CHAPTER 1

Patterns and Induction

Goals

- Experiment with examples and discover patterns.
- Express conjectures precisely and completely.
- Learn proof by induction.
- Begin to write and communicate proofs.

1.1. Mathematics, Patterns, and Computers

Finding Patterns. In the experience of many people, mathematics is a collection of facts and procedures for solving some very specialized problems. It seems that these procedures have been around for a long time, and it is not really clear how anyone would have come up with them. In this view, learning mathematics is the same as learning to apply the appropriate algorithm to a mathematics problem to get the right answer. As such mathematics becomes the antithesis of creativity and not unlike memorizing poems. But the experience of mathematicians is quite at odds with the above description. When most mathematicians think of doing mathematics they envision doing something very creative, and fun. It is thrilling to work on new problems that no one has yet solved, and it is exciting to make progress in solving them. There are many aspects of the mathematics enterprise, and maybe the most difficult is to formulate fruitful questions. However, the starting point of much of mathematics is the exploration of patterns. These could be patterns of numbers, structures, figures, data, or almost anything else. It is exciting to find patterns, but it is even more interesting when you figure out why the pattern exists. Mathematics differs from many disciplines in its very high standards of proof. Not only do we have to solve the puzzle and find the appropriate patterns, we also have to give a mathematically rigorous reason for why it works. We will learn about some methods of proof as we go along, but first we will explore and find patterns. The key is to try, experiment, and keep an open mind. Don't try to remember a formula. Instead look at your data and find any regularity. This can be frustrating, but after a while you will learn to enjoy the challenge.

Example 1.1. We want to find the following sum:

$$1 + 3 + 5 + 7 + \cdots + 1243.$$

Can you think of any method other than just adding them?

When solving problems that we have not seen before, there are no sure-fire methods that always lead to success. However, there are a number of vague ideas that often help. Two of them are:

- Look at examples.
- Try solving simpler problems.

(4.2.18) Assume that the function $y = f(x)$ is differentiable at $x = \alpha$. What is the relationship between the derivative of f at α and the symmetric derivative of f at α? Given your answer to this question and the result of Problem 4.2.17, comment on the relationship between the set of differentiable functions and the set of functions that have a symmetric derivative at all points.

4.3. Linearization and Approximation of Functions

Introduction. Consider the following questions:

QUESTION 4.11. Approximate $\sin(\frac{\pi}{45}) + \cos(\frac{\pi}{180})$.

QUESTION 4.12. Consider a function $y = f(t)$ for which we know $f(0) = 2$ and for some values of t we know the approximate value for the rate of change of the function:

t	0	.5	1	1.5	2
$f'(t)$	1	2/3	1/2	2/5	1/3

Can you find m and M such that $m \leq f(2) \leq M$ and $M - m$ is not too big?

QUESTION 4.13. Assume that a function $y = f(t)$ walks into your life. What you know about y is that $f(0) = 1$ and that the rate of change of y with respect to t satisfies a pleasing relation:

$$\frac{dy}{dt} = t^2 + y^2. \tag{4.4}$$

Can you approximate $f(0.1)$?

QUESTION 4.14. Recall that, for any positive real number x, $\pi(x)$ denotes the number of prime numbers up to (and including) x (see Definition 3.2 and the related problems in Chapter 3). Can we somehow approximate $\pi(10^9)$ without actually finding all the primes up to 10^9?

In both Questions 4.12 and 4.13 we want to approximate the value of a function for some value of x. However, we do not have a tidy formula for the function under consideration. What we do have is some information about the rate of change of the function. In Question 4.11, we are asked to approximate a number. You most probably turn to your calculator to answer such a question. But how does the calculator approximate this number? This question can easily be turned to a question about approximating functions as well. If we let $f(x) = \sin(4x) + \cos(x)$, then Question 4.11 is asking for an approximate value for $f(\frac{\pi}{180})$. This question will be particularly useful for us because we can compare any answers that we get with those of a calculator (or a computer program like Maple) and use the process to understand how to approach problems where it may not be clear how to use a calculator.

In all these problems our aim is to see the power of calculus in approximating functions. Question 4.14 will be the toughest test case. As soon as you experiment with primes (see Chapter 3) you will be convinced that there is no rhyme or reason to the distribution of primes. Sometimes several primes appear soon after each other and at other times we have to wait for a long period before a new prime appears. In addition, studying primes seems not to have much in common with the study of calculus. The latter is concerned with continuous functions and derivatives while the former is a distinctively discrete process.

QUESTION 4.15. How would you use derivatives and calculus to study the distribution of primes?

One of the aims of this book is to convince you of the power of calculus by giving you a glimpse of the answer.

REMARK 4.16. Question 4.11 is the main example in the remainder of the current chapter. Question 4.12 is Problem 13.1.7, Question 4.13 is Problem 4.4.12, and Questions 4.14 and 4.15 are treated in Chapter 6, Section 7.1, and Chapter 12.

