from Counting to Calculus

John Taylor Finn

1 January 2001

Part I

Elementary Arithmetic and Geometry

Chapter 1

Whole number arithmetic

1.1 Counting

If you can remember back to when you first learned to count, chances are that you found it much harder to learn to get from 1 to 20 than to go on from 20 to 100. The reason is that what we *call* the numbers between 10 and 20 is more complicated than what we call the ones between 21 and 30, and 31 and 40, and so on.

To get from 1 to 20, first you have to memorize 13 entirely different number names, then a couple that may or may not make you think of 4 and 5. You will probably hear the resemblance between 16 to 19 and 6 to 9, so that's some help, but then when you get to 20, you've got another entirely unfamiliar sound. That's a lot to keep straight, and you may well think this is going to get really hard.

Luckily, though, going from 21 to 29 is a snap, as is going from 30 to 39, 40 to 49, and so on. You may have some trouble with the names thirty and fifty, but you're more likely to recognize these as coming from 3 and 5 than you are 13 and 15.

Once you learn to read numbers, you can see that the irregularities which make counting hard are only in the spoken names for the numbers; the written forms are completely regular, and use just the digits 1 to 9, and 0.

Our way of writing numbers is based on ten because we have ten fingers. If we were to call the full ten fingers a "load", say (oddly enough, we don't really have a word for two hands full), then the number 23 means 2 loads and 3 more fingers. As you keep counting, what do you do after getting to 9 loads and 9 fingers? The answer is that you reach 10 loads; that is 10 loads and 0 fingers or 10 0. But we don't write this number like this, nor do we call it "tenty", although we could. Instead, that we call it one hundred shows that we now think of a load of loads as another kind of thing; a "hundred".

Counting organizes things like we would if we were counting out money, but always exchanging ten singles for a \$10 bill, ten \$10's for a \$100, and, as we keep going, ten \$100's for a \$1000, and so forth.

Thus \$2345 means 2 \$1000 bills, 3 \$100 bills, 4 \$10 bills, and 5 singles.

A nice mechanical example way of our way of counting is a car's odometer. Each wheel has the numbers 0 through 9, and when it has completed a full cycle and is passing from 9 to 0, it causes the next wheel to the left to advance one notch.

If we lived on a planet where we had 6 fingers instead of 10, our odometer wheels would each have the numbers 0 through 6, and our counting would go

$$\begin{array}{rrrrrr} & 1, & 2, & 3, & 4, & 5, \\ 10, & 11, & 12, & 13, & 14, & 15, \\ 20, & 21, & 22, & 23, & 24, & 25, \\ 30, & 31, & 32, & 33, & 34, & 35, \\ 40, & 41, & 42, & 43, & 44, & 45, \\ 50, & 51, & 52, & 53, & 54, & 55, \\ 100, & & & & & \end{array}$$

and so on.

Suppose that you and your friends decided, just for fun, to use this base 6 way of counting and naming numbers for a week. Your first thought might well be "How confusing! —you don't know what any of the bigger numbers really *means*. Sure, you can figure out that 555 really means

$$\begin{array}{rcrcrcl} 5 \times 6^2 & + & 5 \times 6 & + & 5 & = & \\ 5 \times 36 & + & 5 \times 6 & + & 5 & = & \\ 180 & + & 30 & + & 5 & = & 215, \end{array}$$

and so on, but otherwise you don't really know."

But the truth is that we just *think* we know what 215 "really" means! None of us can tell by looking whether a bunch of marbles spread out on the floor is 215, or 200, or 230; the biggest number most humans can identify by sight is 7; any more than that and we have to count. This remains true no matter what counting or numbering method we use.

My point here is that base 10 is not really somehow naturally better for our way of counting than other bases. But that's OK; it's also no worse, and we do have 10 fingers, so why not use them?

A final point about counting is how to name big numbers, like $2,345,567,342$. This, according to the American way, is two billion, three hundred forty-five million, five hundred sixty seven thousand, three hundred forty-two (the English usage books say "three hundred forty-two" is preferred to "three hundred *and* forty-two", so I'm writing it their way, but most of us really include the "and"). The naming scheme starts like this

Number	Name
100	one hundred
1,000	one thousand
10,000	ten thousand
100,000	one hundred thousand
1,000,000	one million
10,000,000	ten million
100,000,000	one hundred million
1,000,000,000	one trillion
10,000,000,000	ten trillion
100,000,000,000	one hundred trillion

A 1 with 12 zeros is a quadrillion, and with 15 zeros it's a quintillion, I think; nobody really much uses these names past a trillion. Astronomers deal with huge numbers, like the number of miles across the galaxy, but they use *scientific notation* instead of bothering with with this-illions and that-illions. They'd write $2,345,567,342$ as 2.346×10^9.

We'll look into this scientific notation again later, in the section on powers and roots, and the one on decimals.

1.2 Adding

In this section we'll only look at the simplest kind of addition, involving just the counting numbers $1, 2, 3, \ldots$, and 0. The whole key to adding, though, even in the loftiest and most abstract levels of math, is the same as it is here in this most elementary setting. This key is that you always and only

Add LIKE to LIKE!

If I have 5 apples and 3 oranges and you have 2 apples and 2 oranges, then between us we have

Picture here: 3.92 truein by .97 truein (adding1 scaled 1000)

In the same way, if I have a couple of $100 dollar bills, 3 $10 bills, and 2 singles, and you have 3 $100's, 5$10's and 4 $1's, then between us we have

Picture here: .96 truein by .76 truein (adding2 scaled 1000)

Notice that even if we couldn't remember the simplest addition facts, and had to count on our fingers, we could still do this addition problem. We'd have to count out that $2 + 3 = 5$, $3 + 5 = 8$, and $2 + 4 = 6$, but that's all.

If we *did* want to memorize some addition facts, to make things quicker, all we'd need to know would be $0 + 0$ through $9 + 9$. Why are these enough? Because our way of writing numbers lets us just do one place at a time, and each place has just a single-digit number in it. Here's another example, where we can add two huge numbers without even figuring how how to say them:

Picture here: 1.58 truein by .8 truein (adding3 scaled 1000)

I've been cheating here, though, by picking numbers where the sum of the digits in each place is less than 10. When that's not the case we have to use *carrying*, which works like this: suppose we want to add $47 and $36. Adding the singles gives us $6 + 7 = 13$, but we can't stick a 13 in the singles place, because each place only has room for a single digit. But that's OK; our 13 is 1 ten and 3 singles, so we could write:

Picture here: .76 truein by 1.18 truein (adding4 scaled 1000)

and $83 is the right answer

Likewise, to add 387 and 876, we could first do $3 + 8 = 11$, then $8 + 7 = 15$, and then $7 + 6 = 13$, and write it down like this

Picture here: .76 truein by 1.39 truein (adding5 scaled 1000)
or maybe like this,
Picture here: .76 truein by 1.39 truein (adding6 scaled 1000)

to make it clear that the 3 plus 8 is really 3 hundred plus 8 hundred, or 11 hundred, and that the 8 and 7 are really 8 tens and 7 tens.

The only problem in doing it this way is that it takes up a lot of room, vertically, and we might start running out of paper. To save room, someone long ago came up with a short-hand way of writing down our intermediate steps, and this is what's called carrying. This carrying shorthand requires us to work from right to left, adding our singles first, then our tens, and so on. Without the carrying shorthand, we could write our sum, worked out from right to left, like this,
Picture here: 2.39 truein by 1.40 truein (adding7 scaled 1000)
With the shorthand, we write this
Picture here: .76 truein by .97 truein (adding8 scaled 1000)

We say that those 1's put in up top have been "carried"; if you were talking to yourself while doing this exampe, you'd say "7 and 6 is 13, so I write down the 3, and caryy the one. Now 1 plus 8 plus 7 is 16, so I put down the 6 and carry the 1. The 1 and 3 and 8 is 12."

Here's another example, written the long way right to left, and then written in carrying shorthad:
Picture here: 3.50 truein by 2.03 truein (adding9 scaled 1000)

Exercise. This carrying business works for adding 3 or more numbers, too. Try it for 2,364 plus 54,478, plus 19,997. Check your answer with a calculator, or by doing it again.

Earlier I said that if we *did* want to memorize some addition facts, to make things quicker, all we'd need to know would be $0+0$ through $9+9$. They used to make you memorize these 100 single digit additons in grade school, using flash cards or whatever to try to get you fast at it. How can I say these would be "all" we'd need, when there's a hundred of them? Memorizing a hundred things sounds like a lot! In fact, if you randomly scrambled up the answers, and asked me to memorize that scrambled table, I'd give up without even trying. But the table of single digit sums is full of nice patterns; look:

Picture here: 4.18 truein by 2.61 truein (addtable scaled 1000)

So if you did start to memorize these 100 simple additions, all the nice patterns would make things easer for you. And you wouldn't have to have anyone explicitly point them out to you, or have to articulate them yourself; look at the 9 row, for example.

Exercise. Actually, there are only about 50 facts in the table, not 100. Explain this, saying exactly how many facts there are.

Anyhow, you *don't* have to memorize the table; you can do addition with carrying even if you have to count the sums in each column on your fingers. Of course, you wouldn't be very fast at adding big numbers, or a lot of numbers, if you did have to count on your fingers, but most of us really just don't have any need to be fast adders. If you have some kind of job where you have to do a lot of adding, you'll probably have some kind of calculator to use. If not, just the practice would make you get faster at it, if you had any enthusiasm.

1.3 Common sense and estimating

Actually, situations do come up for many of us where we do need to add some numbers fairly quickly, but not exactly. Suppose you and three friends are trying to figure out how much money you have among yourselves, and the individual amounts are \$36, \$23, \$17, and \$54. Instead of writing this down and doing a careful addition, you could say to yourself that this is about 40 bucks, plus 20 twice, plus 50. Mentally figuring that 40 plus 20 is 60, plus another 20 makes 80, is no harder than doing 4 plus 2 plus 2. Adding the 80 and 50 is a bit harder, like adding 8 and 5, but you could be lazy and say to yourself that 80 plus 50 is certainly less than 100 plus 50, which is 150. So you've got somewhere between 100 and 150 dollars. If you were careful about seeing that 80 plus 50 is 130, then \$130 is a better estimate.

Being able to do this ball-park estimating is a more practical skill than being a whiz at getting all your adding and carrying just right. We'll come back to this issue of making good rough estimates after we've seen decimals, since if you're in the store getting groceries and not wanting to spend more than you have, you want to be able to come up with reasonable estimate for adding decimal amounts, like \$.67.

Exercise. How good was our quick estimate of $36+23+17+54$? Add these numbers exactly, and compare. Check yourself with a calculator.

Exercise. Make a rough estimate of

$$467 + 234 + 117$$

and check your answer by careful addition.

1.4 Subtracting

Subtracting is the opposite of adding.

Suppose I have just mastered addition, and have done 100 addition problems as my homework, but that some of my work has gotten smudged in such a way that I see my answers, but only part of each question. On my paper I see

$$\begin{aligned} 7 + \bigcirc &= 10 \\ \bigcirc + 9 &= 13 \\ 23 + \bigcirc &= 35, \end{aligned}$$

and I want to figure out what the smudges are, so I can recopy everything neatly and completely. Looking at the first problem, I see that I added *something* to 7, and got 10. Since I'm good at adding, I can figure out by trial and error that the smudge in $7 + * = 10$ has to be 3. That is, I can answer the question "what do you have to add to 7, to get 10?" And this is what the expression $10 - 7$ (pronounced "10 minus 7") means.

Likewise, figuring out what the questions were in the second and third problems amounts to doing the subtractions $13 - 9$ and $35 - 23$.

I remember it taking me a while to get the hang of this, when I was a kid. An addition problem, like $13+12$, just told me what to do. But a subtraction problem, like $13-9$, didn't seem so straightforward; instead of just telling me something to do, instead of just asking the question, it seemed to be telling me the answer, and part of the question, and asking me to figure out what the rest of the question was.

This situation arises again and again in math; you know how to do a certain operation, and when you first learn about, and try to do, the opposite operation , it seems like you're being given the answer and asked to come up with the question. It seems like you're doing something backwards.

After a while, though, with a little practice, you start to see subtraction problems as just as direct as addition problems, and you come to know by heart simple subtractions like $9 - 5$ and $7 - 3$.

We handle more complicated subtractions, like 345 − 214, in much the same way as we do longer addition problems, dealing with the 1's place, the 10's place and so on , one at a time. And sometimes everything is easy:

Picture here: 1 truein by 1 truein (Subtr00)

But what about a subtraction problem like

Picture here: 1 truein by 1 truein (Subtr01)

Here it seems that we first need to subtract 7 from 5, and you can't do that. We have to use *borrowing*, which is to subtracting what carrying is to adding. The 45 in 345 is the same as 30 plus 15, so we can think of 345 as being "33[15]"; that is, 3 hundreds, 3 tens, and 15 ones. Thinking of it this way, we say we've *borrowed* 1 from the 4 to turn 5 into 15. If we now think of 345 as 33[15], our subtraction problem becomes easy:

Picture here: 1 truein by 1 truein (Subtr02)

We don't really rewrite 345, but just indicate our borrowing like this:

Picture here: 1 truein by 1 truein (Subtr03)

We can check that this is correct by adding 128 to 217, and when we do, we see that we carry a 1 after adding 8 and 7. This 1 we carry in adding is the same 1 we borrow in subtracting.

Here's another example. To subract 2787 from 5432 we go like this:

Picture here: 1 truein by 1 truein (Subtr04)

which amounts to thinking of 5432 as 542[12], then as 53[12][12] and finally as 4[13][12][12], as we move from right to left, since we have to borrow 1 three times. Our marks indicating the borrowing actually show 4[13][12][12] clearly.

When we check, by adding 2787 and 2645, we see that we three times carry 1:

Picture here: 1 truein by 1 truein (Subtr05)

It's even easier to see that the subraction problem and the addition problem are opposites if we indicate our carrying the long way:

Picture here: 2 truein by 2 truein (Subtr06)

See? The weirdly written number 4[13][12][12] shows up cleary this way.

Exercise. Work out 5432 minus 2717, and check your answer by adding, doing the carrying the long way. What's the weird way of writing 5432 that you're really using here?

If the number we're subtratcing from has 0's in it, we have to sort of pass the buck. Suppose for instnace that we want to subtract 2467 from 6203. We can't subtract 7 from 3, so we can thing of the guy who's got the 3 ones as turining to the guy who holds the tens and saying "let me borrow a ten." But the tens-guy hasn't got any tens, so he has to borrow a hundred from the hundreds guy, and then give a ten to the ones-guy. Thus, we're thinking of 6203 as 619[13]. Is this right? Well, ignoring the thousands-keeper, we start witn 2 hundreds, 0 tens, and 3 ones, and we're saying that's the same as 1 hundred, 9 tens, and 13 ones. That's correct; they're the same amount of money. Our subtraction now looks like this:

Picture here: 1 truein by 1.3 truein (Subtr07)

If the top number in a subtraction has several 0's, we have to pass the buck several times, like this:

Picture here: 1 truein by 1 truein (Subtr08)

Notice that here I didn't write in the borrowing, but just did it in my head. In fact, I don't ever write the little borrowing numbers, but always do write the carrying. Like me, you can decide which works best for you.

Exercise. Do 50001 minus 22222, check your answer by adding, and write out the weird version of 50001 that we're really using here.

1.5 Multiplying

Multiplication is a short way of dealing with repeated addition; 3 times 5, written 3×5, means $5 + 5 + 5$ (you can also write this as $3 \cdot 5$, but we'll save that notation for the second volume). Notice right away that 5×3 does not mean the same repeated addition as 3×5; 5×3 means $3 + 3 + 3 + 3 + 3$. However, you can see that the two answers are the same, since both give the total number of marbles in 3-by-5 rectangular arrangement of them.

It's important for us to realize that we wouldn't even *have* multiplication unless it was easier to do than the actual repeated addition; if I want to buy 10 cans of cat food at 75ą can, it's going to cost me \$7.50. We can get that answer in our heads in a split second (don't worry if you're not yet sure just how—we'll go over it soon), but if we try to actually carry out the addition

$$\begin{array}{r} 75 \\ +75 \\ +75 \\ +75 \\ +75 \\ +75 \\ +75 \\ +75 \\ +75 \\ +75 \end{array}$$

not only will it take a long time, we're very likely to get it wrong – at least I am.

