13 Sums and Asymptotics

Sums and products arise regularly in the analysis of algorithms, financial applications, physical problems, and probabilistic systems. For example, according to Theorem 2.2.1,

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}. \tag{13.1}$$

Of course, the lefthand sum could be expressed concisely as a subscripted summation

$$\sum_{i=1}^{n} i$$

but the right hand expression $n(n+1)/2$ is not only concise but also easier to evaluate. Furthermore, it more clearly reveals properties such as the growth rate of the sum. Expressions like $n(n+1)/2$ that do not make use of subscripted summations or products—or those handy but sometimes troublesome sequences of three dots—are called *closed forms*.

Another example is the closed form for a *geometric sum*

$$1 + x + x^2 + x^3 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x} \tag{13.2}$$

given in Problem 5.4. The sum as described on the left hand side of (13.2) involves n additions and $1 + 2 + \cdots + (n-1) = (n-1)n/2$ multiplications, but its closed form on the right hand side can be evaluated using fast exponentiation with at most $2 \log n$ multiplications, a division, and a couple of subtractions. Also, the closed form makes the growth and limiting behavior of the sum much more apparent.

Equations (13.1) and (13.2) were easy to verify by induction, but, as is often the case, the proofs by induction gave no hint about how these formulas were found in the first place. Finding them is part math and part art, which we'll start examining in this chapter.

Our first motivating example will be the value of a financial instrument known as an annuity. This value will be a large and nasty-looking sum. We will then describe several methods for finding closed forms for several sorts of sums, including those for annuities. In some cases, a closed form for a sum may not exist, and so we will provide a general method for finding closed forms for good upper and lower bounds on the sum.

The methods we develop for sums will also work for products, since any product can be converted into a sum by taking its logarithm. For instance, later in the chapter we will use this approach to find a good closed-form approximation to the *factorial function*

$$n! ::= 1 \cdot 2 \cdot 3 \cdots n.$$

We conclude the chapter with a discussion of asymptotic notation, especially "Big Oh" notation. Asymptotic notation is often used to bound the error terms when there is no exact closed form expression for a sum or product. It also provides a convenient way to express the growth rate or order of magnitude of a sum or product.

13.1 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 firms on Wall Street that specialize in trading annuities.[1]

A key question is, "What is an annuity 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? Suddenly, it's not clear which option is better.

13.1.1 The Future Value of Money

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% rate[2] for the rest of the discussion, so $p = 0.08$.

[1]Such trading ultimately led to the subprime mortgage disaster in 2008–2009. We'll talk more about that in a later chapter.

[2]U.S. interest rates have dropped steadily for several years, and ordinary bank deposits now earn around 1.0%. 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.

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 is only really worth $1/(1+p) \cdot 10 \approx 9.26$ dollars today, because if we 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.

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:

$$
\begin{aligned}
V &= \sum_{i=1}^{n} \frac{m}{(1+p)^{i-1}} \\
&= m \cdot \sum_{j=0}^{n-1} \left(\frac{1}{1+p}\right)^j && \text{(substitute } j = i-1) \\
&= m \cdot \sum_{j=0}^{n-1} x^j && \text{(substitute } x = 1/(1+p)). \qquad (13.3)
\end{aligned}
$$

The goal of the preceding substitutions was to get the summation into the form of a simple geometric sum. This leads us to an explanation of a way you could have discovered the closed form (13.2) in the first place using the *Perturbation Method*.

13.1.2 The Perturbation Method

Given a sum that has a nice structure, it is often useful to "perturb" the sum so that we can somehow combine the sum with the perturbation to get something much simpler. For example, suppose

$$S = 1 + x + x^2 + \cdots + x^n.$$

An example of a perturbation would be

$$xS = x + x^2 + \cdots + x^{n+1}.$$

The difference between S and xS is not so great, and so if we were to subtract xS from S, there would be massive cancellation:

$$
\begin{aligned}
S &= 1 + x + x^2 + x^3 + \cdots + x^n \\
-xS &= - x - x^2 - x^3 - \cdots - x^n - x^{n+1}.
\end{aligned}
$$

The result of the subtraction is

$$S - xS = 1 - x^{n+1}.$$

Solving for S gives the desired closed-form expression in equation 13.2, namely,

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

We'll see more examples of this method when we introduce *generating functions* in Chapter 15.

13.1.3 A Closed Form for the Annuity Value

Using equation 13.2, we can derive a simple formula for V, the value of an annuity that pays m dollars at the start of each year for n years.

$$V = m\left(\frac{1 - x^n}{1 - x}\right) \quad \text{(by equations 13.3 and 13.2)} \quad (13.4)$$

$$= m\left(\frac{1 + p - (1/(1+p))^{n-1}}{p}\right) \quad \text{(substituting } x = 1/(1+p)\text{).} \quad (13.5)$$

Equation 13.5 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.