"Good" approximations and errors. The numbers 3.14 and 20 are two approximations for the number π. It is just that one is a much better approximation than the other and, the point is, that whenever you approximate unknown quantities you have to have some idea of how good your approximation is. Otherwise, there will be no way to use your approximation wisely. One (and certainly not the only) way is to formulate an approximation question as in Question 4.12. There instead of asking for an approximation for $f(2)$, we asked for numbers m and M that sandwich $f(2)$ and such that $M-m$ is not too big. Certainly, if we know that the exact value of $f(2)$ satisfies $m \le f(2) \le M$, then we can use $\frac{M+m}{2}$, the value half-way between m and M, as an approximation for $f(2)$ and $\frac{M-m}{2}$ will be the largest possible error in the approximation.

Say that we approximate the quantity A with $\hat{A}$. That is, A is the quantity, $f(2)$ in Question 4.12, whose exact value we do not know and $\hat{A}$ is our suggested approximate value for A. The difference $A-\hat{A}$ will be called the *error* in approximation. If this is a positive number, then $\hat{A}$ is an *underestimate* for A and a negative error implies an *overestimation.* The tricky point is that, unless we know the precise value of A, we do not know the error either and we would have to approximate the error!

The error without the sign, that is $|A-\hat{A}|$ is called[10] the *absolute error* (in approximation). For most scientific purposes it is more appropriate to focus on some kind of a relative error. An error of a few yards when measuring the distance between stars would be quite remarkable and possibly unachievable. The same absolute error would be disastrous in most home remodeling projects. As long as $A \neq 0$, we call

$$\frac{A-\hat{A}}{|A|}$$

the *relative error* (in approximation). If you multiply the relative error by 100, then you get the *percent error* .

The absolute error is related to decimal place accuracy and relative error is related to significant-digit accuracy (see, for example, Maron and Lopez [**56**]). We will explore some of these issues in the problems.

Most of the time, when considering an approximation method, we first try the method on some sample cases where we know the exact values and so can calculate the error. If the results give us some confidence, then we use it on others. However, mathematicians will not be satisfied unless they can prove something about the

[10] The terminology introduced in this section is fairly standard but not uniform. Some authors use slight variations of these terms.

largest possible error. Knowing how big the error could be—even if the actual error could be much smaller—provides a safeguard and is sought after for any good approximation method.

Linearization. A general approach for approximating a function f is to pick, based on the problem at hand, a collection of well-understood and well-behaved functions, $g_1, g_2, \ldots, g_n, \ldots$ and then approximate (or in certain cases find an exact expression for) f with a linear combination of the g's:

$$f \approx c_1 g_1 + c_2 g_2 + \cdots + c_n g_n$$

where $c_1, \ldots, c_n$ are constants. An expression of the form $5g_1 + 2g_2 - \sqrt{2}g_3$ is called a *linear combination* of g_1, g_2, and g_3.

Sometimes for the g's we use sines and cosines (Fourier analysis), and sometimes we use "special functions" (e.g., Bessel functions, Legendre polynomials)[11]. Here we work with the simplest possibility: polynomials. In other words, we try approximating given functions with linear combinations of 1, x, x^2, ...

This is where calculus comes in. One of the main aims of calculus is to provide us with a way of approximating functions with polynomials. The simplest polynomials are linear functions (lines) and approximating functions with lines is called *linearization.* What makes calculus so useful is that linearization and the higher-order approximations can be found in many circumstances where we have incomplete information about the function.

Consider the function $f(x) = \sin(4x) + \cos(x)$. We begin the discussion of approximating functions with this example, where we pretty much have complete information, with the hope that the analysis will point us in the right direction in the more complicated situations. Moreover, Question 4.11 asks that we approximate $f(\frac{\pi}{180})$. To be clear about our goal we will formulate it as follows:

QUESTION 4.17. Let $f(x) = \sin(4x) + \cos(x)$. Find polynomial approximations for $f(x)$ that could be used to approximately evaluate values such as $f(\frac{\pi}{180})$, $f(\frac{9\pi}{40})$, and $f(\frac{\pi}{12})$.

Before we begin, it may be a good idea to look at a graph of the function $f(x) = \sin(4x) + \cos(x)$. Figure 4.7 is a graph of $f(x)$ for $-12 \le x \le 12$. You may want to check some points (e.g., $x = 0, \frac{\pi}{2}, \pi$) to see if you believe the (computer generated) graph.

Now $\frac{\pi}{180} \approx .0175$, $\frac{\pi}{12} \approx .262$, and $\frac{9\pi}{40} \approx .707$. So we don't really need the graph for such a wide range of x's. Zooming in and limiting x between -3 and 3 gives Figure 4.8.

We will come back to Question 4.17 in the next section, but for now we tackle an easier version:

QUESTION 4.18. A line has an equation of the form $y = a + bx$ where a and b are constants. Among all possible lines which one best approximates $f(x) = \sin(4x) + \cos(x)$?

If you look at Figure 4.7 or 4.8, then the most intuitive answer would be that *no* line can approximate this curve. The interesting features of this curve cannot be captured by a straight line. It is not at all clear what we accomplish by answering

[11]We will not discuss special functions in this text. Books and courses on differential equations as well as a number of Physics courses often include an introduction to this subject.