Just as we can add any whole numbers if we only know the sum of any two single digit ones, we can do any multiplication if we only know the times table for 1×1 through 9×9. So the first thing to do is to memorize the times table, right? —wrong! Again, if you do a lot of multiplying, you'll remember your times table *without* expressly memorizing it. And if you don't do lots of multiplying, you don't need to know in a flash what 8 times 7 is.

Let's do have a look at the multiplication table, like we did the addition table. But first notice that 0 times 7 means how many marbles you have in no piles of 7 marbles, which is still no marbles at all, so $0 \times 7 = 0$. And in fact 0 times anything is just 0. And 1 times any number is just that number back, lie $1 \times 9 = 9$. So let's not even bother to include 0 and 1 in our multiplication table; they're too easy.

Here it is, then:

Picture here: 3truein by 3truein (MultTable)

Please look over this table for a while, noticing patters. Then close the book and go fill out one for yourself.

We can think of multiplication as having a geometric meaning; it's about rectangles. For instance, if I have 4×5 marbles, I can arrange them as a 4-by-5 rectangle:

Picture here: 3truein by 3truein (Mult01)

1.6 Dividing

Just as subtraction is the opposite of addition, dividing is the opposite of multiplication. And since multiplication is repeated addition, its opposite must be repeated subtraction.

Suppose 6 boys have 100 penies they wantto dive evenly among them. Let's further suppose that these boys "don't know any math at all". They can still divide up the pennies by passing them around; "one for Jim, one Joe, one for John, . . . ". What will happen?

Well, each time the go once around, passing to each boy one penny from the pileleft, the subtract 6 from it. They keep subtracting 6 until they wind up with a pile with fewer than 6 pennies, or no pennies at all. We call the number of times they were able to go around, which is the same asthe number of pennies each boy gets, the *quotient*, and the number remaining that they can't divvy up, we call the *remainder*, callingthe remainder 0 if they manage to divvy up the pennies entirely.

If we actually carry this out, we'll see that 6 divides 100, or goes into 100, 16 times, with a remainder of 4. Once we have this answer, we can check it using multiplication and division: we take 16 times 6, and add 4 to it, and sure enough

$$16 \times 6 + 4 = 96 + 4 = 100.$$

If you're goingto have to do a lot of dividing, of course, the question is how do do it quickly. After all, we can do the multiplication 16×6 by adding 6 to itself 16 times (if we *really* want to make it hard!), or by adding 16 to itself 6times (to makethings a little easier. But the really easy way to figure 16 times 6 is to do it by long multiplication. So the question is, what's the way to dovide that's like long multiplication?

Well, it's called long division, and it goes like this:

show l.d. of 6 into 16

What we've done here is the same sort of divde and conquer that qorks for long addition and long multiplication. We first ask how many times does 6 go into 10. This we'll know if we know our 6 timeses, and the answer is 1. So we put down the 1 up in the place for our quotient, and then take 1 times 6 and write the answer, 6, below the 10.

1.7 Powers and roots

Just as multiplication is really a short-hand way to indicate repeated addition, we have *powers* to indicate repeated multiplication.

For instance, we write 3^5, pronounced “three raised to the fifth power”, or “three to the fifth power”, or “three raised to the fifth”, or just “three to the fifth”, as a shorthand for 3 multiplied by itself 5 times:

$$3^5 \quad \text{means} \quad 3 \times 3 \times 3 \times 3 \times 3.$$

People don’t really use powers much in doing ordinary arithmetic, but you use them all the time in doing algebra. Outside of algebra, the most common place you’ll find them is in *scientific notation*, where we write 1.23×10^6 to mean 1.23 times 10 to the 6th power. Since 10 to the 6th power is $1,000,000$ or 1 million, this comes out to $1,230,000$; one million, two hundred thirty thousand. In general 10 raised to a power is just a 1 with that many 0’s after it; $10^9 = 1,000,000,000$, for instance. It’s a lot easier to say “10 to the fifteenth” than it is to try to figure out whether $1,000,000,000,000,000$ is a trillion, or a quadrillion, or whatever. There’s a well-known constant in chemistry called *Avogadro’s number*, which is the number of atoms or molecules in a certain amount of a chemical compound. For instance, in 12 grams of carbon 12 there are Avogadro’s number of atoms. This number is approximately 6.0225×10^{23}; that is, about $602,250,000,000,000,000,000,000$. You can work out that this is called “602 sextillion, 250 quintillion”, but nobody ever calls it that. We’ll look more at this scientific notation in the *Approximating and estimating* section.

Another place you see powers is in “higher arithmetic”, or number theory. A *prime* number is one that nothing divides into except itself and 1, like 2, 3, 5, 7, and 11. If you have a prime number of marbles, you can’t arrange them into a rectangle. An important fact in number theory is that every whole number can be written in exactly one way as a product of powers of primes. In fact, this is called the *Fundamental Theorem of Arithmetic*, and we’ll have a closer look at it in the appendix on lowest terms, primes and factorization.

A couple of examples of prime factorization are

$$1,000,000 = 2^6 \times 5^6, \quad \text{and} \quad 576 = 2^6 \times 3^2.$$

The first of these is easy to figure out, as $1,000,000 = 10^6 = (2 \times 5)^6 = 2^6 \times 5^6$. For the second, we can see that 576 is even, so we divide it by 2 and get $576 = 2 \times 288$. Since 288 us even, we divide by 2 again, and so on:

$$576 = 2 \times 288 = 2 \times 2 \times 144 = 2 \times 2 \times 2 \times 72 = 2 \times 2 \times 2 \times 2 \times 36 = \\ 2 \times 2 \times 2 \times 2 \times 2 \times 18 = 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 9 = 2^6 \times 9,$$

and now we see that $9 = 3^2$, so we're done. There are other ways to go about breaking down 576, but however we proceed, we'll wind up with $2^6 \times 3^2$.

The powers 2 and 3 have special names; we usually say "5 *squared*" for 5^2, instead of "5 raised to the second", and say "2 *cubed*" for 2^3, instead of "2 raised to the third".

To see why this is so, notice that if you have a square 5 feet on a side, you can break it up into $25 = 5^2$ squares 1 foot on a side: Thus,

Picture here: 3truein by 1.5truein (*********** scaled 600)

the area of the sqare is 25 square feet. And if you make a cube 2 feet on a side, you can break it up into $8 = 2^3$ 1-foot cubes; its volume is 2^3 cubic feet.

Exercise. Make a model of a cube 2 feet on a side from modelling clay. Cut it up into little cubes 1 "foot" on a side, and check that you get 8 of them.

If you work (or play) with powers much at all, you can't help but notice that there are some nice patterns that come up. For instance,

$$7^3 \times 7^2 = 7^5, \quad \text{and} \quad 11^4 \times 11^5 = 11^9;$$

multiplying involves adding exponents. Now, there's no need to make a big-deal "rule" out of this, or to memorize anything. If I have 7 times itself 3 times, and then another 2 times, I have it times itself 5 times in all:

$$7^3 \times 7^2 = 7 \times 7 \times 7 \ \times \ 7 \times 7 = 7 \times 7 \times 7 \times 7 \times 7 = 7^5.$$

That is, if you really understand what powers mean, this adding exponents business is a snap.

Likewise, taking powers of powers has a nice pattern:

Exercise. The expression $(7^3)^4$ means 7^3 raised to the 4th power. Figure out what power of just plain 7 this is. Now use the pattern you see to figure these out:

1. $(8^5)^3 = 8^?$,
2. $\left((11^2)^3\right)^4 = 11^?$.

The notion of *root* is the opposite of that of power. We can express the fact that

$$5^4 = 625$$

by saying "5 is the 4th root of 625", just as well as by saying "5 to the 4th power is 625." Similarly, 3 is the square root of 9, and 4 is the cube root of 64. The notation for these roots is

$$\sqrt[2]{9} = \sqrt{9} = 3; \qquad \sqrt[3]{64} = 4.$$

What about something like the square root of 5? Since $2^2 = 4$ and $3^2 = 9$, 2 is too small to be the square root of 5, and 3 is too big. Thus 5 doesn't have a whole-number square root. We can get a decimal approximation by trial and error (or by using a calculator!), but let's wait until after the section on decimals to look into that.

If someone asks you for the square root, or cube root, or fifth root of a whole number, with the understanding that there is a whole number answer, how do you find it? You guess, and use trial and error. For example, suppose someone asks us, as a brain-teaser, to find the cube root of 4913. Since $10^3 = 1000$, which is less than 4913 we know that $\sqrt[3]{4913}$ is bigger than 10. On the other hand, 100 cubed is $(10^2)^3 = 10^6 = 1,000,000$, or a million, which is way bigger than 4913. So $\sqrt[3]{4913}$ is between 10 and 100. Now we keep guessing. If we cube 20 we get 8000; which is too big. And $15^3 = 3375$; too small, but getting closer. No point in trying 16^3, because since 6 is even it'll end in an even digit. How about 17^3? That comes out to —hey! That's it! $\sqrt[3]{4913} = 17$.

Don't ever let anybody make you think that there's anyhting wrong with using trial and error in math. After all, how do you think all the big shots discovered their great results? —by just immediately knowing the answer from the start? Not likely.

Chapter 2

Fractions and their companions

2.1 Fractions

Fractions seem to be the most uncomfortable part of elementary math for most people, the first place in math where they get the sinking feeling that they just really don't understand the reasons behind the rules. This is no mere coincidence; there *are* some subtle points about fractions that you have to come to grips with. In this section we'll bring these points out in the open, and have a good look at them.

Let's start by saying what we do and don't mean by a fraction. What we don't mean is a small length, as in "I was only off by a fraction of an inch." What we do mean is something like 3/5, which we can call

- 3 over 5, or
- 3 divided by 5, or
- the quotient of 3 divided by 5, or
- the ratio of 3 to 5,

it being important here that the top and bottom, 3 and 5, are whole numbers. That is, a fraction has a top and bottom that are whole numbers, and is

- the top over the bottom, or
- the top divided by the bottom, or
- the quotient of the top divided by the bottom, or
- the ratio of the top to the bottom,

these being different ways of describing the same thing.

Since a fraction is the *ratio* of one whole number to another, the official name for a fraction is a *rational number*. Here rational means *ratio*-nal; it does not mean reasonable, as in "it's hard to have a rational discussion with a fanatic."

The mathematical symbol for the collection of all fractions is $\mathbb{Q}$, which stands for Quotient. You might think that $\mathbb{R}$, for Rational, would be a better name, but convention has it that we use $\mathbb{R}$ to mean the Real numbers. We needn't worry now about what a real number is; we'll look into the distinction between rationals and reals in the *What's a number* section.

The fancy names for the top and bottom of a fraction are numerator and denominator. Why bother with these polysyllabic Latin terms? Well, we'll see in a minute that they actually are descriptive, and sometimes helpful to use. We'll stick with the simpler terms top and bottom, though, except when the fancier ones make things clearer.

We've alreay seen above that there are at least four ways to name a fraction, which express the way we write it. But one of the first difficulties with fractions is that is that the way we usually say them does *not* express the way we write them. If I have 3 fifths of whiskey, and you have another 4 fifths, then we don't even have to think about fractions to see that between us we have 7 fifths. In fact, for most of us, a fifth of whiskey just means a certain size bottle, and has nothing to do with fractions at all. So it's easy to understand the equation

$$3 \text{ fifths} + 4 \text{ fifths} = 7 \text{ fifths},$$

even if we've been consuming the objects of our study.

On the other hand, the equation

$$\frac{3}{5} + \frac{4}{5} = \frac{7}{5}$$

may look like Greek to us even if we're stone sober. We usually pronounce $\frac{3}{5}$ as "3 fifths", but that isn't what it looks like; it looks like "3 over 5", or "3 divided by 5", or "the quotient of 3 divided by 5", or "the ratio of 3 to 5", like we listed above.

Imagine trying this at working-class bar. You ask the guy next to you, "here, can you solve this?", and hand him a slip of paper with $\frac{3}{5} + \frac{4}{5} = \_\_\_$ written on it. What do you want to bet he's going to say something like "Get outta here with that &%$#@ math &%$#@! I can't do that &%$#@!" On the other hand, if you ask "suppose you could take just 7 fifths of liquor with you on a long sea voyage; what combination would you make them?", you'll probably get a perfectly articulate expression of an adding-fractions equation like the one you wanted him to write down, or maybe even something more involved, like

$$\frac{2}{5} + \frac{2}{5} + \frac{2}{5} + \frac{1}{5} = \frac{7}{5}$$

Is this really true? Are we really talking fractions when we talk fifths of liquor? Yes, because a fifth here really means one fifth of a gallon; what you get if you divvy up a gallon into 5 equal parts. And a quart of liquor is really a quarter, or 1 fourth, of a gallon, so we're also talking fractions if we talk quarts. (A fifth is really a short quart; whoever thought up bottling fifths must have figured people would think they were getting quarts, and would pay the same price.)

Exercise. Here are a few good questions that we'll figure out exactly later on. For now, even if you can work them out exactly, instead just give your hunches. Just how short of a quart *is* a fifth? What fraction of a quart is a fifth? What percentage? If a quart of the good stuff costs $10, what should we pay for a fifth? If I'm passing off fifths as quarts, by what percent am I ripping off my customers? Maybe we'll even figure out how much of an additional rip-off this "750 ml" nonsense is, since that's a short fifth.

What can we do about this problem of fractions not being written the way we say them? We can use it to our advantage, that's what! Our fifths of whisky example can show us how to do this, and how those fancy Latin terms numerator and denominator can help us out. We start by looking at our liquor store's cash register.

A cash register has different compartments for bills of different *denominations*; one for the singles, one for the 5-dollar bills, and so on. Likewise, we can say that a fifth is kind of denomination of whiskey, as is a quart and a half-gallon. When we talk about 3 fifths of whisky, then, the 3 tells us how many units of this particular denomination we have. That is, the "fifth" in "3 fifths" is the denomination, and the 3 tells how many bottles of this denomination we have, which is to say it counts, or *enumerates*,the fifths. So in "3 fifths" the 3 is the numerator and the fifth is the denomination. And in 3/5, we call the 3 numerator, and the 5 the denominator.

Thus, you don't have to memorize that the numerator is the top, and the denominator the bottom, of a fraction. If you just *say* the fraction, it will be clear which part is telling how many things you have, and which is telling what the things are. Thus when we say 4/7 as "4 sevenths", it's clear that 4 is the numerator, and 7 the denominator.

Notice also that if we want to know the total amount in our cash register, and we see it holds 4 fivers and 3 singles, we don't just add 4 and 3. It's true that we do have $4 + 3 = 7$ bills, but to get the total dollar amount, we have to convert things in terms of a *common denomination*. We do this by noting that the 4 fivers amount to the same as 20 singles, so we have

$$4 \text{ fivers} + 3 \text{ singles} = 20 \text{ singles} + 3 \text{ singles} = 23 \text{ singles} = \$23.$$

Note that here we're using the key we introduced back in the *Adding* section: always and only add LIKE to LIKE.

Exercise. Suppose I tell you that I have 7 bills, all fivers and singles. What are the possible amounts of money I have, from least ($7) to most ($35)?

Now you may well ask what this adding money has to do with *fractions*. Well, to add money of different denominations, we have to covert our amounts into a single common denomination. Likewise, to add fractions with different denominators, we have to convert them to fractions with a common denominator.

To take an example that's still about money, let's see how to add a fifth and a fourth of a dollar. That is, we want to solve

$$\frac{\$1}{5} + \frac{\$1}{4} = \text{how much?}$$

Now a fourth of a dollar is the same as a quarter of a dollar, or a plain old quarter. And a fifth of a dollar is 20 cents, since if we split up a dollar equally among 5 of us, each person will get 20 cents.

One possible common denomination, then, for for $\frac{1}{4}$ of a dollar and $\frac{1}{4}$ dollar, is cents; 25 cents and 20 cents. But another common denomination is nickels; 5 nickels and 4 nickels.

We can get a good picture of using nickels to make up fourths and fifths of a dollar by taking the 20 nickels that make up the dollar, and arranging them in a 4-by-5 rectangle, like this:

Note that each of the 4 rows of the rectangle of the rectangle is a quarter, and each of the 5 columns is a fifth.

