

Chapter 14

Sums & Asymptotics

14.1 The Value of an Annuity

Would you prefer a million dollars today or \$50,000 a year for the rest of your life? On the one hand, instant gratification is nice. On the other hand, the total dollars received at \$50K per year is much larger if you live long enough.

Formally, this is a question about the value of an annuity. An *annuity* is a financial instrument that pays out a fixed amount of money at the beginning of every year for some specified number of years. In particular, an n -year, m -payment annuity pays m dollars at the start of each year for n years. In some cases, n is finite, but not always. Examples include lottery payouts, student loans, and home mortgages. There are even Wall Street people who specialize in trading annuities.

A key question is what an annuity is worth. For example, lotteries often pay out jackpots over many years. Intuitively, \$50,000 a year for 20 years ought to be worth less than a million dollars right now. If you had all the cash right away, you could invest it and begin collecting interest. But what if the choice were between \$50,000 a year for 20 years and a *half* million dollars today? Now it is not clear which option is better.

In order to answer such questions, we need to know what a dollar paid out in the future is worth today. To model this, let's assume that money can be invested at a fixed annual interest rate p . We'll assume an 8% rate¹ for the rest of the discussion.

Here is why the interest rate p matters. Ten dollars invested today at interest rate p will become $(1 + p) \cdot 10 = 10.80$ dollars in a year, $(1 + p)^2 \cdot 10 \approx 11.66$ dollars in two years, and so forth. Looked at another way, ten dollars paid out a year from now are only really worth $1/(1 + p) \cdot 10 \approx 9.26$ dollars today. The reason is that if we

¹U.S. interest rates have dropped steadily for several years, and ordinary bank deposits now earn around 1.5%. But just a few years ago the rate was 8%; this rate makes some of our examples a little more dramatic. The rate has been as high as 17% in the past thirty years.

In Japan, the standard interest rate is near zero%, and on a few occasions in the past few years has even been slightly negative. It's a mystery why the Japanese populace keeps any money in their banks.

had the \$9.26 today, we could invest it and would have \$10.00 in a year anyway. Therefore, p determines the value of money paid out in the future.

14.1.1 The Future Value of Money

So for an n -year, m -payment annuity, the first payment of m dollars is truly worth m dollars. But the second payment a year later is worth only $m/(1+p)$ dollars. Similarly, the third payment is worth $m/(1+p)^2$, and the n -th payment is worth only $m/(1+p)^{n-1}$. The total value, V , of the annuity is equal to the sum of the payment values. This gives:

$$\begin{aligned} V &= \sum_{i=1}^n \frac{m}{(1+p)^{i-1}} \\ &= m \cdot \sum_{j=0}^{n-1} \left(\frac{1}{1+p} \right)^j && \text{(substitute } j ::= i - 1) \\ &= m \cdot \sum_{j=0}^{n-1} x^j \quad \left(\text{substitute } x = \frac{1}{1+p} \right). \end{aligned} \tag{14.1}$$

The summation in (14.1) is a geometric sum that has a *closed form*, making the evaluation a lot easier, namely²,

$$\sum_{i=0}^{n-1} x^i = \frac{1-x^n}{1-x}. \tag{14.2}$$

(The phrase “closed form” refers to a mathematical expression without any summation or product notation.)

Equation (14.2) was proved by induction in problem 6.1, but, as is often the case, the proof by induction gave no hint about how the formula was found in the first place. So we’ll take this opportunity to explain where it comes from. The trick is to let S be the value of the sum and then observe what $-xS$ is:

$$\begin{array}{rcccccccc} S & = & 1 & +x & +x^2 & +x^3 & + \dots & +x^{n-1} \\ -xS & = & & -x & -x^2 & -x^3 & - \dots & -x^{n-1} - x^n. \end{array}$$

Adding these two equations gives:

$$S - xS = 1 - x^n,$$

so

$$S = \frac{1-x^n}{1-x}.$$

We’ll look further into this method of proof in a few weeks when we introduce *generating functions* in Chapter 16.

²To make this equality hold for $x = 0$, we adopt the convention that $0^0 ::= 1$.

14.1.2 Closed Form for the Annuity Value

So now we have a simple formula for V , the value of an annuity that pays m dollars at the start of each year for n years.

$$V = m \frac{1 - x^n}{1 - x} \quad (\text{by (14.1) and (14.2)}) \quad (14.3)$$

$$= m \frac{1 + p - (1/(1 + p))^{n-1}}{p} \quad (x = 1/(1 + p)). \quad (14.4)$$

The formula (14.4) is much easier to use than a summation with dozens of terms. For example, what is the real value of a winning lottery ticket that pays \$50,000 per year for 20 years? Plugging in $m = \$50,000$, $n = 20$, and $p = 0.08$ gives $V \approx \$530,180$. So because payments are deferred, the million dollar lottery is really only worth about a half million dollars! This is a good trick for the lottery advertisers!

14.1.3 Infinite Geometric Series

The question we began with was whether you would prefer a million dollars today or \$50,000 a year for the rest of your life. Of course, this depends on how long you live, so optimistically assume that the second option is to receive \$50,000 a year *forever*. This sounds like infinite money! But we can compute the value of an annuity with an infinite number of payments by taking the limit of our geometric sum in (14.2) as n tends to infinity.

