Excursions in Computing Science: Book 8d. Rocket Science. Appendix. Trigonometry and calculus.

T. H. Merrett*
McGill University, Montreal, Canada

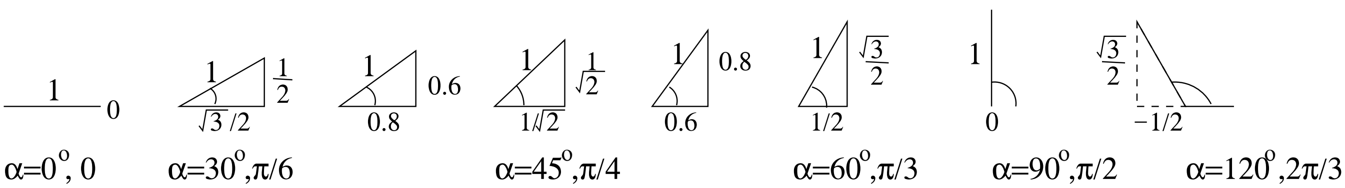
July 31, 2021

29. Microgravity.
30. Radiation.
31. Space debris.
32. Space elevator.
33. Ecology.
34. Population.
35. Genetics.
36. History.
37. Self-reproducing probes.
38. "Where Are They?"

Part IV Spaceship Earth.

39. Speeds.
40. Extinctions.
41. Herd science.
42. Climate.

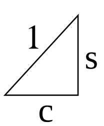
43. Trigonometry. Trigonometry (tri three, gon angle, metry measurement) is the study of right-angled triangles of hypotenuse 1. (The *hypotenuse* is the longest side of a right triangle.) Here are some examples.



The unfamiliar measure of angles, radians, gives us simple results when we come to calculus. The reason we normally measure angles in degrees is because of the divisibility of 360: 180, 120, 90, 72, 60, 45, 30 and 15 degrees are all simple fractions of a full turnaround: 1/2, 1/3, 1/4, 1/5, 1/6, 1/12 and 1/24. There is no reason we couldn't measure angles alternatively, simply as fractions of a full turnaround.

Radians are another measure of angles. A *radian* is the angle you get by marking off along the circumference of a circle the distance equal to the radius of the circle. There are 2π of these in a full turnaround. Hence the fractions of π shown above. In some sense, radians are the "natural" way to measure angles.

We can use letters to stand for *any* numbers. I'll use c and s for the base and height of the triangles, respectively. It's conventional to use Greek letters for angles, so I've used α, alpha, the Greek a, for any angle. Pi, π, the Greek p, of course stands for one particular angle: how many degrees?



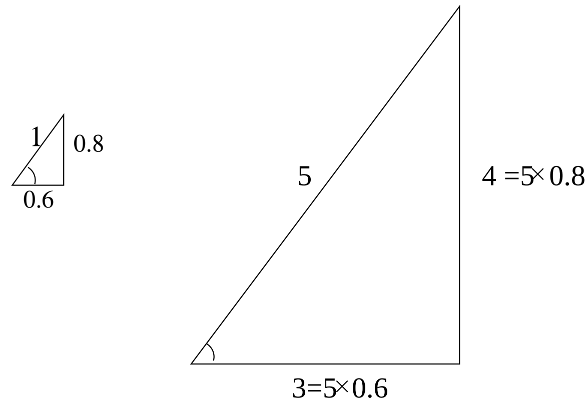
Angle (degrees)	0	30	45	60	90	120	135	150	···
Angle (radians)	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	$2\pi/3$	$3\pi/4$	$5\pi/6$	···
c	1	$\sqrt{3}/2 = 0.87$	$1/\sqrt{2} = 0.70$	$1/2$	0	$-1/2$	$-1/\sqrt{2}$	$-\sqrt{3}/2$	···
s	0	$1/2$	$1/\sqrt{2}$	$\sqrt{3}/2$	1	$\sqrt{3}/2$	$1/\sqrt{2}$	$1/2$	···

Evidently, c and s cannot stand for quite *any* number. There's a constraint that each must lie between -1 and 1. Indeed we can refine these constraints to a simple restriction.

$$c^2 + s^2 = 1$$

This follows from a result on right-angled triangles attributed to Pythagoras, that the square of the hypotenuse equals the sum of the squares of the other two sides.

The two triangles above for which I did not state the angles can be stretched to give right-angled triangles with integer sides. Here is one of them.