13.1.4 Infinite Geometric Series

We began this chapter by asking 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 equation 13.2 as n tends to infinity.

Theorem 13.1.1. *If* $|x| < 1$, *then*

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

Proof.

$$\begin{aligned}
\sum_{i=0}^{\infty} x^i &::= \lim_{n\to\infty} \sum_{i=0}^{n} x^i \\
&= \lim_{n\to\infty} \frac{1 - x^{n+1}}{1-x} && \text{(by equation 13.2)} \\
&= \frac{1}{1-x}.
\end{aligned}$$

The final line follows from the fact that $\lim_{n\to\infty} x^{n+1} = 0$ when $|x| < 1$. ■

In our annuity problem, $x = 1/(1+p) < 1$, so Theorem 13.1.1 applies, and we get

$$\begin{aligned}
V &= m \cdot \sum_{j=0}^{\infty} x^j && \text{(by equation 13.3)} \\
&= m \cdot \frac{1}{1-x} && \text{(by Theorem 13.1.1)} \\
&= m \cdot \frac{1+p}{p} && (x = 1/(1+p)).
\end{aligned}$$

Plugging in $m = \$50{,}000$ and $p = 0.08$, we see that the value V is only \$675,000. It seems amazing that a million dollars today is worth much more than \$50,000 paid every year for eternity! But on closer inspection, 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 as it turns out, this answer really isn't so amazing after all.

13.1.5 Examples

Equation 13.2 and Theorem 13.1.1 are incredibly useful in computer science.

Here are some other common sums that can be put into closed form using equation 13.2 and Theorem 13.1.1:

$$1 + 1/2 + 1/4 + \cdots = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^i = \frac{1}{1-(1/2)} = 2 \tag{13.6}$$

$$0.99999\cdots = 0.9 \sum_{i=0}^{\infty} \left(\frac{1}{10}\right)^i = 0.9\left(\frac{1}{1-1/10}\right) = 0.9\left(\frac{10}{9}\right) = 1 \tag{13.7}$$

$$1 - 1/2 + 1/4 - \cdots = \sum_{i=0}^{\infty} \left(\frac{-1}{2}\right)^i = \frac{1}{1-(-1/2)} = \frac{2}{3} \tag{13.8}$$

$$1 + 2 + 4 + \cdots + 2^{n-1} = \sum_{i=0}^{n-1} 2^i = \frac{1-2^n}{1-2} = 2^n - 1 \tag{13.9}$$

$$1 + 3 + 9 + \cdots + 3^{n-1} = \sum_{i=0}^{n-1} 3^i = \frac{1-3^n}{1-3} = \frac{3^n - 1}{2} \tag{13.10}$$

If the terms in a geometric sum grow smaller, as in equation 13.6, then the sum is said to be *geometrically decreasing*. If the terms in a geometric sum grow progressively larger, as in equations 13.9 and 13.10, then the sum is said to be *geometrically increasing*. In either case, the sum is usually approximately equal to the term in the sum with the greatest absolute value. For example, in equations 13.6 and 13.8, the largest term is equal to 1 and the sums are 2 and 2/3, both relatively close to 1. In equation 13.9, the sum is about twice the largest term. In equation 13.10, the largest term is 3^{n-1} and the sum is $(3^n - 1)/2$, which is only about a factor of 1.5 greater. You can see why this rule of thumb works by looking carefully at equation 13.2 and Theorem 13.1.1.

13.1.6 Variations of Geometric Sums

We now know all about geometric sums—if you have one, life is easy. But in practice one often encounters sums that cannot be transformed by simple variable substitutions to the form $\sum x^i$.

A non-obvious but useful way to obtain new summation formulas from old ones is by differentiating or integrating with respect to x. As an example, consider the following sum:

$$\sum_{i=1}^{n-1} ix^i = x + 2x^2 + 3x^3 + \cdots + (n-1)x^{n-1}$$

This is not a geometric sum. The ratio between successive terms is not fixed, and so our formula for the sum of a geometric sum cannot be directly applied. But differentiating equation 13.2 leads to:

$$\frac{d}{dx}\left(\sum_{i=0}^{n-1} x^i\right) = \frac{d}{dx}\left(\frac{1-x^n}{1-x}\right). \tag{13.11}$$

The left-hand side of equation 13.11 is simply

$$\sum_{i=0}^{n-1} \frac{d}{dx}(x^i) = \sum_{i=0}^{n-1} i x^{i-1}.$$

The right-hand side of equation 13.11 is

$$\begin{aligned}\frac{-nx^{n-1}(1-x) - (-1)(1-x^n)}{(1-x)^2} &= \frac{-nx^{n-1} + nx^n + 1 - x^n}{(1-x)^2} \\ &= \frac{1 - nx^{n-1} + (n-1)x^n}{(1-x)^2}.\end{aligned}$$

Hence, equation 13.11 means that

$$\sum_{i=0}^{n-1} i x^{i-1} = \frac{1 - nx^{n-1} + (n-1)x^n}{(1-x)^2}.$$

Incidentally, Problem 13.2 shows how the perturbation method could also be applied to derive this formula.