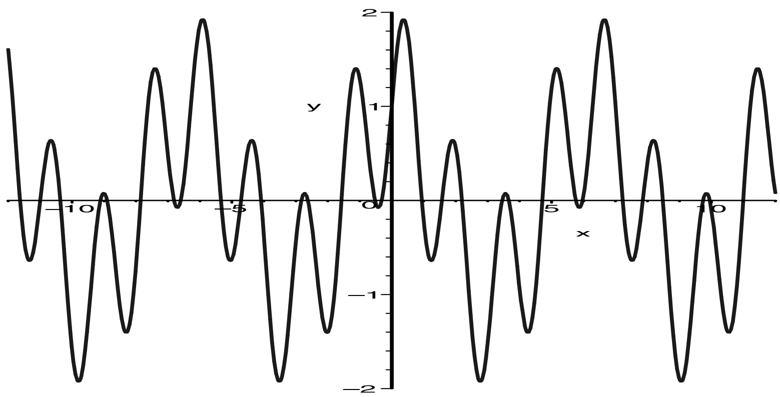

FIGURE 4.7. $y = \sin(4x) + \cos(x)$ for $-12 \leq x \leq 12$

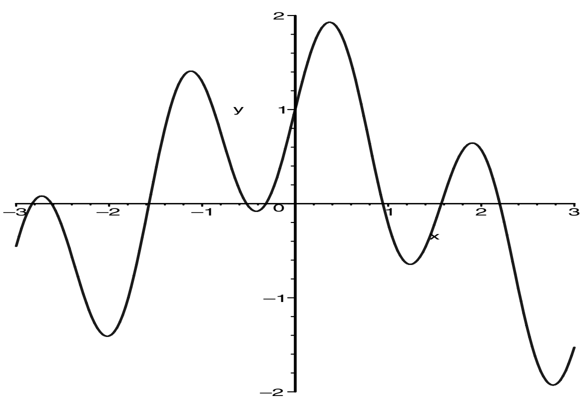

FIGURE 4.8. $y = \sin(4x) + \cos(x)$ for $-3 \leq x \leq 3$

Question 4.18, because any line approximating $f(x)$ would be totally off at most of the points.

Here is the real point of calculus: **The function $f(x) = \sin(4x) + \cos(x)$ is differentiable for all values of x which means that, *sufficiently close to any point,* $f(x)$ is very much like a line.** We can see this graphically if we zoom close enough. Figure 4.9 zooms near the point $x = 0$ (and $y = 1$). In this picture it is clear that the function *can* be approximated by a line as long as we are interested in a local approximation near a point.

By paying careful attention to the definition of the derivative we can make this more precise. We will leave further discussion of this point to the problems (see Problems 4.3.10 through 4.3.13) but we do want to be clear that being like a line when you zoom in at a point is the same as being differentiable at that point.

As another example consider the function

$$y = \frac{5}{4}(x-1)^2 + 2|x| - \frac{5}{4}. \tag{4.5}$$

Figure 4.10 is a graph of the function y given in (4.5). What happens when we zoom in $x = 0$? Figure 4.11 is just such a zoom and notice that at $x = 0$ the

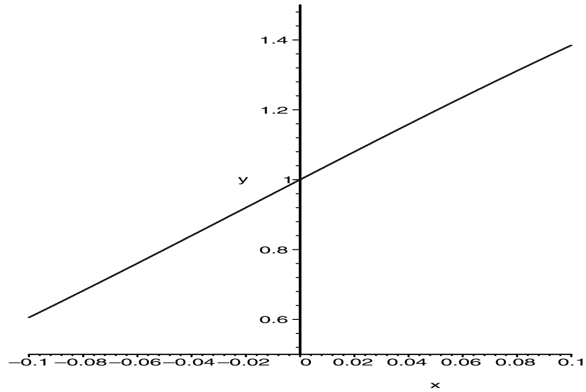

FIGURE 4.9. $y = \sin(4x) + \cos(x)$ for $-.1 \le x \le .1$

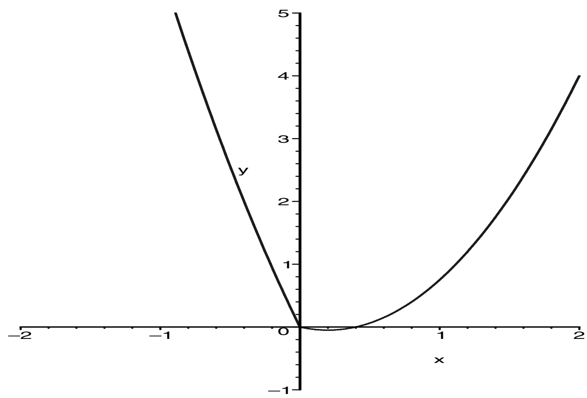

FIGURE 4.10. $y = \frac{5}{4}(x-1)^2 + 2|x| - \frac{5}{4}$ for $-2 \le x \le 2$

graph does not look like a line. It looks like two lines with different slopes coming together. We could not do a good job approximating this function with a single line near $x = 0$. In Problem 4.2.11 you were asked to prove that the function given in (4.5) is not differentiable at $x = 0$.

