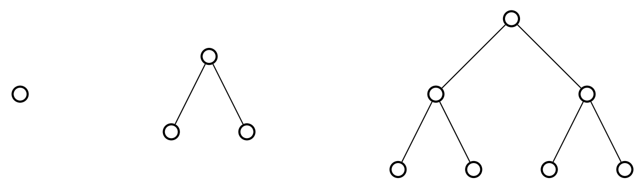
A *perfect binary tree* of depth d is a tree constructed like this:

- A tree of depth zero consists of a single vertex called the root.
- A tree of depth $d+1$ is obtained by taking two perfect binary trees T_1 and T_2 of depth d, a new root r, and adding edges from r to the roots of T_1 and T_2.

Here are the perfect binary trees of depth 0, 1, and 2:



How many vertices $N(d)$ does a perfect k-ary tree of depth d have? When $d \geq 1$, there is one vertex for the two subtrees of depth $d-1$, plus the root vertex. This gives the formula

$$N(d) = 2 \cdot N(d-1) + 1$$

for $d \geq 1$, with the "base case" $N(0) = 1$. Plugging in small values of d, this gives

$$\begin{aligned} N(1) &= 2 \cdot N(0) + 1 = 2 + 1 \\ N(2) &= 2 \cdot N(1) + 1 = 2(2+1) + 1 = 2^2 + 2 + 1 \\ N(3) &= 2 \cdot N(2) + 1 = 2(2^2 + 2 + 1) + 1 = 2^3 + 2^2 + 2 + 1. \end{aligned}$$

In general, $N(d) = 2^d + 2^{d-1} + \cdots + 1$. This type of sum is called a geometric sum.

1 Geometric sums

We can evaluate a sum of the form

$$S = x^d + x^{d-1} + \cdots + 1$$

for every real number x and positive integer d like this: If we multiply both sides by x, we obtain

$$xS = x^{d+1} + x^d + \cdots + x$$

If we now subtract the first expression from the second one, almost all the right hand sides terms cancel out:

$$xS - S = x^{d+1} - 1$$

which simplifies to $(x-1)S = x^{d+1} - 1$. When $x \neq 1$, we can do a division and obtain the formula

$$x^d + x^{d-1} + \cdots + 1 = \frac{x^{d+1} - 1}{x - 1} \qquad \text{for every real number } x \neq 1.$$

A sum of this form is called a *geometric sum*.

So the number of vertices in a perfect k-ary tree of depth d is $(k^{d+1} - 1)/(k-1)$. In particular, for a perfect ternary tree, this number is $(3^{d+1} - 1)/2$. A perfect binary (2-ary) tree of depth d has $2^{d+1} - 1$ vertices.

Annuities

You won a prize and you have two options for the prize money. Option A is that you are paid \$5000 per year for the rest of your life. Option B is that you are paid \$80000 today. Which one would you choose?

To answer this question we need to model how the value of money changes over time. If you keep your money in the bank at no interest then option A will pay off for you in twenty years time. If, on the other hand, you want to throw a lavish party right now then option B would make more sense for you. Now suppose that, as a savvy investor, you are quite confident in making a reliable return of $p = 7\%$ per year. How show this affect your choice?

To answer this question we'll calculate how much option A is worth in today's money. The 5K that you will be getting in your zeroth year are worth... well, 5K. For next year's 50K you can reason like this. If you had invested x dollars this year, they would be worth $(1+p)x$ dollars next year. So today's value of next year's 5K dollars is the amount x for which $(1+p)x = 5\text{K}$, namely $x = 5\text{K}/(1+p)$. By the same reasoning, the 5K you would be getting in two years' time are worth $5\text{K}/(1+p)^2$ today. Continuing this reasoning, you conclude that the value of option A in today's money is

$$5\text{K} + \frac{5\text{K}}{1+p} + \frac{5\text{K}}{(1+p)^2} + \cdots$$

By the geometric sum formula, the contribution from years zero up to d equals

$$5\text{K} \cdot \frac{1/(1+p)^{d+1} - 1}{1/(1+p) - 1}$$

In the large n limit, the term $1/(1+p)^{d+1}$ vanishes and the value converges to $5\text{K} \cdot (1+p)/p$. For $p = 7\%$, the value of option A is about $\$76,429$. So option B is better.