Often, differentiating or integrating messes up the exponent of x in every term. In this case, we now have a formula for a sum of the form $\sum i x^{i-1}$, but we want a formula for the series $\sum i x^i$. The solution is simple: multiply by x. This gives:

$$\sum_{i=1}^{n-1} i x^i = \frac{x - nx^n + (n-1)x^{n+1}}{(1-x)^2} \tag{13.12}$$

and we have the desired closed-form expression for our sum. It seems a little complicated, but it's easier to work with than the sum.

Notice that if $|x| < 1$, then this series converges to a finite value even if there are infinitely many terms. Taking the limit of equation 13.12 as n tends to infinity gives the following theorem:

Theorem 13.1.2. *If* $|x| < 1$, *then*

$$\sum_{i=1}^{\infty} i x^i = \frac{x}{(1-x)^2}. \tag{13.13}$$

As a consequence, suppose that there is an annuity that pays im dollars at the end of each year i, forever. For example, if $m = \$50{,}000$, then the payouts are \$50,000 and then \$100,000 and then \$150,000 and so on. It is hard to believe that the value of this annuity is finite! But we can use Theorem 13.1.2 to compute the value:

$$\begin{aligned} V &= \sum_{i=1}^{\infty} \frac{im}{(1+p)^i} \\ &= m \cdot \frac{1/(1+p)}{(1-\frac{1}{1+p})^2} \\ &= m \cdot \frac{1+p}{p^2}. \end{aligned}$$

The second line follows by an application of Theorem 13.1.2. The third line is obtained by multiplying the numerator and denominator by $(1+p)^2$.

For example, if $m = \$50{,}000$, and $p = 0.08$ as usual, then the value of the annuity is $V = \$8{,}437{,}500$. Even though the payments increase every year, the increase is only additive with time; by contrast, dollars paid out in the future decrease in value exponentially with time. The geometric decrease swamps out the additive increase. Payments in the distant future are almost worthless, so the value of the annuity is finite.

The important thing to remember is the trick of taking the derivative (or integral) of a summation formula. Of course, this technique requires one to compute nasty derivatives correctly, but this is at least theoretically possible!

13.2 Sums of Powers

In Chapter 5, we verified the formula (13.1), but the source of this formula is still a mystery. Sure, we can prove that it's true by using well ordering or induction, but where did the expression on the right come from in the first place? Even more inexplicable is the closed form expression for the sum of consecutive squares:

$$\sum_{i=1}^{n} i^2 = \frac{(2n+1)(n+1)n}{6}. \tag{13.14}$$

It turns out that there is a way to derive these expressions, but before we explain it, we thought it would be fun—OK, our definition of "fun" may be different than yours—to show you how Gauss is supposed to have proved equation 13.1 when he was a young boy.

Gauss's idea is related to the perturbation method we used in Section 13.1.2. Let

$$S = \sum_{i=1}^{n} i.$$

Then we can write the sum in two orders:

$$\begin{aligned} S &= 1 + \quad 2 \quad + \ldots + (n-1) + n, \\ S &= n + (n-1) + \ldots + \quad 2 \quad + 1. \end{aligned}$$

Adding these two equations gives

$$\begin{aligned} 2S &= (n+1) + (n+1) + \cdots + (n+1) + (n+1) \\ &= n(n+1). \end{aligned}$$

Hence,

$$S = \frac{n(n+1)}{2}.$$

Not bad for a young child—Gauss showed some potential....

Unfortunately, the same trick does not work for summing consecutive squares. However, we can observe that the result might be a third-degree polynomial in n, since the sum contains n terms that average out to a value that grows quadratically in n. So we might guess that

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

If our 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:

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

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 13.14.

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, all the coefficients of the polynomial can be found automatically.

Be careful! This method lets 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 by induction or some other method. If the initial guess at the solution was not of the right form, then the resulting formula will be completely wrong! A later chapter will describe a method based on generating functions that does not require any guessing at all.

13.3 Approximating Sums

Unfortunately, it is not always possible to find a closed-form expression for a sum. For example, no closed form is known for

$$S = \sum_{i=1}^{n} \sqrt{i}.$$

In such cases, we need to resort to approximations for S if we want to have a closed form. The good news is that there is a general method to find closed-form upper and lower bounds that works well for many sums. Even better, the method is simple and easy to remember. It works by replacing the sum by an integral and then adding either the first or last term in the sum.

