

PROFESSOR: So we figured out that you can get the book stack, the overhang of the books, to be half the harmonic sum. With n books, you can get out H_n over 2 where H_n is this harmonic sum, or harmonic number. The question is, how are we going to estimate or calculate what this sum is? Now it turns out there is no simple formula for exactly what this sum is, but there is a simple formula that estimates it quite accurately. And we get the estimation by bounding the sum by integrals. And so let's look at this integral method for estimating sums.

Remember what I wanted was the sum of 1 plus $1/2$ plus $1/3$ down to $1/n$. So let's form some unit width rectangles of heights equal to the amount that I want. So here's a rectangle of height 1, rectangle of height $1/2$, rectangle of height $1/3$. I'm going out here, if you actually count to 8, but let's propose that this is a height $1/n$. And what I know is that the total area of these rectangles is actually equal to the number that I want. The total area of these rectangles is the harmonic number. And I'm interested in a lower bound for H_n , because I want to know how far I can get out. I want a tight lower bound. It says H_n is larger than a certain amount. That's the amount that I'm sure that I can stack out in books.

So the way I'm going to get a lower bound on this number H_n is by looking at this curve that goes through the corners of the rectangles. And if you check it, that curve is $1/(x+1)$. That is, the point here is-- when x is 0, I'm at $1/1$. When x is 1, I'm at $1/2$, the height of the second rectangle, and so on. So $1/(x+1)$ is a curve that is strictly below the boundaries of all these rectangles. That means that the area under $1/(x+1)$ going from 0 to n is a lower bound on H_n , because it's a lower bound on the area of the rectangles. So H_n equals the area of the rectangles. It's greater than the area of $1/(x+1)$, which of course is equal to the integral from 0 to n of $1/(x+1)$, which shifting variables is the same as the integral from 1 to $n+1$ of $1/x$ dx, which of course we know is a natural logarithm of $n+1$.

So there we have it. The overhang that you need for three books, which is B_n greater than or equal to 3, means that H_n has to be greater than or equal to 6. So by this estimate, I need \log of $n+1$ greater than or equal to 6 in order to get three books out, that the back end of the top book is two books past the edge of the table and the right end of the furthest out book is three book lengths past the edge of the table. So my bound tells me that I need n books such that \log of $n+1$ is greater than or equal to 6. Well, exponentiating both sides, the right-

hand side becomes e to the 6th. And I figure out that as long as n is greater than or equal to e to the 6th minus 1 books-- rounded up, of course, because there's only-- you can't have fractions of a book-- you get an estimate that with 403 books, I can actually get my stack to stick out three book lengths past the edge of the table.

Well if you do the actual calculation instead of the estimate, it turns out that 227 books are enough. So this estimate's a little off. But for our purposes, it tells us a dramatic fact, which is that we know that \log of n plus 1 approaches infinity as n approaches infinity. And that means that, with enough books, I can get out as far as I want. You tell me how many book lengths you want to be out, I'll use the $\log n$ formula to calculate how many books I need to get that far out.

So here's an example of some students in the class some years ago decided to do this as an experiment. Now when we used to do this in class, we first tried to do it with the big, heavy textbook that we used. And we kept trying to get them to balance and go out over the edge of the table, and they failed, because it turns out that textbooks are heavy and they compress. They're not the rigid rectangular cross sections that our model was based on. But CD cases work very well. They are more rigid. They don't compress easily, and they're very lightweight so that they don't cause problems with distortions because of the size of the stack. And so you can actually get CD cases to stick out pretty far.

This is an example where it's 43 CD cases high, and the top four cases are completely past the edge of the table. The leftmost edge is about 1.81 or 1.91 case lengths past the table. There's another view of it from the guy who made the stack. And of course, they were right on the edge of stability in trying to get the CDs to stick out as far as possible. So if you notice these little spaces there, in terms of the balancing, it's really just on the brink of falling over. If you sneeze at it, it'll tip. But if you don't sneeze at it, it's stable, and you get the top CD out that far.

So while we're at it, let's get an upper bound for H_n . We just got a lower bound of H_n , but the same kind of logic of using an integral will give you an upper bound for H_n . What I do now is I run a curve from the upper right corners of the rectangles. And that curve is simply 1 over x . So an upper bound for the harmonic number H_n is the area under 1 over x out to n plus this 1 . And so I get an upper bound that says that the harmonic number H_n is less than the integral from 1 to n of 1 over x dx plus 1 , or it's equal to 1 plus \log of n .

So combining those two bounds that I got by looking at a curve that's a lower bound on the area and a curve that's an upper bound on the area and integrating, I discover that H_n is bracketed between the natural log of n plus 1 and 1 plus the natural log of n . Now these two numbers, $\log n + 1$ and $1 + \log n$, are very close, and they get closer and closer as n grows. So it turns out that what we can say pretty accurately is that H_n is asymptotically equal to $\log n$. It's approximately equal to $\log n$. And the precise definition of this tilde symbol that I've highlighted in magenta is called asymptotically equal.

And the general definition of asymptotically equal, that H_n is asymptotically equal to $\log n$, is that a function f of n is asymptotically equal to a function g of n when the limit of their quotient goes to 1. That is to say, as a multiplicative factor, each is within a constant one of the other in the limit. $1 + \epsilon$ of g_n is going to bracket f_n , and vice versa. Asymptotic equivalents, or asymptotic equality.

Let's do an example. So the remark would be, for example, that $n^2 + n$ is asymptotically equal to n^2 . Why? Well let's look at the limit as n approaches infinity of $n^2 + n$ over n^2 . It's the same as simplifying algebraically the limit of $1 + 1/n$. As n approaches infinity, that term goes to zero. Sure enough, the limit is 1, and so these two terms are asymptotically equal. The idea of asymptotically equal is that all we care about is the high order term. The low order terms will disappear, and we're looking at the principal term that controls the growth rate of the functions when we look at them up to asymptotic equivalence.

So let's step back for just a moment and generalize what we've done here with estimating the harmonic sum. There's a general method called the integral method where, in this particular case, suppose we have a function f , the positive real valued function that's weakly decreasing, like $1/x$. Then let's let S be the sum from $i = 1$ to n of $f(i)$. So I was interested where $f(x)$ was $1/x$, and I wanted the sum from 1 to n of $1/i$, which was the n th harmonic number. And I was comparing that to the integral from 1 to n of $f(x)$, or $1/x$ in our example.

So if I is the integral and S is the sum that I want, what we can conclude really is that, in general, the sum is bracketed between the integral plus the first term of f of 1 in the sum, and the integral plus the last term of the sum. Remember, f is weakly decreasing, so $f(n)$ is smaller than $f(1)$. There's a similar theorem actually, just reverse $f(1)$ and $f(n)$, to use an integral estimate to get a bound on a weakly increasing function. And that gives us a general

tool for estimating the growth rate of sums.