Here is another way to see the wisdom of option B over option A without evaluating the geometric sum. With a budget of 80K, I can always take out x dollars for spending this year and invest the remaining $80\text{K} - x$ dollars aiming to grow them back to 80K by next year. To do this, I need to choose x so that

$$(1+p) \cdot (80\text{K} - x) = 80\text{K},$$

which solves to

$$x = 80\text{K} \cdot \left(1 - \frac{1}{1+p}\right) = 80\text{K} \cdot \frac{p}{1+p}.$$

When p equals 7% we get that x is about 5234 dollars. This improves on the annual 5K in option A.

2 Polynomial sums

In Lecture 3 we proved that

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

for every integer $n \geq 0$. How did I come up with the expression on the right? Instead of going back to something we already know, let's work out a new one:

What is $1^2 + 2^2 + \cdots + n^2$?

We have to do some guesswork. The sum $1+2+\cdots+n$ was a quadratic function in n, perhaps $1^2+2^2+\cdots+n^2$ might equal some cubic? Let's make a guess: For all n, there exist real numbers a, b, c, d such that

$$1^2 + 2^2 + \cdots + n^2 = an^3 + bn^2 + cn + d.$$

Suppose our guess was correct. Then what are the numbers a, b, c, d? We can get an idea by evaluating both sides for different values of n:

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

I solved this system of equations on the computer and obtained $a = 1/3$, $b = 1/2$, $c = 1/6$, $d = 0$. This suggests the formula

$$1^2 + 2^2 + \cdots + n^2 = \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n$$

for all integers $n \geq 0$. Let us see if we can prove its correctness by induction on n.

We already worked out the base case $n = 0$, so let us do the inductive step. Fix $n \geq 0$ and assume that the equality holds for n. Then

$$1^2 + 2^2 + \cdots + (n+1)^2 = \left(\tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n\right) + (n+1)^2 = \tfrac{1}{3}n^3 + \tfrac{3}{2}n^2 + \tfrac{13}{6}n + 1.$$

This indeed equals $\frac{1}{3}(n+1)^3 + \frac{1}{2}(n+1)^2 + \frac{1}{6}(n+1)$. So we have discovered and proved a new theorem:

Theorem 1. *For every integer $n \geq 0$, $1^2 + 2^2 + \cdots + n^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$.*

3 Approximating sums

Exact "closed-form" expressions for sums are rather exceptional. Often we have to resort to approximations. As an example, let us look at the sum

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

What can we say about this sum? At the very least, we can say that it is always non-negative, namely $S(n) \geq 0$ for all n. We can also say that each of the individual terms does not exceed $\sqrt{n}$, so the sum can be at most $n \cdot \sqrt{n} = n^{3/2}$:

$$0 \leq \sqrt{1} + \sqrt{2} + \cdots + \sqrt{n} \leq \sqrt{n} + \sqrt{n} + \cdots + \sqrt{n} = n \cdot \sqrt{n} \leq n^{3/2}.$$

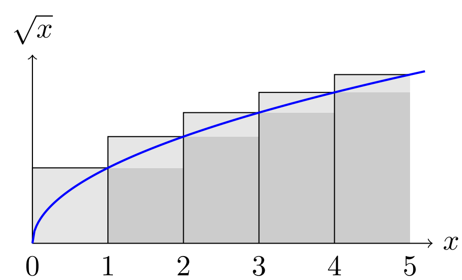
We can therefore be sure that the sum is always "sandwiched" between 0 and $n^{3/2}$:

$$0 \leq S(n) \leq n^{3/2} \qquad \text{for all } n.$$

This is an example of an *approximation*: It tells us, for example, that $S(100)$ is between 0 and 1000 and $S(1000)$ is between 0 and $1000^{3/2} \approx 31623$.

This approximation is not very informative leaves a large range of possibilities for the value of $S(n)$. We can narrow down the range by working harder. An effective method works by comparing the sum with a related integral.