Theorem 14.1.1. *If $|x| < 1$, then*

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1 - x}.$$

Proof.

$$\begin{aligned} \sum_{i=0}^{\infty} x^i &::= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} x^i \\ &= \lim_{n \rightarrow \infty} \frac{1 - x^n}{1 - x} && (\text{by (14.2)}) \\ &= \frac{1}{1 - x}. \end{aligned}$$

The final line follows from that fact that $\lim_{n \rightarrow \infty} x^n = 0$ when $|x| < 1$. □

In our annuity problem, $x = 1/(1+p) < 1$, so Theorem 14.1.1 applies, and we get

$$\begin{aligned} V &= m \cdot \sum_{j=0}^{\infty} x^j && \text{(by (14.1))} \\ &= m \cdot \frac{1}{1-x} && \text{(by Theorem 14.1.1)} \\ &= m \cdot \frac{1+p}{p} && (x = 1/(1+p)). \end{aligned}$$

Plugging in $m = \$50,000$ and $p = 0.08$, the value, V , is only \$675,000. Amazingly, a million dollars today is worth much more than \$50,000 paid every year forever! Then again, if we had a million dollars today in the bank earning 8% interest, we could take out and spend \$80,000 a year forever. So on second thought, this answer really isn't so amazing.

14.1.4 Problems

Class Problems

Problem 14.1.

You've seen this neat trick for evaluating a geometric sum:

$$\begin{aligned} S &= 1 + z + z^2 + \dots + z^n \\ zS &= z + z^2 + \dots + z^n + z^{n+1} \\ S - zS &= 1 - z^{n+1} \\ S &= \frac{1 - z^{n+1}}{1 - z} \end{aligned}$$

Use the same approach to find a closed-form expression for this sum:

$$T = 1z + 2z^2 + 3z^3 + \dots + nz^n$$

Homework Problems

Problem 14.2.

Is a Harvard degree really worth more than an MIT degree?! Let us say that a person with a Harvard degree starts with \$40,000 and gets a \$20,000 raise every year after graduation, whereas a person with an MIT degree starts with \$30,000, but gets a 20% raise every year. Assume inflation is a fixed 8% every year. That is, \$1.08 a year from now is worth \$1.00 today.

(a) How much is a Harvard degree worth today if the holder will work for n years following graduation?

(b) How much is an MIT degree worth in this case?

(c) If you plan to retire after twenty years, which degree would be worth more?

Problem 14.3.

Suppose you deposit \$100 into your MIT Credit Union account today, \$99 in one month from now, \$98 in two months from now, and so on. Given that the interest rate is constantly 0.3% per month, how long will it take to save \$5,000?

14.2 Book Stacking

Suppose you have a pile of books and you want to stack them on a table in some off-center way so the top book sticks out past books below it. How far past the edge of the table do you think you could get the top book to go without having the stack fall over? Could the top book stick out completely beyond the edge of table?

Most people’s first response to this question—sometimes also their second and third responses—is “No, the top book will never get completely past the edge of the table.” But in fact, you can get the top book to stick out as far as you want: one booklength, two booklengths, any number of booklengths!

14.2.1 Formalizing the Problem

We’ll approach this problem recursively. How far past the end of the table can we get one book to stick out? It won’t tip as long as its center of mass is over the table, so we can get it to stick out half its length, as shown in Figure 14.1.

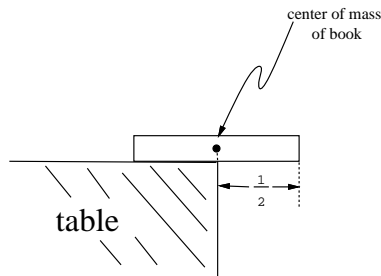


Figure 14.1: One book can overhang half a book length.

Now suppose we have a stack of books that will stick out past the table edge without tipping over—call that a *stable* stack. Let’s define the *overhang* of a stable stack to be the largest horizontal distance from the center of mass of the stack to the furthest edge of a book. If we place the center of mass of the stable stack at the edge of the table as in Figure 14.2, that’s how far we can get a book in the stack to stick out past the edge.

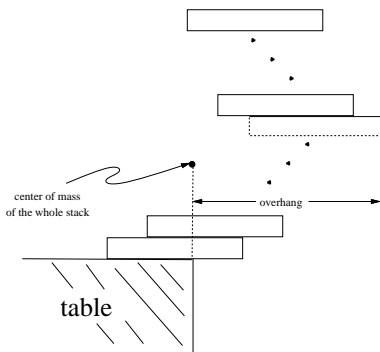


Figure 14.2: Overhanging the edge of the table.

So we want a formula for the maximum possible overhang, B_n , achievable with a stack of n books.

We've already observed that the overhang of one book is $1/2$ a book length. That is,

$$B_1 = \frac{1}{2}.$$

Now suppose we have a stable stack of $n + 1$ books with maximum overhang. If the overhang of the n books on top of the bottom book was not maximum, we could get a book to stick out further by replacing the top stack with a stack of n books with larger overhang. So the maximum overhang, B_{n+1} , of a stack of $n + 1$ books is obtained by placing a maximum overhang stable stack of n books on top of the bottom book. And we get the biggest overhang for the stack of $n + 1$ books by placing the center of mass of the n books right over the edge of the bottom book as in Figure 14.3.

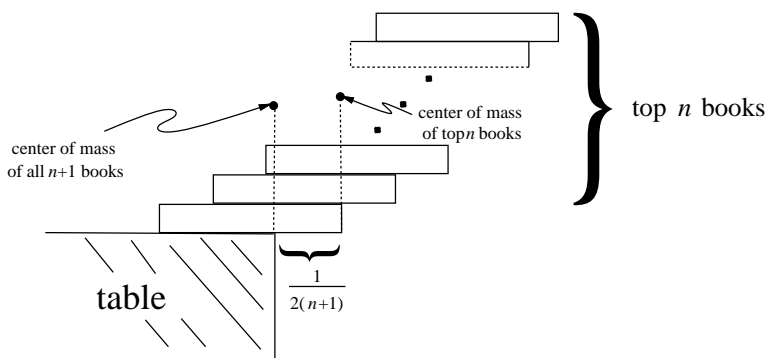
So we know where to place the $n + 1$ st book to get maximum overhang, and all we have to do is calculate what it is. The simplest way to do that is to let the center of mass of the top n books be the origin. That way the horizontal coordinate of the center of mass of the whole stack of $n + 1$ books will equal the increase in the overhang. But now the center of mass of the bottom book has horizontal coordinate $1/2$, so the horizontal coordinate of center of mass of the whole stack of $n + 1$ books is

$$\frac{0 \cdot n + (1/2) \cdot 1}{n + 1} = \frac{1}{2(n + 1)}.$$

In other words,

$$B_{n+1} = B_n + \frac{1}{2(n + 1)}, \quad (14.5)$$

as shown in Figure 14.3.

