_k approximating the value of $y(x_k)$. Obviously, a nice formula or equation for $y(x)$ would be usually be preferred over a list of approximate values, but, when obtaining that nice formula or equation is not practical, a numerical solution is better than nothing.

The method we will study in this chapter is "Euler's method". It is but one of many methods for generating numerical solutions to differential equations. We choose it as the first numerical method to study because is relatively simple, and, using it, you will be able to see many of the advantages and the disadvantages of numerical solutions. Besides, most of the other methods that might be discussed are refinements of Euler's method, so we might as well learn this method first.

9.1 Deriving the Steps of the Method

Euler's method is based on approximating the graph of a solution $y(x)$ with a sequence of tangent line approximations computed sequentially, in "steps". Our first task, then, is to derive a useful formula for the tangent line approximation in each step.

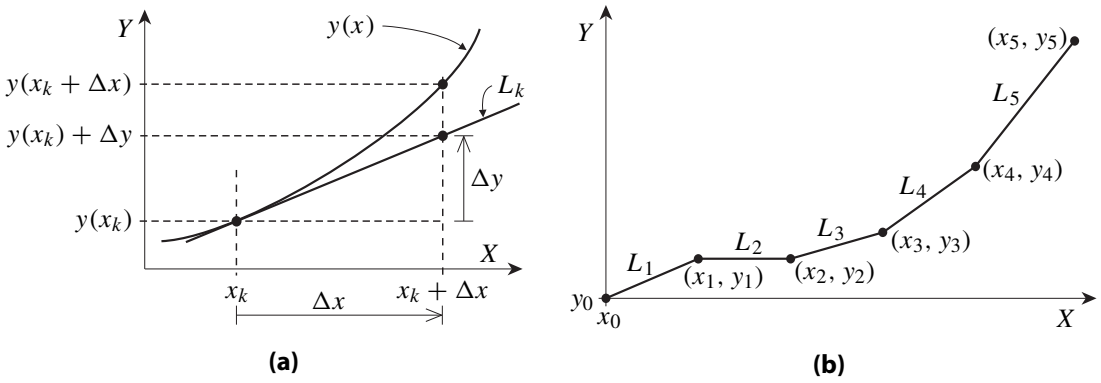


Figure 9.1: **(a)** A single tangent line approximation for the Euler method, and **(b)** the approximation of the solution curve generated by five steps of Euler's method.

The Basic Step Approximation

Let $y = y(x)$ be the desired solution to some first-order differential equation

$$\frac{dy}{dx} = f(x, y) \quad ,$$

and let x_k be some value for x on the interval of interest. As illustrated in figure 9.1a, $(x_k, y(x_k))$ is a point on the graph of $y = y(x)$, and the nearby points on this graph can be approximated by corresponding points on the straight line tangent at point $(x_k, y(x_k))$ (line L_k in figure 9.1a). As with the slope lines in the last chapter, the differential equation can give us the slope of this line:

$$\text{the slope of the approximating line} = \frac{dy}{dx} \text{ at } (x_k, y(x_k)) = f(x_k, y(x_k)) \quad .$$

Now let Δx be any positive distance in the X direction. Using our tangent line approximation (again, see figure 9.1a), we have that

$$y(x_k + \Delta x) \approx y(x_k) + \Delta y$$

where

$$\frac{\Delta y}{\Delta x} = \text{slope of the approximating line} = f(x_k, y(x_k)) \quad .$$

So,

$$\Delta y = \Delta x \cdot f(x_k, y(x_k))$$

and

$$y(x_k + \Delta x) \approx y(x_k) + \Delta x \cdot f(x_k, y(x_k)) \quad . \tag{9.1}$$

Approximation (9.1) is the fundamental approximation underlying each basic step of Euler's method. However, in what follows, the value of $y(x_k)$ will usually only be known by some approximation y_k. With this approximation, we have

$$y(x_k) + \Delta x \cdot f(x_k, y(x_k)) \approx y_k + \Delta x \cdot f(x_k, y_k) \quad ,$$

which, combined with approximation (9.1), yields the approximation that will actually be used in Euler's method,

$$y(x_k + \Delta x) \approx y_k + \Delta x \cdot f(x_k, y_k) \quad . \tag{9.2}$$

The distance Δx in the above approximations is called the *step size*. We will see that choosing a good value for the step size is important.

Generating the Numerical Solution (Generalities)

Euler's method is used to solve first-order initial-value problems. We start with the point (x_0, y_0) where $y_0 = y(x_0)$ is the initial data for the initial-value problem to be solved. Then, repeatedly increasing x by some positive value Δx, and computing corresponding values of y using a formula based on approximation (9.2), we will obtain those two sequences

$$\{ x_0 , x_1 , x_2 , x_3 , \ldots , x_N \} \qquad \text{and} \qquad \{ y_0 , y_1 , y_2 , y_3 , \ldots , y_N \}$$

with $y_k \approx y(x_k)$ for each k. Plotting the (x_k, y_k) points, and connecting the resulting dots with short straight lines leads to a piecewise straight approximation to the graph of the solution $y(x)$ as illustrated in figure 9.1b. For convenience, let us denote this approximation generated by the Euler method by $y_{E,\Delta x}$.

As already indicated, N will denote the number of steps taken. It must be chosen along with Δx to ensure that x_N is the maximum value of x of interest. In theory, both N and the maximum value of x can be infinite. In practice, they must be finite.

The precise steps of Euler's method are outlined and illustrated in the next section.

9.2 Computing Via Euler's Method (Illustrated)

Suppose we wish to find a numerical solution to some first-order differential equation with initial data $y(x_0) = y_0$, say,

$$5\frac{dy}{dx} - y^2 = -x^2 \qquad \text{with} \quad y(0) = 1 \quad . \tag{9.3}$$

(As it turns out, this differential equation is not easily solved by any of the methods already discussed. So if we want to find the value of, say, $y(3)$, then a numerical method may be our only choice.)

To use Euler's method to find our numerical solution, we follow the steps given below. These steps are grouped into two parts: the main part in which the values of the x_k's and y_k's are iteratively computed, and the preliminary part in which the constants and formulas for those iterative computations are determined.

The Steps in Euler's Method
Part I (Preliminaries)

1. Get the differential equation into derivative formula form,

$$\frac{dy}{dx} = f(x, y) \quad .$$

For our example, solving for the derivative formula form yields

$$\frac{dy}{dx} = \frac{1}{5}\left[y^2 - x^2\right] \quad .$$

2. Set x_0 and y_0 equal to the x and y values of the initial data.

In our example, the initial data is $y(0) = 1$. *So*

$$x_0 = 0 \qquad \text{and} \qquad y_0 = 1 \quad .$$

3. Pick a distance Δx for the step size, a positive integer N for the maximum number of steps, and a maximum value desired for x , $x_{\max}$. These quantities should be chosen so that

$$x_{\max} = x_0 + N\Delta x \quad .$$

Of course, you only choose two of these values, and compute the third. Which two are chosen depends on the problem.