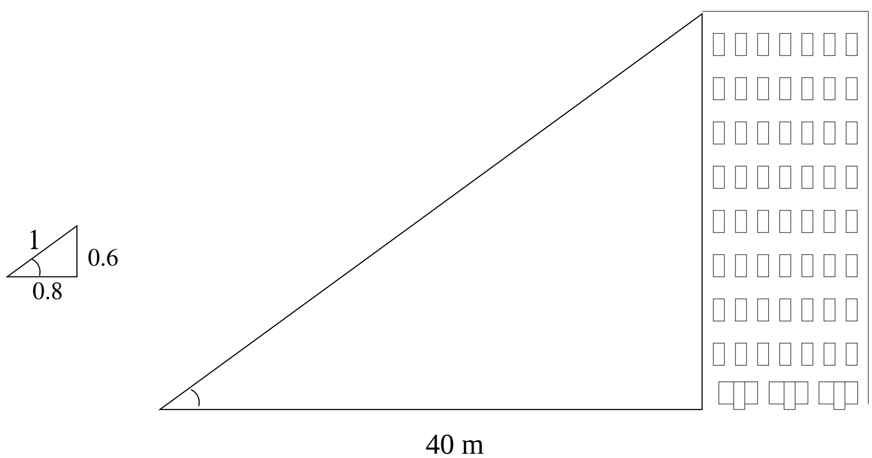


You can check that $3^2 + 4^2 = 5^2$ in the big triangle, and, more importantly, that

$$0.6^2 + 0.8^2 = 0.36 + 0.64 = 1 = \frac{9+16}{25} = \frac{9}{25} + \frac{16}{25} = \frac{3^2}{5^2} + \frac{4^2}{5^2} = \left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2$$

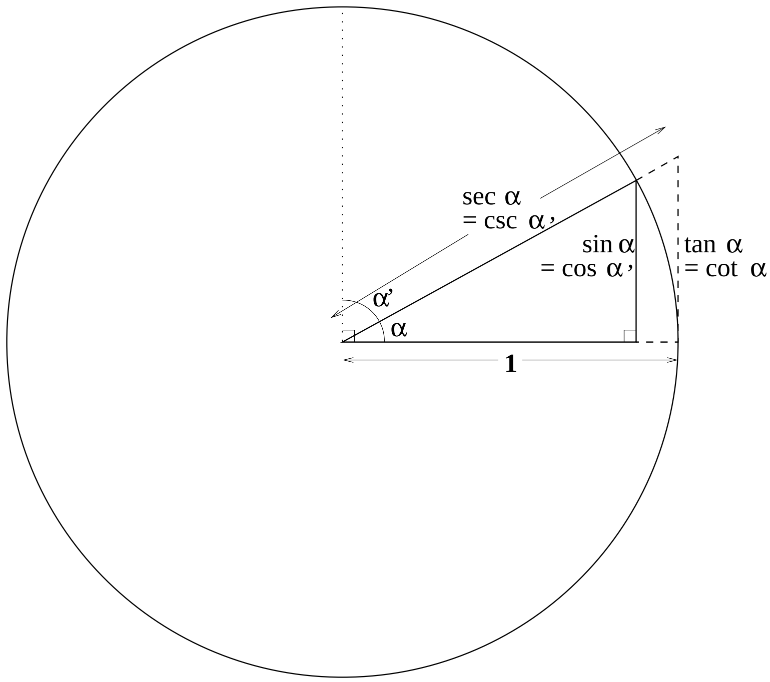
What makes trig really useful is this fact that the hypotenuse-1 right triangles can be stretched to any size. The c and the s need only then be multiplied by the same number, which is the size of the new hypotenuse. Or, the *ratio* of the horizontal and vertical sides will be the same as the ratio of c and s.

So we can measure the height of a building by knowing how far we are away from it and the angle to the top.