Now going back to the function $f(x) = \sin(4x) + \cos(x)$ we see that Question 4.18 was not a good question. We *can* approximate $f(x)$ with a line but only *locally*. In other words, as long as we focus very near a given point then we can find a line that approximates $f(x)$ well. Given the fact that we are ultimately interested in the value of the function at $x = .0175, .262, .707$ and we know something about the function at $x = 0$ we can reformulate Question 4.18 as:

QUESTION 4.19. Among all possible lines which one best approximates $f(x) = \sin(4x) + \cos(x)$ "near" $x = 0$?

With your previous knowledge of calculus you probably can answer this question quickly. But let us together work through a possible solution. A line has an equation of the form $y = a + bx$ where a and b are constants. So we basically have to decide on a and b. If we want this line to be a good approximation for $f(x)$ near $x = 0$, then what conditions would we want to impose on this line?

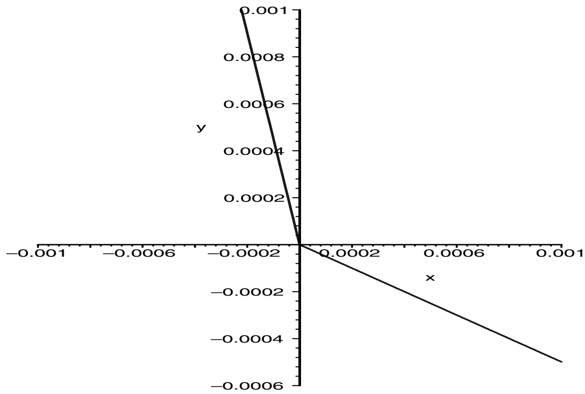

FIGURE 4.11. $y = \frac{5}{4}(x-1)^2 + 2|x| - \frac{5}{4}$ for $-.001 \le x \le .001$

Write our line as $g(x) = a + bx$. We want for $f(x)$ and $g(x)$ to behave similarly near $x = 0$. So it would be reasonable to require that

$$g(0) = f(0). \tag{4.6}$$

Since $g(0) = a$, we should have

$$a = g(0) = f(0) = \sin(0) + \cos(0) = 1.$$

So far we have only required that the line go through $(0, f(0)) = (0, 1)$ which is certainly reasonable. At least we now know that at $x = 0$ the two functions $f(x)$ and $g(x)$ are exactly equal. But we want them to be not that different *near* $x = 0$ as well. One way to assure this is to require that at $x = 0$ both f and g change at the same rate. Then we could argue that these two functions have the same value at $x = 0$, and they change at the same rate, and hence near $x = 0$ they could not be much different.

Of course, a line continues to have the same rate of change as we move off $x = 0$ while the rate of change of $f(x)$ will change as we get off $x = 0$. This is going to mean that eventually $f(x)$ and $g(x)$ will separate off but at least we have made sure that they start at the same place and at the starting point they are changing at the same rate as well. Thus we also require that

$$g'(0) = f'(0). \tag{4.7}$$

Now $g'(x) = b$ and so $g'(0) = b$. Thus we now also know what b should be. We should pick b to be $f'(0)$. $f'(x) = 4\cos(4x) - \sin(x)$ and hence $f'(0) = 4$. We now have our line.

The best line that approximates $f(x) = \sin(4x) + \cos(x)$ near $x = 0$ is

$$y = 1 + 4x. \tag{4.8}$$

As you knew all along, this is the line tangent to the curve at $x = 0$, and, not surprisingly, we could have guessed its equation by looking at Figure 4.9.

Also note that our procedure was very general and hence if $f(x)$ is any function which has a derivate at $x = 0$, then the best line that approximates f near $x = 0$ is given by

$$g(x) = f(0) + f'(0)x. \tag{4.9}$$

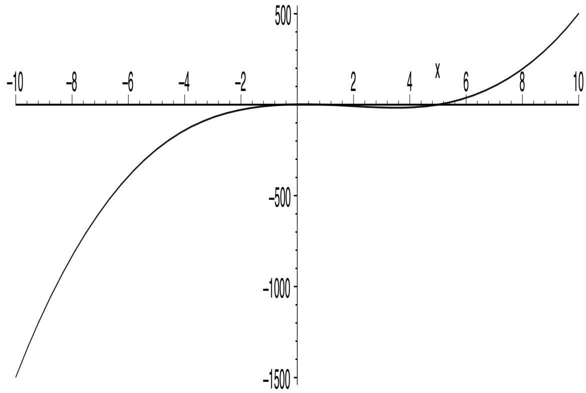

FIGURE 4.12. $y = x^3 + 2\sin(x) - 5x^2 + 2$ for $-10 \le x \le 10$

Since we can use our calculators to evaluate $f(x) = \sin(4x) + \cos(x)$ and it is easy to evaluate $g(x) = 1 + 4x$, we can compare some of their values to see whether we have a good approximation:

x	$f(x)$	$g(x)$	absolute error	percent error
$\frac{\pi}{180} \approx .017453$	1.0696	1.0698	0.0002	0.02%
$\frac{\pi}{12} \approx .26180$	1.8320	2.0472	0.2152	10.51%
$\frac{9\pi}{40} \approx .70686$	1.0694	3.8274	2.7580	72.06%

So it seems that multiplying by four and adding one does give a good approximation for $\sin(4x) + \cos(x)$ as long as x is very close to zero (e.g., when $x = \pi/180$) but by the time x is as far as .2 from 0 the approximation does not work that well. If we want to get better approximations at these further away points, then we have to use something more than a line. For this you have to wait until the next section.