One answer to our problem, then, is

$$\frac{\$1}{5} + \frac{\$1}{4} = 4 \text{ nickels} + 5 \text{ nickels} = 9 \text{ nickels}$$

But wait! We were supposed to be doing fractions, and "nickel" doesn't look like a fraction. But we can write nickels as fractions:

$$1 \text{ nickel} = \frac{\$1}{20}.$$

What about 4 nickels? This is 4 twentieths of a dollar, which means

$$4 \times \frac{\$1}{20},$$

which still isn't in the form of one whole number over another. Can we write

$$4 \times \frac{\$1}{20} = \frac{\$4}{20}?$$

This is not a dumb question, because the two sides of this equation *mean* something different; the left-hand side means 4 nickels, while the right-hand side, as it is *written*, is \$4 divided by 4, which means how much you have if you split 4 dollars into 20 equal parts. If we have our 4 dollars in nickels, that's 80 nickels, and if we arrange these 80 nickels into 20 equal stacks, each stack will have 4 nickels. So the two sides are indeed equal in value.

Let me repeat that the question was not a dumb one. And this issue comes up constantly when we're working with fractions. When we write

$$3 \times \frac{1}{7} = \frac{3}{7},$$

we're correct, but the two sides do em not mean the same thing, not as they are *written* —or said, if we call the right-hand side "3 over 7" or "3 divided by 7". The fact of the matter, though, is that we probably call the right-hand side "3 sevenths", which *does* mean the same thing as the left! So the crux of the matter is that calling $\frac{3}{7}$ by the name "3 sevenths" amounts to saying that the equation above is correct. It is correct, and I don't want you to think you have to make a big fuss about this matter. But you should at least a couple of times think it through, and convince yourself that all is well.

Exercise. Draw a pizza cut into 7 equal slices; it'll probably take you more than one try to do this well. Now, in terms of pizzas, explain the difference in meaning between 3 sevenths and 3 divided by 7, and explain why these are the same amount of pizza.

To finish up our problem of adding a fifth and a fourth of a dollar, we write

$$\frac{\$1}{5} + \frac{\$1}{4} = \frac{\$4}{20} + \frac{\$5}{20} = \frac{\$9}{20}.$$

And of we take it as understood that we're talking about dollars, we have

$$\frac{1}{5} + \frac{1}{4} = \frac{4}{20} + \frac{5}{20} = \frac{9}{20}.$$

What about if we're talking not dollars, but donuts? Is a fifth of a donut the same as 4 twentieths? And is this the same as 4 donuts divided into 20 equal parts? Yes:

Exercise. Use picture to explain why, for donuts,

$$\frac{1}{5} = 4 \times \frac{1}{20} = \frac{4}{20}.$$

Having now mastered adding a single fourth to a single fifth, let's go on to bigger things, like

$$\frac{2}{5} + \frac{3}{4} = \text{how much?}$$

If we again think of working with a fifth and a fourth of a dollar, so that they're 4 nickels and 5 nickels, what we have here is $2 \times 4 = 8$ nickels and $3 \times 5 = 15$ nickels, which comes to 23 nickels. That is, we have

$$\frac{2}{5} + \frac{3}{4} = 2 \times \frac{4}{20} + 3 \times \frac{5}{20} = \frac{8}{20} + \frac{15}{20} = \frac{23}{20}.$$

This little string of two equations has several points that are worth taking some time to think about.

For one thing, does $\frac{23}{20}$ bother you? Some people don't like fractions where the top is bigger than the bottom; in fact, some call them *improper fractions*. There's nothing really wrong or improper about them, but math is full of bad-sounding names like that, so you have to get used to it. If we talk about 23 nickels, no one's going to object; there's nothing wrong with 23 nickels. On the other hand, if you do have a pile of 23 nickels, you may well separate off 20 of them as a dollar's worth. The mathematical description of doing this is

$$\frac{23}{20} = 1 + \frac{3}{20};$$

in words, 23 twentieths is equal to 1 and $\frac{3}{20}$. Doing this is called coverting an improper fraction to a mixed number, and grade-school texts probably still have pages of exercises where you're supposed to do that, or go the other direction. Here; try a few:

Exercise. Convert as indicated:

- $\frac{11}{7} = \rule{3cm}{0.4pt}$;
- $\frac{23}{7} = \rule{3cm}{0.4pt}$;
- $\rule{3cm}{0.4pt} = 4\frac{3}{4}$;
- $\rule{3cm}{0.4pt} = 5\frac{2}{3}$.

Here's a word about notation. You do sometimes see mixed numbers like $1+3/4$ written as 13/4, particularly in newspapers and magazines. But we'll keep the "+" sign in there.

Another thing to notice about

$$\frac{2}{5}+\frac{3}{4} = 2\times\frac{4}{20}+3\times\frac{5}{20} = \frac{8}{20}+\frac{15}{20} = \frac{23}{20}$$

is that in converting 2/5 to 8/20 we have multiplied the top and the bottom by 4:

$$\frac{2}{5} = \frac{8}{20} = \frac{4\times 2}{4\times 5},$$

and likewise

$$\frac{3}{4} = \frac{15}{20} = \frac{5\times 3}{5\times 4}.$$

Does this always work? If we multiply the top and bottom of a fraction by the same number will we always get another fraction that's equal in value to the first? Let's see. Suppose we have 2/3 and want to convert it to twelfths. To get 12 in the bottom we need to multiply by 4, and if we also multiply the top by 4 we get 8. So our question is whether

$$\frac{2}{3} = \frac{4\times 2}{4\times 3} = \frac{8}{12}$$

is correct.

To picture 2/3, imagine some object divided into 3 equal parts, with 2 of them set aside, like this slab of modelling clay:

Picture here: 3truein by 3truein (*********** scaled 600)

Now, to get twelfths, that is, to have the orginal chopped up into 12 equal parts, we can just chop each of the 3 thirds up into 4 parts, like this:

Picture here: 3truein by 3truein (*********** scaled 600)

Sure enough, the 2 thirds that we have set aside amounts to exactly 8 twelfths.

Exercise. Draw pictures to convince yourself that $\frac{4}{7} = \frac{12}{21}$. Now write a simple explanation of why this conversion amounts to multiplying both the top and the bottom of $\frac{4}{7}$ by 3.

At this stage what we should understand quite well is that

- to add two fractions that don't start out with the same denominator we need to convert them to fractions that do the same denominator, and
- to convert a fraction we multiply its top and bottom by the same amount.

But how do we figure out what the common denominator, the one we want to convert to, should be? The examples we've done suggest that one choice is always to take it to be the product of the denominators we started with. For instance, to add

$$\frac{4}{7} + \frac{5}{11}$$

we can use $7 \times 11 = 77$ as the common denominator, and get

$$\frac{4}{7} + \frac{5}{11} = \frac{44}{77} + \frac{35}{77} = \frac{79}{77}.$$

Instead of just rushing off now, let's take our time and make sure this answer makes sense. 4/7 is a little more than 1/2, as you can easily check by drawing a picture:

Picture here: 3truein by 3truein (*********** scaled 600)

And 5/11 is a bit less than 1/2. So 4/7 plus 5/11 had better be pretty close to 1. Sure enough,

$$\frac{79}{77} = 1 + \frac{2}{77},$$

which is a bit more than 1 (You could carefully draw a picture of a slab of clay cut into 77 equal pieces, with 2 extra pieces stuck on at the end).

other choices of common denom pennies insted of nickels

2.1.1 multiplying fractions

How much is 1/2 of 1/3? If you cut a slab of clay into 3 equal pieces, and then cut one of those in two, you have 1/2 of 1/3. If you cut all 3 thirds in two, you see that you have now 6 equal pieces. So 1/2 of 1/3 is 1/6.

Picture here: 3truein by 3truein (*********** scaled 600)

Likewise, you can check yourself that 1/4 of 1/5, say, is 1/20. This makes it clear how to take one fraction *of* another.

But now notice that 4 stacks *of* 5 pennies gives a total of 4 *times* 5, or 20, pennies. That is, with whole numbers,*of* and *times* really mean the same thing. This is true for fractions, too; 1/2 *times* 1/3 is the same as 1/2 *of* 1/3. One way to see this is suppose you make 50 cents a a minute by thinking real hard about math problems. If you work 5 minutes, you make 5 times 50 cents, or $2.50. Since the 50 cents is the same as 1/2 of a dollar, this shows that

$$\frac{1}{2} \times 5 = 2 + \frac{1}{2} = \frac{5}{2},$$

which is the same as saying

$$\frac{1}{2} \times \frac{5}{1} = \frac{5}{2}.$$

Now suppose you only work for half a minute. You then make 1/2 of 50 cents, or 1/2 times 50 cents, or 25 cents. Since 25 cents is 1/4 of a dollar, this means that 1/2 times 1/2 of a dollar is 1/4 of a dollar.

2.2 Decimals

Probably the hardest thing about fractions is adding them. You have to go through all that common denominator stuff, which many people find a real pain. Decimals give us another way to handle fractions, and have the advantage of being easy to add. They also have disadvantages, though, as we will see.

The number 234 means $2 \times 10^2 + 3 \times 10^1 + 4 * 1$; the further to the left, the higher the power of 10. The decimal idea is to do the same thing with powers of 1/10 as we go to the right. Thus

$$234.567 = 2 \times 10^2 + 3 \times 10^1 + 4 * 1 + 5 \times \frac{1}{10} + 6 \times \frac{1}{10^2} + 7 \times \frac{1}{10^3}.$$

Thus, the decmal part, .567 comes out to

$$\begin{aligned} 5 \times \frac{1}{10} + 6 \times \frac{1}{10^2} + 7 \times \frac{1}{10^3} &= \frac{5}{10} + \frac{6}{100} + \frac{7}{1000} \\ &= \frac{500 + 60 + 7}{1000} = \frac{567}{1000} \end{aligned}$$

Percent is shorthand for the latin phrase *per centum*, meaning per 100. Thus

$$50 \text{ percent} = 50\% = \frac{50}{100} = \frac{5}{10} = \frac{1}{2}.$$

There's nothing really *mathematically* special about 100—our fondness for it comes from our having 10 fingers—and mathmatics could get along just fine without ever having heard of percent. So why do we bother with it at all?

The answer is that if we express fractions as percents, we're really giving them all a common denominator—100—and having *some* common benchmark like this makes comparing fractions easer. Suppose, for instance, that we want to compare 3/4, 5/6 and 7/10. As percents, these are 75%, about 83%, and 70%. You get each percent from the fraction by dividing the bottom into the top, which you can do by hand or using a calculator; we get

$$\frac{3}{4} = .75, \tag{2.1}$$

$$\frac{5}{6} = .8333\ldots, \tag{2.2}$$

$$\frac{7}{10} = .7 = .70. \tag{2.3}$$

This example shows that you can just read off what percent a number is from what it is as a decimal, and vice versa. It also shows that percents have the same disadvantage as decimals: lots simple fractions don't come out as exact percents. 5/6 is *about* about 83%, but not exactly; 1/3 is about 33%, but not exactly.

Now, 1/3 is exactly $33\frac{1}{3}\%$, but this may seem circular; in order to express 1/3 exactly as a percent, I have to use 1/3 itself, so why bother? The answer is that the percent value, even if it still uses 1/3 itslef, still makes comparing easier; 1/3 is a bit smaller than 7/20, since

$$\frac{1}{3} = 33\frac{1}{3}\% < 35\% = \frac{7}{20}.$$

Anyhow, when we talk in terms of percents, we usually round off, and are often talking about aproximationns. If 14 out of the 23 students in my French class got an A on the last quiz, I'm going to tell you 61% of us got A's.

********exercise: what is 14/23 as an exact percet?

Since 61% is the same as .61, 75% the same as ./75, 153% the same as 1.53, we can also ask: why not just say these as decimals/ The answer is that as a matter of English usgae, the decimal version sounds stuffier

2.3 Ratios and proportion

A ratio is a comparison in size between to quantities. Or something like that; I guess I'm not quite sure how to *define* ratio, but a couple of examples will make clear what the term means.

If there are twice as many girls as boys in my English class, then the ratio of girls to boys is 2 to 1. If I do a survey of surgeons, and find that 355 of those who respond prefer green operating gowns, and 113 prefer blue, then the ratio of the greens to the blues is 355 to 113. If I were reporting this fact in a newspaper, though, I'd probably say that about 7 out of 9 surgeons prefer green gowns, or that the ratio of greens to blues is about 7 to 2. How'd I come up with this? Well, 355 and 113 are about 350 and 100, and the ratio of 350 to 100 really means the value of the fraction

$$\frac{350}{100} = \frac{35}{10} = \frac{7}{2}.$$

But if a ratio is really just the value of a fraction, why do I call the first one above "2 to 1", instead of just "2"? Why don't I say "the ratio of girls to boys is 2"? Beacuse this isn't the way English usage works, that's why! That is, the reason we speak of ratios the way we do isn't beacuse of math, it's because of English.

2.4 Putting fractions into lowest terms

We'll call a number *rectangular* if you can arrange that many marbles into a rectangle. Thus 6 is rectangular:

Picture here: 3truein by 2truein (*********** scaled 600)

but 7 is not; no fair saying you can arrange 7 marbles in a 1-by-7 rectangle!

Picture here: 3truein by 1truein (*********** scaled 600)

Exercise.

1. Draw 12 as a rectangular number in two really different ways. (We'll say a 3-by-4 and a 4-by-3 rectangle *aren't* really different, since one is just the other turned sideways.
2. Find, and draw, the next 4 numbers after 12 that are rectangular in at least two really different ways. Make sure you don't miss the one that's got *three* really different ways.

A number that isn't rectangular is called *prime*, except that we don't count 1 as prime. The first few primes are

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29.$$

Exercise. Figure out the rest of the primes up to 100. (There are 25 primes between 1 and 100, so you'd better get 15 more).

******************* stuff goes here

$$1,000,000 = 2^6 \times 5^6, \quad \text{and} \quad 559017 = 3^2 \times 179 \times 347.$$

The first of these is easy to figure out, as $1,000,000 = 10^6 = (2 \times 5)^6 = 2^6 \times 5^6$. The second would take you quite a while, even with a calculator.

You can start by just trying each prime in turn, and seeing if it goes into 559017. Since this is an odd number, you know 2 won't go into it, so you try 3 and find that 3 goes into 559017 exactly 186339 times. Now we start again with 186339. Since 2 doesn't go into it, we try 3 and get lucky again; $186339/3 = 62113$. Now we start again with 62113. Again we don't need to bother trying 2, so we try 3, then 5, then 7, then 11, and so on. Our calculator keeps telling us no; for instance it says $62113/3 = 20704.333\ldots$ (and since $3 \times 20704 = 62112$, we see that 3 goes into 62113 20704 times with a remainder of 1, if we want to be exact about it). No need to try 5, really, since anything 5 goes into has to end in 0 or 5. As we keep plugging away, though, none of the primes we have at hand —those less than 100— goes into 62113. Past 100, we don't yet know which numbers are prime, but we can just try them one after another; except we don't need to bother with the even ones (why not?). It'd take us a while to get to 179, and then we'd see that $62113 = 179 \times 347$. Now we'd still need to check both 179 and 347, to see if anything goes into them. How far do we have to go to see whther 179 is prime? You might think we'd have to try all the odds up to 179 itself, but, mercifully, we don't have to.

2.5 Powers and roots revisited

0 and negative powers fractional powers

We can get an approximation by trial and error. 2 is too small, and 3 too big, as $2^2 = 4$ and $3^2 = 9$. So let's try $2 + \frac{1}{2}$, which we can do in either fraction or decimal form:

$$\left(2 + \frac{1}{2}\right)^2 = \left(\frac{5}{2}\right)^2 = \frac{25}{4} = 6 + \frac{1}{4}, \quad \text{or} \quad 2.5^2 = 6.25;$$

$2 + \frac{1}{2}$ is too big. We could get as good an approximation as we like by keeping up this trial and error. If I just ask my calculator, it says the answer is 2.236068. But if I check by multiplying 2.236068 by itself, doing it out painstakingly by hand, I actually get 5.000000100624, which is too big. It turns out, as we'll see in the section *What's a Number*, that *no* simple fraction is exactly the square root of 5.

They used to, in grade schol or high school, make you learn a method for cranking out square root approximations. Maybe they still do. At any rate, I don't have the faintest idea how to do it any more, and neither you nor I *need* to know how. If the unlikely event should ever occur that you actually need to approximate a square root, or cube root or something, and you have no calculator available, then use trial and error. But believe me; it'll never happen.

As we'll see in *What's a Number*, if someone asks you for the square root, or cube root, or fifth root, or whatever, of a whole number, then the answer is either a whole number, or else it's not a simple fraction or decimal.

Chapter 3

Shapes and areas

3.1 Familiar shapes

Probably the most familiar geometric shapes are the rectangle the triangle, and the circle:

Picture here: 3truein by 3truein (ShapeArea1 scaled 600)

The "rect" in rectangle comes from the Latin word for "right", and refers to the fact that the angles in a rectangle are right angles. The "tri" in "triangle" means "three"; a triangle is a figure with three angles.