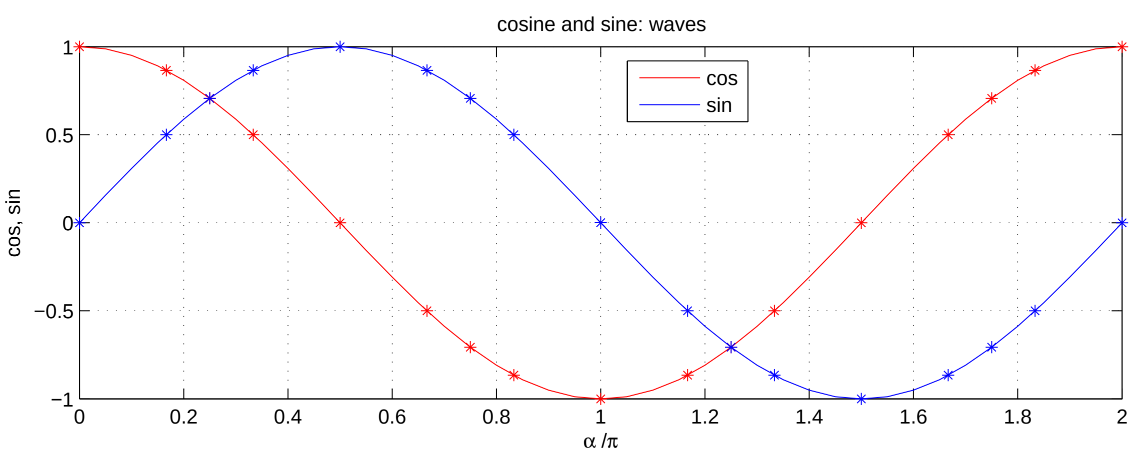
If the triangle has base 40 meters and the same angle as the (1,0.8,0.6) trigonometric triangle (which is 36.87 degrees) then we can conclude that the building is 30 meters tall.

(Now you know what all those engineering students are up to on the university campus in the spring surveying course. Well, there are side-to-side angles as well as up-and-down angles to be measured, too.)



(The c and s stand for "cosine" and "sine" which are trigonometric functions which depend on the angle α. The "co" in "cosine" means apply the sine to the complementary angle, $\alpha' = 90 - \alpha$, or $\alpha' = \pi/2 - \alpha$. Trigonometry has four other functions which depend simply on sine and cosine. The above diagram essentially captures all of trigonometry.)

Returning to cosine and sine, we can make a picture.



Here, the asterisks mark the values I gave in the table above for $\alpha = 0,\ \pi/6,\ \pi/4,\ \pi/3,\ \pi/2$ and their extensions all the way up to 2π.

The lines connect the asterisks in the smoothest possible way, without kinks, and so show the values of c and s for all possible values of α between 0 and 2π. Notice that they go negative in places, and that they do not go beyond ± 1.

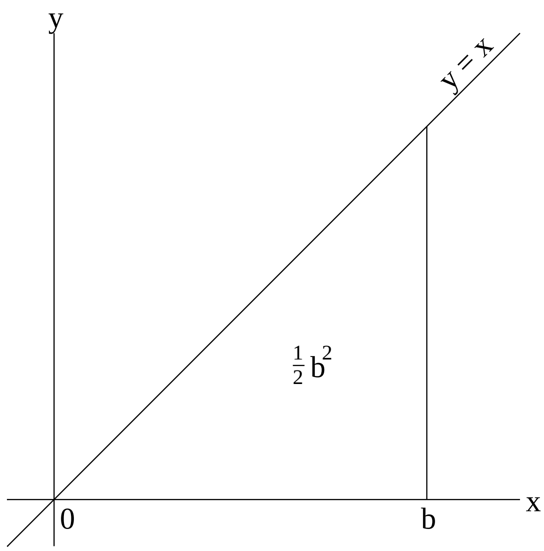
These curves are called "sine waves"—we see they are the same curve, just one displaced from the other by $\pi/2$ radians (90 degrees)—but here called cosine and sine waves to distinguish them. Evidently the curves repeat beyond 2π (and, in the opposite direction, below 0): they are *periodic*.

The footnote to Note 8, Part II mentions trig in the context of ellipses. Note 9, Part II: use trig to express x and y in terms of r and an angle. Pythagoras' theorem is proved in Note 10, Part II and used in Note 27 of Part II to find distances from Moon to Sun and Earth. Radians are used in Note 11, Part II. Note 26, Part II discusses components of force.

44. Integral calculus. What the integral calculus essentially does is to find the area under a curve. For instance, we might want to calculate the area under the sine curve in the previous Note.