Maple and calculus. Maple can do almost all routine calculus problems. Any calculus problem that has a clear algorithm for its solution can be attempted by Maple. Usually you first have to enter the function. An expression can be entered as:

```
>f:=x^3+2*sin(x)-5*x^2+2;
```

f is now defined to be the expression $x^3 + 2\sin(x) - 5x^2 + 2$. Take the derivative of f with respect to x and call the result g by:

```
>g:=diff(f,x);
```

g is now defined to be the expression $3x^2 + 2\cos(x) - 10x$ which is the derivative of f.

Graphing functions is very easy also:

```
>plot(f,x=-10..10);
```

The above command will give a plot of the function $y = f(x)$ from $x = -10$ to $x = 10$ (see Figure 4.12). Usually you have to experiment with various ranges to find a range for x that gives a reasonable graph of your function. Figure 4.13 is Maple's drawing of the same function but for $-3 \le x \le 6$.

It is just as easy to plot two functions on the same graph. For example, to graph f and its derivative g together, you would use:

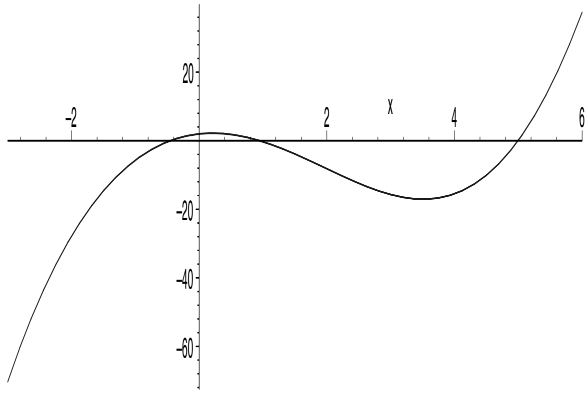

FIGURE 4.13. $y = x^3 + 2\sin(x) - 5x^2 + 2$ for $-3 \leq x \leq 6$

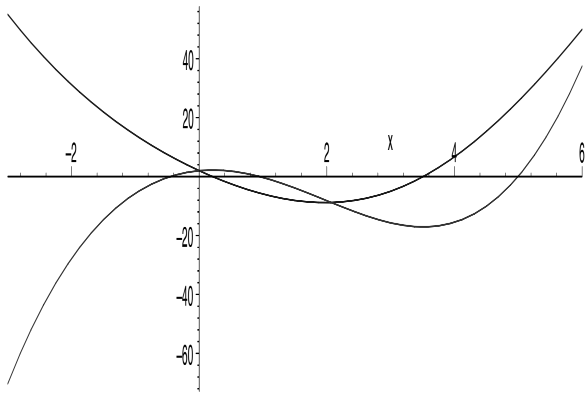

FIGURE 4.14. $f(x) = x^3 + 2\sin(x) - 5x^2 + 2$ and its derivative for $-3 \leq x \leq 6$

```
>plot({f,g},x=-3..6);
```

The result is Figure 4.14.

In all of the above Maple considers f an expression and not a function. If you want to define a function (so as, for example, to substitute expressions for the variable), you do it as follows:

```
>T:=x -> x^2 - 2;
```

Now $T(x) = x^2 - 2$ is the rule for a function. You can now ask Maple for $T(3)$ or $T(y)$ or $T(T(x))$ as follows:

```
>T(3);
>T(y);
>T(T(x));
```

Maple will give you $(x^2 - 2)^2 - 2$ for $T(T(x))$. You can expand this expression by the command `expand(%);` In Maple `%` always stands for the result of the last computation performed.

Problems

(4.3.1) Is rounding transitive? For example, assume that I have a number A and I round it to 5 decimal places and get the number B and then round B to 4 decimal places and get the number C. If I round A to 4 decimal places directly, then do I have to get C? Either prove that you always do or give a counterexample.

Recall that if x is a real number with n digits to the right of the decimal point and if $m < n$, then to *round* x to m decimal places, you look at the value of the $m+1$st digit (to the right of the decimal point). If this digit is less than 5, then you leave the mth digit unchanged and all the following digits are discarded. On the other hand, if the $m+1$st digit is greater or equal to 5, then you add one unit to the mth digit (which may result in changes in other digits as well) and discard the following digits.

(4.3.2) Is it possible to find two numbers A and $\hat{A}$ such that $|A-\hat{A}| < 10^{-5}$ but A and $\hat{A}$ differ in the first digit to the right of the decimal point?