3.2 Rectangles

If we have a rectangle that's 3 feet wide and 5 feet high, then we can divide it up into 15 square pieces, each a foot on each side. This shows that the area of a 3-by-5 rectangle is 15 square feet.

Picture here: 3truein by 3truein (ShapeArea2 scaled 600)

If our rectangle is instead $3\frac{1}{2}$ feet by $5\frac{1}{2}$ feet, we can divide each side up into $\frac{1}{2}$-foot pieces, and again break the figure up into a bunch of little squares. This time there are $7 \times 11 = 77$ little squares, each $\frac{1}{2}$-foot on a side. Since it takes 4 of these little squares to fill up a single square a foot on each side, each little square has area $\frac{1}{4}$ of a square foot, and so our total area is $\frac{77}{4} = 19\frac{1}{4}$ square feet. Thus in both cases we figure out the area of the rectangle by multiplying its width times its length, which is the same as saying its base times its height:

$$\text{area of a rectangle} = \text{base} \times \text{height};$$

$$15 = 5 \times 7;$$
$$\frac{77}{4} = \frac{7}{2} \times \frac{11}{2}.$$

3.3 Triangles and parallelograms

How about the area of a triangle? You can tell by looking that you're not going to be able to cut the triangle up into a bunch of square pieces; they just won't fit right. So we've got to be more clever. If we're lucky enough to have a *right* triangle, we might notice that it's exactly half of a rectangle,

Picture here: 3truein by 3truein (ShapeArea3 scaled 600)

and since we know the base and height of the rectangle, we can figure out its area, and the area of our triangle will be exactly half that. But what if the triangle doesn't have a right angle in it?

In this case we can still try putting together two copies of our triangle, and we get something that looks like a rectangle leaning over:

Picture here: 3truein by 3truein (ShapeArea4 scaled 600)

The kind of figure is called a *parallelogram*, because its opposite sides are parallel (parallel lines are ones that are the same distance apart everywhere). This doesn't seem to be any help though, because we can't fill up the parallelogram with little squares any more than we could the triangle we started with. However, there's a nick trick that we can do to turn this parallelogram into a rectangle; we cut a triangle off one side and move it over to the other:

Picture here: 3truein by 3truein (ShapeArea5 scaled 600)

Now we do have a rectangle, and its area is its base times its height, which is the same as the *base* of the parallelogram times its *height*. Thus we have

$$\text{area of a parallelogram} = \text{base} \times \text{height},$$

and since any triangle is half a parallelogram,

$$\text{area of a triangle} = \frac{1}{2}\text{base} \times \text{height}.$$

There's another way to see that the area of a parallelogram is equal to its base times its height. Think of slicing the parallelogram up into lots of tiny strips, and then pushing the figure over until the sides are as straight as you can get them. The result will be approximately the rectangle with the the same base and height; it's not exactly a rectangle, because the sides are a little bit jagged. But the finer the strips we slice the parallelogram into, the more nearly our almost-rectangle approaches a true rectangle.

Picture here: 3truein by 2truein (ShapeArea6 scaled 600)

We can also do this the other way around, and get an almost-parallelogram by pushing over a rectangle that's been cut into strips. The picture you see here is what you'd see if you were to look at a deck of cards:

Picture here: 3truein by 2truein (ShapeArea7 scaled 600)

In fact, we can make plenty of other shapes besides a parallelogram by pushing around our sliced-up rectangle, and they'll all have the same area: all that matters is that the sides, whether straight or curvey, are the same distance apart (that is, are *parallel* from top to bottom. Any two shapes with parallel sides like this will have the same area as long as they have the same base and the same height. Here are a few examples:

Picture here: 3truein by 2truein (ShapeArea8 scaled 600)

This idea for figuring out areas of shapes other than rectangles goes back at least to Archimedes, but is called Cavalieri's Principle after the Italian mathematician of that name, who more explicitly presented it. Cavalieri's Principle works for solids just as well; pushing a deck of cards around doesn't change the volume of the deck; thus the two solids below, which are both cylinders, have the same volume. We can't now say what that volume is, but whatever it is, it's the same for both solids. We'll figure it out in a later section.

Picture here: 3truein by 3truein (ShapeArea9 scaled 600)

Now suppose we want to figure out the area of a triangular piece of cloth that's 1 foot on each side. What do we do? Unfortunately the answer for the time being is that we can't solve this problem; we need to know the *height*, and that's not given to us.

Picture here: 3truein by 2truein (ShapeArea10 scaled 600)

Later, though, we'll see that we can get the height by knowing the sides, using the famous Pythagorean Theorem.

Exercise: The figure shown, called a regular hexagon, is made up of 6 triangles each 1 foot on a side. The height of each of these is approximately $\frac{19}{11}$. Find the approximate area of the hexagon.

Picture here: 3truein by 2truein (ShapeArea11 scaled 600)

Exercise: Find the areas of the figures shown.

Picture here: 3truein by 2truein (ShapeArea12 scaled 600)

3.4 Circles

In Archimedes' explanation of how to determine the area of a circle, he says that this area is the same as the area of a right triangle:

Picture here: 3truein by 2truein (ShapeArea13 scaled 600)

The base of this triangle is equal to the circumference (distance around) the circle; its height is equal to the radius (distance from the center to the edge) of the circle. You can see that the circle's area is at least close to that of the triangle if you do this: make a circle out of concentric rings of string, or thin strips of modeling clay, as in the figure below. Now take a knife and make a straight cut along a radius of the circle, and then straighten out the strips:

Picture here: 3truein by 2truein (ShapeArea14 scaled 600)

You get a triangle as advertised.

Of course, if you try to do this by starting with a circle on a piece of paper, you *can't* straighten out the strips. So we'd better be a little more careful.

So this time just make a paper circle, and cut it up into a bunch of pieces like you'd slice a pie, and now stack the pieces back and forth as shown.

Picture here: 3truein by 3truein (ShapeArea15 scaled 600)

The stack of slices is approximately a rectangle whose base is the radius of the circle, and whose height is half the circumference (half the pie crust makes up one side of the almost-rectangle; the other half the other side. The more slices you cut, the closer this almost-rectangle approaches a true rectangle.

And how does the area of the true rectangle compare to that of Archimedes' triangle? Both are equal to one-half the radius of the circle times the circumference.

Picture here: 3truein by 2truein (ShapeArea16 scaled 600)

We conclude from this that

$$\text{area of a circle} = \frac{1}{2} \times \text{radius} \times \text{circumference}.$$

3.5 More on circles: what is π?

Suppose we just knew the radius of a circle, and wanted to figure its area. After all, if we just had a ruler, it'd be easy to measure the radius, as long as we knew where the center was. But it's be hard to measure the circumference, because we'd be trying to measure a curved line with our straight ruler.

However, if you examine a few circles, and use a piece of string to measure (approximately) their circumferences, you'll find that the *ratio* of the circumference to the radius is the same for all the circles, and is about 6. This means that the ratio of the circumference to the diameter (the distance across the circle; twice the radius) is about 3. Some time around the mid-1700's the Swiss mathematician Euler started calling this ratio π. This is the Greek letter pi, and he used it to stand for the *perimeter* of the circle, which is another word for the circumference (In Greek "peri" means around, "meter" means measure; in Latin "circum" means around, "ferre" means to carry).

Thus, if we know the radius of a circle, its circumference is 2π times its radius, or π times its diameter. And its area is thus

$$\text{area of a circle} = \frac{1}{2}\times\text{radius}\times\text{circumference} = \frac{1}{2}\times\text{radius}\times 2\pi\times\text{radius} = \pi\times\text{radius}\times\text{radius}.$$

If we use the letters A, C and r to mean the area, circumference, and radius, we now have these two formulas:

$$C = 2\pi r; \quad A = \pi r^2,$$

where r^2 (pronounced "r squared"). is shorthand for $r \times r$. Note, though, that the first formula really just tells us what π means, while the second one really says that the area of a circle is the same as that of Archimedes' triangle.

Exercise: Using that $\pi \approx 3$, find the approximate areas of the circles whose radii are 1 foot, 2 feet, and 3 feet.

Part II

Understanding Functions

Chapter 4

An introduction to functions

4.1 The function idea

If we drop a stone from the top of a 40-foot building, how high the stone is depends on how long it's been falling; that is, the stone's height is a *function* of the time elapsed.

We can draw a graph, to describe visually how the height depends on time, by plotting time horizontally and height vertically. As a first guess we might and we might expect to get something like this:

Picture here: 3truein by 3truein (fcnIntro1 scaled 600)

This particular graph says that at time 0 seconds, the height is 40 feet, and then it decreases steadily to 0 feet, hitting the ground after 4 seconds. According to this graph, the height after 1 second is 30 feet, after 2 seconds it's 20 feet, and so on; the stone is dropping at a steady speed of 10 feet each second. Please don't just read this description and say "unh hunh"; take a few minutes, or even longer, right now and *study* the graph, making sure that you thoroughly understand what it's saying about the stone's behavior as it falls.

If you really think about the graph, your own common sense and experience should make you see that something's wrong with it; no stone is really going to fall this way! For if it fell at a steady speed, it would be going just as fast when it hit the ground as it would be when it's only dropped a couple of feet. And if you believe that's true, you shouldn't hesitate to be the stone yourself, and jump off a 40 building! —because you'd you'd be going no faster when you hit than you would be if you just hopped down a couple of feet.

OK, so our first rough guess as to our function's shape turns out to be wrong, so we have to think some more. There's already an important lesson here: we'll be seeing that straight line functions are the very simplest, and that a lot of real-life functions really are this simple. This makes it common for people to assume that some function they're thinking about is like this —is *linear*, as we say; but plenty of real-life functions, like our stone's behavior, *aren't* linear. If you start listening for it, you can often hear people making this assumption; they're probably not going to be picturing the graph, and seeing a straight line, but they'll be saying things that amount to this. So the lesson is that

lots of people tend to think that everything is linear

A second important lesson here is that we've just figured out that our stone's behavior is not linear just by using common sense; we didn't need to analyze a formula or anything, since we haven't even looked at any formulas; that is,

common sense alone can tell us a lot about functions.

Exercise. Sketch the graph of a real-life function that *is* linear, and explain why, using only common sense; no formulas allowed!

Exercise. Give an example of another function people might assume linear, and explain why it isn't. If you can, sketch the graph.

To get a second estimate of what our function looks like, let's use the fact that we know that when we drop a stone it doesn't fall at a steady rate, but instead picks up speed. What does that mean our graph ought to really look like? We know the height at time 0 is 40 feet, so our graph does start out where we had it in our first try, and we know from experience that the stone has to hit the ground after a few seconds —in fact, it'll hit after a little more than $1\frac{1}{2}$ seconds. So we know where the graph starts, and about where it ends, but what does it look like in between? Let's try drawing some graphs connecting the start to the finish, and see if we can get one that says the stone picks up speed as it falls. Here are four examples

Picture here: 5truein by 2.5truein (fcnIntro2 scaled 600)

2a 2b 2c 2d

What can we say about these? See if you can figure this out yourself before reading the answer. Figure 2a is really the same idea as we had before, that the stone falls at a steady speed, only this time has it hitting the ground at the right time. So that's not what we want. What does 2d say? It shows the stone going down, and then going back up, and then down again. This might happen if there were a gust of wind from below, but it's not the right picture for what would really happen under ordinary circumstances. What about 2b and 2c? One of them shows what we want; what does the other one mean? You can answer this by seeing what each says about

- how high we are when half the time has gone by, or
- how long it takes to get halfway down.

I'll leave this question as an exercise for the time being, plus I'll ask you

Exercise: what would the graph look like if

1. the stone got halfway down and then hovered in mid-air for a while before falling the rest of the way?

This example shows that we can get a pretty good idea of what the graph is going to look like *without* having a formula.

If we did lots of experiments, and made careful ,measurements, we could get a more accurate graph, which we could use to answer the question "how high is the stone at a given time?" We'd just find that time on the time axis, then go straight up to the graph, and then over to the left, and see how high we are. We could do the same thing by making a table that lists the heights at at, say .01, .02, .03 seconds, but the graph is a much better way of *seeing* the behavior.

What we really do, though, is to use the laws of physics, and a little calculus, to figure out the formula for height as a function of time. This formula turns out to be

$$h = 40 - 16t^2$$

For a given time t, we plug that value into this formula, and get the height h:

t	h
0	40
$\frac{1}{2}$	36
1	24
$1\frac{1}{2}$	4

Can you figure out how I got these answers by plugging in? If not, don't worry about it, or get someone to show you.

The point is that what we're studying is *functions*, and if we're lucky, we have a formula for the function we're studying. In fact, in real life our job is often done once we figure out the formula. Formulas are written using algebra — that is, using letter like h and t, or x, y, and z, to represent values of the quantities we're studying. So part of understanding functions involves learning how to do algebra — how to manipulate these symbols that stand for numbers. In fact, as far as I know, this is the the only reason you'd ever want to understand algebra — so you can use it to understand functions.

Since we get the best overall grasp of our function by drawing its graph, which is a geometric figure, we need to understand some geometry as well, in order to be able to understand functions.

What else will we need before we come back and take a more detailed look at functions? The answer is in the table of contents of this volume; besides making sure we're familiar with some geometry and some algebra, we'll need to have another look at what we mean by a "number", and we'll want to make sure that our reasoning is sound.

The rest of this book is about understanding functions, which is what a high school math program ought to be about; you learn algebra – x,y,z, and all that – so you can write formulas for functions; you study geometry and trigonometry to get to know some important elementary functions.

Calculus itself is really just bringing in two new tools for understanding functions; namely, the slope of a curve and the area under it.

Unfortunately, far too many people start their study of calculus, or start memorizing calculus formulas, without first having much understanding at all about functions, which is what's really the problem. Listen to what Euler said: (Euler quote).

The truth, apparently unknown to or concealed by generations of high school texts and teachers, is that there's a whole lot you can know about functions *before* you bring in the two new tools of calculus, and that understanding functions can be fun, and easy, and really on common sense.

4.2 Logical Thinking

4.3 What's logical?

4.4 *and* vs. *or*

4.5 $A \to B$

4.6 Guessing and trial and error

Contradiction sucks. Induction sucks.

Chapter 5

What's a number?

5.1 Counting numbers and fractions

For our first sections, on counting, adding, subtracting and dividing, we used the word "number" just to mean the counting numbers $1, 2, 3, \ldots$, and the number 0. The counting numbers are the humblest instance of the number concept, the oldest historically, and the one we all encounter earliest in our childhood. The number 0, though, is actually quite sophisticated. We may tend to take 0 for granted, since it's so familiar to us as part of our way of writing numbers as ordinary as 10 or 20, but the discovery of using 0 as a place-holder changed not just the world of mathematics, but the course of human history. If you ever have the misfortune of of having to work with Roman numerals, where $303 + 323 = 626$ comes out as $\mathrm{CCCIII} + \mathrm{CCCXXIII} = \mathrm{DCXXVI}$, you'll realize how fortunate we are that the Arabs gave us 0 in our number system.

In our section on fractions, we saw that fractions, too, are numbers, and in fact that the fractions include the whole numbers. This is what most normal people mean when they say "number"; they're thinking of the positive fractions (actually the "nonnegative fractions", if we include 0). What's the picture to represent these numbers? Most people would say it's this:

Picture here: 5truein by 1truein (WhaNum1 scaled 1000)

That is, we have a line, or really a half-line, and all the points on it. That is, the idea that most people have is that fractions are the same thing as lengths, which means that fractions nicely tie together arithmetic (number) and geometry (length).

Legend has it that the ancients, particularly the Pythagoreans, were very pleased with this set-up, and marvelled that God in His infinite wisdom would have all lengths be expressible as ratios of nice whole numbers. Legend also has it that they were so upset when one of their members showed this was false, they threw him off a boat to drown!

5.2 Irrational numbers

How can we see that there are lengths that aren't fractions? Imagine a square 1 foot on a side growing steadily. At the start, its area is 1 square foot. As it grows to having side 3 feet, say, we can certainly imagine its area taking all values between 1 square foot and 9 square feet; the starting and finishing area. Thus, at some time during the growth its area will be 2 square feet, then later 3 square feet, then later 4, 5, 6, 7, and 8 square feet. Now let's ask how long the sides of the square are when its area has these nice simple values. That is, we want to figure out the actual values, as fractions, of $\sqrt{2}$, $\sqrt{3}$, $\sqrt{4}$, and so forth. Heck, let's just use our calculators to make a nice little table:

Area	Side
1	1
2	1.4142136
3	1.7320508
4	2
5	2.236068
6	2.4494897
7	2.6457513
8	2.8284271
9	3

Exercise Here's a chance to practice your long multiplication! Carefully multiply out the value given for $\sqrt{5}$, and show the answer comes out to 5.000000100624.