Figure 14.3: Additional overhang with $n + 1$ books.

Expanding equation (14.5), we have

$$\begin{aligned}
 B_{n+1} &= B_{n-1} + \frac{1}{2n} + \frac{1}{2(n+1)} \\
 &= B_1 + \frac{1}{2 \cdot 2} + \cdots + \frac{1}{2n} + \frac{1}{2(n+1)} \\
 &= \frac{1}{2} \sum_{i=1}^{n+1} \frac{1}{i}.
 \end{aligned} \tag{14.6}$$

The n th *Harmonic number*, H_n , is defined to be

Definition 14.2.1.

$$H_n ::= \sum_{i=1}^n \frac{1}{i}.$$

So (14.6) means that

$$B_n = \frac{H_n}{2}.$$

The first few Harmonic numbers are easy to compute. For example, $H_4 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{25}{12}$. The fact that H_4 is greater than 2 has special significance; it implies that the total extension of a 4-book stack is greater than one full book! This is the situation shown in Figure 14.4.

14.2.2 Evaluating the Sum—The Integral Method

It would be nice to answer questions like, “How many books are needed to build a stack extending 100 book lengths beyond the table?” One approach to this question

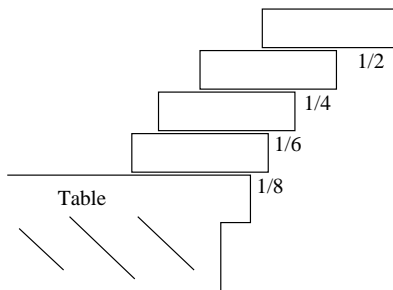


Figure 14.4: Stack of four books with maximum overhang.

would be to keep computing Harmonic numbers until we found one exceeding 200. However, as we will see, this is not such a keen idea.

Such questions would be settled if we could express H_n in a closed form. Unfortunately, no closed form is known, and probably none exists. As a second best, however, we can find closed forms for very good approximations to H_n using the Integral Method. The idea of the Integral Method is to bound terms of the sum above and below by simple functions as suggested in Figure 14.5. The integrals of these functions then bound the value of the sum above and below.

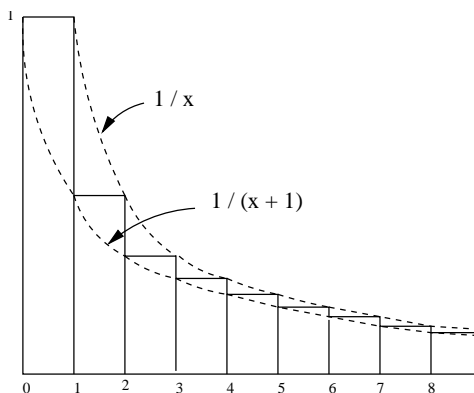


Figure 14.5: This figure illustrates the Integral Method for bounding a sum. The area under the “stairstep” curve over the interval $[0, n]$ is equal to $H_n = \sum_{i=1}^n 1/i$. The function $1/x$ is everywhere greater than or equal to the stairstep and so the integral of $1/x$ over this interval is an upper bound on the sum. Similarly, $1/(x+1)$ is everywhere less than or equal to the stairstep and so the integral of $1/(x+1)$ is a lower bound on the sum.

The Integral Method gives the following upper and lower bounds on the har-

monic number H_n :

$$H_n \leq 1 + \int_1^n \frac{1}{x} dx = 1 + \ln n \quad (14.7)$$

$$H_n \geq \int_0^n \frac{1}{x+1} dx = \int_1^{n+1} \frac{1}{x} dx = \ln(n+1). \quad (14.8)$$

These bounds imply that the harmonic number H_n is around $\ln n$.

But $\ln n$ grows—slowly—but without bound. That means we can get books to overhang *any distance* past the edge of the table by piling them high enough! For example, to build a stack extending three book lengths beyond the table, we need a number of books n so that $H_n \geq 6$. By inequality (14.8), this means we want

$$H_n \geq \ln(n+1) \geq 6,$$

so $n \geq e^6 - 1$ books will work, that is, 403 books will be enough to get a three book overhang. Actual calculation of H_6 shows that 227 books is the smallest number that will work.

14.2.3 More about Harmonic Numbers

In the preceding section, we showed that H_n is about $\ln n$. An even better approximation is known:

$$H_n = \ln n + \gamma + \frac{1}{2n} + \frac{1}{12n^2} + \frac{\epsilon(n)}{120n^4}$$

Here γ is a value 0.577215664 . . . called *Euler's constant*, and $\epsilon(n)$ is between 0 and 1 for all n . We will not prove this formula.

Asymptotic Equality

The shorthand $H_n \sim \ln n$ is used to indicate that the leading term of H_n is $\ln n$. More precisely:

Definition 14.2.2. For functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$, we say f is *asymptotically equal* to g , in symbols,

$$f(x) \sim g(x)$$

iff

$$\lim_{x \rightarrow \infty} f(x)/g(x) = 1.$$

It's tempting to might write $H_n \sim \ln n + \gamma$ to indicate the two leading terms, but it is not really right. According to Definition 14.2.2, $H_n \sim \ln n + c$ where c is *any constant*. The correct way to indicate that γ is the second-largest term is $H_n - \ln n \sim \gamma$.