For no good reason whatsoever, let us pick

$$\Delta x = \frac{1}{2} \qquad \text{and} \qquad N = 6 \quad .$$

Then

$$x_{max} = x_0 + N\Delta x = 0 + 6 \cdot \frac{1}{2} = 3 \quad .$$

4. Write out the equations

$$x_{k+1} = x_k + \Delta x \tag{9.4a}$$

and

$$y_{k+1} = y_k + \Delta x \cdot f(x_k, y_k) \tag{9.4b}$$

using the information from the previous steps.

For our example,

$$f(x, y) = \frac{1}{5}\left[y^2 - x^2\right] \qquad \text{and} \qquad \Delta x = \frac{1}{2} \quad .$$

So, for our example, equation set (9.4) becomes

$$x_{k+1} = x_k + \frac{1}{2} \tag{9.4a$'$}$$

and

$$\begin{aligned} y_{k+1} &= y_k + \frac{1}{2} \cdot \frac{1}{5}\left[y^2 - x^2\right] \\ &= y_k + \frac{1}{10}\left[{y_k}^2 - {x_k}^2\right] \quad . \end{aligned} \tag{9.4b$'$}$$

Formula (9.4b) for y_{k+1} is based on approximation (9.2). According to that approximation, if $y(x)$ is the solution to our initial-value problem and $y_k \approx y(x_k)$, then

$$y(x_{k+1}) = y(x_k + \Delta x) \approx y_k + \Delta x \cdot f(x_k, y_k) = y_{k+1} \quad .$$

Because of this, each y_k generated by Euler's method is an approximation of $y(x_k)$.

Part II of Euler's Method (Iterative Computations)

1. Compute x_1 and y_1 using equation set (9.4) with $k = 0$ and the values of x_0 and y_0 from the initial data.

   For our example, using equation set (9.4′) with $k = 0$ and the initial values $x_0 = 0$ and $y_0 = 1$ gives us

$$x_1 = x_{0+1} = x_0 + \Delta x = 0 + \frac{1}{2} = \frac{1}{2} \quad ,$$

   and

$$\begin{aligned} y_1 = y_{0+1} &= y_0 + \Delta x \cdot f(x_0, y_0) \\ &= y_0 + \frac{1}{10}\left[{y_0}^2 - {x_0}^2\right] \\ &= 1 + \frac{1}{10}\left[1^2 - 0^2\right] = \frac{11}{10} \quad . \end{aligned}$$

2. Compute x_2 and y_2 using equation set (9.4) with $k = 1$ and the values of x_1 and y_1 from the previous step.

   For our example, equation set (9.4′) with $k = 1$ and the above values for x_1 and y_1 yields

$$x_2 = x_{1+1} = x_1 + \Delta x = \frac{1}{2} + \frac{1}{2} = 1 \quad ,$$

   and

$$\begin{aligned} y_2 = y_{1+1} &= y_1 + \Delta x \cdot f(x_1, y_1) \\ &= y_1 + \frac{1}{10}\left[{y_1}^2 - {x_1}^2\right] \\ &= \frac{11}{10} + \frac{1}{10}\left[\left(\frac{11}{10}\right)^2 - \left(\frac{1}{2}\right)^2\right] = \frac{290}{250} \quad . \end{aligned}$$

3. Compute x_3 and y_3 using equation set (9.4) with $k = 2$ and the values of x_2 and y_2 from the previous step.

   For our example, equation set (9.4′) with $k = 2$ and the above values for x_2 and y_2 yields

$$x_3 = x_{2+1} = x_2 + \Delta x = 1 + \frac{1}{2} = \frac{3}{2} \quad ,$$

   and

$$\begin{aligned} y_3 = y_{2+1} &= y_2 + \frac{1}{10}\left[{y_2}^2 - {x_2}^2\right] \\ &= \frac{29}{250} + \frac{1}{10}\left[\left(\frac{29}{250}\right)^2 - 1^2\right] = \frac{774{,}401}{625{,}000} \quad . \end{aligned}$$

   For future convenience, note that

$$y_3 = \frac{774{,}401}{625{,}000} \approx 1.2390 \quad .$$

(d), (e), ... In each subsequent step, increase k by 1, and compute x_{k+1} and y_{k+1} using equation set (9.4) with the values of x_k and y_k from the previous step. Continue until x_N and y_N are computed.

For our example (omitting many computational details):

With $k+1=4$,

$$x_4 = x_{3+1} = x_3 + \Delta x = \frac{3}{2} + \frac{1}{2} = 2 \quad ,$$

and

$$y_4 = y_{3+1} = y_2 + \frac{1}{10}\left[{y_3}^2 - {x_3}^2\right] = \cdots \approx 1.1676 \quad .$$

With $k+1=5$,

$$x_5 = x_{4+1} = x_4 + \Delta x = 2 + \frac{1}{2} = \frac{5}{2} \quad ,$$

and

$$y_5 = y_{4+1} = y_4 + \frac{1}{10}\left[{y_4}^2 - {x_4}^2\right] = \cdots \approx 0.9039 \quad .$$

With $k+1=6$,

$$x_6 = x_{5+1} = x_5 + \Delta x = \frac{5}{2} + \frac{1}{2} = 6 \quad ,$$

and

$$y_6 = y_{5+1} = y_5 + \frac{1}{10}\left[{y_5}^2 - {x_5}^2\right] = \cdots \approx 0.3606 \quad .$$

Since we had earlier chosen N, *the maximum number of steps, to be* 6, *we can stop computing.*

Using the Results of the Method

What you do with the results of your computations in depends on why you are doing these computations. If N is not too large, it is usually a good idea to write the obtained values of

$$\{x_0, x_1, x_2, x_3, \ldots, x_N\} \qquad \text{and} \qquad \{y_0, y_1, y_2, y_3, \ldots, y_N\}$$

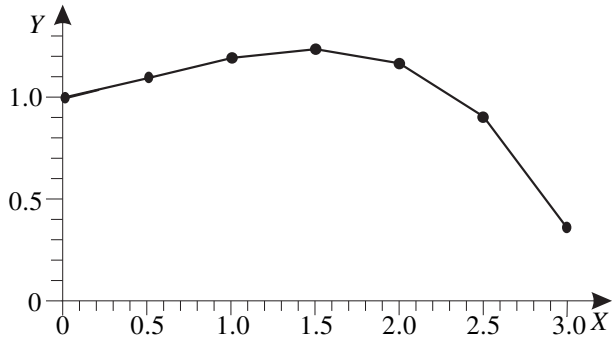
in a table for convenient reference (with a note that $y_k \approx y(x_k)$ for each k) as done in figure 9.2a for our example. And, whatever the size of N, it is always enlightening to graph — as done in figure 9.2b for our example — the corresponding piecewise straight approximation $y = y_{E,\Delta x}(x)$ to the graph of $y = y(x)$ by drawing straight lines between each (x_k, y_k) and (x_{k+1}, y_{k+1}).