As the exercise shows, the value our calculator gives for $\sqrt{5}$ isn't quite correct; when we square it, we get something that's extremely close to 5, but just a bit more. Likewise, you could check, by long multiplication, that the other messy answers in the table are just a little bit off.

Now, the Pythagoreans didn't have our easy decimal system or calculators, but they had their own ways —trial and error would work, for instance— of getting good fraction approximations for lengths like $\sqrt{5}$. And they, too, had to have noticed that no matter how good their approximation, it was never perfect. This must have led that unlucky member to suspect that *no* fractions were really exactly equal to these lengths. He then figured out how to show this was true, which is what we'll do now.

Remember that decimals really are fractions; for instance 2.236068 is really

$$\frac{2236068}{1000000} = \frac{559017}{250000}.$$

Here I reduced the fraction by dividing top and bottom by 2 twice. And this is as far as we can go with reducing it. Why? Well,

$$1,000,000 = 10^6 = (2 \times 5)^6 = 2^6 \times 5^6,$$

which means that to reduce the fraction we only have to divide out 2's and 5's, since they'll let us reduce 1000000 as much as we want to. But 5 doesn't go into 2236068, so all we can do is divide top and bottom by 2, and after doing this a couple of times we get the odd number 559017, and have to stop. Maybe we can still divide 559017 by *something*, but certainly not by 2 or 5. Actually, it turns out that $559017 = 3 \times 3 \times 179 \times 347$, and this is as far down as you can go, since 179 and 347 are prime numbers (nothing goes into them), and so is 3, of course.

What this means is that if we write 2.236068 as a fraction, and then reduce it to lowest terms, we get

$$\frac{2236068}{1000000} = \frac{559017}{250000} = \frac{3^2 \cdot 179 \cdot 347}{2^4 \cdot 5^6}.$$

Now notice that we don't even *have* to square this fraction to see that if we do we can't possibly get a whole number. Why? Because the primes on top $(3, 179, 347)$ are all different from the primes on bottom (2 and 5). But when we square this fraction, we'll still have the same primes on top, and the same ones on bottom, only all to twice the power. That is,

$$\left(\frac{3^2 \cdot 179 \cdot 347}{2^4 \cdot 5^6}\right)^2 = \frac{3^4 \cdot 179^2 \cdot 347^2}{2^8 \cdot 5^{12}},$$

and we can't reduce the fraction after squaring any more than we can before!

To see more generally what's happening here, let's use the term *honest-to-goodness* fraction to mean one that isn't really just a whole number; so $\frac{3^2 \cdot 179 \cdot 347}{2^4 \cdot 5^6}$ is an honest-to-goodness fraction, but $\frac{8}{2} = \frac{4}{1} = 4$ is not. If you take an honest-to-goodness fraction, then, and reduce it to lowest terms, its bottom isn't 1. So in lowest terms it has some primes on top, and a bunch of completely different primes on bottom. When we then square it, or cube it, or raise it to any power, we still have the same primes on top, and the same ones on bottom, and don't get any cancellation at all. This means that

Every power of an honest-to-goodness fraction is still an honest-to-goodness fraction,

or, in other words,

We won't ever get a whole number by raising an honest-to-goodness fraction to any power.

Now think about the square whose area is 6, so whose side is $\sqrt{6}$. This square, in area, is between the square with area 4 and the one with area 9, so its side has to be between 2 and 3. This means that the length $\sqrt{6}$ can't possibly be a whole number; it's got to be between 2 and 3. But every fraction between 2 and 3 is an honest-to-goodness fraction, and therefore isn't going to give a whole number when we square it. This means we don't have to multiply the answer our calculator gives for $\sqrt{6}$, namely 2.4494897, by itself to see that it's a little bit off. 2.4494897 as a fraction is an honest-to-goodness one, so when we square it we can't get a whole number. Period, the end. And it won't do any good to get a better calculator to give us a better decimal approximation to $\sqrt{6}$;

whatever it is, even if it has 100 decimal places, it's going to be just an approximation, and won't quite come out to 6 when we square it.

The same goes for the other numbers in our list, except for 4 and 9: there just isn't any fraction whose square is 5, or 6, or 7, or 8. Of course, there are fractions whose squares are 4 and 9, but they're really the whole numbers 2 and 3, and not honest-to-goodness fractions.

Now the same thing goes for cube roots; either a whole number is a perfect cube, like 8 and 27, and has a whole-number cube root, or else there's no fraction that is equal to its cube root.

When the Greeks discovered that there were squares or cubes whose sides weren't equal to any fraction, I believe they really thought of the situation as being that there are some lengths that aren't numbers. For to them, number *meant* fraction, just as I believe it does to ordinary people today. But mathematicians prefer to extend their idea of "number" to include these lengths. Thus mathematicians call a quantity like $\sqrt{2}$ an *irrational number*. Here "irrational" just means not rational, and a *ratio*nal number just means one that is a *ratio* of integers, that is, a fraction.

Exercise Given that 205,097 is a prime that divides $26,457,513$, find 2.6457513 (which is approximately $\sqrt{7}$) as fraction in lowest terms.

Now, we've already seen that a fraction, or rational number, has a decimal expansion that eventually repeats itself; 1/7, for instance, is $.142857142857\ldots$. It's not hard to show that the converse is true; that is, a repeating decimal is equal to some fraction. Thus a real number has an eventually repeating decimal expansion if and only if the number is rational. Equivalently, a real number is irrational if and only if it has a never-repeating decimal expansion.

Notice, however, that we can never apply this criterion to check whether a number is irrational. For instance, if we were to carry the decimal expansion of $\sqrt{2}$ out to a million places, we wouldn't see it start to repeat. But for all we know, maybe it repeats after 2 million places, or after 50 million. I mention this just because students seem to get the idea from textbooks that you can tell that a number is irrational by examining a few places of its decimal expansion; but nothing could be further from the truth.

5.3 Negative numbers

5.4 Complex numbers

Chapter 6

Geometry

6.1 Introduction

These sections on geometry have turned out to be the hardest part of this book for me to figure out and write down. This isn't what I expected, since when I made a list of what facts of geometry we need in order to be able to do calculus, the list was short:

- the Pythagorean theorem;
- basic facts about similar triangles;
- areas of triangles and polygons;
- the area of a circle;
- volumes of cylinders and cones.

Why should this be hard? I took a standard geometry course in high school, and have gone on in math, so I should know all this stuff cold, right? And if I want to give explanations for these few simple facts, I can just follow the approach in a standard high-school geometry text, so that for many readers it'll just be a matter of remembering, right?

Wrong. The standard high school geometry text is supposed to be based on the *Elements* of Euclid, which has for centuries been regarded as *the* geometry book. Euclid, a Greek mathematician, wrote the *Elements* around 300 B.C., and in it derived hundreds of results of geometry, pretty much everything anyone could ever want to know, from just five *postulates*, or results he assumed to be true. The opening sentences of the preface to Heath's edition of the *Elements* (1925) indicates how important Euclid's work is considered:

> "There has never been, and till we see it we never shall believe that there can be, a system of geometry worthy of the name, which has any material departures (we do not speak of *corrections* or *extensions* or *developments*) from the plan laid down by Euclid." De Morgan wrote thus in 1848 (*Short supplementary remarks on the first six books of Euclid's Elements* in the *Companion to the Almanac* for 1849); and I do not think that, if he had been living to-day, he would have seen any reason to revise the opinion so deliberately pronounced sixty years ago.

This all sounds great. The simple geometric facts we need to do calculus have to be right there in Euclid, and hence anyone who's had a high school geometry course will know them, and, if they're any good at math, be able to derive them from Euclid's postulates. I think this is the attitude that most calculus teachers, and most calculus texts, take —whatever geometry we need for calculus is something that anyone with high-school math will know. I can imagine a calculus student objecting "but, Sir, our text's two-page proof that $\frac{\sin x}{x}$ goes to 1 uses that the area of a circle is πr^2, and the book doesn't even comment on that, much less prove it", and I sincerely believe that most teachers would respond "well, that's something you're just supposed to know from high-school geometry." But figuring out the area of a circle isn't in Euclid, at least not explicitly; it's in Archimedes (287–212 B.C.), as are a number of other things that everybody thinks of as being in Euclid. And Archimedes' proof the area of a circle *assumes* something that amounts to what the calculus book is trying to "prove".

And they wonder why students have trouble with math!

So what's the real story? First of all, as most mathematicians know, Euclid doesn't really derive everything from the postulates he sets down. As Bertrand Russell puts it,

> His definitions do not always define, his axioms are not always indemonstratable, his demonstrations require many axioms of which he is quite unconscious.... The value of his work as a masterpiece of logic has been very grossly exaggerated.

We don't need to take Russell's word for this; here are Euclid's postulates, paraphrased in modern terms:

1. Two points determine a unique straight line.
2. A straight line can be extended indefinitely far.
3. A given center and radius determine a unique circle.
4. All right angles are equal.
5. The parallel postulate.

And here are his "common notions", which accompany the postulates:

1. Things equal to the same thing are equal to each other.
2. Equals added to equals are equal.
3. Equals subtracted from equals are equal.
4. Things that coincide are equal.
5. The whole is greater than the part.

The full wording of the parallel postulate is long, so I've left it for here. Euclid defines parallel lines to be lines in the same plane that never meet, no matter how far they're extended. We call two angles *supplementary* if they add to a straight angle, or flat angle, like this:

Picture here: 3truein by 2truein (GeoIntro1 scaled 600)

The Parallel Postulate: If a line crossing two lines makes the interior angles on one side less than supplementary, the two lines meet on that side.

Picture here: 3truein by 2truein (GeoIntro2 scaled 600)

Now here's Euclid's Proposition 4, as translated from the Greek:

> If two triangles have the two sides equal to two sides respectively, and have the angles contained by the equal straight lines equal, they will also have have the base equal to the base, the triangle will be equal to the triangle, and the remaining angles will be equal to the remaining angles respectively, namely those which the equal sides subtend.

Please don't just sit there and tell yourself that you can't figure out what this is saying, especially with words like "subtend"; draw some pictures and figure it out! This is what is known to generations of high-school students as SAS or Side-Angle-Side. Euclid's proof goes like this: let ABC and DEF be our two triangles, having equal sides and angles as marked:

Picture here: 3truein by 2truein (GeoIntro3 scaled 600)

He then begins by saying

> For, if the triangle ABC be applied to the triangle DEF, and if the point A be placed on the point D and the straight line AB on DE, then the point B will also coincide with E, because AB is equal to DE.

He goes on the like this, pointing out that all parts of the two triangles match up when one is put on top of ("applied to") the other. This is fine, and it's certainly common sense, but Euclid is *not* just using his postulates and common notions! —they say nothing about putting one triangle on top of another. This is exactly the sort of thing Russell is talking about. It's really only in getting started, though, that Euclid doesn't quite deliver what he's promised. As Heath goes on to say in his preface,

> It is true that in the interval much valuable work has been done on the continent in the investigation of the first principles, including the formulation and classification of axioms or postulates which are necessary to make good the deficiencies of Euclid's own explicit postulates and axioms and to justify the

> further assumptions which he tacitly makes in certain propositions, content apparently to let their truth be inferred from observation of the figures as drawn; but, once the first principles are disposed of, the body of the doctrine contained in the recent textbooks of elementary geometry does not, and from the nature of the case cannot, show any substantial differences from that set forth in the the *Elements*.

OK, so one difficulty is that no high-school graduate can be very sure of which things from geometry were assumptions (or axioms and postulates) and which were propositions —which is which all depends on which book you use. You might think, then, that calculus books would have a section saying exactly which facts from geometry they're assuming, but I have yet to see one. Instead, they all act like everybody should just somehow know what's what, which simply isn't possible. That is, the claim that there even is such a thing as "what every high-school graduate should know about geometry" is false.

Even if we agree, as I do, that the problem so far is just one of getting started, and that once we get going things are pretty much the same, no matter which geometry text we use, there's still another problem. This is that you may never have even seen, in your geometry course, some of the results we want. And even if you have seen them, they were probably so mixed in with dozens, even hundreds of others, that you'd have little chance of remembering them. I have in my hand, for instance, the standard geometry text used in a local high school. It has over 700 pages, and exactly 130 theorems. The section on similar triangles starts after page 300.

And does the text follow Euclid's approach to similar triangles? Not very likely! Euclid starts off his results on similar triangles with the proposition that triangles with the same height have areas proportional to their bases. And to show this, he applies his definition of what it means to be proportional, which in his words is this:

> Magnitudes are said to be in the same ratio, the first to the second and the third to the fourth, when, if any equimultiples whatever be taken of the first and third, and any equimultiples whatever of the second and fourth, the former equimultiples

> alike exceed, are alike equal to, or alike fall short of, the latter equimultiples respectively taken in corresponding order.

Yet a calculus student is supposed to automatically know, in solving a problem about the rate of water flowing out of a conical tank, that you need to use similar triangles.

Listen to what the well known French mathematician, Jean Dieudonné, suggested in 1961 as a way to improve the high-school math curriculum:

> Some elements of calculus, vector algebra and a little analytic geometry have recently been introduced for the last two or three years of secondary school. But such topics have always been relegated to a subordinate position, the center of interest remaining as before "pure geometry taught more or less according to Euclid, with a little algebra and number theory".
>
> I think the day of such patchwork is over, and we are now committed to a much deeper reform-unless we are willing to let the situation deteriorate to the point where it will seriously impede further scientific progress. And if the whole program I have in mind had to be summarized in one slogan it would be: Euclid must go!
>
> This statement may perhaps shock some of you, but I would like to show you in some detail the strong arguments in its favor. Let me first say that I have the deepest admiration for the achievements of the Greeks in mathematics: I consider their creation of geometry perhaps the most extraordinary intellectual accomplishment ever realized by mankind. It is thanks to the Greeks that we have been able to erect the towering structure of modern science.
>
> But in so doing, the basic notions of geometry itself have been deeply scrutinized, especially since the middle of the 19th century. This has made it possible to reorganize the Euclidean corpus, putting it on simple and sound foundations, and to re-evaluate its importance with regard to modern mathematics — separating what is fundamental from a chaotic heap of results with no significance except as scattered relics of clumsy methods or an obsolete approach.

> The result may perhaps be a bit startling. Let us assume for the sake of argument that one had to teach plane Euclidean geometry to mature minds from another world who had never heard of it, or having only in view its possible applications to modern research. Then the whole course might, I think, be tackled in two or three hours —one of them being occupied by the description of the axiom system, one by its useful consequences and possibly a third one by a few mildly interesting exercises.
>
> Everything else which now fills volumes of "elementary geometry" —and by that I mean, for instance, everything about triangles (it is perfectly feasible and desirable to describe the whole theory without even defining a triangle !), almost everything about inversion, systems of circles, conics, etc. —has just as much relevance to what mathematicians (pure and applied) are doing today as magic squares or chess problems!

I found this passage of Dieudonné's in 1973 and thought "yeah; come to think of it, I seem to remember theorem after theorem about triangle this and triangle that in high-school geometry. But I've never used or even seen any of these since, certainly not in going through calculus."

On the other hand, while the most common response from regular people, when I tell them I teach math, is "yuck; math was my worst subject", the most common thing they go on to say, if anything at all, is "except geometry; I really liked geometry."

Now, if I press them to tell me just why they liked geometry, the usual answer is something like "well, it helped me to think logically." It's only been in the last couple of years that a friend suggested to me "they're just repeating their teachers' propaganda with that 'helped me to think logically' stuff; the real reason high school kids like geometry is because it's about *pictures*, and they can understand the pictures, and understand what the theorem they're supposed to prove says, and why it's true."

Well, I guess I've gone on long enough with these introductory remarks. Everything I've just said has just been to show you, Dear Reader, why my original idea —that I'd simply set down the few geometry results we need for calculus, in terms of the standard Euclidean geometry that every high-school kid has a course in— has turned out to be much easier said than done.

So don't expect to find here a list of formal definitions of geometric terms, or a list of undefined terms. And likewise don't expect a list of axioms and postulates, and an airtight proof of each and every result in terms of these and these alone. There ain't no such animal, anyhow, so we shouldn't feel guilty about using common sense, and letting terms define themselves from the context.