The reason that the \sim notation is useful is that often we do not care about lower order terms. For example, if $n = 100$, then we can compute $H(n)$ to great precision using only the two leading terms:

$$|H_n - \ln n - \gamma| \leq \left| \frac{1}{200} - \frac{1}{120000} + \frac{1}{120 \cdot 100^4} \right| < \frac{1}{200}.$$

14.2.4 Problems

Class Problems

Problem 14.4.

An explorer is trying to reach the Holy Grail, which she believes is located in a desert shrine d days walk from the nearest oasis. In the desert heat, the explorer must drink continuously. She can carry at most 1 gallon of water, which is enough for 1 day. However, she is free to create water caches out in the desert.

For example, if the shrine were $2/3$ of a day's walk into the desert, then she could recover the Holy Grail with the following strategy. She leaves the oasis with 1 gallon of water, travels $1/3$ day into the desert, caches $1/3$ gallon, and then walks back to the oasis—arriving just as her water supply runs out. Then she picks up another gallon of water at the oasis, walks $1/3$ day into the desert, tops off her water supply by taking the $1/3$ gallon in her cache, walks the remaining $1/3$ day to the shrine, grabs the Holy Grail, and then walks for $2/3$ of a day back to the oasis—again arriving with no water to spare.

But what if the shrine were located farther away?

(a) What is the most distant point that the explorer can reach and then return to the oasis if she takes only 1 gallon from the oasis?

(b) What is the most distant point the explorer can reach and still return to the oasis if she takes only 2 gallons from the oasis? No proof is required; just do the best you can.

(c) The explorer will travel using a recursive strategy to go far into the desert and back drawing a total of n gallons of water from the oasis. Her strategy is to build up a cache of $n - 1$ gallons, plus enough to get home, a certain fraction of a day's distance into the desert. On the last delivery to the cache, instead of returning home, she proceeds recursively with her $n - 1$ gallon strategy to go farther into the desert and return to the cache. At this point, the cache has just enough water left to get her home.

Prove that with n gallons of water, this strategy will get her $H_n/2$ days into the desert and back, where H_n is the n th Harmonic number:

$$H_n ::= \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}.$$

Conclude that she can reach the shrine, however far it is from the oasis.

(d) Suppose that the shrine is $d = 10$ days walk into the desert. Use the asymptotic approximation $H_n \sim \ln n$ to show that it will take more than a million years for the explorer to recover the Holy Grail.

Problem 14.5.

There is a number a such that $\sum_{i=1}^{\infty} i^p$ converges iff $p < a$. What is the value of a ? Prove it.

Homework Problems

Problem 14.6.

There is a bug on the edge of a 1-meter rug. The bug wants to cross to the other side of the rug. It crawls at 1 cm per second. However, at the end of each second, a malicious first-grader named Mildred Anderson *stretches* the rug by 1 meter. Assume that her action is instantaneous and the rug stretches uniformly. Thus, here's what happens in the first few seconds:

- The bug walks 1 cm in the first second, so 99 cm remain ahead.
- Mildred stretches the rug by 1 meter, which doubles its length. So now there are 2 cm behind the bug and 198 cm ahead.
- The bug walks another 1 cm in the next second, leaving 3 cm behind and 197 cm ahead.
- Then Mildred strikes, stretching the rug from 2 meters to 3 meters. So there are now $3 \cdot (3/2) = 4.5$ cm behind the bug and $197 \cdot (3/2) = 295.5$ cm ahead.
- The bug walks another 1 cm in the third second, and so on.

Your job is to determine this poor bug's fate.

(a) During second i , what *fraction* of the rug does the bug cross?

(b) Over the first n seconds, what fraction of the rug does the bug cross altogether? Express your answer in terms of the Harmonic number H_n .

(c) The known universe is thought to be about $3 \cdot 10^{10}$ light years in diameter. How many universe diameters must the bug travel to get to the end of the rug?

14.3 Finding Summation Formulas

The Integral Method offers a way to derive formulas like those for the sum of consecutive integers,

$$\sum_{i=1}^n i = n(n+1)/2,$$

or for the sum of squares,

$$\begin{aligned}\sum_{i=1}^n i^2 &= \frac{(2n+1)(n+1)n}{6} \\ &= \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}.\end{aligned}\tag{14.9}$$

These equations appeared in Chapter 2 as equations (2.2) and (2.3) where they were proved using the Well-ordering Principle. But those proofs did not explain how someone figured out in the first place that these were the formulas to prove.

Here's how the Integral Method leads to the sum-of-squares formula, for example. First, get a quick estimate of the sum:

$$\int_0^n x^2 dx \leq \sum_{i=1}^n i^2 \leq \int_0^n (x+1)^2 dx,$$

so

$$n^3/3 \leq \sum_{i=1}^n i^2 \leq (n+1)^3/3 - 1/3.\tag{14.10}$$

and the upper and lower bounds (14.10) imply that

$$\sum_{i=1}^n i^2 \sim n^3/3.$$

To get an exact formula, we then guess the general form of the solution. Where we are uncertain, we can add parameters a, b, c, \dots . For example, we might make the guess:

$$\sum_{i=1}^n i^2 = an^3 + bn^2 + cn + d.$$

If the guess is correct, then we can determine the parameters $a, b, c,$ and d by plugging in a few values for n . Each such value gives a linear equation in $a, b, c,$ and d . If we plug in enough values, we may get a linear system with a unique solution. Applying this method to our example gives:

$$\begin{aligned}n = 0 &\text{ implies } 0 = d \\ n = 1 &\text{ implies } 1 = a + b + c + d \\ n = 2 &\text{ implies } 5 = 8a + 4b + 2c + d \\ n = 3 &\text{ implies } 14 = 27a + 9b + 3c + d.\end{aligned}$$

Solving this system gives the solution $a = 1/3, b = 1/2, c = 1/6, d = 0$. Therefore, if our initial guess at the form of the solution was correct, then the summation is equal to $n^3/3 + n^2/2 + n/6$, which matches equation (14.9).