On Doing the Computations

The first few times you use Euler's method, attempt to do all the computations by hand. If the numbers become too awkward to handle, use a simple calculator and decimal approximations. This will help you understand and appreciate the method. It will also help you appreciate the tremendous value of programming a computer to do the calculations in the second part of the

k	x_k	y_k
0	0	1
1	0.5	1.1000
2	1.0	1.1960
3	1.5	1.2390
4	2.0	1.1676
5	2.5	0.9039
6	3.0	0.3606

(a)



(b)

Figure 9.2: Results of Euler's method to solve $5y' - y^2 = -x^2$ with $y(0) = 1$ using $\Delta x = {}^1/_2$ and $N = 6$: **(a)** The numerical solution in which $y_k \approx y(x_k)$ (for $k \geq 3$, the values of y_k are to the nearest 0.0001). **(b)** The graph of the corresponding approximate solution $y = y_{E,\Delta x}(x)$.

method. That, of course, is how one should really carry out the computations in the second part of Euler's method.

In fact, Euler's method may already be one of the standard procedures in your favorite computer math package. Still, writing your own version is enlightening, and is highly recommended for the good of your soul.

9.3 What Can Go Wrong

Do not forget that Euler's method does not yield exact answers. Instead, it yields values

$$\{x_0, x_1, x_2, x_3, \ldots, x_N\} \qquad \text{and} \qquad \{y_0, y_1, y_2, y_3, \ldots, y_N\}$$

with

$$y_k \approx y(x_k) \qquad \text{for} \quad k > 0 \quad .$$

What's more, each y_{k+1} is based on the approximation

$$y(x_k + \Delta x) \approx y(x_k) + \Delta x \cdot f(x_k, y(x_k))$$

with $y(x_k)$ being replaced with approximation y_k when $k > 0$. So we are computing approximations based on previous approximations.

Because of this, the accuracy of the approximation $y_k \approx y(x_k)$, especially for larger values of k, is a serious issue. Consider the work done in the previous section: Just how well can we trust the approximation

$$y(3) \approx 0.3606$$

obtained for the solution to initial-value problem (9.3)? In fact, it can be shown that

$$y(3) = -.23699 \quad \text{to the nearest } 0.00001 \quad .$$

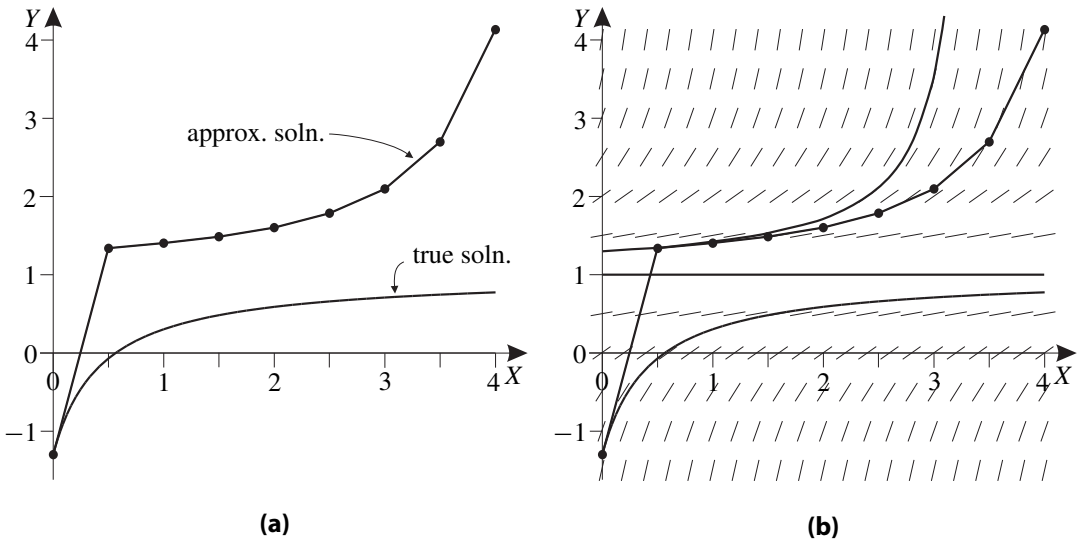


Figure 9.3: Catastrophic failure of Euler's method in solving $y' = (y-1)^2$ with $y(0) = -1.3$: **(a)** Graphs of the true solution and the approximate solution. **(b)** Same graphs with a slope field, the graph of the equilibrium solution, and the graph of the true solution to $y' = (y-1)^2$ with $y(x_1) = y_1$.

So our approximation is not very good!

To get an idea of how the errors can build up, look back at figure 9.1a on page 192. You can see that, if the graphs of the true solutions to the differential equation are generally concave up (as in the figure), then the tangent line approximations used in Euler's method lie below the true graphs, and yield underestimates for the approximations. Over several steps, these underestimates can build up so that the y_k's are significantly below the actual values of the $y(x_k)$'s.

Likewise, if the graphs of the true solutions are generally concave down, then the tangent line approximations used in Euler's method lie above the true graphs, and yield overestimates for the approximations.

Also keep in mind that most of the tangent line approximations used in Euler's method are not based on lines tangent to the true solution, but on lines tangent to solution curves passing through the (x_k, y_k)'s. This can lead to the "catastrophic failure" illustrated in figure 9.3a. In this figure, the true solution to

$$\frac{dy}{dx} = (y-1)^2 \qquad \text{with} \quad y(0) = -\frac{13}{10} \quad ,$$

is graphed along with the graph of the approximate solution generated from Euler's method with $\Delta x = {}^1\!/_2$. Exactly why the graphs appear so different becomes apparent when we superimpose the slope field in figure 9.3b. The differential equation has an unstable equilibrium solution $y = 1$. If $y(0) < 1$, as in the above initial-value problem, then the true solution $y(x)$ should converge to 1 as $x \to \infty$. Here, however, one step of Euler's method overestimated the value of y_1 enough that (x_1, y_1) ended up above equilibrium and in the region where the solutions diverge away from the equilibrium. The tangent lines to these solutions led to higher and higher values for the subsequently computed y_k's. Thus, instead of correctly telling us that