What I hope you you *will* find here is an honest account, and one you can follow, of the few basic geometric results you'll need to do calculus.

6.2 The Pythagorean Theorem

The Pythagorean Theorem is by far the best-known theorem in all mathematics; in fact, it's about the only math theorem that most non-mathematicians have even heard of.

Here's what it says: take a right triangle, and construct squares on the 3 sides. Then the area of the biggest square is equal to the combined areas of the 2 smaller squares:

Picture here: 3.6 truein by 3.4 truein (Pyth01 scaled 600)

In this example the sides of the triangle are 3, 4, and 5, and the areas of the squares on these sides are 9, 16, and 25, and sure enough

$$9 + 16 = 25.$$

To see why the Pythagorean Theorem is true, we start by drawing a general right triangle with squares on the sides, and then add copies of the original triangle, like this:

Picture here: 3.6truein by 2.5truein (Pyth02 scaled 1000)

Here we've added three copies around the big square, and two copies nested between the two smaller squares. If we now add a body and a couple of antennas, you see that our figure is a butterfly, with two wings exactly the same size and shape (or *congruent*); at least, they sure *look* like they're exactly the same. Assuming for now that the wings are indeed congruent, which we'll soon show, we see that they certainly have the same area. And if we remove three triangles from the upper wing, and three from the lower (the two copies, and the original), the areas remaining have to be the same.

Picture here: 3.2truein by 2.5truein (Pyth03 scaled 1000)

Now let's be a little more careful; sure, the two wings look the same, but how do we really know they really are exactly the same? Each appears to be a 5-sided figure, or *pentagon*, and if we start from the butterfly's head and go around the two wings in opposite directions, it seems that each side of one is equal to the corresponding side of the other, and the same for each angle. How do we know this is really true? If you think about this question, really think about it while examining the picture, you'll see that what we need to be sure of is that the angles of a triangle add up to a straight line; that's how we know that the two more heavily lines sides in this picture really are straight lines:

Picture here: 3.2truein by 2.8truein (Pyth04 scaled 1000)

To see that the angles of a triangle add up to a straight line, we need to start by seeing that certain angles formed by crossing lines are equal. If you have two lines crossing at a point, imagine pivoting the lines at that point so that you bring them together, and then opening them up again:

Picture here: 3truein by 1.5 truein (Pyth05 scaled 600)

Try it! Take two drinking straws, line them up, and stick a straight pin through both of them and into a piece of cardboard beneath them. Notice that as we spread apart the two lines, the *opposite* angles A and B remain the same, starting as 0 angles, and spreading apart at exactly the same rate until they both become straight angles and the two lines once again coincide, which we see about to happen here:

Picture here: 1.5truein by 3truein (Pyth06 scaled 600)

We express this fact about angles by saying that

opposite (or "vertical") angles are equal.

In case you're curious, here's how Euclid arrives at this result from his postulates: two right angles combined make a straight angle, and since all right angles are equal, so are all straight angles. Now consider a pair of opposite angles A and B, and note that both A and B are *supplements* of C, meaning that A plus C makes a stright line, and so does B plus C:

Picture here: 4.61truein by 1.10truein (Pyth07 scaled 1000)

Since all straight angles are equal, we thus have that $A + C = B + C$, and subtracting C from both sides yields $A = B$, which is what we want.

We only paid attention to one of the two pairs of opposite angles formed when two lines cross, but of course the same is true for both pairs. Thus whenever we have crossing lines, we can indicate as shown that we have two pairs of equal opposite angles:

Picture here: 4.58truein by 1.11truein (Pyth08 scaled 1000)

Now suppose we have two parallel lines crossed by a third line:

Picture here: 4.08truein by 2.08truein (Pyth09 scaled 1000)

That the lines are parallel means they're the same distance apart everywhere, so we can get this last figure by starting with the parallel lines lying atop one another, so that they coincide. Now we slide one copy away from the other, keeping the distance between them the same everywhere. This keeps the angles that each makes with the crossing line exactly the same:

Picture here: 4.06truein by 2.22truein (Pyth10 scaled 1000)

Here angles A and A' are equal, and so on. We express this briefly by saying that

corresponding angles are equal.

(Again, we can derive this result from Euclid's postulates; specifically, from the parallel postulate. I leave it to you to do so if you like this sort of thing.)

Combining the facts that both opposite and corresponding angles are equal, we see that when one line crosses two others that are parallel, we have a whole bunch of equal angles:

Picture here: 4.08truein by 2.18truein (Pyth11 scaled 1000)

In particular, we see that in this situation

corresponding zig-zag angles are equal.

(The high-falutin' way to say zig-zag angles is to call them alternate interior angles. You'll need to know that to read about them in standard geometry books.)

O.K. Now, take a triangle, and put a line through its top corner (or "vertex"; most math books just love to refuse to call anything by an ordinary name), making it parallel to the base. Using that zig-zag angles are equal shows us that the angles of the triangle do indeed add up to a straight line!

Picture here: 3.56truein by 1.42truein (Pyth12 scaled 1000)

6.3 More on Triangles

While I don't think that congruent triangles really come up very often in calculus, I think that understanding a few simple results about them will help us understand the analogous results about similar triangles. Also, I think that the part of high-school geometry a student is most likely to remember is how to spot congruent triangles, and how to use the basic results about them to figure out other elementary facts. For instance, as we'll soon see, if a triangle has two sides equal (is *isosceles*), then the angles opposite these sides are also equal, and vice-versa. We'll be able to see this using congruent triangles.

Two triangles are congruent if they have the same size and shape. We've already seen that Euclid's Proposition 4 says that SAS, or Side-Angle-Side, is enough to insure congruence. Side-Angle-Side is our shorthand for saying that the two triangles have a side, an angle, and another side, in that order, equal. Euclid's proof, remember, really just says that if we put one triangle on top of the other, with the equal sides and angle matching up, then the remaining parts of the triangles have to match up as well.

The other propositions about congruent triangles are SSS, ASA, and AAS.

Exercise Explain why SSA does *not* work.

Why does SSS work? At first I thought that we could use the same argument as for SAS: just put one triangle on top of the other. But two 4-sides figures (*quadrilaterals*) whose sides are equal don't have to be the same:

Picture here: 3truein by 3truein (Fig1 scaled 600)

What's the difference between triangles and quadrilaterals that's showing here? It's that a triangle with given sides is rigid, which isn't true for a quadrilateral. This is why the frameworks of bridges is made up of triangles. Can we give a common sense argument why this is true? Sure - if we have a triangle with two sides specified, then to change the angle between them we have to either squish or stretch the third side:

Picture here: 3truein by 3truein (Fig2 scaled 600)

So if the third side is also fixed, we can't change the angle between the first two. This means that's there's only one possible shape for a triangle whose sides are specified, which is what SSS means.

Is this how Euclid shows SSS? Not at all; he makes two propositions out of it, and uses that an isosceles triangle has equal angles. Do working mathematicians know Euclid's proof? According to my own informal survey, they do not. Still, it is kind of neat, so let's make it an exercise. Your hint is the outline of what you need to do: start by putting the bases of the two triangles together. Then all you need to show is that the top vertices coincide. Suppose, as Euclid did, that they *don't* coincide. Draw a picture of this, and you'll see that connecting the two different top vertices gives you a base of couple of downward pointing triangles, which turn out to be isosceles by our SSS assumptions. But now if you look at the base angles in each of these isosceles triangles, which ought to be equal, you'll see that this is impossible. Thus the two vertices *can't* be different, since assuming they are makes us wind up in an impossible situation.

Exercise Fill in the details, including of course the picture, of Euclid's proof of SSS¿

One way to see that ASA and AAS are true is to use that the angles of a triangle add to a straight line. Thus if two of the angles are the same in my triangle as in yours, so are the third angles in each the same. This means that having ASA or AAS automatically gives us SAS, and we already know that SAS means our triangles are congruent.

Again, this isn't the way Euclid did it, but then Euclid was very careful not to use the parallel postulate if he could get around it. And our result that the angles of a triangle add to a straight line does use the parallel postulate.

Exercise Explain ASA by putting one triangle on top of the other.

How can we use these triangle congruence theorems? Let's look at a couple of examples. To see that an isosceles triangle has equal angles, for instance, let ABC be isosceles, as shown. Then note that ABC is congruent to its own mirror image, ACB, by ASA:

Picture here: 3truein by 3truein (Fig* scaled 600)

This means that angles B and C are indeed the same.

Exercise Use the same mirror image idea to explain why a triangle with equal angles has equal sides.

We have already seen that if two lines are parallel, then a line falling across them has corresponding and zig-zag angles equal. In fact, the *converse* is also true: if the zig-zag angles are equal, or if the corresponding angles are equal, then the lines are parallel.

To see that the zig-zag angles being equal implies that the lines are parallel, we can show the equivalent result that if the lines aren't parallel, then the zig-zags can't be equal (remember? these two statements are *contrapositives* of one another). But if the two lines aren't parallel, they meet on one, and thus form a triangle with the crossing line:

Picture here: 3truein by 3truein (Fig* scaled 600)

Now the zig-zag angles A and B have to be different, as A is the supplement of C alone, whereas B is the supplement of C plus D.

Exercise Explain why corresponding angles equal means the lines are parallel.

Let's finish this section by seeing that opposite sides of a parallelogram are equal, and that, conversely, if opposite sides of a quadrilateral are equal, then it's a parallelogram.

Suppose first that we have a parallelogram, which means two pairs of parallel lines crossing. If we connect opposite corners of the parallelogram with a diagonal, then we can use zig-zag angles to get that the two triangles formed are congruent, and this gives us what we want. Where's the picture? You have to provide it:

Exercise Provide the picture, and the details, of the argument sketched just above.

Exercise Provide the picture and everything for why opposite sides equal means we have a parallelogram.

6.4 Similar Triangles

We've been calling two triangles congruent if they have the same size and shape. They're *similar* if they just have the same shape. If you take a closer look at just what this means, you'll see that similar triangles have their angles equal and corresponding sides proportional:

Picture here: 3truein by 2truein (Fig1 scaled 600)

Actually there are two ways of thinking about what it means to have corresponding sides proportional, but they really amount to the same thing. In the figure, side A and X are corresponding, as are B and Y, and C and Z. One way to say that corresponding sides proportional is to write

$$\frac{A}{X} = \frac{B}{Y} = \frac{C}{Z};$$

the other way is to say that

$$\frac{A}{B} = \frac{X}{Y}; \quad \frac{A}{C} = \frac{X}{Z}; \quad \frac{B}{C} = \frac{Y}{Z}.$$

To see that these are equivalent, remember that two fractions are equal if and only if the cross-products are equal. That means that $\frac{A}{X} = \frac{B}{Y}$ is the same as $A \times Y = B \times X$, but so is $\frac{A}{B} = \frac{X}{Y}$ the same as $A \times Y = B \times X$. Thus $\frac{A}{X} = \frac{B}{Y}$ and $\frac{A}{B} = \frac{X}{Y}$ really say the same thing.

Exercise Let $A, B, C = 1, 2, 3$ and $X, Y, Z = 3, 6, 9$. Draw the triangles, and verify that all the proportions claimed are indeed equal.

For telling when two triangles are congruent, we have our results SAS, SSS, and so forth. Is there any reason to memorize these? Of course not! You can tell by common sense which possible combinations of the letters S and S really work, and which don't; there's nothing to bother memorizing, since you can always check yourself.

What's neat is that all these SAS-like results also work for similar triangles, if we think of S as indicating that the corresponding sides indicated are proportional instead of equal. Thus SSS for similar triangles would say that if two triangles have corresponding sides proportional, then they're similar. That is, if the corresponding sides are proportional, we get for free that the angles are also equal.

I already mentioned that the high-school geometry text I have at hand doesn't get around to similar triangles until after page 300. Likewise, Euclid doesn't get to similar triangles until the 6th book (out of 13) of the *Elements*. He begins by arguing that triangles with the same height have areas proportional to their bases.

Now back in Volume 1 we argued that the area of a triangle is half its base times its height. If this is really correct then Euclid's proposition, that triangles with the same height have areas proportional to their bases, follows immediately. For we have

$$\frac{\text{my base}}{\text{your base}} = \frac{\text{my base} \times \text{height}}{\text{your base} \times \text{height}},$$

since multiplying the top and bottom of a fraction by the same amount doesn't change its value.

How correct was our argument, then, that the area of a triangle is half its base times its height? If you have another look, from our now more sophisticated perspective, you'll see to make sure that

- two copies of a triangle do fit together to make a parallelogram, and that
- a parallelogram can be dissected and rearranged to make a rectangle, and that
- a rectangle can be divided up into a bunch of little rectangles, like we drew it,

all just requires using the facts we now know about parallel lines.

That part's all fine, then. And as long as the sides of our rectangle are fractions, we can cut both sides up into a whole number of units, all of the same length; just find a common denominator of the sides, and take our unit length to be its reciprocal. But there's a problem here which, legend has it, came as a nasty surprise to the Greeks: there are lengths that aren't fractions. Let's postpone our discussion of lengths that aren't fractions (or "irrational numbers") until the section that treats exactly that topic, though. Suffice it to say for now that any irrational number can be arbitrarily closely approximated by a rational one, and that fussing with this idea will let us give a 100%

legal argument that the area of an arbitrary rectangle is indeed its base times its height.

Thus we can be quite confident of our understanding that triangles with the same height have areas proportional to their bases. From this we can get a result that is the heart of all the stuff about similar triangles:

> *A line parallel to the base of a triangle cuts the sides proportionally.*

Picture here: 3truein by 2truein (Fig3 scaled 600)

That is, if DE is parallel to BC, then

$$\frac{AD}{DB} = \frac{AE}{EC}.$$

Here's Euclid's proof: add in diagonals like this,

Picture here: 3truein by 2truein (Fig* scaled 600)

Now start by thinking of the left side of the triangle as being two bases; that is, we consider AD and BD to be the bases of triangles ADE and BDE, and note that these have the same height, measured from E:

Picture here: 3truein by 2truein (Fig* scaled 600)

Using Euclid's own abbreviated notation, this means that

$$\frac{ADE}{BDE} = \frac{AD}{BD}.$$

Likewise, taking the right side to be two bases, we have that

$$\frac{ADE}{CDE} = \frac{AE}{CE}.$$

Finally, taking the bottom to be a common base, we have that

$$\frac{BDE}{CDE} = \frac{BC}{BC} = 1.$$

This means that BDE and CDE are equal, which means the fractions $\frac{ADE}{BDE}$ and $\frac{ADE}{CDE}$ are really equal, since they have the same top, and their bottoms are equal. But this then means that $\frac{AD}{BD}$ and $\frac{AE}{CE}$ are equal, which is what it means for the line DE to cut the sides proportionally.

Here we have interpreted "cuts the sides proportionally" to mean that the ratio of the top part to the bottom part of each side is the same. But this is the same as saying the the ratio of the top part to the whole side is the same.

Now let's see why AAA works for similar triangles. If we have a little triangle and a big one with the same angles, we can fit the little one into the top corner of the big one, like this:

Picture here: 3truein by 2truein (Fig* scaled 600)

But now the base of the little one makes a line across the big one which is parallel to the big base, as we have corresponding angles equal. Thus the left and right sides of our triangles are in the same proportion.

Now again fit the little triangle into the big one, this time into the left corner. Thinking now of the right side of each as being the base, we again see that the base of the little one makes a line parallel to the base of the big one. This means, then, that the left and bottom sides of the little and big triangles are in the same proportion.

But since both the left and right, and also the left and bottom sides of our triangles are in the same proportion, all three sides are in this same proportion, which is what we wanted.

Now in fact we have AA for similar triangles; if two angles are the same, then the triangles are similar. Why? Because if two angles are the same in two triangles, so are the third angles.

Exercise Explain the last statement.

Now note that when we have a line parallel to the base of a triangle, the little triangle formed is actually similar to the big one, since the two base angles of the two triangles are equal. Thus not only are the left and right sides of the little and big triangles proportional, but their bases as well.

Note also that since AA alone is enough to guarantee that two triangles are similar, we get for free ASA and AAS.

We still need to see that SAS and SSS work for similar triangles.