The point is that if the desired formula turns out to be a polynomial, then once you get an estimate of the *degree* of the polynomial —by the Integral Method or any other way —all the coefficients of the polynomial can be found automatically.

Be careful! This method let's you discover formulas, but it doesn't guarantee they are right! After obtaining a formula by this method, it's important to go back and *prove* it using induction or some other method, because if the initial guess at the solution was not of the right form, then the resulting formula will be completely wrong!

14.3.1 Double Sums

Sometimes we have to evaluate sums of sums, otherwise known as *double summations*. This can be easy: evaluate the inner sum, replace it with a closed form, and then evaluate the outer sum which no longer has a summation inside it. For example,

$$\begin{aligned}
 & \sum_{n=0}^{\infty} \left(y^n \sum_{i=0}^n x^i \right) \\
 &= \sum_{n=0}^{\infty} \left(y^n \frac{1 - x^{n+1}}{1 - x} \right) && \text{(geometric sum formula (14.2))} \\
 &= \frac{\sum_{n=0}^{\infty} y^n}{1 - x} - \frac{\sum_{n=0}^{\infty} y^n x^{n+1}}{1 - x} \\
 &= \frac{1}{(1 - y)(1 - x)} - \frac{x \sum_{n=0}^{\infty} (xy)^n}{1 - x} && \text{(infinite geometric sum, Theorem 14.1.1)} \\
 &= \frac{1}{(1 - y)(1 - x)} - \frac{x}{(1 - xy)(1 - x)} && \text{(infinite geometric sum, Theorem 14.1.1)} \\
 &= \frac{(1 - xy) - x(1 - y)}{(1 - xy)(1 - y)(1 - x)} \\
 &= \frac{1 - x}{(1 - xy)(1 - y)(1 - x)} \\
 &= \frac{1}{(1 - xy)(1 - y)}.
 \end{aligned}$$

When there's no obvious closed form for the inner sum, a special trick that is often useful is to try *exchanging the order of summation*. For example, suppose we want to compute the sum of the harmonic numbers

$$\sum_{k=1}^n H_k = \sum_{k=1}^n \sum_{j=1}^k 1/j$$

For intuition about this sum, we can try the integral method:

$$\sum_{k=1}^n H_k \approx \int_1^n \ln x \, dx \approx n \ln n - n.$$

Now let's look for an exact answer. If we think about the pairs (k, j) over which we are summing, they form a triangle:

		j						
		1	2	3	4	5	...	n
k	1	1						
	2	1	1/2					
	3	1	1/2	1/3				
	4	1	1/2	1/3	1/4			
	...							
	n	1	1/2		...			1/n

The summation above is summing each row and then adding the row sums. Instead, we can sum the columns and then add the column sums. Inspecting the table we see that this double sum can be written as

$$\begin{aligned}
 \sum_{k=1}^n H_k &= \sum_{k=1}^n \sum_{j=1}^k 1/j \\
 &= \sum_{j=1}^n \sum_{k=j}^n 1/j \\
 &= \sum_{j=1}^n 1/j \sum_{k=j}^n 1 \\
 &= \sum_{j=1}^n \frac{1}{j} (n - j + 1) \\
 &= \sum_{j=1}^n \frac{n - j + 1}{j} \\
 &= \sum_{j=1}^n \frac{n + 1}{j} - \sum_{j=1}^n \frac{j}{j} \\
 &= (n + 1) \sum_{j=1}^n \frac{1}{j} - \sum_{j=1}^n 1 \\
 &= (n + 1)H_n - n.
 \end{aligned} \tag{14.11}$$

14.4 Stirling's Approximation

The familiar factorial notation, $n!$, is an abbreviation for the product

$$\prod_{i=1}^n i.$$

This is by far the most common product in Discrete Mathematics. In this section we describe a good closed-form estimate of $n!$ called *Stirling's Approximation*. Unfortunately, all we can do is estimate: there is no closed form for $n!$ —though proving so would take us beyond the scope of 6.042.

14.4.1 Products to Sums

A good way to handle a product is often to convert it into a sum by taking the logarithm. In the case of factorial, this gives

$$\begin{aligned}\ln(n!) &= \ln(1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n) \\ &= \ln 1 + \ln 2 + \ln 3 + \cdots + \ln(n-1) + \ln n \\ &= \sum_{i=1}^n \ln i.\end{aligned}$$

We've not seen a summation containing a logarithm before! Fortunately, one tool that we used in evaluating sums is still applicable: the Integral Method. We can bound the terms of this sum with $\ln x$ and $\ln(x+1)$ as shown in Figure 14.6. This gives bounds on $\ln(n!)$ as follows:

$$\begin{aligned}\int_1^n \ln x \, dx &\leq \sum_{i=1}^n \ln i \leq \int_0^n \ln(x+1) \, dx \\ n \ln\left(\frac{n}{e}\right) + 1 &\leq \sum_{i=1}^n \ln i \leq (n+1) \ln\left(\frac{n+1}{e}\right) + 1 \\ \left(\frac{n}{e}\right)^n e &\leq n! \leq \left(\frac{n+1}{e}\right)^{n+1} e.\end{aligned}$$

The second line follows from the first by completing the integrations. The third line is obtained by exponentiating.

So $n!$ behaves something like the closed form formula $(n/e)^n$. A more careful analysis yields an unexpected closed form formula that is asymptotically exact:

Lemma (Stirling's Formula).