Let's start with something simpler: the "curve" that is the straight line at 45 degrees. Drawn in an x-y diagram—a "Cartesian plane"—this is the line $y = x$. That is, it is made up of all possible points such as

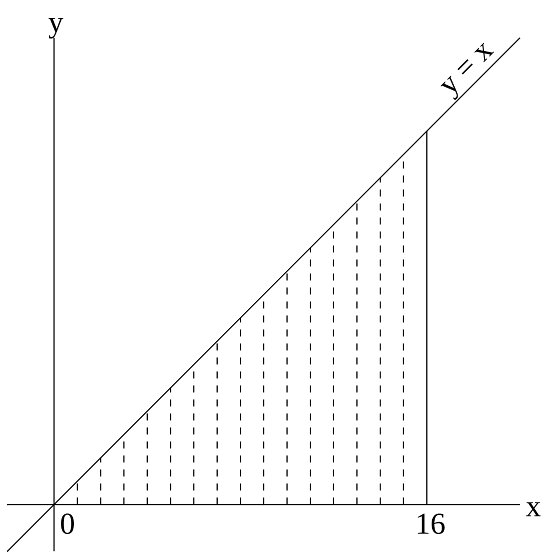
x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	$\cdots$
y	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	$\cdots$



Now we know the area under $y = x$ from $x = 0$ to $x = b$. It is half the base, b, times the height, which in this case is also b: $b^2/2$. (That's because if we doubled the area by completing the square, that would give b^2.)

But if the curve were more complicated, we may not be able to figure it out so easily. So let's look at a new way of calculating the area, one which will work for any shape of curve.

To be really clear, let's suppose b is a particular number, say 16.



This turns the area into the sum of the heights of the dashed lines

$$1+\ 2+\ 3+\ 4+\ 5+\ 6+\ 7+\ 8+\ 9+10+11+12+13+14+15+16$$

or

$$16+15+14+13+12+11+10+\ 9+\ 8+\ 7+\ 6+\ 5+\ 4+\ 3+\ 2+\ 1$$

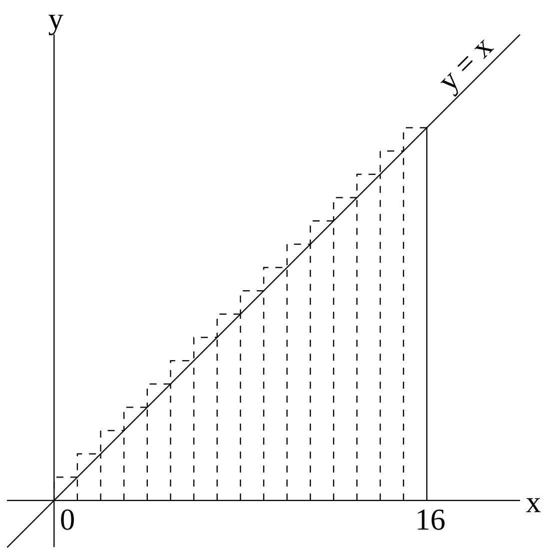
The reason I've written it down the second time—it's the same sum; it doesn't matter if we add them up backwards—is just like the trick of doubling the triangle to get a square.

That trick gives us twice the answer we need, which is

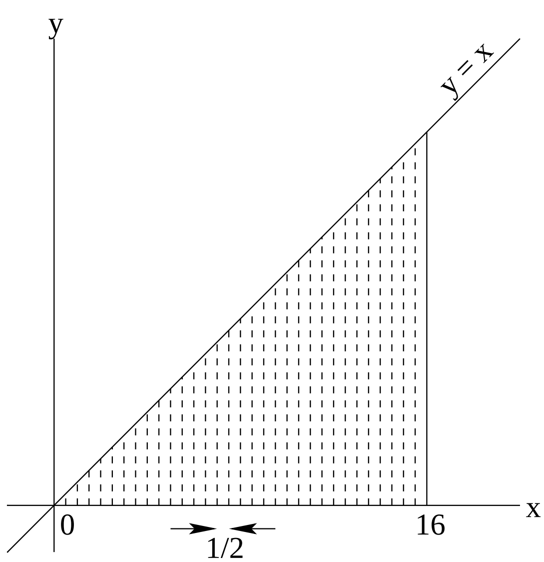
$$(17+17+17+17+17+17+17+17+17+17+17+17+17+17+17+17)/2 = 16\times 17/2 = 8\times 17$$

That result is pretty close, but just a little bigger than $16^2/2 = 8\times 16$ which is what it should be.

That's because what we've actually added up is not the area under $y = x$ but the area under a "staircase" on top of that line.