To see SAS, suppose that ABC and DEF have their left and right sides proportional, and their top angles equal, like this:

Picture here: 3truein by 2truein (Fig* scaled 600)

Choose X on DE so that $DX = AB$, and make XY parallel to EF. Then XY cuts the sides of DEF proportionally, so

$$\frac{DY}{DF} = \frac{DX}{DE} = \frac{AB}{DE} = \frac{AC}{DF},$$

and the equality of the first and last ratios in this chain gives us that $AC = DY$. Thus, by SAS we have that ABC and DXY are congruent. But since we made DXY by putting in a line parallel to the base of DEF, it's similar to DEF, and hence so is ABC.

Picture here: 3truein by 2truein (Fig* scaled 600)

For SSS, suppose that ABC has its sides proportional to those of DEF.

Picture here: 3truein by 2truein (Fig* scaled 600)

Again we choose X on DE so that $DX = AB$, and make XY parallel to EF. Then again DXY is similar to DEF, hence has its sides proportional, and since one of its sides (DX) is equal to a side of ABC, the others must be as well. Thus again DXY is congruent to ABC and similar to DEF, which means that ABC and DEF are similar.

6.5 Constructions

Maybe you don't really need to do any ruler-and-compass constructions in calculus, but they're fun and they show a practical application of some of the theoretical stuff about triangles.

We should really say straightedge-and-compass constructions, since the rules of the construction game don't allow you to use marks on your straightedge. A compass is a sort of two-legged gizmo that has a needle point for one leg, and a pencil for the other, like this:

Picture here: 3truein by 2truein (Fig* scaled 600)

You can spread the compass apart a certain amount, and it'll stay that far apart, so by sticking the needle end at a point you can use the pencil end to draw a circle.

In Euclid's own constructions, he might as well have been using a *collapsible* compass. That is, for his constructions you can pretend that the compass only holds its shape as long as you have the needle stuck in a given point, and that once you pick it up to move it, it gets all floppy until you stick the needle into another point. This means that Euclid didn't just assume you could use the compass to transfer a given length from one place to another, and he managed to figure out a clever way to do this. We'll assume our compass will hold its shape, though.

Here are the constructions we'll work out:

1. To copy a line segment.
2. To copy an angle.
3. To copy a triangle.
4. To make an equilateral triangle, given a side.
5. To make a line parallel to a given line, through a given point.
6. To bisect a line segment.
7. To drop a perpendicular.
8. To bisect an angle.
9. To make a similar triangle on a given base.

line perpendicular to given line and through some point perp bisector of a segment angle bisector copy of an angle circle with given radius through two points

Chapter 7

Special Triangles

Several kinds of triangles have special names:

- **an equilateral triangle** is one with all three sides equal;
- **an isosceles triangle** is one with at least two sides equal;
- **an right triangle** is one with a right angle in it.

Note that every equilateral triangle is also isosceles.

Exercise: All equilateral triangles are similar.

Exercise: Equilateral triangles have equal angles.
A couple of related results are that

- an isosceles triangle has two equal angles (in fact, the angles opposite the equal sides are equal), and that
- a triangle with two equal angles is isosceles (in fact, the sides opposite the equal angles are equal).

The first of these is historically famous as Euclid's Proposition 5. His argument essentially goes like this: given the isosceles triangle ABC, note that it is congruent to its mirror image ACB by SSS. hence corresponding angles are equal, which means that angle ABC is equal to angle ACB.

Picture here: 3truein by 2truein (Fig1 scaled 600)

Now suppose that the base angles in triangle ABC are equal. Then again ABC is congruent to ACB, this time by SAS, so corresponding sides are equal. Thus, side AB is equal to side AC.

Picture here: 3truein by 2truein (Fig2 scaled 600)

We need now to talk about some special angles, and about how we measure angles. We've already talked about right angles, without really fussing about exactly what they are, and we've talked about straight angles, whose meaning is clear. Very likely you've heard of these 90 degree angles and 180 degree angle: if you make an ordinary turn at most intersections, you turn 90 degrees; if you make a U-turn, that's 180 degrees; and if you go completely around in a circle, that's 360 degrees.

But why 360? Why not make it so a degree is 1/100 of a circle?

Part III

The calculus: two new tools

Chapter 8

The calculus: two new tools

Calculus really just means bringing in two new tools to help us understand functions; the *slope*, or derivative, of a function, and its *area*, or integral. These two new tools, which turn out to be opposites sides of the same coin, are a lot more powerful than we might expect, at least in combination; Newton's and Leibnitz's independent discovery of the derivative in the mid 1600's led to revolutionary advances in mathematics, particularly in the mathematics of *change*, which is arguably what calculus is. We've really already seen the integral, or its beginnings, in figuring out areas and volumes of some of the shapes and solids we've looked at. And this goes back to Archimedes (xxxBC).

Chapter 9

The slope of a function

9.1 graphing the slope function

Given a function $f(x)$, imagine a little line segment riding along on its graph like a car on a roller coaster, always staying tangent to the graph, that is, just touching it in one point. The slope of $f(x)$ at a point, which we take to mean the same as the slope of our tangent segment, gives us a new function, which we'll call the slope function, or the derivative. We use the notation $f'(x)$ for this derived function.

Working just from a *picture* of $f(x)$, that is, without any formula for it, we can make a pretty good sketch of the graph of $f'(x)$, and we can figure out some general rules about this derivative.

To draw f' from f, we might first notice that f' is 0 wherever f has a *flat spot*, meaning a place where the roller coaster car is horizontal. Between the places where f' is 0, we just need to figure out whether it's positive or negative, and increasing or decreasing. We may not be able to figure out the exact values of f', but that may well be beside the point anyhow. Here's an example; suppose this is f:

Picture here: 3truein by 3truein (Fig1 scaled 600)

Now, where f starts, at the left of the picture, it looks like the slope is somewhere around 2 or 3; at any rate, it's certainly positive. And at the first high bump on f, the slope is 0. Between where f starts, and this flat spot, you can see that the slope decreases, if you just imagine the car driving along the roller-coaster track.

Chapter 10

The area of a function

10.1 introduction

The concept of the area of a function, meaning the area under its graph, is a little more difficult than the concept of the slope. Why? Because you can't just talk about *the* area at a given point, like you can the slope at a given point; instead you have to say *from where to where* you mean to take the area. So there really isn't such a thing as *the* area function, but a choice of many of them; infinitely many, in fact. We'll see, though, that this doesn't really amount to a big problem; the different area functions turn out not to be all that different after all.

10.2 graphing some area functions

Let's start by delaring an area function that tells us, for a given x, how much area there is under the graph of f between 0 and x. That is, we'll base this area function at 0. Suppose now that f looks like this:

Picture here: 4.5truein by 1.5truein (AreaFcn1 scaled 600)

And for a given x we our area function $A(x)$ we take to be this area:

Picture here: 4.5truein by 1.5truein (AreaFcn2 scaled 600)

What does the graph of $A(x)$ look like? Well, for one thing, $A(0)$ is certainly going to be 0, since the shaded area collapes to nothing when $x = 0$. And since this particular f is positive, $A(x)$ is going to increase as x does. But it's going to increase more and more slowly as x increases, since $f(x)$ is getting smaller and smaller. So $A(x)$ starts at 0 and increases as x does, though more and more slowly. So $A(x)$ looks something like this, for positive x:

Picture here: 4.5truein by 1.5truein (AreaFcn3 scaled 600)

What about $A(x)$ for negative x? Well, since we drew f to be an even function, we can see that the area under the graph of f between -2 and 0, say, is going to be exactly the same as the area under it between 0 and 2, and so on. But we're going to count the area to the *left* of 0 as negative. Here's why: the notation we're going to use for this area function was invented by Leibnitz, and looks like this:

$$A(x) = \int_0^x f(t)dt.$$

Here the $\int$ sign is a fancy s, standing for *sum*, and the 0 and x mean to take the sum for all numbers between 0 and x. We imagine chopping the interval $[0, x]$ up into infinitely many infnitesimal pieces, and for each t in $[0, x]$ we add to our sum the area of the infinitesimally narrow rectangle whose base is dt and whose height is $f(t)$. That's what Leibnitz intended the notation to mean. He thought of dt as being the infnitesimal change between two consecuative values of t. Suppose now that we want

$$A(-2) = \int_0^{-2} f(t)dt.$$

In this sum, the dt's are all negative, since we think ofthe t's as running from 0 *down* to -2.

If you have trouble with this explanation, you can instead settle for this one: we take the area to the left of 0 to be negative because doing so will make things turn out better later, for reasons that we'll be able to see clearly when the time is right.

At any rate, given that we do declare that $A(x)$ is to be negative here for negative values of x, we see that $A(x)$ is an odd function, so it looks like this:

Picture here: 4.5truein by 2truein (AreaFcn4 scaled 600)

Now let's let $B(x)$ the area function ....

Chapter 11

Slope and Area Compared

Chapter 12

Computing Slopes

Suppose we were a 17^{th}-century mathematician, and had discovered the limit definition to get the slope of the tangent line to the graph of a function f at a point x.

Picture here: 2.2 truein by 2truein (Fig1 scaled 600) Picture here: 2.2 truein by 2truein (Fig2 scaled 600)

(Δx positive) (Δx negative)

$$f'(x) = \frac{dy}{dx} := \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

What we would very likely hope to do might well be to find nice simple rules for the derivatives of the functions most familiar to us; *polynomials* like $3x^2+4x+6$, *rational functions* like $(3x^2+4x+6)/(x^3+2x-11)$, *algebraic functions* like $\sqrt{3x^2+4x+6}/(x^3+2\sqrt{x}-11)$, *trigonometric functions*, like $\sin x$ or $\cos x/\sin x$, and *expopnential* and *logarithmic* functions like 2^x and $\log_2 x$ (all these functions and their further combinations comprise the so-called *elementary functions*).

In fact, if you think about it, and take another look at your text, this is exactly what is done in the sections introducing and developing the derivative.

Let's make a quick review of the results we thus discover, recall why they hold, and see how they fit together to enable us to take the derivative of any elementary function. We should bear in mind that we next might hope to accomplish the same sort of results for integrals; but to be honest I may as well tell you that things don't turn out so easy for the integral!

Were we indeed a mathematician who had just discovered the derivative, I think it quite likely that we might begin by looking for a formula for the derivative for the simplest polynomials, like x^7. Here if we just plug into our formula, things pretty much take care of themselves, for we have

$$\frac{d}{dx}x^7 = \lim_{\Delta x \to 0} \frac{(x + \Delta x)^7 - x^7}{\Delta x},$$

and in the numerator, when we expand $(x + \Delta x)^7$ according to the binomial formula, we see that the x^7 terms cancel, leaving us with $7x^6\Delta x$, plus $(\Delta x)^2$ times an expression (the exact details of which we may, thank God, ignore) in x and Δx. Thus we may divide through by the denominator Δx, and get $7x^6$ plus an expression that goes to 0 as Δx does.

Thus we see that $(x^7)' = 7x^6$, and that the same argument easily extends to the general case $(x^n)' = nx^{n-1}$, where n is a positive whole number.

At this point we would probably go on and discover directly that the same result in fact holds for negative and fractional powers as well. The algebraic manipulations for negative and fractional powers are a bit trickier, but would likely occur to us after a bit of trial and error. But a little hindsight is a wonderful thing; let's see that if we first establish a few general rules, we get the result $(x^q)' = qx^{q-1}$ for all rational q, positive or negative.

Now, we surely wouldn't be long in discovering the crucial *linearity* of the derivative;

$$(af(x) + bg(x))' = af'(x) + bg'(x),$$

for arbitrary functions $f(x)$ and $g(x)$, and constants a and b. If you don't think this is a very nice property indeed, consider the following:

what are the nice linear functions on the real numbers? That is, what functions $f(x)$ (continuous, say) have the property that $f(x+y) = f(x) + f(y)$? Note that despite generations of testimony to the contrary, many of our most familiar functions decidedly do *not* have this property; $(x+y)^2 \neq x^2 + y^2$, $\sin(x+y) \neq \sin x + \sin y$, etc., etc. You should be able to determine exactly which continuous functions do have this property, and I think you'll be surprised by their paucity.

Having seen that the derivative so nicely respects addition, we might expect as well to have the trivial rule $(fg)' = f'g'$ for products, but even as simple an example as $f(x) = x = g(x)$ shows this is not the case. Instead we have the more puzzling *product rule*, whose proof is easily seen in pictures; if we consider the product fg, then we see that $\Delta(fg)$ is the darkly shaded area here:

Picture here: 2 truein by 2truein (Fig3 scaled 600)

and that we then have

Picture here: 4 truein by 4truein (Fig4 scaled 600)

thus establishing the product rule:

$$(fg)' = f'g + g'f;$$

Another crucial rule is the *chain rule*, which seems quite clear if we write it as

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx},$$

and indeed the proof essentially consists of just taking the limit of the related trivial assertion,

$$\frac{\Delta z}{\Delta x} = \frac{\Delta z}{\Delta y} \cdot \frac{\Delta y}{\Delta x}.$$

In fact, nearly all of the proof is devoted to being fussy about the possibility that the Δy in the denominator might sometimes be equal to 0.

We are now prepared to derive the power rule for arbitrary *positive* fractional powers. To do so we use the Chain Rule to perform what is known as *implicit differentiation*; to find y' if, say,

$$y = x^{2/3},$$

we first cube both sides to get

$$y^3 = x^2.$$

Now we take d/dx of both sides. But by the Chain Rule, $\frac{d}{dx}(y^3) = 3y^2 \cdot y'$, so we have

$$3y^2 \cdot y' = 2x,$$

and so, solving for y', we have

$$y' = \frac{2x}{3y^2} = \frac{2x}{3(x^{2/3})^2} = \frac{2}{3}x^{-1/3};$$

i.e.,

$$(x^{2/3})' = \frac{2}{3}x^{-1/3};$$

so that we have still "brought down the exponent in front and reduced it by 1", as we do for positive integer powers. You yourself should repeat the exact same argument for the general case and see that, indeed,

$$(y^{\frac{m}{n}})' = \frac{m}{n}y^{\frac{m}{n}-1}.$$

We can also use implicit differentiation to get the *reciprocal rule*, that

$$(\frac{1}{x})' = \frac{-1}{x^2},$$

or, in terms of exponents, that $(x^{-1})' = -1 \cdot x^{-2}$, just as we would hope. For if

$$y = \frac{1}{x},$$

then

$$xy = 1,$$

so that, taking the derivative with resect to x of both sides (and hence using the Product Rule and Chain Rule),

$$y + xy' = 0.$$

If we now solve this for y' we get

$$y' = \frac{-y}{x} = \frac{-1}{x^2},$$

as we wanted.

Now we are ready to take the derivative of *negative* integer power of x. For if E is any expression in x, by the Reciprocal Rule combined with the Chain Rule we have

$$\frac{(1}{E)' = \frac{-1}{E^2} \cdot E'.}$$

Now if

$$y = x^{-5} = (x^5)^{-1} = \frac{1}{x^5},$$

then

$$y' = (\frac{1}{x^5})' = \frac{-1}{(x^5)^2} \cdot (x^5)' = \frac{-5x^4}{x^{10}} = -5x{-}6,$$

which is again what we would hope for.

How about the *quotient rule*? Actually, it's an easy consequence of the the reciprocal rule, the product rule, and the chain rule. For by the reciprocal and chain rules,

$$\left(\frac{1}{g}\right)' = \frac{-1}{g^2} \cdot g' = \frac{-g'}{g^2},$$

(which result we may call the *generalzed reciprocal rule*) so that

$$\left(\frac{f}{g}\right)' = \left(f \cdot \frac{1}{g}\right)' = f'\left(\frac{1}{g}\right) + f \cdot \left(\frac{1}{g}\right)' = \frac{f'}{g} + f \cdot \frac{-g'}{g^2} = \frac{f'}{g} - \frac{fg'}{g^2} = \frac{f'g - fg'}{g^2}.$$

Our final big rule, and the one that is probably least understood by beginners, is the *inverse function theorem.* I have to give credit here to the thousands of math texts that quite successively manage to make this utterly simple result seem so obscure; theirs is a veritable *tour de force* of obfuscation. The simple truth is that if y is a function f of x, then x *is* f^{-1} *of* y; "Di is the wife of Charles" and "Charles is the husband of Di" are just two equivalent ways of expressing exactly the same relationship. Suppose then that $y = f(x)$, for which the picture is

Picture here: 3 truein by 2truein (Fig5 scaled 600)

$$y = f(x)$$

Then the *exact same picture*, turned over, shows the graph of the inverse function, $x = f^{-1}(y)$.

Picture here: 2 truein by 3truein (Fig6 scaled 600)

$$x = f^{-1}(x)$$

Now consider the tangent line shown, which is the same in both views. Its slope in the second view is the reciprocal of that in the first, since the rôles of x and y as *run* and *rise* are reversed. There is absolutely nothing more than this to the inverse function theorem.