(4.3.3) If $1 \le A \le 5$ and we decide that A is approximately 3, then how big can the absolute error be? How big can the relative error be?

(4.3.4) If $-1 \le A \le 3$ and we decide that A is approximately 1, then how big can the absolute error be? How big can the relative error be?

(4.3.5) Assume that m, A, and M are real numbers, and we know that $0 < m \le A \le M$. We approximate A with $\hat{A}$ where $\hat{A} = \frac{M+m}{2}$.

(a) If $M - m < 10^{-d}$, then how big could the absolute error in the approximation be?

(b) If $\frac{M-m}{m} < 10^{-d}$, then how big could the relative error in the approximation be?

(4.3.6) (a) Graph the function $y = \sin x$ for $0 \le x \le \pi$.

(b) I do not like the function $\sin x$ and would like to approximate it with a linear function: $y = mx + b$. Notwithstanding the comments after Question 4.18, in your opinion, which line best approximates $y = \sin x$ for $0 \le x \le \pi$? Why?

(c) What if I want to approximate $y = \sin x$ with a linear function, but I want the best approximation near the point $x = \frac{\pi}{4}$?

(d) Write your approximation in 4.3.6c as $y = mx + b$ where m and b are numbers written in decimal notation with four significant digits. My calculator claims that $\pi/4 = .785398163$. For $x = .7$ compare the value of $\sin x$ with your linear approximation. Do you have an overestimate or an underestimate? Can you explain this using a graph? What about $x = .785$? Is the approximation still reasonable when $x = .1$? For each of the three values of x, find the absolute and percent errors.

(4.3.7) Let $f(x) = \cos x$, where $-\frac{\pi}{2} \le x \le \frac{\pi}{2}$.

(a) What is the best linear approximation for the function $f(x)$, near the point $x = 0$?

(b) I still would like to approximate $f(x)$, but I want to give myself a bit more flexibility. Instead of looking for lines, that is, equations of the form $y = mx + b$, I would like to look for polynomials of second degree (parabolas) that approximate $f(x)$. Thus I want to find a function of the form $y = a + bx + cx^2$ that best approximates $y = f(x)$ near the point $x = 0$. What would you suggest? Using Maple graph both $\cos x$ and your approximating function for $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$.

(c) What if I wanted a function of the form $y = a + bx + cx^2$ that approximates $y = f(x)$ over the whole interval $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$?

(4.3.8) I do not like the function $f(x) = \sin x$. I would like to have a nicer function that behaves like $f(x)$. This might be asking too much, and so I limit myself to points near $x = 0$. Thus I would like a nice function $g(x)$ such that

$$f(0) = g(0),\ f'(0) = g'(0),\ f''(0) = g''(0), \text{ and } f'''(0) = g'''(0).$$

(a) Can you find a $g(x)$ of the form:

$$g(x) = a_0 + a_1x + a_2x^2 + a_3x^3,$$

where a_0, a_1, a_2 and a_3 are constants?

(b) How about:

$$g(x) = a_0 + a_1x + a_2x^2.$$

(c) How about:

$$g(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4.$$

(d) In any of the above could we have imposed any extra conditions on $g(x)$?

(e) Using your function $g(x)$ in part (a), find $g(.01)$ and compare with $\sin .01$. Do the same for $g(2)$. Any comments? Note that when finding $\sin x$ using your calculator, you have to be in the radian mode.

(f) Using Maple draw (on the same graph) both the function $\sin x$ and your function $g(x)$ from part (a).

(4.3.9) Consider the function $f(x) = \frac{1}{1-x}$.

(a) Proceed as in Problem 4.3.8, and find a polynomial of degree 4 that approximates $f(x)$ closely at $x = 0$, i.e., so that the value of the polynomial and an appropriate number of its derivatives at $x = 0$ agree with the value of $f(x)$ and its derivatives at $x = 0$. Graph both functions on the same graph using Maple.

(b) What if you wanted a polynomial of degree n?

(c) Can you use long division to divide 1 by $1 - x$ and get a quotient with 5 terms and a remainder term? Comment.

Linear approximation of differentiable functions. In Problems 4.3.10 through 4.3.13 we explore the contention on pages 76–79 that to say a function $y = f(x)$ is differentiable at $x = \alpha$ is the same as saying that, sufficiently close to $x = \alpha$, $f(x)$ is very much like a line. In Problem 4.3.13, which relies on the intuition developed in Problems 4.3.11–4.3.12, we see that for differentiable functions the error in linearly approximating the function, in a precise sense, is not too great.

(4.3.10) (a) Is the function $y = \sqrt{4x^2 - 16x + 16}$ differentiable at $x = 2$? Why? Graph this function and explain what about the graph confirms your previous answer.

(b) Now consider $y = \sqrt{4x^2 - 16x + 16.0001}$. Using Maple graph this function for $0 < x < 4$. Looking at the graph, do you think this function is differentiable at $x = 2$? Does zooming in to $1.99 < x < 2.01$ help? If you believe that the function is differentiable at $x = 2$, then find the equation of the line tangent to y at this point.