The value $S(n)$ can be visualized as the joint area of n bars $R_1, \ldots, R_n$ where R_x has base $(x-1, x)$ and height $\sqrt{x}$. For example, $S(5)$ equals the area covered by the shaded bars (both light and dark shades) in this plot:



The area under the bars can be lower bounded by the area (i.e., the integral) of the curve $f(x) = \sqrt{x}$ from $x = 0$ to $x = n$:

$$S(5) \geq \int_0^5 \sqrt{x}\, dx.$$

If we remove the area L covered by the lightly shaded bars, the darker shaded area is now dominated by the curve $f(x) = \sqrt{x}$ and so

$$S(5) - L \leq \int_0^5 \sqrt{x}\, dx.$$

The area under L is exactly $\sqrt{5}$: If we stack all of the lightly shaded bars on top of one another, we obtain a column of width 1 and height $\sqrt{5}$. Therefore

$$\int_0^5 \sqrt{x}\, dx \leq S(5) \leq \int_0^5 \sqrt{x}\, dx + \sqrt{5}.$$

By the same reasoning, for every integer $n \geq 1$, we have the inequalities

$$\int_0^n \sqrt{x}\, dx \leq S(n) \leq \int_0^n \sqrt{x}\, dx + \sqrt{n}.$$

We can now use rules from calculus to evaluate the integrals: Recalling that $x^{1/2} = \frac{d}{dx}\frac{2}{3}x^{3/2}$, it follows from the fundamental theorem of calculus that

$$\frac{2}{3}n^{3/2} \leq S(n) \leq \frac{2}{3}n^{3/2} + \sqrt{n}.$$

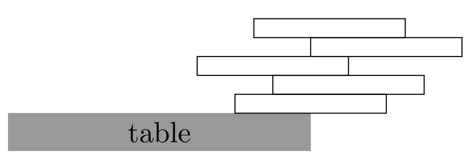
To get a feel for these inequalities, let us plug in a few values of n. (I calculated $S(n)$ by evaluating the sum on the computer.)

n	$\frac{2}{3}n^{3/2}$	$S(n)$	$\frac{2}{3}n^{3/2} + \sqrt{n}$
10	21.082	22.468	24.244
100	666.67	671.46	676.67
1,000	21,081.9	21,097.5	21,113.5
10,000	666,666	666,716	666,766

As n becomes large, the accuracy of these approximations looks quite good.

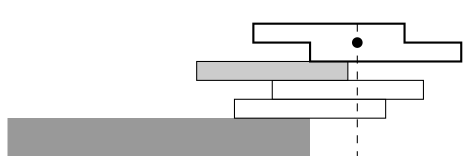
4 Overhang

You have n identical rectangular blocks and you stack them on top of one another at the edge of a table like this:



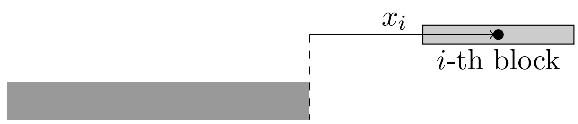
Is this configuration stable, or will it topple over?

In general, a configuration of n blocks is *stable* if for every i between 1 and n, the center of mass of the top i blocks sits over the $(i+1)$st block, where we think of the table as the $(n+1)$st block in the stack. For example, the top stack is not stable because the center of mass of the top two blocks does not sit over the third block:



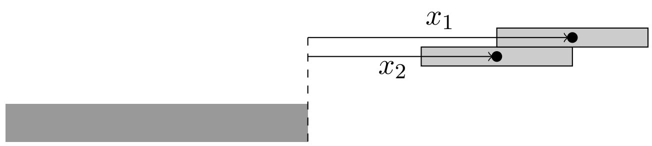
We want to stack our n blocks so that the rightmost block hangs as far over the edge of the table as possible. What should we do? One reasonable strategy is to try to push the top blocks as far as possible away from the table as long as they do not topple over.

We will assume each block has length 2 units and we will use x_i to denote the offset of the center of the i-th block (counting from the top) from the edge of the table:



The offset of a block can be positive, zero, or negative, depending on the position of its center of mass.