Definition 13.3.1. A function $f : \mathbb{R}^+ \to \mathbb{R}^+$ is *strictly increasing* when

$$x < y \text{ IMPLIES } f(x) < f(y),$$

and it is *weakly increasing*[3] when

$$x < y \text{ IMPLIES } f(x) \leq f(y).$$

Similarly, f is *strictly decreasing* when

$$x < y \text{ IMPLIES } f(x) > f(y),$$

and it is *weakly decreasing*[4] when

$$x < y \text{ IMPLIES } f(x) \geq f(y).$$

[3] Weakly increasing functions are usually called *nondecreasing* functions. We will avoid this terminology to prevent confusion between being a nondecreasing function and the much weaker property of *not* being a decreasing function.

[4] Weakly decreasing functions are usually called *nonincreasing*.

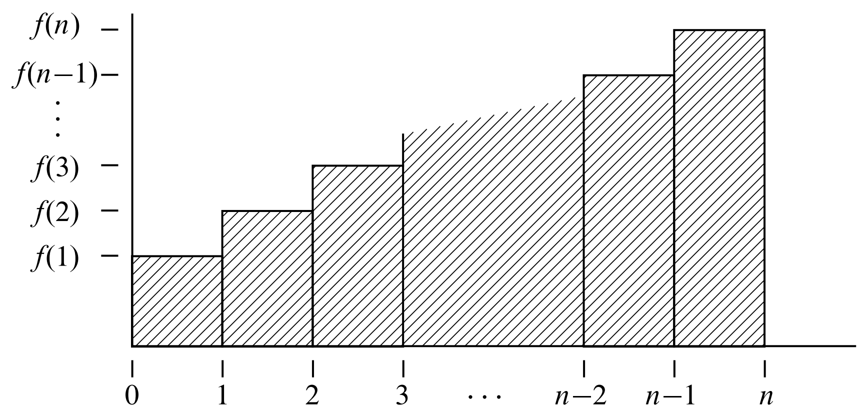


Figure 13.1 The area of the ith rectangle is $f(i)$. The shaded region has area $\sum_{i=1}^{n} f(i)$.

For example, 2^x and $\sqrt{x}$ are strictly increasing functions, while $\max\{x, 2\}$ and $\lceil x \rceil$ are weakly increasing functions. The functions $1/x$ and 2^{-x} are strictly decreasing, while $\min\{1/x, 1/2\}$ and $\lfloor 1/x \rfloor$ are weakly decreasing.

Theorem 13.3.2. *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a weakly increasing function. Define*

$$S ::= \sum_{i=1}^{n} f(i) \tag{13.15}$$

and

$$I ::= \int_1^n f(x)\,dx.$$

Then

$$I + f(1) \le S \le I + f(n). \tag{13.16}$$

Similarly, if f is weakly decreasing, then

$$I + f(n) \le S \le I + f(1).$$

Proof. Suppose $f : \mathbb{R}^+ \to \mathbb{R}^+$ is weakly increasing. The value of the sum S in (13.15) is the sum of the areas of n unit-width rectangles of heights $f(1), f(2), \ldots, f(n)$. This area of these rectangles is shown shaded in Figure 13.1.

The value of

$$I = \int_1^n f(x)\,dx$$

is the shaded area under the curve of $f(x)$ from 1 to n shown in Figure 13.2.

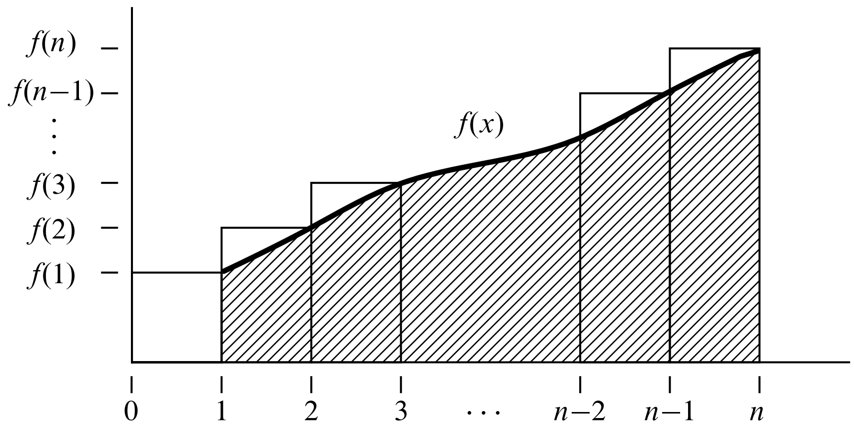


Figure 13.2 The shaded area under the curve of $f(x)$ from 1 to n (shown in bold) is $I = \int_1^n f(x)\,dx$.