(c) For the function in the previous part, how far can you ask Maple to zoom in? Experiment and find Maple's limitations.

(4.3.11) I have two functions $e_1(x)$ and $e_2(x)$. They are both zero at $x = 2$, i.e., $e_1(2) = 0 = e_2(2)$. I want to know, for values of x close to 2, which of the two functions is closer to zero. So I calculate and I find

$$\lim_{x\to 2} \frac{e_2(x)}{e_1(x)} = c.$$

(a) Give an example of $e_1(x)$ and $e_2(x)$ such that $c = 3$.

(b) If $c = 3$, then sufficiently close to $x = 2$, what approximately is the relation between $e_1(x)$ and $e_2(x)$? Which one has smaller values (in absolute value)? What if $c = 1/5$? What if $c = 1/100$?

(c) Give an example of $e_1(x)$ and $e_2(x)$ such that $c = 0$. Graph your two functions near $x = 2$ and decide which one is approaching zero faster.

(d) If $c = 0$, then is it true, that sufficiently close to $x = 2$, we have

$$|e_2(x)| \le \frac{1}{100}|e_1(x)|?$$

Do we need the absolute values in the previous expression? What other constants other than $\frac{1}{100}$ can we put in that expression?

(e) If you are hoping that, near $x = 2$, the value of $|e_2(x)|$ is much smaller than $|e_1(x)|$, then what is the best c that you could hope for?

(4.3.12) I have two continuous functions $e_1(x)$ and $e_2(x)$. I know the following information about them:

$$e_1(2) = 0 = e_2(2), \quad \lim_{x\to 2} e_1(x) = 0, \quad \lim_{x\to 2} \frac{e_1(x)}{x-2} = c \neq 0, \quad \lim_{x\to 2} \frac{e_2(x)}{x-2} = 0.$$

(a) Give examples of two functions that satisfy these conditions.

(b) What can you say about $\lim_{x\to 2} e_2(x)$?

(c) Find $\lim_{x\to 2} \frac{e_2(x)}{e_1(x)}$.

(d) As x approaches 0, does $e_1(x)$ or $e_2(x)$ approach zero faster? Is it a real contest?

(4.3.13) We have a function $y = f(x)$ that we know is differentiable at $x = 2$. We find the line tangent to $y = f(x)$ at $x = 2$ and denote it by $t(x)$. To check the contention that differentiability at $x = 2$ means that, near $x = 2$, the function $f(x)$ is very much like the line $t(x)$(see pages 76–79), we calculate $e(x) = f(x) - t(x)$. We are interested in the behavior of $e(x)$ near $x = 2$.

(a) Using $f(2)$ and $f'(2)$ write an expression for $t(x)$ and for $e(x)$.

(b) What are $t(2)$ and $e(2)$?

(c) What is $\lim_{x\to 2} e(x)$?

(d) What is

$$\lim_{x\to 2} \frac{e(x)}{x-2}?$$

(e) What can you say about the behavior of the function $e(x)$ as x approaches zero?

(f) When we say that, close to $x = 2$, the function $y = f(x)$ is very much like the line $y = t(x)$, do we mean anything stronger than $\lim_{x\to 2} f(x) - t(x) = 0$? What?

4.4. Taylor Polynomials

Continuing our discussion of Questions 4.11, 4.17 and 4.19 of the previous section, and in trying to better approximate the function $f(x) = \sin(4x) + \cos(x)$ we ask

QUESTION 4.20. Consider quadratic functions of the form $y = a + bx + cx^2$ where a, b, and c are constants. Among all possible such functions which one best approximates $f(x) = \sin(4x) + \cos(x)$ near $x = 0$?

Let $g_2(x) = a + bx + cx^2$ be the function we are looking for. As we did in the last section it seems reasonable to require that

$$g_2(0) = f(0) \text{ and } g_2'(0) = f'(0). \tag{4.10}$$

Just like before, from these two conditions, we get $a = g_2(0) = f(0) = 1$ and $b = g_2'(0) = f'(0) = 4$. But these conditions impose no requirements on c. In other words, we have an extra degree of freedom and we can impose another condition on g_2. We want g_2 to best approximate f near $x = 0$ and we already know that f and g_2 have the same value and the same rate of change at $x = 0$. A further condition could be that their rate of changes change with the same rate at the origin! They start at the same point and initially they change with the same rate. If we insist that initially these rates of change also change at the same rate then we have a better chance of keeping the two functions together. An easier and more geometric way of saying this is that we could require that the two functions have the same concavity at the origin. That is,

$$g_2''(0) = f''(0). \tag{4.11}$$

In our particular example, it is now straightforward to find c. Try it yourself and you should get $c = -\frac{1}{2}$ and hence the quadratic function that approximates $f(x) = \sin(4x) + \cos(x)$ near $x = 0$ is

$$g_2(x) = 1 + 4x - \frac{1}{2}x^2. \tag{4.12}$$

This is called the *Taylor*[12] *polynomial of degree two* for $\sin(4x) + \cos(x)$ at $x = 0$.