$$\lim_{x\to\infty} y(x) = 1 \quad ,$$

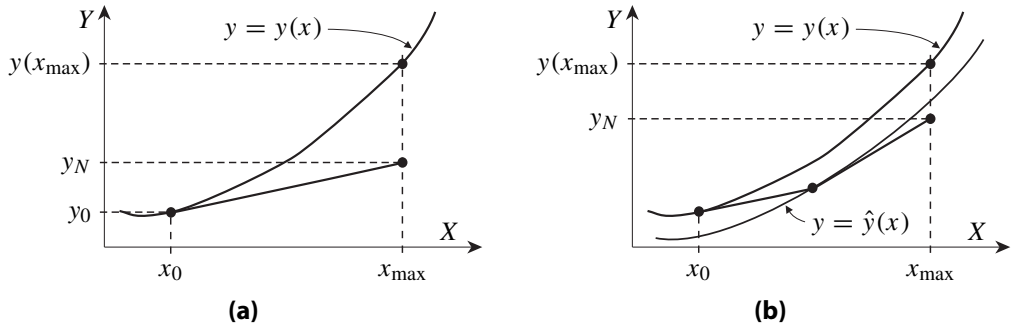


Figure 9.4: Two approximations y_N of $y(x_{\max})$ where y is the solution to $y' = f(x, y)$ with $y(x_0) = y_0$: **(a)** Using Euler's method with Δx equaling the distance from x_0 to $x_{\max}$. **(b)** Using Euler's method with Δx equaling half the distance from x_0 to $x_{\max}$ (Note: $\hat{y}$ is the solution to $y' = f(x, y)$ with $y(x_1) = y_1$.)

this application of Euler's method suggests that

$$\lim_{x \to \infty} y(x) = \infty \quad .$$

A few other situations where blindly applying Euler's method can lead to misleading results are illustrated in the exercises (see exercises 9.6, 9.7, and 9.8, 9.9). And these sorts of problems are not unique to Euler's method. Similar problems can occur with all numerical methods for solving differential equations. Because of this, it is highly recommended that Euler's method (or any other numerical method) be used only as a last resort. Try the methods developed in the previous chapters first. Use a numerical method only if the other methods fail to yield usable formulas or equations.

Unfortunately, the world is filled with first-order differential equations for which numerical methods are the only practical choices. So be sure to skim the next section on improving the method. Also, if you must use Euler's method (or any other numerical method), be sure to do a reality check. Graph the corresponding approximation on top of the slope field for the differential equation, and ask yourself if the approximations are reasonable. In particular, watch out that your numerical solution does not "jump" over an unstable equilibrium solution.

9.4 Reducing the Error

Smaller Step Sizes

Suppose we are applying Euler's method to a given initial-value problem over some interval $[x_0, x_{\max}]$. The one parameter we can adjust is the step size, Δx (or, equivalently, the number of steps, N , in going from x_0 to $x_{\max}$). By shrinking Δx (increasing N), at least two good things are typically accomplished:

1. The error in the underlying approximation

$$y(x_k + \Delta x) \approx y(x_k) + \Delta x \cdot f(x_k, y(x_k))$$

is reduced.

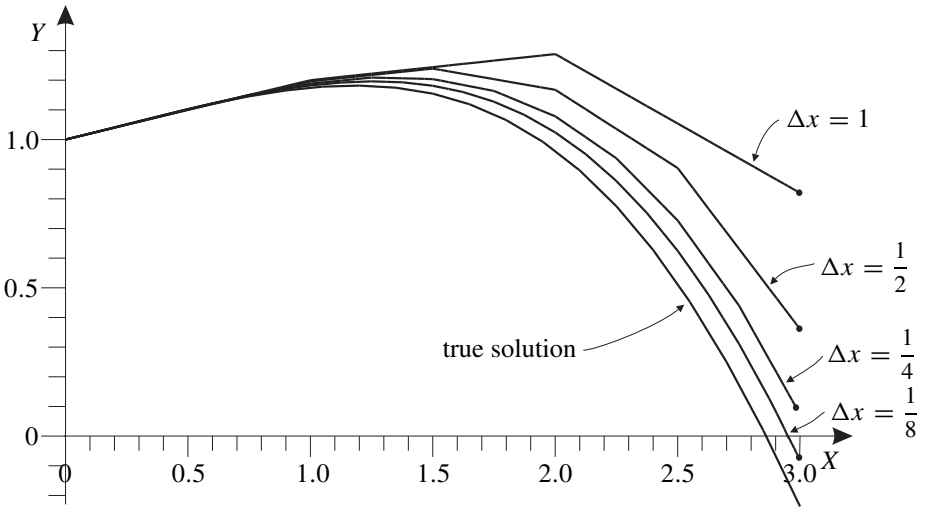


Figure 9.5: Graphs of the different piecewise straight line approximations of the solution to $5y' - y^2 = -x^2$ with $y(0) = 1$ obtained by using Euler's method with different values for the step size $\Delta x = {}^1\!/_2$. Also graphed is the true solution.

2. The slope in the piecewise straight approximation $y = y_{E,\Delta x}(x)$ is recomputed at more points, which means that this approximation can better match the bends in the slope field for the differential equation.

Both of these are illustrated in figure 9.4.

Accordingly, we should expect that shrinking the step size in Euler's method will yield numerical solutions that more accurately approximate the true solution. We can experimentally test this expectation by going back to our initial-value problem

$$5\frac{dy}{dx} - y^2 = -x^2 \qquad \text{with} \quad y(0) = 1 \quad ,$$

computing (as you'll be doing for exercise 9.5) the numerical solutions arising from Euler's method using, say,

$$\Delta x = 1 \quad , \quad \Delta x = \frac{1}{2} \quad , \quad \Delta x = \frac{1}{4} \quad \text{and} \quad \Delta x = \frac{1}{8} \quad ,$$

and then graphing the corresponding piecewise straight approximations over the interval $[0, 3]$ along with the graph of the true solution. Do this, and you will get the graphs in figure 9.5.[1] As expected, the graphs of the approximate solutions steadily approach the graph of the true solution as Δx gets smaller. It's even worth observing that the distance between the true value for $y(3)$ and the approximated value appears to be cut roughly in half each time Δx is cut in half.

In fact, our expectations can be rigorously confirmed. In the next section, we will analyze the error in using Euler's method to approximate $y(x_{\text{max}})$ where y is the solution to a first-order initial-value problem

$$\frac{dy}{dx} = f(x, y) \qquad \text{with} \quad y(x_0) = y_0 \quad .$$

[1] The graph of the "true solution" in figure 9.5 is actually the graph of a very accurate approximation. The difference between this graph and the graph of the true solution is less than the thickness of the curve used to sketch it.

Assuming f is a "reasonably smooth" function of x and y, we will discover that there is a corresponding constant M such that

$$|y(x_{\max}) - y_N| \ < \ M \cdot \Delta x \tag{9.5}$$

where y_N is the approximation to $y(x_{\max})$ generated from Euler's method with step size Δx.