So we need to refine our method. We'll cut the strips in half, making a finer staircase. (I won't show the stairs.)



This time we must add all the integers up to 32. We must also multiply the sum we get by 1/2 because that is the width of each of the new strips. Not only that but we must multiply by a further 1/2 because we are only adding up to $x = 16$, not 32, so we must halve 32 to 16, and so for each of the heights of the strips.

You have seen how to add up $1+\cdots+32$. It is $32 \times 33/2 = 16 \times 33$. Do you begin to see why I avoided actually doing any arithmetic above? Out of the "unarithmetic" emerged a pattern which we can now use to add up to any number.

The area under this refined staircase is a quarter of that, 4×33. This is indeed closer to 8×16 than 8×17 was. You can do the arithmetic if you like, or just compare

$$8 \times 16 = 4 \times 32 < 4 \times 33 < 4 \times 34 = 8 \times 17$$

So we see the way. Next we try 64 strips, dividing the sum by a further 4, to get $64 \times 65/2/4/4 = 2 \times 65$ which is even closer to 2×64 than 2×66 was.

And so on. Clearly the process *converges* on $8 \times 16 = 16^2/2$.

Or, for the area from 0 to *any* b, $b^2/2$.

If we wanted the area from a to b we would just subtract the 0-to-a area from the 0-to-b area to get $(b^2 - a^2)/2$. This works for negative a too, but note that the area "under" $y = x$ where it goes *below* the x-axis is *negative*. So the area from $-a$ to 0 *cancels* the corresponding area from 0 to a, and the result is still $(b^2 - a^2)/2$.

Let's go to a real curve, as opposed to the straight line we've been calling a "curve". How about the area under the curve $y = x^2$, since we've been seeing squares in $b^2/2$? This curve is a *parabola* and you should see what it looks like by plotting, say,

x	-4	-3	-2	-1	0	1	2	3	4
y	16	9	4	1	0	1	4	9	16

We can build on what we have been doing. x^2, if x were a whole number, is double the sum of all the whole numbers from 0 to x. So let's add up the *sums of numbers* to get approximations for the area under $x^2/2$. We can double that later for the area under x^2.

Number b	1	2	3	4	5	$\cdots$	
Sum to b	1	3	6	10	15	$\cdots$	"triangular numbers"
Sum of sums to b	1	4	10	20	35	$\cdots$	"tetrahedral numbers"
$b(b+1)/2$	1	3	6	10	15	$\cdots$	
$b(b+1)(b+2)/6$	1	4	10	20	35	$\cdots$	

There is quite a lot of thinking in this table. Can you see how the second line builds up by just adding its previous number to the number above in the first line? And how the third line builds in the same way on the second line?

Notice that the fourth line is the same as the second. That is because the formula $b(b+1)/2$ captures what we got above when we added all the numbers up to b.

And the fifth line is the same as the third. This time I've guessed at a formula, $b(b+1)(b+2)/6$, which extends the pattern in $b(b+1)/2$ and seems to work. That is not a proof, but let's see where it gets us.

Doubling that guessed formula gives $b(b+1)(b+2)/3$. Just as $b(b+1)/2$ turned out to approximate $b^2/2$ above, we might suppose that $b(b+1)(b+2)/3$ approximates $b^3/3$.

So I am going to say that the area under curve $y = x^2$ from 0 to b is $b^3/3$.

Then the area under $y = x^2$ from a to b is $(b^3 - a^3)/3$. Note that, this time, if a is negative we get a bigger result. That's because the area under the negative side of the parabola is positive and so increases the sum.

There is a pattern in this $b^2/2$ and $b^3/3$. We can extend it (again, not a proof, but proof can be given) to areas under curves given by any power. This is the beginning of a *table of integrals*.

y	x^0	x^1	x^2	x^3	x^4	$\cdots$
Area from 0 to b	b	$b^2/2$	$b^3/3$	$b^4/4$	$b^5/5$	$\cdots$