Comparing the shaded regions in Figures 13.1 and 13.2 shows that S is at least I plus the area of the leftmost rectangle. Hence,

$$S \geq I + f(1) \tag{13.17}$$

This is the lower bound for S given in (13.16).

To derive the upper bound for S given in (13.16), we shift the curve of $f(x)$ from 1 to n one unit to the left as shown in Figure 13.3.

Comparing the shaded regions in Figures 13.1 and 13.3 shows that S is at most I plus the area of the rightmost rectangle. That is,

$$S \leq I + f(n),$$

which is the upper bound for S given in (13.16).

The very similar argument for the weakly decreasing case is left to Problem 13.10. ■

Theorem 13.3.2 provides good bounds for most sums. At worst, the bounds will be off by the largest term in the sum. For example, we can use Theorem 13.3.2 to bound the sum

$$S = \sum_{i=1}^{n} \sqrt{i}$$

as follows.

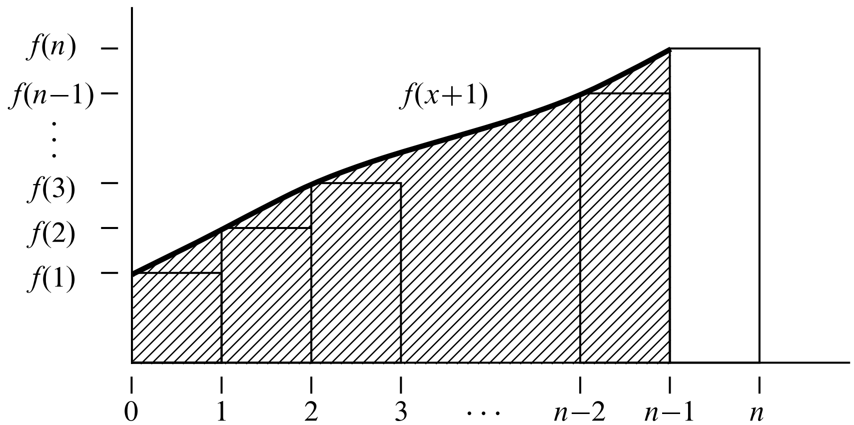


Figure 13.3 This curve is the same as the curve in Figure 13.2 shifted left by 1.

We begin by computing

$$\begin{aligned} I &= \int_1^n \sqrt{x}\, dx \\ &= \left.\frac{x^{3/2}}{3/2}\right|_1^n \\ &= \frac{2}{3}(n^{3/2} - 1). \end{aligned}$$

We then apply Theorem 13.3.2 to conclude that

$$\frac{2}{3}(n^{3/2} - 1) + 1 \;\le\; S \;\le\; \frac{2}{3}(n^{3/2} - 1) + \sqrt{n}$$

and thus that

$$\frac{2}{3}n^{3/2} + \frac{1}{3} \;\le\; S \;\le\; \frac{2}{3}n^{3/2} + \sqrt{n} - \frac{2}{3}.$$

In other words, the sum is very close to $\frac{2}{3}n^{3/2}$. We'll define several ways that one thing can be "very close to" something else at the end of this chapter.

As a first application of Theorem 13.3.2, we explain in the next section how it helps in resolving a classic paradox in structural engineering.

13.4 Hanging Out Over the Edge

Suppose you have a bunch of books and you want to stack them up, one on top of another in some off-center way, so the top book sticks out past books below it without falling over. If you moved the stack to the edge of a table, how far past the edge of the table do you think you could get the top book to go? Could the top book stick out completely beyond the edge of table? You're not supposed to use glue or any other support to hold the stack in place.

Most people's first response to the Book Stacking Problem—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!

13.4.1 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 13.4.

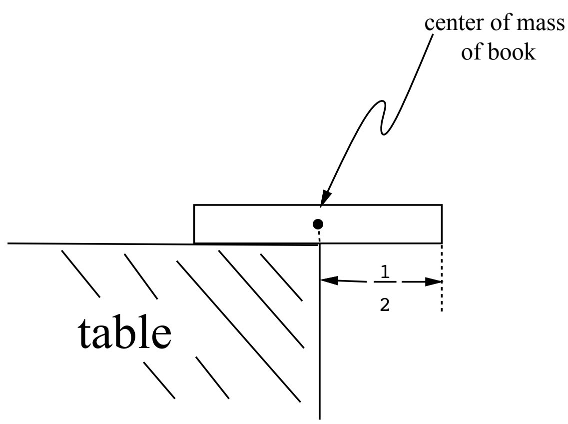


Figure 13.4 One book can overhang half a book length.