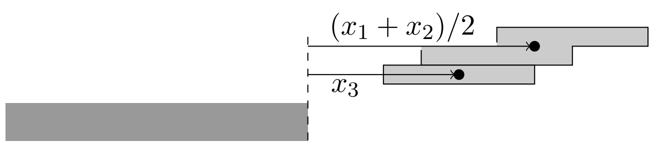
For the top block not to topple over, its center of mass must sit over the second block. To move it as far away from the edge of the table as possible, we should move its center exactly one unit to the right of the center of the second block:



This forces the offsets x_1 and x_2 to satisfy the equation

$$x_1 = x_2 + 1. \tag{1}$$

How about the third block? The center of mass of the first two blocks is at offset $(x_1 + x_2)/2$ from the edge of the table. To push this as far to the right as possible without toppling over the third block



we must set

$$\frac{x_1 + x_2}{2} = x_3 + 1. \tag{2}$$

Continuing our reasoning in this way, for every i between 1 and n, the offset of the center of mass of the top i blocks is $(x_1 + \cdots + x_i)/i$. To push this as far to the right without toppling over the $(i+1)$st block, we must set

$$\frac{x_1 + x_2 + \cdots + x_i}{i} = x_{i+1} + 1 \quad \text{for all } 1 \le i \le n. \tag{3}$$

Finally, when $i = n + 1$, we have reached the table whose offset is zero. Since we are thinking of the table as the $(n + 1)$st block, its centre of mass is one unit left to its edge:

$$x_{n+1} = -1. \tag{4}$$

The overhang of the set of blocks is $x_1 + 1$. To figure out what this number is, we need to solve for x_1 in the system of equations (3-4). Let us develop some intuition first. Equation (1) tells us that $x_2 = x_1 - 1$. Plugging in this formula for x_2 into (2), we get that

$$x_3 = x_1 - \frac{1}{2} - 1.$$

Let's do one more step. Equation (3) tells us that $(x_1 + x_2 + x_3)/3 = x_4 + 1$. Plugging in our formulas for x_2 and x_3 in terms of x_1 we get that

$$\frac{x_1 + (x_1 - 1) + (x_1 - \frac{3}{2})}{3} = x_4 + 1$$

from where

$$x_4 = x_1 - \frac{1+\frac{3}{2}}{3} - 1 = x_1 - \frac{1}{3} - \frac{1}{2} - 1.$$

At this point it is reasonable to guess that x_{i+1} should equal x_1 minus the sum

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{i}.$$

Let us prove that this guess is correct.

Lemma 2. *For all i between 1 and n, $x_i - x_{i+1} = 1/i$.*

Proof. If we multiply both sides of the i-th equation (3) by i we obtain

$$x_1 + x_2 + \cdots + x_{i-1} + x_i = i \cdot (x_{i+1} + 1).$$

Under this scaling the $(i-1)$st equation is

$$x_1 + x_2 + \cdots + x_{i-1} = (i-1) \cdot (x_i + 1).$$

Subtracting the two we obtain that

$$x_i = i(x_{i+1} - 1) - (i-1)(x_i - 1) = ix_{i+1} - (i-1)x_i + 1$$

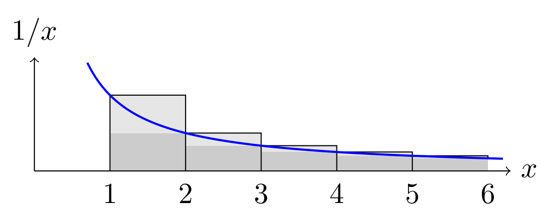
from where, after moving the variables around, we conclude that

$$x_i = x_{i+1} + \frac{1}{i}.$$ □

It follows immediately from this Lemma that

$$x_1 - x_{n+1} = (x_1 - x_2) + (x_2 - x_3) + \cdots + (x_{n+1} - x_n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n}.$$

Since $x_{n+1} = -1$, the overhang $x_1 + 1$ equals exactly this number, which is called the *n-th harmonic number* and is denoted by $H(n)$. There is no closed-form expression for $H(n)$, but we can obtain an excellent approximation using the integral method. To do this, we compare $H(n)$ with the integral of the function $1/x$:



By similar reasoning as before, the sum $H(n) = 1 + 1/2 + \ldots 1/n$ is given by the area of the first n shaded bars. This area is larger than the integral of $1/x$ from 1 to $n+1$:

$$H(n) \geq \int_1^{n+1} \frac{1}{x}\, dx.$$

On the other hand, if we subtract from $H(n)$ the area of the lightly shaded bars, then the integral becomes larger. This area equals $1 - 1/(n+1)$:

$$H(n) - 1 + \frac{1}{n+1} \leq \int_1^{n+1} \frac{1}{x}\, dx.$$

Combining this two inequalities gives the approximation

$$\int_1^{n+1} \frac{1}{x}\, dx \le H(n) \le \int_1^{n+1} \frac{1}{x}\, dx + 1 - \frac{1}{n+1}.$$

The antiderivative of $1/x$ is $\ln x$. By the fundamental theorem of calculus it follows that

$$\ln(n+1) \le H(n) \le \ln(n+1) + 1 - \frac{1}{n+1}. \tag{5}$$

The left hand side of this inequality tells us that our method of stacking blocks achieves overhang at least $\ln(n+1)$. The logarithm function is unbounded; given enough blocks, we can grow our stack all the way to New York!