Inequality (9.5) is an *error bound*. It describes the worst theoretical error in using y_N for $y(x_{\max})$. In practice, the error may be much less than suggested by this bound, but it cannot be any worse (unless there are other sources of error). Since this bound shrinks to zero as Δx shrinks to zero, we are assured that the approximations to $y(x_{\max})$ obtained by Euler's method will converge to the correct value of $y(x_{\max})$ if we repeatedly use the method with step sizes shrinking to zero. In fact, if we know the value of M and wish to keep the error below some small positive value, we can use error bound (9.5) to pick a step size, Δx, that will ensure the error is below that desired value. Unfortunately,

1. M can be fairly large.
2. In practice (as we will see), M can be difficult to determine.
3. Error bound (9.5) does not take into account the round-off errors that normally arise in computations.

Let's briefly consider the problem of round-off errors. Inequality (9.5) is only the error bound arising from the theoretically best implementation of Euler's method. In a sense, it is an "ideal error bound" because it is based on all the computations being done with infinite precision. This is rarely practical, even when using a computer math package that can do infinite precision arithmetic — the expressions for the numbers rapidly become too complicated to be usable, even by the computer math packages, themselves. In practice, the numbers must be converted to approximations with finite precision, say, decimal approximations accurate to the nearest 0.0001 as done in the table on page 197.

Don't forget that the computations in each step involve numbers from previous steps, and these computations are affected by the round-off errors from those previous steps. So the ultimate error due to round-off will increase as the number of steps increases. With modern computers, the round-off error resulting from each computation is usually very small. Consequently, as long as the number of steps N remains relatively small, the total error due to round-off will usually be insignificant compared to the basic error in Euler's method. But if we attempt to reduce the error in Euler's method by taking the step size very, very small, then we must take many, many more steps to go from x_0 to the desired $x_{\max}$. It is quite possible to reach a point where the accumulated round-off error will negate the theoretic improvement in accuracy of the Euler method described by inequality (9.5).

Better Methods

Be aware that Euler's method is a relatively primitive method for numerically solving first-order initial-value problems. Refinements on the method can yield schemes in which the approximations to $y(x_{\max})$ converge to the true value much faster as the step size decreases. For example, instead of using the tangent line approximation in each step,

$$y_{k+1} \ = \ y_k \ + \ \Delta x \cdot f(x_k, y_k) \quad ,$$

we might employ a "tangent parabola" approximation that better accounts for the bend in the graphs. (However, writing a program to determine this "tangent parabola", can be tricky.)

In other approaches, the $f(x_k, y_k)$ in the above equation is replaced with a cleverly chosen weighted average of values of $f(x, y)$ computed at cleverly chosen points near (x_k, y_k). The idea is that this yields a straight a line approximation with the slope adjusted to reduce the over- or undershooting noted a page or two ago. At least two of the more commonly used methods, the "improved Euler method" and the "fourth-order Runge-Kutta method", take this approach.

Numerous other methods may also worth learning if you are going to make extensive use of numerical methods. However, an extensive discussion of numerical methods beyond Euler's would take us beyond the brief introduction to numerical methods intended by this author for this chapter. So let us save a more complete discussion of these alternative methods for the future.

9.5 Error Analysis for Euler's Method*

The Problem and Assumptions

Throughout this section we will be concerned with the accuracy of numerical solutions to some first-order initial-value problem

$$\frac{dy}{dx} = f(x, y) \qquad \text{with} \quad y(x_0) = y_0 \quad . \tag{9.6}$$

The precise results will be given in theorem 9.1, somewhere below. For this theorem, L is some finite length, and we will assume there is a corresponding rectangle in the XY–plane

$$\mathcal{R} = \{(x, y) : x_0 \le x \le x_0 + L \quad \text{and} \quad y_{\min} < y < y_{\max}\}$$

such that all of the following holds:

1. f and its first partial derivatives are continuous, bounded functions on $\mathcal{R}$. This "boundedness" means there are finite constants A, B and C such that, at each point in $\mathcal{R}$,

$$|f| \le A \quad , \quad \left|\frac{\partial f}{\partial x}\right| \le B \quad \text{and} \quad \left|\frac{\partial f}{\partial y}\right| \le C \quad . \tag{9.7}$$

2. There is a unique solution, $y = y(x)$, to the given initial-value problem valid over the interval $[x_0, x_0 + L]$. (We'll refer to $y = y(x)$ as the "true solution" in what follows.)

3. The rectangle $\mathcal{R}$ contains the graph over the interval $[x_0, x_0 + L]$ of the true solution.

4. If $x_0 \le x_k \le x_0 + L$ and (x_k, y_k) is any point generated by any application of Euler's method to solve our problem, then (x_k, y_k) is in $\mathcal{R}$.

The rectangle $\mathcal{R}$ may be the entire vertical strip

$$\{(x, y) : x_0 \le x \le x_0 + L \quad \text{and} \quad -\infty < y < \infty\}$$

* Another one of those optional sections for the "interested reader".

if f and its partial derivatives are bounded on this strip. If f and its partial derivatives are not bounded on this strip, then finding the appropriate upper and lower limits for this rectangle is one of the challenges in using the theorem.

Theorem 9.1 (Error bound for Euler's method)
Let f, x_0, y_0, L and $\mathcal{R}$ be as above, and let $y = y(x)$ be the true solution to initial-value problem (9.6). Then there is a finite constant M such that

$$|y(x_N) - y_N| \ < \ M \cdot \Delta x \tag{9.8}$$

whenever

$$\{ x_0 , x_1 , x_2 , x_3 , \ldots , x_N \} \qquad \text{and} \qquad \{ y_0 , y_1 , y_2 , y_3 , \ldots , y_N \}$$

is a numerical solution to initial-value problem (9.6) obtained from Euler's method with step spacing Δx and total number of steps N satisfying

$$0 \ < \ \Delta x \cdot N \ \leq \ L \quad . \tag{9.9}$$

This theorem is only concerned with the error inherent in Euler's method. Inequality (9.8) does not take into account errors arising from rounding off numbers during computation. For a good discussion of round-off errors in computations, the interested reader should consult a good text on numerical analysis

To prove this theorem, we will derive a constant M that makes inequality (9.8) true. (The impatient can look ahead to equation (9.16) on page 207.) Accordingly, for the rest of this section, $y = y(x)$ will denote the true solution to our initial-value problem, and

$$\{ x_0 , x_1 , x_2 , x_3 , \ldots , x_N \} \qquad \text{and} \qquad \{ y_0 , y_1 , y_2 , y_3 , \ldots , y_N \}$$

will be an arbitrary numerical solution to initial-value problem (9.6) obtained from Euler's method with step spacing Δx and total number of steps N satisfying inequality (9.9).