Now suppose we have a stack of books that will not tip over if the bottom book rests on the table—call that a *stable stack*. Let's define the *overhang* of a stable stack to be the horizontal distance from the center of mass of the stack to the furthest edge of the top book. So the overhang is purely a property of the stack, regardless of its placement on the table. If we place the center of mass of the stable stack at the edge of the table as in Figure 13.5, the overhang is how far we can get the top

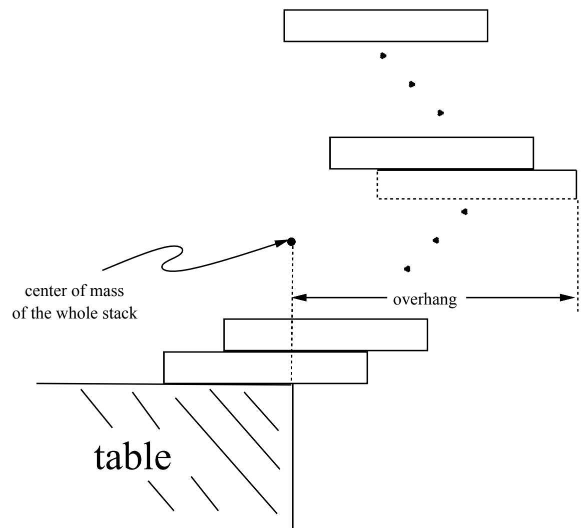


Figure 13.5 Overhanging the edge of the table.

book in the stack to stick out past the edge.

In general, a stack of n books will be stable if and only if the center of mass of the top i books sits over the $(i+1)$st book for $i = 1, 2, \ldots, n-1$.

So we want a formula for the maximum possible overhang, B_n, achievable with a stable 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 13.6.

So we know where to place the $n+1$st book to get maximum overhang. In fact, the reasoning above actually shows that this way of stacking $n+1$ books is the *unique* way to build a stable stack where the top book extends as far as possible. All we have to do is calculate what this extension is.

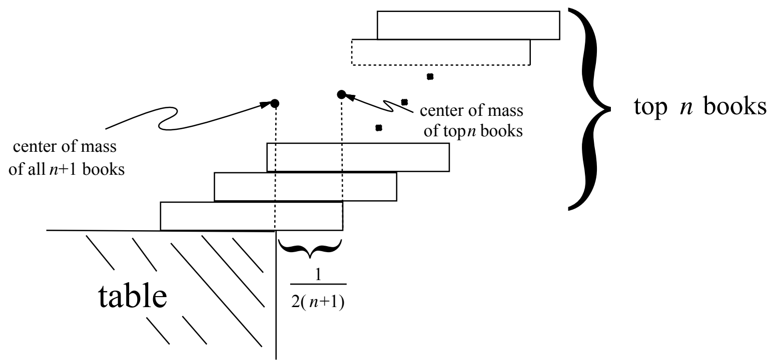


Figure 13.6 Additional overhang with $n+1$ books.

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,

$$B_{n+1} = B_n + \frac{1}{2(n+1)}, \tag{13.18}$$

as shown in Figure 13.6.

Expanding equation (13.18), we have

$$\begin{aligned} B_{n+1} &= B_{n-1} + \frac{1}{2n} + \frac{1}{2(n+1)} \\ &= B_1 + \frac{1}{2 \cdot 2} + \cdots + \frac{1}{2n} + \frac{1}{2(n+1)} \\ &= \frac{1}{2} \sum_{i=1}^{n+1} \frac{1}{i}. \end{aligned} \tag{13.19}$$

So our next task is to examine the behavior of B_n as n grows.

13.4.2 Harmonic Numbers

Definition 13.4.1. The nth *harmonic number*, H_n, is

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

So (13.19) means that

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

The first few harmonic numbers are easy to compute. For example, $H_4 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{25}{12} > 2$. 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 13.7.

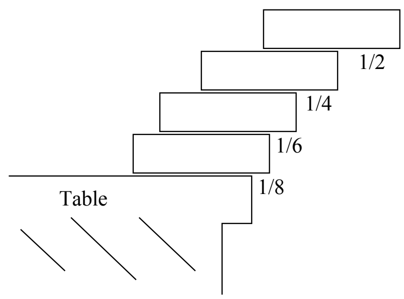


Figure 13.7 Stack of four books with maximum overhang.

There is good news and bad news about harmonic numbers. The bad news is that there is no known closed-form expression for the harmonic numbers. The good news is that we can use Theorem 13.3.2 to get close upper and lower bounds on H_n. In particular, since

$$\int_1^n \frac{1}{x}\,dx = \ln(x) \Big|_1^n = \ln(n),$$

Theorem 13.3.2 means that

$$\ln(n) + \frac{1}{n} \le H_n \le \ln(n) + 1. \tag{13.20}$$

In other words, the nth harmonic number is very close to $\ln(n)$.