5 Order of growth

Theorem 1 gives us an exact formula for the sum of the first n squares: It equals $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$. Sometimes precise like this one give us too much information. For example, say we want to know which of these two numbers is bigger:

$$A = 1 + 2 + \cdots + 10000 \qquad \text{or} \qquad B = 1^2 + 2^2 + \cdots + 1000^2?$$

In other words, we want to compare the numbers

$$A = \tfrac{1}{2}n^2 + \tfrac{1}{2}n \text{ when } n = 10000 = 10^4 \qquad \text{and} \qquad B = \tfrac{1}{3}m^3 + \tfrac{1}{2}m^2 + \tfrac{1}{6}m \text{ when } m = 1000 = 10^3.$$

When n and m are large, the two expressions are "dominated" by their leading terms $\frac{1}{2}n^2$ and $\frac{1}{3}m^3$, which evaluate to $\frac{1}{2}(10^4)^2 = 0.5 \cdot 10^8$ and $\frac{1}{3}(10^3)^3 \approx 0.33 \cdot 10^9$. The second one is an order of magnitude larger, which suggests that B should be the bigger number. Indeed, B is bigger than A by the same order of magnitude:

$$A = 50,005,000 = 1.0001 \cdot \tfrac{1}{2}(10^4)^2 \qquad \text{and} \qquad B = 333,833,500 \approx 1.002 \cdot \tfrac{1}{3}(10^3)^3.$$

This example suggests that when comparing expressions like $\frac{1}{2}n^2 + \frac{1}{2}n$ and $\frac{1}{3}m^3 + \frac{1}{2}m^2 + \frac{1}{6}m$, what really matters are the leading terms $\frac{1}{2}n^2$ and $\frac{1}{3}m^3$, or even more crudely just n^2 and m^3.

There is a special notation for expressing "the relevant part of a function" when the input is large. To understand how it works let's think about what "the relevant part of $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$ is n^3 for large n" should mean. It means two things: that the actual value is never much larger than n^3, and that is also never much smaller than n^3. We'll start with the "not much larger" part.

Big-oh

Definition 3. For two real-valued functions f and g (defined over the positive reals, or over the positive integers), we say f is $O(g)$ (big-oh of g) if there exists a constant $C > 0$ such that for every sufficiently large input x, $f(x) \le C \cdot g(x)$.

For example, $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$ is $O(n^3)$ because $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n \le n^3 + n^3 + n^3 \le 3n^3$ when n is large (specifically, at least 1).

In Section 3 we showed that $\sqrt{1} + \cdots + \sqrt{n} \le \frac{2}{3}n^{3/2} + \sqrt{n}$. By the same reasoning we can say that $\sqrt{1} + \cdots + \sqrt{n}$ is $O(n^{3/2})$ because $\frac{2}{3}n^{3/2} + \sqrt{n} \le n^{3/2} + n^{3/2} \le 2n^{3/2}$ (when $n \ge 1$).

This type of reasoning leads to the following general rule: Every polynomial (even one with fractional exponent) is big-oh of its leading monomial.

In Section 4 we showed that $1 + 1/2 + \cdots + 1/n \leq \ln(n+1) + 1 - 1/(n+1)$. What is the "leading term" in big-oh notation here? We can say that

$$\ln(n+1) + 1 - 1/(n+1) \leq \ln(2n) + 1 - 1/(n+1) = \ln n + \ln 2 + 1 - 1/(n+1).$$

When n is large (say bigger than 2) then $\ln 2$ and 1 are both at most $\ln n$ so the whole expression is $O(\ln n)$. The base of the logarithm is irrelevant in big-oh notation as it changes the value by a multiplicative constant, so we can say that $\ln n + \ln 2 + 1 - 1/(n+1)$ is $O(\log n)$. In computer science, the "default" base of the logarithm is two.