This can be generalized to a rule which is worth memorizing. The area from 0 to b under curve $y = x^n$ is

$$\int_0^b dx\ x^n = \frac{b^{n+1}}{n+1}$$

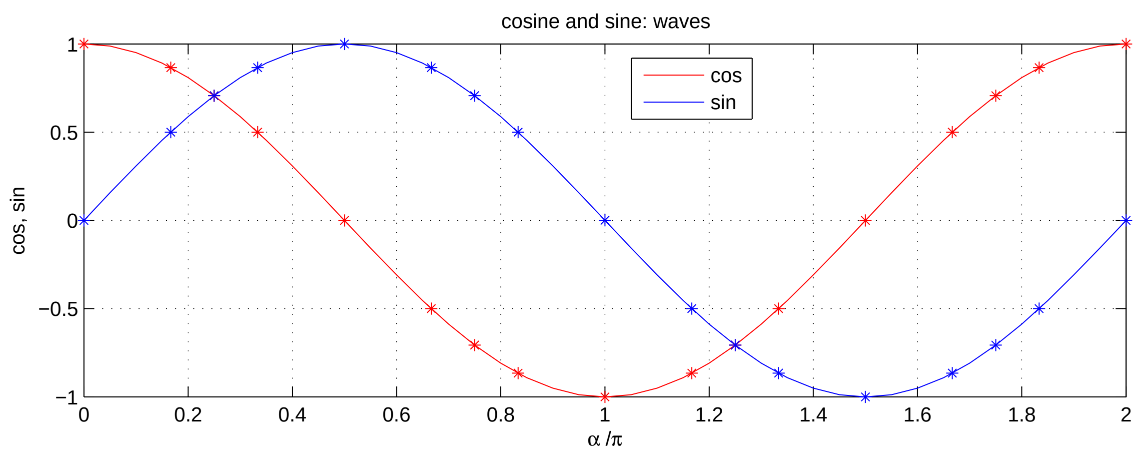
I've given the proper mathematical notation so that you'll know what it is next time you see it. (The dx is the width of the strips, conceived of as arbitrarily small, which must be multiplied by the "height" of the curve to get an area. The $\int$ indicates the sum over all the strips.)

It works for any n, not just non-negative integers, except, obviously, for $n = -1$.

We started this Note by wondering about the area under the sine curve. We can give a pictorial argument for the cosine curve, that

$$\int_0^b d\alpha\ \cos(\alpha) = \sin(b)$$

if the angle α is measured in radians. (Otherwise the scale of the strip widths is changed and the result is multiplied by this scale.)



In the red plot of $\cos(\alpha)$ we must visualize b as moving from left to right, 0 to 2π. When b is at 0 on the extreme left, the area from 0 to b under cos() is 0, which is what the blue sin() curve is. As b moves rightwards the area grows—in constant steps at first, because cos() is 1 or almost for a while, but then in decreasing steps until it levels off and does not increase. That's what sin() does, approaching and at its maximum.

After this, cos() goes negative so the area decreases, as does sin(), reaching 0 at the minimum of cos() at $b = \pi$. cos() remains negative until $b = 3\pi/2$ so the area goes negative and reaches a minimum there. Then the area climbs back to 0 as the now positive cos() overcomes the negative minimum.

It is trickier to argue that

$$\int_0^b d\alpha\ \sin(\alpha) = 1 - \cos(b)$$

but looking just at the values for $b = \pi$ shows us that the area under the blue curve must reach a maximum just before the blue curve goes negative at π. $-\cos()$, the mirror image of $\cos()$ in the horizontal axis, would have a maximum there.

The 1 stops the area under the blue curve from being negative at 0 and 2π, which is what the flipped red ($\cos()$) curve would give us.

This is an almost symmetric exchange between $\sin()$ and $\cos()$. We can bring out the symmetry by taking the endpoints off the integral sign and saying.

$$\begin{aligned} \int d\alpha\ \cos(\alpha) &= \sin(\alpha) + \text{const} \\ \int d\alpha\ \sin(\alpha) &= -\cos(\alpha) + \text{const} \end{aligned}$$

and so

$$\begin{aligned} \int_a^b d\alpha\ \cos(\alpha) &= \sin b - \sin a \\ \int_a^b d\alpha\ \sin(\alpha) &= -(\cos b - \cos a) \end{aligned}$$

This gives us the above since $\sin(0) = 0$ and $\cos(0) = 1$.

Note that I've added arbitrary constants to the two indefinite integrals, because this is what is done, and the constants cancel once we put limits back on.

Now you know how to find areas under power curves x^n and under the two trigonometric waves $\sin()$ and $\cos()$.

The reference to calculus in Note 6 of Part I is the special case

$$\int dx\ x^{-2} = -x^{-1}$$

as is part of the sum of the pieces of the space elevator in Note 32 of Part III.