Does this work for any function f or was the solution particular to our example? For this line of argument to work the function has to have a second derivative. Now many everyday functions are *infinitely differentiable* meaning that you can find derivatives of any order, but, there are functions (see Problem 4.2.12) that have a first derivative but not a second derivative. For such functions we cannot find Taylor polynomials of degree 2.

[12]Brook Taylor (1685–1731) published his work on power series, which later became known as Taylor series, in 1715. A number of other mathematicians had already found versions of this result before, and Taylor's own work followed the footsteps of Isaac Newton (1642–1727). Taylor was born in a wealthy family, had private tutors, and, in addition to doing mathematics, was an accomplished musician and painter. Even so, his life was tragic. He died at the age of 46 and lost two wives to childbirth.

CHAPTER 12

The Prime Number Theorem

Goals

- Use calculus ideas to explore the consequences of the Prime Number Theorem.
- Is there a prime between n and $2n$? Can analytic thinking help to answer this question?
- Develop the logarithmic integral from the Prime Number Theorem.
- Informally discuss the Riemann Hypothesis.

12.1. The Prime Number Theorem

Calculus is a powerful tool for approximating functions. This has been the theme of this book and the test case was the attempt to approximate $\pi(x)$, the number of primes up to x.

After experimenting with primes enough to see that their distribution is quite erratic (Chapter 3 and particularly Problem 3.1.8 on prime deserts and the Twin Prime Conjecture 3.3), we proceeded by looking at data to see how the function $\frac{x}{\pi(x)}$ is behaving (see Chapter 6). The function $\frac{x}{\pi(x)}$ is not even a continuous function, but—because we wanted to use calculus and analytic thinking—we decided to approximate it with a differentiable function. We argued—using the data—that if $f(x)$ is the function that is going to approximate $\frac{x}{\pi(x)}$, then we should have $f(ab) = f(a) + f(b)$. Using this functional equation, and an initial condition, we were able to find that the derivative of $f(x)$ had to be $1/x$. Hence, we conjectured that the function $\frac{x}{\pi(x)}$ should behave like $\int_1^x \frac{1}{t} dt$ (Problem 6.1.4). We then called this function, defined by an integral, the natural log function and found its properties (Section 7.1). Finally, putting it all together, by finding $\pi(x)$ in terms of $f(x)$, we conjectured, in Problem 7.1.3, the Prime Number Theorem which we restate here:

THEOREM 12.1 (The Prime Number Theorem). *Let x be a positive real number, and let $\pi(x)$ denote the number of primes up to and including x. Then*

$$\lim_{x \to \infty} \frac{\pi(x)}{x/\ln x} = 1.$$

At the turn of the previous century, the proof of the Prime Number Theorem stood as one of the major achievements of mathematics to date[1].

[1]The Prime Number Theorem was conjectured by Carl Friedrich Gauss (1777–1855) in 1792. However, Gauss did not publish his work on the distribution of primes. The first published statement that is similar to the Prime Number Theorem is due to Adrien-Marie Legendre (1752–1833) who in 1798 conjectured that $\pi(x)$ is of the form $x/(A\ln(x)+B)$ for some constants A and B. The first important result in the direction of a proof was obtained by the Russian mathematician

CHAPTER 13

Local Approximation of Functions and Integral Estimations

Goals

- To see relationships between Taylor polynomial approximations and estimations of integrals using rectangles.
- Appreciate the fact that methods developed from different perspectives may indeed be related.
- Develop systematic methods, including the trapezoid and Simpson rules, for approximating integrals, and compare their effectiveness.
- Use the local approximation paradigm to define curvature of plane curves and Padé approximants.

13.1. Taylor Polynomials and Approximations of Integrals. Are they related?

Consider the following question:

QUESTION 13.1. A mystery function $y = f(x)$ is under consideration. We know that $f'(x) = \sqrt{1 + x^3}$, and $f(0) = 0$. Can you approximate $f(2)$?

Of course, you could try to find an antiderivative for f, and then you would know that f is that antiderivative plus a suitably chosen constant (see Section 5.4). You would then use the initial condition $f(0) = 0$ to find the constant. At that point you would have an exact expression for f and you could plug in $x = 2$ to find $f(2)$. The only catch here is that, even though $\sqrt{1 + x^3}$ looks straightforward enough, its antiderivative is not so easy to find. You may actually want to use Maple to see what it is. Finding antiderivatives is not always practical. In addition, if f' is known only numerically or as a function of x and y, then finding an exact formula could be impossible.

As we have seen in previous chapters, calculus provides many tools and approaches for approximating, and we can approximate $f(2)$ in several ways.

Paradigm I: In the first approach—you can think of it as the local approximation of functions paradigm—we say that we have local information about f at $x = 0$ and we can use this information to approximate f with a (Taylor) polynomial (see Chapter 4). This can be used to approximate $f(2)$. In particular, we can try the following:

- Find a linear approximation to f at zero and use it to approximate $f(2)$.
- Find a linear approximation to f at zero, use it to approximate $f(1)$, then find a linear approximation at $x = 1$ and use it to approximate $f(2)$. You can make the step size even smaller and approximate $f(2)$