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}, \quad (14.12)$$

Stirling's Formula describes how $n!$ behaves in the limit, but to use it effectively, we need to know how close it is to the limit for different values of n . That information is given by the bounding formulas:

Fact (Stirling's Approximation).

$$\sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{1/(12n+1)} \leq n! \leq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{1/12n}.$$

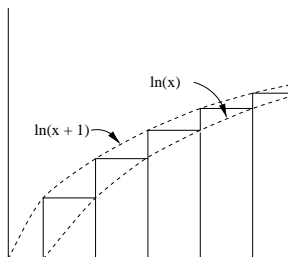


Figure 14.6: This figure illustrates the Integral Method for bounding the sum $\sum_{i=1}^n \ln i$.

Stirling's Approximation implies the asymptotic formula (14.12), since $e^{1/(12n+1)}$ and $e^{1/12n}$ both approach 1 as n grows large. These inequalities can be verified by induction, but the details are nasty.

The bounds in Stirling's formula are very tight. For example, if $n = 100$, then Stirling's bounds are:

$$100! \geq \sqrt{200\pi} \left(\frac{100}{e}\right)^{100} e^{1/1201}$$

$$100! \leq \sqrt{200\pi} \left(\frac{100}{e}\right)^{100} e^{1/1200}$$

The only difference between the upper bound and the lower bound is in the final term. In particular $e^{1/1201} \approx 1.00083299$ and $e^{1/1200} \approx 1.00083368$. As a result, the upper bound is no more than $1 + 10^{-6}$ times the lower bound. This is amazingly tight! Remember Stirling's formula; we will use it often.

14.5 Asymptotic Notation

Asymptotic notation is a shorthand used to give a quick measure of the behavior of a function $f(n)$ as n grows large.

14.5.1 Little Oh

The asymptotic notation, \sim , of Definition 14.2.2 is a binary relation indicating that two functions grow at the *same* rate. There is also a binary relation indicating that one function grows at a significantly *slower* rate than another. Namely,

Definition 14.5.1. For functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$, with g nonnegative, we say f is

asymptotically smaller than g , in symbols,

$$f(x) = o(g(x)),$$

iff

$$\lim_{x \rightarrow \infty} f(x)/g(x) = 0.$$

For example, $1000x^{1.9} = o(x^2)$, because $1000x^{1.9}/x^2 = 1000/x^{0.1}$ and since $x^{0.1}$ goes to infinity with x and 1000 is constant, we have $\lim_{x \rightarrow \infty} 1000x^{1.9}/x^2 = 0$. This argument generalizes directly to yield

Lemma 14.5.2. $x^a = o(x^b)$ for all nonnegative constants $a < b$.

Using the familiar fact that $\log x < x$ for all $x > 1$, we can prove

Lemma 14.5.3. $\log x = o(x^\epsilon)$ for all $\epsilon > 0$ and $x > 1$.

Proof. Choose $\epsilon > \delta > 0$ and let $x = z^\delta$ in the inequality $\log x < x$. This implies

$$\log z < z^\delta / \delta = o(z^\epsilon) \quad \text{by Lemma 14.5.2.} \quad (14.13)$$

□

Corollary 14.5.4. $x^b = o(a^x)$ for any $a, b \in \mathbb{R}$ with $a > 1$.

Proof. From (14.13),

$$\log z < z^\delta / \delta$$

for all $z > 1, \delta > 0$. Hence

$$\begin{aligned} (e^b)^{\log z} &< (e^b)^{z^\delta / \delta} \\ z^b &< \left(e^{\log a (b / \log a)} \right)^{z^\delta / \delta} \\ &= a^{(b / \delta \log a) z^\delta} \\ &< a^z \end{aligned}$$

for all z such that

$$(b / \delta \log a) z^\delta < z.$$

But choosing $\delta < 1$, we know $z^\delta = o(z)$, so this last inequality holds for all large enough z . □

Lemma 14.5.3 and Corollary 14.5.4 can also be proved easily in several other ways, for example, using L'Hopital's Rule or the McLaurin Series for $\log x$ and e^x . Proofs can be found in most calculus texts.

14.5.2 Big Oh

Big Oh is the most frequently used asymptotic notation. It is used to give an upper bound on the growth of a function, such as the running time of an algorithm.

Definition 14.5.5. Given nonnegative functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$, we say that

$$f = O(g)$$

iff

$$\limsup_{x \rightarrow \infty} f(x)/g(x) < \infty.$$

This definition³ makes it clear that

Lemma 14.5.6. *If $f = o(g)$ or $f \sim g$, then $f = O(g)$.*

Proof. $\lim f/g = 0$ or $\lim f/g = 1$ implies $\lim f/g < \infty$. □

It is easy to see that the converse of Lemma 14.5.6 is not true. For example, $2x = O(x)$, but $2x \not\sim x$ and $2x \neq o(x)$.

The usual formulation of Big Oh spells out the definition of \limsup without mentioning it. Namely, here is an equivalent definition:

Definition 14.5.7. Given functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$, we say that

$$f = O(g)$$

iff there exists a constant $c \geq 0$ and an x_0 such that for all $x \geq x_0$, $|f(x)| \leq cg(x)$.

This definition is rather complicated, but the idea is simple: $f(x) = O(g(x))$ means $f(x)$ is less than or equal to $g(x)$, except that we're willing to ignore a constant factor, namely, c , and to allow exceptions for small x , namely, $x < x_0$.

We observe,

Lemma 14.5.8. *If $f = o(g)$, then it is not true that $g = O(f)$.*

Proof.

$$\lim_{x \rightarrow \infty} \frac{g(x)}{f(x)} = \frac{1}{\lim_{x \rightarrow \infty} f(x)/g(x)} = \frac{1}{0} = \infty,$$

so $g \neq O(f)$. □

Proposition 14.5.9. $100x^2 = O(x^2)$.