In fact, the log of any polynomial in n is $O(\log n)$. For example,

$$\log(16n^5 + 3n + 11) \leq \log(16n^5 + 3n^5 + 11n^5) \leq \log(30n^5) \leq \log(n^6) \leq 6 \log n.$$

One important property of big-oh is that it is *transitive*: If f is $O(g)$ and g is $O(h)$ then f is $O(h)$. For example, n^3 is $O(n^4)$ (because $n^3 \leq n^4$ when $n \geq 1$), so $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$ is also $O(n^4)$. Thus big-oh provides an *upper bound* on asymptotic growth, but does not "pin down" the exact rate. That is accomplished by big-theta.

Big-theta

Big-Theta says that two functions have the same order of growth:

Definition 4. We say f is $\Theta(g)$ if f is $O(g)$ and g is $O(f)$.

We saw that $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$ is $O(n^3)$. We can also say that n^3 is $O(\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n)$ because $n^3 \leq 3 \cdot \frac{1}{3}n^3 \leq 3 \cdot (\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n)$. We can therefore conclude that $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$ is $\Theta(n^3)$.

We also saw that $S(n) = \sqrt{1} + \cdots + \sqrt{n}$ is $O(n^{3/2})$. Is $n^{3/2}$ also $O(S(n))$? As $S(n) \geq \frac{2}{3}n^{2/3}$, indeed we get that $n^{2/3} \leq \frac{2}{3}S(n)$ which is $O(S(n))$. Therefore $S(n)$ is $O(n^{3/2})$.

By similar logic we can say that

$$\boxed{\text{A polynomial is big-theta of its leading term.}}$$

So is every function that is sandwiched between two polynomials with the same leading term.

The big-theta definition says that both $f(x) \leq Cg(x)$ and $g(x) \leq Cf(x)$ for some constant C when x is sufficiently large, from where

$$C^{-1}g(x) \leq f(x) \leq Cg(x).$$

Namely, f is sandwiched between a small and a large multiple of g.

Therefore $\log(n^3 + 16n)$ is $\log \Theta(n^3)$, which is $\Theta(\log n)$: The function $n^3 + 16n$ is sandwiched between $C^{-1}n^3$ and Cn^3 for large n, so $\log(n^3 + 16n)$ is sandwiched between $3 \log n + \log C^{-1}$ and $3 \log n + \log C$. Both are $\Theta(\log n)$. By the same logic,

$$\boxed{\text{The log of any polynomial in } n \text{ is } \Theta(\log n).}$$

Big-theta is transitive and symmetric (f is $\Theta(g)$ iff g is $\Theta(f)$) so it is an *equivalence relation*: It splits all functions into "classes" in which every function is big-theta of every other. For example, $\frac{1}{2}n^2 + \frac{1}{2}n$, $0.1n^2 - 10n$, and n^2 are all big-theta of one another. The objective of big-theta notation is to express complicated functions in terms of more basic ones, so if we had to pick a big-theta representative among these n^2 would be the most sensible choice.

In computer science the most basic functions are the powers n^a (e.g., $n, n^2, \sqrt{n}$), exponentials B^n (e.g., $2^n, e^n, 2^{-n}$), and the logarithm $\log n$. Not every function can be represented as big-theta of these. As we will see shortly, $(\log n)^2$ grows strictly faster than $\log n$ but strictly slower than any power n^a. To understand why we need the last piece of asymptotic notation.

Little-oh

The little-oh notation says that asymptotically, one function grows at a significantly slower rate than another one.

Definition 5. For two real-valued functions f and g, we say f is $o(g)$ (little-oh of g) if *for every* constant $c > 0$ and every sufficiently large input x, $f(x) \leq c \cdot g(x)$.

If f is $o(g)$, then f is also $O(g)$, but not necessarily the other way. For example, $\frac{1}{2}n^2 + \frac{1}{2}n$ is $O(n^2)$, but it is not $o(n^2)$ because as n gets large, $(\frac{1}{2}n^2 + \frac{1}{2}n)$ is *not* less than, say, $\frac{1}{4}n^2$. On the other hand, $\frac{1}{2}n^2 + \frac{1}{2}n$ is $o(n^3)$ because $\frac{1}{2}n^2 + \frac{1}{2}n \leq n^2 = (1/n) \cdot n^3$ and as n becomes large, $1/n$ is smaller than any constant c.