Also, to simplify discussion, let us agree that, in all the following, k always denotes an arbitrary nonnegative integer less than than N.

Preliminary Bounds

Our derivation of a value for M will be based on several basic inequalities and facts from calculus. These include the inequalities

$$|A + B| \ \leq \ |A| \ + \ |B| \qquad \text{and} \qquad \left| \int_a^b \psi(s)\, ds \right| \ \leq \ \int_a^b |\psi(s)| \, ds$$

when $a < b$. Of course, if $|\psi(s)| \leq K$ for some constant K, then, whether or not $a < b$,

$$\int_a^b |\psi(s)| \, ds \ \leq \ K \, |b - a|$$

Also remember that, if $\phi = \phi(x)$ is continuous and differentiable, then

$$\phi(a) \ - \ \phi(b) \ = \ \int_a^b \frac{d\phi}{ds} \, ds \quad .$$

Combining the above, we get

Corollary 9.2
Assume ϕ is a continuous differentiable function on some interval. Assume further that $\phi' \le K$ on this interval for some constant K. Then, for any two points a and b in this interval,

$$|\phi(a) - \phi(b)| \le K|b-a| \quad .$$

We will use this corollary twice.

First, we apply it with $\phi(x) = f(x, y(x))$. Recall that, by the chain rule in chapter 7,

$$\frac{d}{dx} f(x, y(x)) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} \quad ,$$

which we can rewrite as

$$\frac{d}{dx} f(x, y(x)) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} f(x, y)$$

whenever $y = y(x)$ is a solution to $y' = f(x, y)$. Applying bounds (9.7), this then yields

$$\left|\frac{df}{dx}\right| \le \left|\frac{\partial f}{\partial x}\right| + \left|\frac{\partial f}{\partial y}\right| |f(x,y)| \le B + CA \qquad \text{at every point in } \mathcal{R} \quad .$$

The above corollary (with $\phi(x) = f(x, y(x))$ and $K = B + CA$) then tells us that

$$|f(a, y(a)) - f(b, y(b))| \le (B + CA)(b - a) \tag{9.10}$$

whenever $x_0 \le a \le b \le x_0 + L$.

The second application of the above corollary is with $\phi(y) = f(x_k, y)$. Here, y is the variable, x remains constant, and $\phi' = {}^{\partial f}/_{\partial y}$. Along with the fact that $\left|{}^{\partial f}/_{\partial y}\right| < C$ on rectangle $\mathcal{R}$, this corollary immediately gives us

$$|f(x_k, b) - f(x_k, a)| \le C|b - a| \tag{9.11}$$

whenever a and b are any two points in the interval $[x_0, x_0 + L]$.

Maximum Error in the Underlying Approximation

Now consider the error in the underlying approximation

$$y(x_k + \Delta x) \approx y(x_k) + \Delta x \cdot f(x_k, y(x_k)) \quad .$$

Let ϵ_{k+1} be the difference between $y(x_k + \Delta x)$ and the above approximation,

$$\epsilon_{k+1} = y(x_k + \Delta x) - [y(x_k) - \Delta x \cdot f(x_k, y(x_k))] \quad .$$

Note that this can be rewritten both as

$$y(x_{k+1}) = y(x_k) + \Delta x \cdot f(x_k, y(x_k)) + \epsilon_{k+1} \tag{9.12}$$

and as

$$\epsilon_{k+1} = [y(x_k + \Delta x) - y(x_k)] - f(x_k, y(x_k)) \cdot \Delta x \quad .$$

From basic calculus, we know that

$$f(x_k, y(x_k)) \cdot \Delta x \;=\; f(x_k, y(x_k)) \int_{x_k}^{x_k+\Delta x} dx \;=\; \int_{x_k}^{x_k+\Delta x} f(x_k, y(x_k))\, dx \quad .$$

We also know $y = y(x)$ satisfies $y' = f(x, y)$. Hence,

$$y(x_k + \Delta x) - y(x_k) \;=\; \int_{x_k}^{x_k+\Delta x} \frac{dy}{dx}\, dx \;=\; \int_{x_k}^{x_k+\Delta x} f(x, y(x))\, dx \quad .$$

Taking the absolute value of ϵ_{k+1} and applying the last three observations yields

$$\begin{aligned}
|\epsilon_{k+1}| \;&=\; |[y(x_k + \Delta x) - y(x_k)] \;-\; f(x_k, y(x_k)) \cdot \Delta x| \\
&=\; \left| \int_{x_k}^{x_k+\Delta x} f(x, y(x))\, dx \;-\; \int_{x_k}^{x_k+\Delta x} f(x_k, y(x_k))\, dx \right| \\
&=\; \left| \int_{x_k}^{x_k+\Delta x} f(x, y(x)) - f(x_k, y(x_k))\, dx \right| \\
&\leq\; \int_{x_k}^{x_k+\Delta x} |f(x, y(x)) - f(x_k, y(x_k))|\, dx \quad .
\end{aligned}$$

Remarkably, we've already found an upper bound for the integrand in the last line (inequality (9.10), with $a = x$ and $b = x_k$). Replacing this integrand with this upper bound, and then doing a little elementary integration yields

$$|\epsilon_{k+1}| \;\leq\; \int_{x_k}^{x_k+\Delta x} (B + CA)(x - x_k)\, dx \;=\; \frac{1}{2}(B + CA)(\Delta x)^2 \quad .$$

This last inequality combined with equation (9.12) means that we can rewrite the underlying approximation more precisely as

$$y(x_{k+1}) \;=\; y(x_k) \;+\; \Delta x \cdot f(x_k, y(x_k)) \;+\; \epsilon_{k+1} \tag{9.13a}$$

where

$$|\epsilon_{k+1}| \;\leq\; \frac{1}{2}(B + CA)(\Delta x)^2 \quad . \tag{9.13b}$$

Ideal Maximum Error in Euler's Method

Now let E_k be the difference between $y(x_k)$ and y_k,

$$E_k \;=\; y(x_k) \;-\; y_k \quad .$$

Because $y_0 = y(x_0)$:

$$E_0 \;=\; y(x_0) \;-\; y_0 \;=\; 0 \quad .$$

More generally, using formula (9.13a) for $y(x_k + \Delta x)$ and the formula for y_{k+1} from Euler's method, we have

$$\begin{aligned}
E_{k+1} \;&=\; y(x_{k+1}) \;-\; y_{k+1} \\
&=\; y(x_k + \Delta x) \;-\; y_{k+1} \\
&=\; \big[y(x_k) \;+\; \Delta x \cdot f(x_k, y(x_k)) \;+\; \epsilon_{k+1}\big] \;-\; [y_k \;+\; \Delta x \cdot f(x_k, y_k)] \quad .
\end{aligned}$$