³

$$\limsup_{x \rightarrow \infty} h(x) ::= \lim_{x \rightarrow \infty} \text{lub}_{y \geq x} h(y).$$

We need the \limsup in the definition of $O()$ because if $f(x)/g(x)$ oscillates between, say, 3 and 5 as x grows, then $f = O(g)$ because $f \leq 5g$, but $\lim_{x \rightarrow \infty} f(x)/g(x)$ does not exist. However, in this case we would have $\limsup_{x \rightarrow \infty} f(x)/g(x) = 5$.

Proof. Choose $c = 100$ and $x_0 = 1$. Then the proposition holds, since for all $x \geq 1$, $|100x^2| \leq 100x^2$. \square

Proposition 14.5.10. $x^2 + 100x + 10 = O(x^2)$.

Proof. $(x^2 + 100x + 10)/x^2 = 1 + 100/x + 10/x^2$ and so its limit as x approaches infinity is $1+0+0 = 1$. So in fact, $x^2 + 100x + 10 \sim x^2$, and therefore $x^2 + 100x + 10 = O(x^2)$. Indeed, it's conversely true that $x^2 = O(x^2 + 100x + 10)$. \square

Proposition 14.5.10 generalizes to an arbitrary polynomial:

Proposition 14.5.11. For $a_k \neq 0$, $a_k x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0 = O(x^k)$.

We'll omit the routine proof.

Big Oh notation is especially useful when describing the running time of an algorithm. For example, the usual algorithm for multiplying $n \times n$ matrices requires proportional to n^3 operations in the worst case. This fact can be expressed concisely by saying that the running time is $O(n^3)$. So this asymptotic notation allows the speed of the algorithm to be discussed without reference to constant factors or lower-order terms that might be machine specific. In this case there is another, ingenious matrix multiplication procedure that requires $O(n^{2.55})$ operations. This procedure will therefore be much more efficient on large enough matrices. Unfortunately, the $O(n^{2.55})$ -operation multiplication procedure is almost never used because it happens to be less efficient than the usual $O(n^3)$ procedure on matrices of practical size.

14.5.3 Theta

Definition 14.5.12.

$$f = \Theta(g) \quad \text{iff} \quad f = O(g) \text{ and } g = O(f).$$

The statement $f = \Theta(g)$ can be paraphrased intuitively as “ f and g are equal to within a constant factor.”

The value of these notations is that they highlight growth rates and allow suppression of distracting factors and low-order terms. For example, if the running time of an algorithm is

$$T(n) = 10n^3 - 20n^2 + 1,$$

then

$$T(n) = \Theta(n^3).$$

In this case, we would say that T is of order n^3 or that $T(n)$ grows cubically.

Another such example is

$$\pi^2 3^{x-7} + \frac{(2.7x^{113} + x^9 - 86)^4}{\sqrt{x}} - 1.08^{3x} = \Theta(3^x).$$

Just knowing that the running time of an algorithm is $\Theta(n^3)$, for example, is useful, because if n doubles we can predict that the running time will *by and large*⁴ increase by a factor of at most 8 for large n . In this way, Theta notation preserves information about the scalability of an algorithm or system. Scalability is, of course, a big issue in the design of algorithms and systems.

14.5.4 Pitfalls with Big Oh

There is a long list of ways to make mistakes with Big Oh notation. This section presents some of the ways that Big Oh notation can lead to ruin and despair.

The Exponential Fiasco

Sometimes relationships involving Big Oh are not so obvious. For example, one might guess that $4^x = O(2^x)$ since 4 is only a constant factor larger than 2. This reasoning is incorrect, however; actually 4^x grows much faster than 2^x .

Proposition 14.5.13. $4^x \neq O(2^x)$

Proof. $2^x/4^x = 2^x/(2^x 2^x) = 1/2^x$. Hence, $\lim_{x \rightarrow \infty} 2^x/4^x = 0$, so in fact $2^x = o(4^x)$. We observed earlier that this implies that $4^x \neq O(2^x)$. \square

Constant Confusion

Every constant is $O(1)$. For example, $17 = O(1)$. This is true because if we let $f(x) = 17$ and $g(x) = 1$, then there exists a $c > 0$ and an x_0 such that $|f(x)| \leq cg(x)$. In particular, we could choose $c = 17$ and $x_0 = 1$, since $|17| \leq 17 \cdot 1$ for all $x \geq 1$. We can construct a false theorem that exploits this fact.

False Theorem 14.5.14.

$$\sum_{i=1}^n i = O(n)$$

False proof. Define $f(n) = \sum_{i=1}^n i = 1 + 2 + 3 + \dots + n$. Since we have shown that every constant i is $O(1)$, $f(n) = O(1) + O(1) + \dots + O(1) = O(n)$. \square

Of course in reality $\sum_{i=1}^n i = n(n+1)/2 \neq O(n)$.

The error stems from confusion over what is meant in the statement $i = O(1)$. For any *constant* $i \in \mathbb{N}$ it is true that $i = O(1)$. More precisely, if f is any constant function, then $f = O(1)$. But in this False Theorem, i is not constant but ranges over a set of values $0, 1, \dots, n$ that depends on n .

And anyway, we should not be adding $O(1)$'s as though they were numbers. We never even defined what $O(g)$ means by itself; it should only be used in the context " $f = O(g)$ " to describe a relation between functions f and g .

⁴Since $\Theta(n^3)$ only implies that the running time, $T(n)$, is between cn^3 and dn^3 for constants $0 < c < d$, the time $T(2n)$ could regularly exceed $T(n)$ by a factor as large as $8d/c$. The factor is sure to be close to 8 for all large n only if $T(n) \sim n^3$.