Because the harmonic numbers frequently arise in practice, mathematicians have worked hard to get even better approximations for them. In fact, it is now known that

$$H_n = \ln(n) + \gamma + \frac{1}{2n} + \frac{1}{12n^2} + \frac{\epsilon(n)}{120n^4} \tag{13.21}$$

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

We are now finally done with our analysis of the book stacking problem. Plugging the value of H_n into (13.19), we find that the maximum overhang for n books is very close to $1/2\ln(n)$. Since $\ln(n)$ grows to infinity as n increases, this means that if we are given enough books we can get a book to hang out arbitrarily far over the edge of the table. Of course, the number of books we need will grow as an exponential function of the overhang; it will take 227 books just to achieve an overhang of 3, never mind an overhang of 100.

Extending Further Past the End of the Table

The overhang we analyzed above was the furthest out the *top* book could extend past the table. This leaves open the question of if there is some better way to build a stable stack where some book other than the top stuck out furthest. For example, Figure 13.8 shows a stable stack of two books where the bottom book extends further out than the top book. Moreover, the bottom book extends 3/4 of a book length past the end of the table, which is the same as the maximum overhang for the top book in a two book stack.

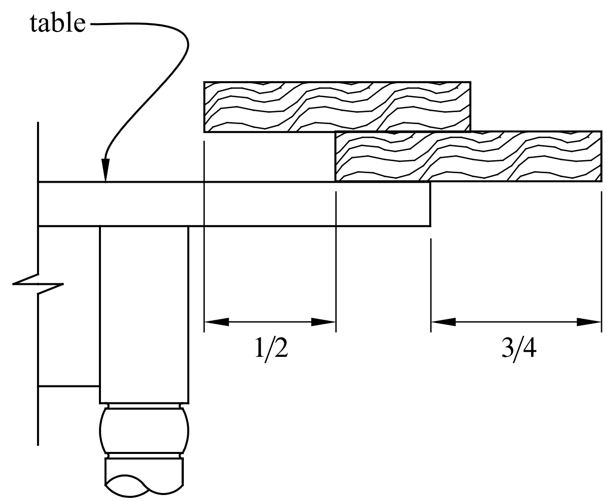
Since the two book arrangement in Figure 13.8(a) ties the maximum overhang stack in Figure 13.8(b), we could take the unique stable stack of n books where the top book extends furthest, and switch the top two books to look like Figure 13.8(a). This would give a stable stack of n books where the second from the top book extends the same maximum overhang distance. So for $n > 1$, there are at least two ways of building a stable stack of n books which both extend the maximum overhang distance—one way where the top book is furthest out, and another way where the second from the top book is furthest out.

It turns out that there is no way to beat these two ways of making stable stacks. In fact, it's not too hard to show that these are the *only* two ways to get a stable stack of books that achieves maximum overhang.

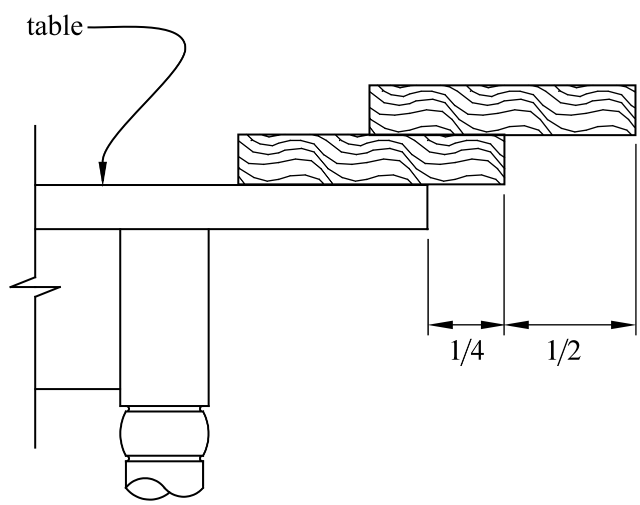
But there is more to the story. All our reasoning above was about stacks in which *one* book rests on another. It turns out that by building structures in which more than one book rests on top of another book—think of an inverted pyramid—it is possible to get a stack of n books to extend proportional to $\sqrt[3]{n}$—much more than $\ln n$—book lengths without falling over. See [13], *Maximum Overhang*.