By similar reasoning,

$$\boxed{n^a \text{ is } o(n^b) \text{ if } a < b.}$$

More generally, every degree-a polynomial $p(n)$ is $o(n^b)$ if $a < b$.

Little-oh is transitive and *asymmetric*: If f is $o(g)$, then g is not $o(f)$. Owing to these features, little-oh orders all (sufficiently nice) functions in terms of their asymptotic growth. For example, $\sqrt{n}$, n^2, and $n + \log n$ are ordered as $\sqrt{n}, n + \log n, n^2$ because $\sqrt{n}$ is $o(n + \log n)$ which is in turn $o(n^2)$.

Big-oh, big-theta, and little-oh are the asymptotic analogues of "less than or equal", "equal", and "strictly less than". The reasoning we use when ordering numbers applies to ordering functions as well. For example, we can say that because $n + \log n$ is $\Theta(n)$ and $\sqrt{n}$ is $o(n)$, $\sqrt{n}$ must also be $o(n + \log n)$. Specifically, when we want to order a sequence of functions it is legitimate to replace each with its simplest big-theta representative.

One incomplete but often successful method for figuring out the relative order of growth of f and g is to take the limit of the ratio $f(x)/g(x)$ as x tends to infinity. Assuming this limit exists we can say that

$$\begin{aligned} f \text{ is } O(g) &\text{ if } \lim_{x\to\infty} f(x)/g(x) < \infty, \\ f \text{ is } \Theta(g) &\text{ if } 0 < \lim_{x\to\infty} f(x)/g(x) < \infty, \\ f \text{ is } o(g) &\text{ if } \lim_{x\to\infty} f(x)/g(x) = 0 \end{aligned}$$

So $17n^4 + 5n^3$ is $\Theta(n^4)$ because $(17n^4 + 5n^3)/n^4$ tends to 17 in the limit, while $n^{3/2}$ is $o(n^2)$ because $n^{3/2}/n^2 = 1/\sqrt{n}$ tends to zero in the limit. In particular, exponentials are strictly ordered by their basis

$$\boxed{B^n \text{ is } o(C^n) \text{ when } 0 < B < C}$$

because $B^n/C^n = (B/C)^n$ is an exponential of base less than one, so it tends to zero in the limit.

To summarize the asymptotic growth of "simple" functions, powers are little-oh ordered in increasing exponent, exponentials are little-oh ordered in increasing basis, and logarithms are big-theta equivalent to one another. What is the relative ordering of logarithms, powers, and exponentials?

Exponential growth

Theorem 6. *For all constants $a, b > 0$, $(\log x)^a$ is $o(x^b)$.*

In the proof we may assume that the logarithm is a natural logarithm because $\log x$ is $\Theta(\ln x)$.

Proof. When $a = 1$, we can calculate $\lim_{x\to\infty}(\ln x)/x^b$ using L'Hôpital's rule from calculus. Both numerator and denominator grow to infinity, but this is not true for their derivatives: $\frac{d}{dx}\ln x = 1/x$, while $\frac{d}{dx}x^b = bx^{b-1}$. The ratio of these two numbers is $1/(bx^b)$, which tends to zero as x grows. Therefore

$$\lim_{x\to\infty} \frac{\ln x}{x^b} = \lim_{x\to\infty} \frac{1}{bx^b} = 0$$

and so $\log x \leq cx^b$ for every constant $c > 0$ and sufficiently large x.

Now let $a > 0$ be arbitrary and $c > 0$ be an arbitrary constant. By what we just proved,