Cleverly rearranging the last line and taking the absolute value leads to

$$\begin{aligned} |E_{k+1}| &= |\epsilon_{k+1} + [y(x_k) - y_k] + \Delta x \cdot [f(x_k, y(x_k)) - f(x_k, y_k)]| \\ &= |\epsilon_{k+1} + E_k + \Delta x \cdot [f(x_k, y(x_k)) - f(x_k, y_k)]| \\ &\le |\epsilon_{k+1}| + |E_k| + |\Delta x \cdot [f(x_k, y(x_k)) - f(x_k, y_k)]| \quad . \end{aligned}$$

Fortunately, from inequality (9.13b), we know

$$|\epsilon_{k+1}| \le \frac{1}{2}(B + CA)(\Delta x)^2 \quad ,$$

and from inequality (9.11) and the definition of E_k, we know

$$|f(x_k, y(x_k)) - f(x_k, y_k)| \le C\,|y(x_k) - y_k| = C\,|E_k| \quad .$$

Combining the last three inequalities, we get

$$\begin{aligned} |E_{k+1}| &\le |\epsilon_{k+1}| + |E_k| + |\Delta x \cdot [f(x_k, y(x_k)) - f(x_k, y_k)]| \\ &\le \frac{1}{2}(B + CA)(\Delta x)^2 + |E_k| + \Delta x \cdot C\,|E_k| \\ &\le \frac{1}{2}(B + CA)(\Delta x)^2 + (1 + \Delta x \cdot C)\,|E_k| \quad . \end{aligned}$$

This is starting to look ugly. So let

$$\alpha = \frac{1}{2}(B + CA) \qquad \text{and} \qquad \beta = 1 + \Delta x \cdot C \quad ,$$

just so that the above inequality can be written more simply as

$$|E_{k+1}| \le \alpha(\Delta x)^2 + \beta\,|E_k| \quad .$$

Remember, $E_0 = 0$. Repeatedly applying the last inequality, we then obtain the following:

$$|E_1| = |E_{0+1}| = \alpha(\Delta x)^2 + \beta\,|E_0| = \alpha(\Delta x)^2 \quad .$$

$$\begin{aligned} |E_2| = |E_{1+1}| &\le \alpha(\Delta x)^2 + \beta\,|E_1| \\ &\le \alpha(\Delta x)^2 + \beta\alpha(\Delta x)^2 \le (1 + \beta)\,\alpha(\Delta x)^2 \quad . \end{aligned}$$

$$\begin{aligned} |E_3| = |E_{2+1}| &\le \alpha(\Delta x)^2 + \beta\,|E_2| \\ &\le \alpha(\Delta x)^2 + \beta\,(1 + \beta)\,\alpha(\Delta x)^2 \\ &\le \alpha(\Delta x)^2 + \left(\beta + \beta^2\right)\alpha(\Delta x)^2 \le \left(1 + \beta + \beta^2\right)\alpha(\Delta x)^2 \quad . \end{aligned}$$

$$\vdots$$

Continuing, we eventually get

$$|E_N| \le S_N\,\alpha(\Delta x)^2 \qquad \text{where} \quad S_N = 1 + \beta + \beta^2 + \cdots + \beta^{N-1} \quad . \tag{9.14}$$

You may recognize S_N as a partial sum for a geometric series. Whether you do or not, we have

$$\begin{aligned}(\beta-1)S_N &= \beta S_N - S_N \\ &= \beta\left[1+\beta+\beta^2+\cdots+\beta^{N-1}\right] - \left[1+\beta+\beta^2+\cdots+\beta^{N-1}\right] \\ &= \left[\beta+\beta^2+\cdots+\beta^k\right] - \left[1+\beta+\beta^2+\cdots+\beta^{N-1}\right] \\ &= \beta^N - 1 \quad .\end{aligned}$$

Dividing through by β and recalling what α and β represent then gives us

$$\begin{aligned}S_N\alpha &= \frac{\beta^N-1}{\beta-1}\alpha \\ &= \frac{(1+\Delta x\cdot C)^N-1}{1+\Delta x\cdot C-1}\cdot\frac{B+CA}{2} = \frac{\left[(1+\Delta x\cdot C)^N-1\right](B+CA)}{\Delta x\cdot 2C} \quad .\end{aligned}$$

So inequality (9.14) can be rewritten as

$$|E_N| \le \frac{(1+\Delta x\cdot C)^N-1}{\Delta x\cdot C}\alpha(\Delta x)^2$$

Dividing out one Δx leaves us with

$$|E_N| \le M_{N,\Delta x}\cdot\Delta x \qquad \text{where} \quad M_{N,\Delta x} = \frac{\left[(1+\Delta x\cdot C)^N-1\right](B+CA)}{2C} \quad . \tag{9.15}$$

The claim of theorem 9.1 is almost proven with inequality (9.15). All we need to do now is to find a single constant M such that $M_{N,\Delta x} \le M$ for all possible choices of M and Δx. To this end, recall the Taylor series for the exponential,

$$e^X = \sum_{n=0}^{\infty}\frac{1}{n!}X^n = 1 + X + \frac{1}{2}X^2 + \frac{1}{6}X^3 + \cdots \quad .$$

If $X > 0$ then

$$1 + X < 1 + X + \frac{1}{2}X^2 + \frac{1}{6}X^3 + \cdots = e^X \quad .$$

Cutting out the middle and letting $X = \Delta x\cdot C$, this becomes

$$1 + \Delta x\cdot C < e^{\Delta x\cdot C} \quad .$$

Thus,

$$(1+\Delta x\cdot C)^N < \left[e^{\Delta x\cdot C}\right]^N = e^{N\Delta x\cdot C} \le e^{LC}$$

where L is that constant with $N\Delta x \le L$. So

$$M_{N,\Delta x} = \frac{\left[(1+\Delta x\cdot C)^N-1\right](B+CA)}{2C} < M$$

where

$$M = \frac{(e^{LC}-1)(B+CA)}{2C} \quad . \tag{9.16}$$

And this (finally) completes our proof of theorem 9.1 on page 203.