Lower Bound Blunder

Sometimes people incorrectly use Big Oh in the context of a lower bound. For example, they might say, “The running time, $T(n)$, is at least $O(n^2)$,” when they probably mean something like “ $O(T(n)) = n^2$,” or more properly, “ $n^2 = O(T(n))$.”

Equality Blunder

The notation $f = O(g)$ is too firmly entrenched to avoid, but the use of “=” is really regrettable. For example, if $f = O(g)$, it seems quite reasonable to write $O(g) = f$. But doing so might tempt us to the following blunder: because $2n = O(n)$, we can say $O(n) = 2n$. But $n = O(n)$, so we conclude that $n = O(n) = 2n$, and therefore $n = 2n$. To avoid such nonsense, we will never write “ $O(f) = g$.”

14.5.5 Problems

Homework Problems

Problem 14.7. (a) Prove that $\log x < x$ for all $x > 1$ (requires elementary calculus).

(b) Prove that the relation, R , on functions such that $f R g$ iff $f = o(g)$ is a strict partial order.

(c) Prove that $f \sim g$ iff $f = g + h$ for some function $h = o(g)$.

Problem 14.8.

Indicate which of the following holds for each pair of functions $(f(n), g(n))$ in the table below. Assume $k \geq 1$, $\epsilon > 0$, and $c > 1$ are constants. Pick the four table entries you consider to be the most challenging or interesting and justify your answers to these.

$f(n)$	$g(n)$	$f = O(g)$	$f = o(g)$	$g = O(f)$	$g = o(f)$	$f = \Theta(g)$	$f \sim g$
2^n	$2^{n/2}$						
\sqrt{n}	$n^{\sin n\pi/2}$						
$\log(n!)$	$\log(n^n)$						
n^k	c^n						
$\log^k n$	n^ϵ						

Problem 14.9.

Let f, g be nonnegative real-valued functions such that $\lim_{x \rightarrow \infty} f(x) = \infty$ and $f \sim g$.

(a) Give an example of f, g such that NOT($2^f \sim 2^g$).

(b) Prove that $\log f \sim \log g$.

(c) Use Stirling's formula to prove that in fact

$$\log(n!) \sim n \log n$$

Class Problems

Problem 14.10.

Give an elementary proof (without appealing to Stirling's formula) that $\log(n!) = \Theta(n \log n)$.

Problem 14.11.

Recall that for functions f, g on the natural numbers, \mathbb{N} , $f = O(g)$ iff

$$\exists c \in \mathbb{N} \exists n_0 \in \mathbb{N} \forall n \geq n_0 \quad c \cdot g(n) \geq |f(n)|. \quad (14.14)$$

For each pair of functions below, determine whether $f = O(g)$ and whether $g = O(f)$. In cases where one function is $O()$ of the other, indicate the *smallest natural number*, c , and for that smallest c , the *smallest corresponding natural number* n_0 ensuring that condition (14.14) applies.

(a) $f(n) = n^2, g(n) = 3n$.

$f = O(g)$ YES NO If YES, $c = \underline{\hspace{2cm}}$, $n_0 = \underline{\hspace{2cm}}$

$g = O(f)$ YES NO If YES, $c = \underline{\hspace{2cm}}$, $n_0 = \underline{\hspace{2cm}}$

(b) $f(n) = (3n - 7)/(n + 4), g(n) = 4$

$f = O(g)$ YES NO If YES, $c = \underline{\hspace{2cm}}$, $n_0 = \underline{\hspace{2cm}}$

$g = O(f)$ YES NO If YES, $c = \underline{\hspace{2cm}}$, $n_0 = \underline{\hspace{2cm}}$

(c) $f(n) = 1 + (n \sin(n\pi/2))^2, g(n) = 3n$

$f = O(g)$ YES NO If yes, $c = \underline{\hspace{2cm}}$ $n_0 = \underline{\hspace{2cm}}$

$g = O(f)$ YES NO If yes, $c = \underline{\hspace{2cm}}$ $n_0 = \underline{\hspace{2cm}}$

Problem 14.12.

False Claim.

$$2^n = O(1). \quad (14.15)$$

Explain why the claim is false. Then identify and explain the mistake in the following bogus proof.

Bogus proof. The proof by induction on n where the induction hypothesis, $P(n)$, is the assertion (14.15).

base case: $P(0)$ holds trivially.

inductive step: We may assume $P(n)$, so there is a constant $c > 0$ such that $2^n \leq c \cdot 1$. Therefore,

$$2^{n+1} = 2 \cdot 2^n \leq (2c) \cdot 1,$$

which implies that $2^{n+1} = O(1)$. That is, $P(n+1)$ holds, which completes the proof of the inductive step.

We conclude by induction that $2^n = O(1)$ for all n . That is, the exponential function is bounded by a constant.

□

Problem 14.13.

- (a) Define a function $f(n)$ such that $f = \Theta(n^2)$ and NOT($f \sim n^2$).
- (b) Define a function $g(n)$ such that $g = O(n^2)$, $g \neq \Theta(n^2)$ and $g \neq o(n^2)$.

Index

$\Theta()$, 303

\sim , 300

\sim (asymptotic equality), 293

annuity, 285

approximations, 292

asymptotically equal, 293

asymptotically smaller, 301

Big Oh, 304

Book Stacking, 289

closed form, 286, 292

double summations, 297

Euler's constant, 293

factorial, 298

geometric sum, 286

Harmonic number, 291

Integral Method, 291, 299

interest rate, 289

logarithm, 299

matrix multiplication, 303

$o()$, asymptotically smaller, 301

$O()$, big oh, 302

$o()$, little oh, 300

overhang, 289

Stirling's Approximation, 299, 300

Stirling's Formula, 299

value of an annuity, 287