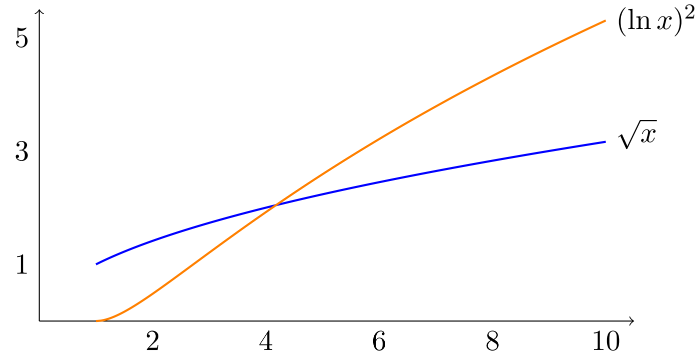
$$\ln x \leq c^{1/a} x^{b/a}$$

for x sufficiently large, from where

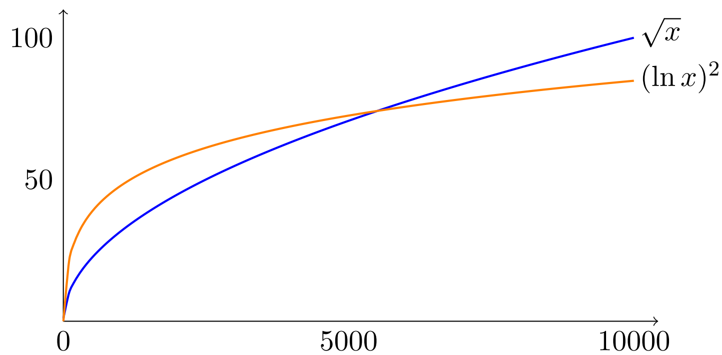
$$(\ln x)^a \leq (c^{1/a} x^{b/a})^a \leq c x^b$$

for x sufficiently large. $\square$

This theorem says that $\sqrt{x}$ outgrows $(\ln x)^2$ when x is large. Here is a plot of the two functions:



It looks like the log takes a lead around $x = 4$. The lead eventually disappears:



If we set $x = e^y$ and $B = e^b$, we get the following corollary:

Corollary 7. *For all constants $a > 0$ and $B > 1$, y^a is $o(B^y)$.*

In conclusion, we now know how to little-oh order the basic functions: The logarithm, then all powers in increasing order of exponent, then all exponentials in increasing order of basis. Using this we can reason out the relative order of other functions.

Example How are the functions 2^x, 2^{x^2}, x^2, x^x ordered in terms of their asymptotic growth?

By Corollary 7 we know that x^2 is both $o(2^x)$ and $o(2^{x^2})$. How do 2^x and 2^{x^2} compare to each other? As x is $o(x^2)$ we would expect that 2^x should also be $o(2^{x^2})$. In fact the ratio $2^x/2^{x^2}$ equals 2^{x-x^2} and this goes to zero as the exponent $x - x^2$ tends to $-\infty$. So we know that

$$x^2 \text{ is } o(2^x) \quad \text{and} \quad 2^x \text{ is } o(2^{x^2}).$$

Where does x^x fit in? To answer this it is usually a sensible strategy to identify whether x^x is a polynomial or an exponential-type function. In this example x^x grows faster than x, x^2, x^{100} and any polynomial in x, so x^2 is certainly $o(x^x)$. To compare x^x against 2^x and 2^{x^2} it is sensible to rewrite it as a base-2 exponential, namely $x^x = 2^{x \log x}$. Now we see that x is $o(x \log x)$ and $x \log x$ is $o(x^2)$ so x^x should fit between 2^x and 2^{x^2}:

$$x^2 \text{ is } o(2^x), \quad 2^x \text{ is } o(x^x), \quad \text{and } x^x \text{ is } o(2^{x^2}).$$

A warning about asymptotic notation It is customary to abuse the equality sign when talking about order of growth. In books you often see "$f = O(g)$" instead of "f is $O(g)$". Technically, this is incorrect because f and $O(g)$ are objects of different types: f is a single function while $O(g)$ is not. It is okay to use this notation as long as you are aware of what it means. What you should *not* do is write "equations" like

$$1 + 2 + \cdots + n = O(1) + O(1) + \cdots + O(n) = (n-1) \cdot O(1) + O(n) = O(n)$$

because it is not clear what they mean and may lead to incorrect conclusions.

References

This lecture is based on Chapter 13 of the text *Mathematics for Computer Science* by E. Lehman, T. Leighton, and A. Meyer. The variant of the integral methods described in the textbook is slightly different from the one in these notes.

Surprisingly, if we allow for the blocks to be stacked not only on top of one another but also side by side, the overhang can be much improved. If you are interested, see the amazing work *Overhang* by Mike Paterson and Uri Zwick.