Additional Exercises

9.1. *Several initial-value problems are given below, along with values for two of the three parameters in Euler's method: step size Δx, number of steps N, and maximum variable of interest x_{max}. For each, find the corresponding numerical solution using Euler's method with the indicated parameter values. Do these problems without a calculator or computer.*

a. $\dfrac{dy}{dx} = \dfrac{y}{x}$ *with* $y(1) = -1$; $\Delta x = \dfrac{1}{3}$ *and* $N = 3$

b. $\dfrac{dy}{dx} = -8xy$ *with* $y(0) = 10$; $x_{max} = 1$ *and* $N = 4$

c. $4x + \dfrac{dy}{dx} = y^2$ *with* $y(0) = 2$; $x_{max} = 2$ *and* $\Delta x = \dfrac{1}{2}$

d. $\dfrac{dy}{dx} + \dfrac{y}{x} = 4$ *with* $y(1) = 8$; $\Delta x = \dfrac{1}{2}$ *and* $N = 6$

9.2. *Again, several initial-value problems are given below, along with values for two of the three parameters in Euler's method: step size Δx, number of steps N, and maximum variable of interest x_{max}. For each, find the corresponding numerical solution using Euler's method with the indicated parameter values. Do these problems with a (nonprogramable) calculator.*

a. $\dfrac{dy}{dx} = \sqrt{2x + y}$ *with* $y(0) = 0$; $\Delta x = \dfrac{1}{2}$ *and* $N = 6$

b. $(1 + y)\dfrac{dy}{dx} = x$ *with* $y(0) = 1$; $N = 6$ *and* $x_{max} = 2$

c. $\dfrac{dy}{dx} = y^x$ *with* $y(1) = 2$; $\Delta x = 0.1$ *and* $x_{max} = 1.5$

d. $\dfrac{dy}{dx} = \cos(y)$ *with* $y(0) = 0$; $\Delta x = \dfrac{1}{5}$ *and* $N = 5$

9.3 a. *Using your favorite computer language or computer math package, write a program or worksheet for finding the numerical solution to an arbitrary first-order initial-value problem using Euler's method. Make it easy to change the differential equation and the computational parameters (step size, number of steps, etc.).*[2,3]

b. *Test your program/worksheet by using it to re-compute the numerical solutions for the problems in exercise 9.2, above.*

9.4. *Using your program/worksheet from exercise 9.3 a with each of the following step sizes, find an approximation for $y(5)$ where $y = y(x)$ is the solution to*

$$\frac{dy}{dx} = \sqrt[3]{x^2 + y^2 + 1} \qquad \text{with} \quad y(0) = 0 \quad .$$

[2] If your computer math package uses infinite precision or symbolic arithmetic, you may have to include commands to ensure your results are given as decimal approximations.

[3] It may be easier to compute all the x_k's first, and then compute the y_k's.

a. $\Delta x = 1$ **b.** $\Delta x = 0.1$ **c.** $\Delta x = 0.01$ **d.** $\Delta x = 0.001$

9.5. *Let y be the (true) solution to the initial-value problem considered in section 9.2,*

$$5\frac{dy}{dx} - y^2 = -x^2 \qquad \text{with} \quad y(0) = 1 \quad .$$

For each step size Δx given below, use your program/worksheet from exercise 9.3 a to find an approximation to $y(3)$. Also, for each, find the magnitude of the error (to the nearest 0.0001) in using the approximation for $y(3)$, assuming the correct value of $y(3)$ is -0.23699.

a. $\Delta x = 1$ **b.** $\Delta x = \frac{1}{2}$ **c.** $\Delta x = \frac{1}{4}$ **d.** $\Delta x = \frac{1}{8}$

e. $\Delta x = 0.01$ **f.** $\Delta x = 0.001$ **g.** $\Delta x = 0.0001$

9.6. *Consider the initial-value problem*

$$\frac{dy}{dx} = (y-1)^2 \qquad \text{with} \quad y(0) = -\frac{13}{10} \quad .$$

This is the problem discussed in section 9.3 in the illustration of a "catastrophic failure" of Euler's method.

a. *Find the exact solution to this initial-value problem using methods developed in earlier chapters. What, in particular, is the exact value of $y(4)$?*

b. *Using your program/worksheet from exercise 9.3 a, find the numerical solution to the above initial-value problem with $x_{max} = 4$ and step size $\Delta x = {}^1\!/_2$. (Also, confirm that this numerical solution has been properly plotted in figure 9.3 on page 198.)*

c. *Find the approximation to $y(4)$ generated by Euler's method with each of the following step sizes (use your answer to the previous part or your program/worksheet from exercise 9.3 a). Also, compute the magnitude of the error in using this approximation for the exact value found in the first part of this exercise.*

i. $\Delta x = 1$ ***ii.*** $\Delta x = \frac{1}{2}$ ***iii.*** $\Delta x = \frac{1}{4}$ ***iv.*** $\Delta x = \frac{1}{10}$

9.7. *Consider the following initial-value problem*

$$\frac{dy}{dx} = -4y \qquad \text{with} \quad y(0) = 3 \quad .$$

The following will illustrate the importance of choosing appropriate step sizes.

a. *Find the numerical solution using Euler's method with $\Delta x = {}^1\!/_2$ and N being any large integer (this will be more easily done by hand than using calculator!). Then do the following:*

i. *There will be a pattern to the y_k's. What is that pattern? What happens as $k \to \infty$?*

ii. *Plot the piecewise straight approximation corresponding to your numerical solution along with a slope field for the above differential equation. Using these plots, decide whether your numerical solution accurately describes the true solution, especially as x gets large.*

iii. *Solve the above initial-value problem exactly using methods developed in earlier chapters. What happens to* $y(x)$ *as* $x \to \infty$ *? Compare this behavior to that of your numerical solution. In particular, what is the approximate error in using* y_k *for* $y(x_k)$ *when* x_k *is large?*

b. *Now find the numerical solution to the above initial-value problem using Euler's method with* $\Delta x = {}^1/_{10}$ *and* N *being any large integer (do this by hand, looking for patterns in the* y_k*'s). Then do the following:*

i. *Find a relatively simple formula describing the pattern in the* y_k*'s .*

ii. *Plot the piecewise straight approximation corresponding to this numerical solution along with a slope field for the above differential equation. Does this numerical solution appear to be significantly better (more accurate) than the one found in part 9.7 a?*

9.8. *In this problem we'll see one danger of blindly applying a numerical method to solve an initial-value problem. The initial-value problem is*

$$\frac{dy}{dx} = \frac{3}{7-3x} \qquad \text{with} \quad y(0) = 0 \quad .$$

a. *Find the numerical solution to this using Euler's method with step size* $\Delta x = {}^1/_2$ *and* $x_{\text{max}} = 5$ *. (Use your program/worksheet from exercise 9.3 a).*

b. *Sketch the piecewise straight approximation corresponding to the numerical solution just found.*

c. *Sketch the slope field for this differential equation, and find the exact solution the above initial-value problem by simple integration.*

d. *What happens in the true solution as* $x \to {}^7/_3$ *?*

e. *What can be said about the approximations to* $y(x_k)$ *obtained in the first part when* $x_k > {}^7/_3$ *?*

9.9. *What goes wrong with attempting to find a numerical solution to*

$$(y-1)^{2/3}\frac{dy}{dx} = 1 \qquad \text{with} \quad y(0) = 0$$

using Euler's method with, say, step size $\Delta x = {}^1/_2$ *?*