As in the case of the chain rule, this rule looks trivial when written in dx-notation:

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

As an application of this rule, we can get the derivative of $y = \sqrt[3]{x} = x^{1/3}$. For then $x = y^3$, so that by the power rule for positive integer powers we know $\frac{dx}{dy} = 3y^2$. Hence by the inverse function theorem

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{3y^2} = \frac{1}{3x^{2/3}},$$

where the last equality is got by plugging in $x^{1/3}$ for y. If we rewrite this purely in terms of exponents it becomes

$$\left(x^{1/3}\right)' = \frac{1}{3}x^{-2/3},$$

extending perfectly our power rule.

Now for an arbitrary positive fractional power we can combine this result with the chain rule; for instance,

$$\left(x^{2/3}\right)' = \left(\left(x^{1/3}\right)^2\right)' = 2\left(x^{1/3}\right)^1 \cdot \left(x^{1/3}\right)',$$

and so on, the rest of this derivation being clear in terms of what we have already established.

Finally, for a negative fractional power, like $y = x^{-2/3} = \frac{1}{x^{2/3}}$, we can apply what we have just derived along with the general reciprocal rule to verify that the power rule works in this case as well.

With what we have developed so far we can take the derivative of any algebraic expression whatever, at least in theory, without resorting to any limits. It remains to extend our prowess to *transcendental* (just a fancy word for 'not algebraic') functions.

Since all the trigonometric functions may ultimately be defined in terms of $\sin x$ ($\cos x = \sin(\frac{\pi}{2} - x)$; $\tan x = \frac{\sin x}{\cos x}$, etc.), to handle all of them we need only directly establish the derivative of $\sin x$. Here again we are forced to consider first principles, and must resort to the limit definition. We have

$$\frac{d}{dx}\sin x := \lim_{\Delta x \to 0} \frac{\sin(x + \Delta x) - \sin x}{\Delta x}.$$

To proceed we need to know the formula for $\sin(A + B)$. But as a 17^{th}-century mathematician interested in furthering her study of these familiar functions, we surely know this result well. We then have

$$\frac{d}{dx}\sin x := \lim_{\Delta x \to 0} \frac{(\sin x \cos \Delta x + \cos x \sin \Delta x) - \sin x}{\Delta x} =$$

$$\lim_{\Delta x \to 0} \frac{\sin x(\cos \Delta x - 1)}{\Delta x} + \lim_{\Delta x \to 0} \cos x \frac{\sin \Delta x}{x} =$$

$$\sin x \left[\lim_{\Delta x \to 0} \frac{\cos \Delta x - 1}{\Delta x} \right] + \cos x \left[\lim_{\Delta x \to 0} \frac{\sin \Delta x}{\Delta x} \right].$$

At this point we now have to convince ourselves of the familiar limit

$$\lim_{t \to 0} \frac{\sin t}{t} = 1,$$

which is fairly straightfoward with the usual picture (and from this we can derive the related $\lim\limits_{t \to 0} \frac{\cos t - 1}{t} = 0$, by multipling numerator and denominator by $\cos x + 1$). Well, this all takes a bit more work than our basic algebraic derivative — the power rule for positive integers, which is a simple application of the binomial formula. But then $\sin x$ itself is a much less obvious function than x^n.

If we now try to finish by determining the derivative, using the original limit definition, of an exponential function like 2^x, we can't quite push it the all the way through. For by definition we have

$$\frac{d}{dx}2^x := \lim_{\Delta x \to 0} \frac{2^{x+\Delta x} - 2^x}{\Delta x} = \lim_{\Delta x \to 0} \frac{2^x(2^{\Delta x} - 1)}{\Delta x} = 2^x \left[\lim_{\Delta x \to 0} \frac{2^{\Delta x} - 1}{\Delta x} \right],$$

and if we reflect for a moment we see that the last limit is precisely the definition of the derivative at the point $x = 0$. Thus the derivative of 2^x is just 2^x back again, *times whatever its derivative is at 0.* Thus we can rest assured that all is well provided only that we can convince ourselves that the derivative does indeed exist at 0, i.e., that the last limit exists. But here we're stuck; while a picture of the graph certainly makes us believe with all our heart and soul that there is a well-defined tangent line at the point $(0, 1)$, the limit is still elusive.

Suppose however that as a *reasonable* 17^{th}-century mathematician, we decide to accept for the time being the existence of the derivative of 2^x at $x = 0$. Then we have that $(2^x)' = 2^x \cdot C_2$, where the constant C_2 is the value of the derivative at 0. We would of course have as well that $(3^x)' = 3^x \cdot C_3$, etc. and we could hardly fail to conjecture that for some base a we should have $C_a = 1$ so that $(a^x)' = a^x$, i.e., so that a^x is its own derivative. Numerical experimentation would lead us to see that a lies somewhere between 2 and 3. In fact, of course, this a is the well-known $e = 2.71828\ldots$. The point here is that we probably wouldn't know about e before having started our developement of calculus, since it may be defined as above as that basis a so that a^x is exactly its own derivative.

You may recall that the derivative of the exponential function is finally resolved after the introduction of the integral, and the fundamental theorem of calculus. For since the function $\frac{1}{x}$ is nice for positve reals, the function

$$\log x = \int_1^x \frac{1}{t} dt$$

is well-defined for all x greater than 0. Then by the fundamental theorem, we have

$$\frac{d}{dx} \log x = \frac{1}{x}.$$

Moreover, since the definition makes it clear that $\log x$ is a strictly increasing function, we know that it has a single-valued inverse. If we temporarily denote this inverse function by "exp", so that $y = \log x$ if and only if $x = \exp y$, then by the inverse function theorem we have

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{\frac{1}{x}} = x,$$

i.e.,

$$\frac{d}{dy} \exp y = \exp y.$$

Now what we would discover is that $\log x$ does in fact behave like a logarithm, which is equivalent to saying that its inverse behaves like an exponential; i.e.

$$\log(xy) = \log x + \log y$$

and

$$\exp(a) \cdot \exp(b) = \exp(a + b)$$

are equivalent statements. And from the latter we see that if we put $e = \exp 1$ then $\exp q = e^q$ for all rational q. Now since $\exp y$ is differentiable — hence continuous — and since we want e^y to be continuous as well, we see that $\exp y$ must be equal to e^y for all real y.

With this last one of our basic tools made legitimate, we can now take the derivative of any elementary function whatever without ever again having to resort to the original limit definition of the derivative. To summarize, let us list our basic rules, i.e., those that are derived from first principles:

Specific

- crudest power rule: $(x^n)' = nx^{n-1}$ for positive integer n;
- reciprocal rule: $(\frac{1}{x})' = \frac{-1}{x^2}$;
- sine rule: $(\sin x)' = \cos x$;
- exponential rule: $(e^x)' = e^x$.

General

- linearity: $(af(x) + bg(x))' = af'(x) + bg'(x)$;
- product rule:$(fg)' = f'g + fg'$;
- chain rule: $\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$;
- inverse function rule: $\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$.

12.1 Infite sequences and sums

1. Sequences and limits

- If s_n converges then *the difference of successive terms must go to 0*;

$$s_{n+1} - s_n \to 0.$$

  Warning: The converse is false (see the harmonic series below).

- The possible kinds of behavior of a sequence are well illustrated by x^n:

$$x^n \quad \begin{cases} \text{diverges (to } \infty\text{) if } x > 1; \\ \to 1 \text{ if } x = 1; \\ \to 0 \text{ if } -1 < x < 1; \\ \text{diverges (with bounded oscillation) if } x = -1; \\ \text{diverges (with unbounded oscillation) if } x < -1; \end{cases}$$

  only divergence to $-\infty$ is not included here.

- If a_n and b_n are of the *same order of magnitude* ($a_n \asymp b_n$), in particular if a_n and b_n are *asymptotic* ($a_n \sim b_n$), then a_n and b_n *behave alike*; either both converge, both diverge to ∞, both diverge with bounded oscillation, etc. Example:

$$(-1)^n + \frac{1}{n} \sim (-1)^n, \quad \text{hence diverges with bounded oscillation.}$$

- In determining the behavior of ratios of sequences, *you can ignore all but the highest order terms*; i.e.

$$\frac{f + o(f)}{g + o(g)} \sim \frac{f}{g}.$$

  For example,

$$\frac{2k^2 + 3k + 1}{3k + 2} \sim \frac{2k^2}{3k} = \frac{2}{3}k, \quad \text{hence diverges to } \infty; \tag{12.1}$$

$$\frac{2k^2+3k+1}{5k^3+7} \sim \frac{2k^2}{5k^3} = \frac{2}{5k}, \quad \text{hence converges to } 0; \tag{12.2}$$

$$\frac{2k^2+3k+1}{7k^2+3k-2} \sim \frac{2k^2}{7k^2} = \frac{2}{7}, \quad \text{hence converges to } \frac{2}{7}; \tag{12.3}$$

$$\frac{2+\sin k+\frac{3}{k}}{3+\cos k+\frac{5}{\log k}} \sim \frac{2+\sin k}{3+\cos k}, \quad \text{hence diverges.} \tag{12.4}$$

$$\tag{12.5}$$

- An increasing, bounded sequence must converge.

- Important order of magnitude comparisons: for fixed (large) n we have

$$\lim_{x\to\infty} \frac{(\log x)^n}{x} = 0, \qquad \text{(i.e., } (\log x)^n = o(x)); \tag{12.6}$$

$$\lim_{x\to\infty} \frac{x^n}{e^x} = 0, \qquad (x^n = o(e^x)); \tag{12.7}$$

$$\lim_{k\to\infty} \frac{\left(e^k\right)^n}{k!} = 0, \qquad (e^k = o(k!)). \tag{12.8}$$

2. Infinite series

- $\sum^{\infty} a_k := \lim_{n\to\infty} s_n$, where $s_n = \sum^{n} a_k$ is the n^{th} *partial sum*. Thus infinite series are really limits of sequences.

- If Σa_k converges, then the a_k's *must converge to 0*. This follows immediately from 1(a), and again the converse is false (see the harmonic series below).

- The **harmonic series**:

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \infty.$$

- *Only tails count* in deciding convergence: $\sum_{k=1}^{\infty} a_k$ converges $\iff$ $\sum_{k=10^6}^{\infty} a_k$ converges, etc. (Of course, the entire series counts in determining the value.)

- **Geometric series**:

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \quad \text{if } |x| < 1\text{; diverges otherwise.}$$

- (f) **Telescoping series**: sometimes we can rewrite a series so that the new terms can be seen to cancel: we can then easily sum the series:

$$\frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \cdots = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k(k+1)} = \lim_{n\to\infty} \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) =$$

$$\lim_{n\to\infty} \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right) =$$

$$\lim_{n\to\infty} 1 - \frac{1}{n+1} = 1.$$

3. Nonnegative series.

- (a) **Integral test**: If $f(x) \downarrow 0$, then $\Sigma f(k)$ and $\int f(x)dx$ converge or diverge together.

- (b) p**-series**: By the integral test, $\sum \frac{1}{k^p}$ converges $\iff p > 1$.

- (c) **Comparison test**: for *nonnegative* a_k and b_k, if $a_k = O(b_k)$, then Σa_k converges if Σb_k does.

- Corollary: if $a_k \asymp b_k$, in particular if $a_k \sim b_k$, then Σa_k and Σb_k converge or diverge together. Examples:

$\sum \frac{1}{2k^2 - 47k - 23}$ converges by comparison with $\sum \frac{1}{k^2}$, since $\frac{1}{2k^2 - 47k - 23} \asymp \frac{1}{k^2}$.

-

- $\sum \frac{1}{(2k^2+47k+23)^{1/3}}$ diverges by comparison with $\sum \frac{1}{k^{2/3}}$.

4. Alternating series.

- (a) **Alternating series test**. If $a_k \downarrow 0$ then $\Sigma(-1)^k a_k$ converges. Example: the alternating harmonic series:

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k} \quad \text{converges.}$$

- (b) Alternating series error estimate: if $a_k \downarrow 0$ then $\left|\sum^{\infty}(-1)^k a_k - \sum^{n}(-1)^k a_k\right| \le |a_{n+1}|$; i.e., *the error is no more in magnitude than the first omitted term.*

5. Absolute and Conditional Convergence.

- (a) Σa_k *converges absolutely* means that $\Sigma|a_k|$ converges.

- (b) If Σa_k converges absolutely, then it converges; i.e., convergence of $\Sigma|a_k|$ implies that of Σa_k. Moreover, by the *triangle inequality*, $|\Sigma a_k| \le \Sigma|a_k|$

- (c) Examples:

- $\sum(-1)^k \frac{1}{\sqrt{k+1}}$ does not converge absolutely (by comparison with the p-series $\sum \frac{1}{k^{1/2}}$),

- but does converge (by the alternating series test), hence converges conditionally.

- $\sum(-1)^k \frac{1}{2k^2-15}$ converges absolutely by comparison with the p-series $\sum \frac{1}{k^2}$

- (d) **The Ratio Test**. Suppose that $\dfrac{|a_{k+1}|}{|a_k|} \to q$. Then

- (i) If $q < 1$, the series Σa_k converges absolutely.

- (ii) If $q > 1$, the series Σa_k diverges.

- (iii) If $q = 1$, the ratio test tells us nothing.

  Note that the ratio test can *only* tell us about *absolute* convergence, since it concerns the absolute values of the a_k's.

- (e) **Infinite Arithmetic**. You can do with *absolutely convergent* series everything you can do with finite sums: regrouping terms, rearranging terms, long multiplication, long division, etc.

- (f) **Remarks**

- For a series of *nonnegative* terms, absolute convergence is the same as convergence; there's no such thing as a conditionally convergent series of nonnegative terms.

- **A good rule of thumb** for a series with both positive and negative terms is to *test first for absolute convergence*, since you're going to have to do this anyhow.

- The **only test for conditional convergence** that we know of, in this course, is the alternating series test.

6. Power Series: $\sum a_k x^k$.

- (a) A power series *converges absolutely* on the *interior* of its interval of convergence, and diverges at every point outside this interval.

- (b) The **radius of convergence** is determined by the *ratio test*; is fact,

$$\rho = \frac{1}{q},$$

where q is as in the ratio test (with the understanding that $q = 0$ means $\rho = \infty$).

- (c) The behavior of a power series at the **endpoints** must be determined *ad hoc*.

7. Functions Defined by Power Series: $f(x) = \sum a_k x^k$.

- (a) A function, all its derivatives, and all its integrals, have the same *radius* of convergence. The *interval* of convergence need not be the same; as a general rule, a derivative is less likely to converge at the endpoints, and an integral more likely.

- (b) You can differentiate and integrate power series term-by-term on the interior of the interval of convergence.

- (c) If $f(x)$ has the power series expansion $f(x) = \sum a_k(x-a)^k$ around the point $x = a$, then the a_k's are determined by the derivatives of $f(x)$; to wit,

$$a_k = \frac{1}{k!} f^{(k)}(a).$$

8. Taylor Series.

- (a) **The n^{th} order Taylor polynomial**,

$$T_n(x) = f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n,$$

is the best approximation to $f(x)$ at the point a in that this is the unique polynomial of degree n or less such that

$$T_n(x) - f(x) = o(x-a)^n, \quad \text{i.e.,} \quad \lim_{x\to a} \frac{T_n(x) - f(x)}{(x-a)^n} = 0.$$

- (b) The n^{th} **error** R_n is like the first omitted term;

$$f(x) = T_n(x) + R_n, \quad \text{where} \quad R_n = \frac{1}{(n+1)!} f^{(n+1)}(c)(x-a)^{n+1},$$

for some $c \in (a, x)$. If this error goes to 0 as $n \to \infty$, then $T_n(x) = f(x)$ at the particular point x under consideration. In particular,

- (c) If all the derivatives of $f(x)$ are uniformly bounded in an interval containing a, then $f(x)$ is equal to its Taylor series in that interval.

- **Examples**

- $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$, for all x;

- $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$, for all x;

- $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$, for all x;

- $\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - x^5 + \cdots$, for all $x \in (-1, 1)$;

- $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$, for all $x \in (-1, 1]$; thus

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots = \log 2;$$

- $\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + x^8 - \cdots$, for all $x \in (-1, 1)$;

- $\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$, for all $x \in (-1, 1]$; thus

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \arctan 1 = \frac{\pi}{4},$$

whence the marvelous formula

$$\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} \pm \cdots\right).$$