13.4.3 Asymptotic Equality

For cases like equation 13.21 where we understand the growth of a function like H_n up to some (unimportant) error terms, we use a special notation, $\sim$, to denote the leading term of the function. For example, we say that $H_n \sim \ln(n)$ to indicate that



(a)



(b)

Figure 13.8 Figure (a) shows a stable stack of two books where the bottom book extends the same amount past the end of the table as the maximum overhang two-book stack shown in Figure (b).

the leading term of H_n is $\ln(n)$. More precisely:

Definition 13.4.2. For functions $f, g : \mathbb{R} \to \mathbb{R}$, we say f is *asymptotically equal* to g, in symbols,

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

iff

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

Although it is tempting to write $H_n \sim \ln(n) + \gamma$ to indicate the two leading terms, this is not really right. According to Definition 13.4.2, $H_n \sim \ln(n) + c$ 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:

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

We will spend a lot more time talking about asymptotic notation at the end of the chapter. But for now, let's get back to using sums.

13.5 Products

We've covered several techniques for finding closed forms for sums but no methods for dealing with products. Fortunately, we do not need to develop an entirely new set of tools when we encounter a product such as

$$n! ::= \prod_{i=1}^{n} i. \tag{13.22}$$

That's because we can convert any product into a sum by taking a logarithm. For example, if

$$P = \prod_{i=1}^{n} f(i),$$

then

$$\ln(P) = \sum_{i=1}^{n} \ln(f(i)).$$

We can then apply our summing tools to find a closed form (or approximate closed form) for $\ln(P)$ and then exponentiate at the end to undo the logarithm.

For example, let's see how this works for the factorial function $n!$. We start by taking the logarithm:

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

Unfortunately, no closed form for this sum is known. However, we can apply Theorem 13.3.2 to find good closed-form bounds on the sum. To do this, we first compute

$$\begin{aligned}\int_1^n \ln(x)\,dx &= x\ln(x) - x\Big|_1^n\\ &= n\ln(n) - n + 1.\end{aligned}$$

Plugging into Theorem 13.3.2, this means that

$$n\ln(n) - n + 1 \;\le\; \sum_{i=1}^{n} \ln(i) \;\le\; n\ln(n) - n + 1 + \ln(n).$$

Exponentiating then gives

$$\frac{n^n}{e^{n-1}} \;\le\; n! \;\le\; \frac{n^{n+1}}{e^{n-1}}. \tag{13.23}$$

This means that $n!$ is within a factor of n of n^n/e^{n-1}.

13.5.1 Stirling's Formula

The most commonly used product in discrete mathematics is probably $n!$, and mathematicians have workedto find tight closed-form bounds on its value. The most useful bounds are given in Theorem 13.5.1.

Theorem 13.5.1 (*Stirling's Formula*)**.** *For all $n \ge 1$,*

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\epsilon(n)}$$

where

$$\frac{1}{12n+1} \le \epsilon(n) \le \frac{1}{12n}.$$

Theorem 13.5.1 can be proved by induction (with some pain), and there are lots of proofs using elementary calculus, but we won't go into them.

There are several important things to notice about Stirling's Formula. First, $\epsilon(n)$ is always positive. This means that

$$n! > \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \tag{13.24}$$

for all $n \in \mathbb{N}^+$.

Second, $\epsilon(n)$ tends to zero as n gets large. This means that

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \tag{13.25}$$

which is impressive. After all, who would expect both π and e to show up in a closed-form expression that is asymptotically equal to $n!$?

Third, $\epsilon(n)$ is small even for small values of n. This means that Stirling's Formula provides good approximations for $n!$ for most all values of n. For example, if we use

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

as the approximation for $n!$, as many people do, we are guaranteed to be within a factor of

$$e^{\epsilon(n)} \le e^{\frac{1}{12n}}$$

of the correct value. For $n \ge 10$, this means we will be within 1% of the correct value. For $n \ge 100$, the error will be less than 0.1%.

If we need an even closer approximation for $n!$, then we could use either

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

or

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

depending on whether we want an upper, or a lower, bound. By Theorem 13.5.1, we know that both bounds will be within a factor of

$$e^{\frac{1}{12n}-\frac{1}{12n+1}} = e^{\frac{1}{144n^2+12n}}$$

of the correct value. For $n \ge 10$, this means that either bound will be within 0.01% of the correct value. For $n \ge 100$, the error will be less than 0.0001%.

For quick future reference, these facts are summarized in Corollary 13.5.2 and Table 13.1.

Approximation	$n \geq 1$	$n \geq 10$	$n \geq 100$	$n \geq 1000$
$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$	$< 10\%$	$< 1\%$	$< 0.1\%$	$< 0.01\%$
$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/12n}$	$< 1\%$	$< 0.01\%$	$< 0.0001\%$	$< 0.000001\%$

Table 13.1 Error bounds on common approximations for $n!$ from Theorem 13.5.1. For example, if $n \geq 100$, then $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ approximates $n!$ to within 0.1%.

Corollary 13.5.2.

$$n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot \begin{cases} 1.09 & \text{for } n \geq 1, \\ 1.009 & \text{for } n \geq 10, \\ 1.0009 & \text{for } n \geq 100. \end{cases}$$