Note 18 of Part II uses the same relationship but reversed to go from gravitational potential energy $-GM/r$ to gravitational force GM/r^2

$$\partial(-r^{-1}) = r^{-2}$$

Note 32 of Part III also uses

$$\int dx\ x = \frac{x^2}{2}$$

But the reference in Note 1 of Part I lands on the one exception, which is

$$\int dx\ x^{-1} = \ln(x)$$

and which is best approached through ∂ rather than $\int$, and through the inverse of the logarithm, $\ln()$, which is the exponential, $\exp()$, for which

$$\partial\ \exp(x) = \exp(x)$$

In Note 26 of Part II, calculating the angular momentum of a sphere (the Earth) adds up r^4, effectively, from 0 to R, and you see the $R^5/5$ in the answer.

45. Differential calculus. The integrals we've learned about can be thought of as converting one curve to another. Can these conversions be reversed?

This is a table of the integrals we know how to do (without the arbitrary constants)

$y(x)$	1	x	x^2	x^3	x^n	$\cos(x)$	$\sin(x)$
$\int dx\, y(x)$	x	$x^2/2$	$x^3/3$	$x^4/4$	$x^{n+1}/(n+1)$	$\sin(x)$	$-\cos(x)$

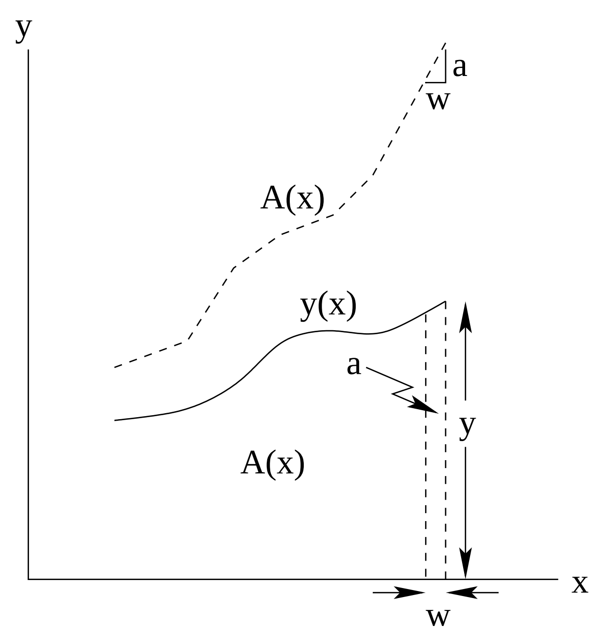
This is how we might reverse it

$y(x)$	1	x	x^2	x^3	x^n	$\cos(x)$	$\sin(x)$
$\partial y(x)$	0	1	$2x$	$3x^2$	nx^{n-1}	$-\sin(x)$	$\cos(x)$

Check this. Can you figure out the exponents and coefficients?

∂ is pronounced "differential" or, more suggestively, "slope". It tells how steep a curve is at any point. In the cosine and sine plot you can see that the blue, sin(), curve is steepest at $\alpha = 0$, and so the red, cos(), curve is highest. cos() goes horizontal at $\alpha = \pi/2$ and so sin() is zero there. And so on.

Since we've constructed *slope* as the inverse of *integral*, we can get a formal definition of slope.



For a function $y(x)$, if $A(x)$ is the area under it, we have (apart from the arbitrary constant, which depends on where we start calculating area)

$$A(x) = \int dx\, y(x)$$

Conversely

$$y(x) = \partial A(x)$$

That is, $\partial \int dx\, y(x) = y(x)$ and $\int dx\, \partial A(x) = A(x)$.

In the figure, $A(x)$ is the area under $y(x)$. The small change in area shown in the right is $a = wy(x)$, with w being the small change in x which added the extra area.

So

$$y = \frac{a}{w}$$

is the slope of A. We can see that this is

$$\text{slope} = \frac{\text{rise}}{\text{run}}$$

which is the classical definition of slope. (The next time you see a warning road sign on a steep hill saying, say, "1 in 10", you can read it as "rise in run"; or if it says, instead, "10%", you can translate it.)

Connecting slopes with integrals is the "fundamental theorem of calculus", so we've come pretty far.

Note 18 of Part II uses the same relationship as Note 6 of Part I but reversed to go from gravitational potential energy $-GM/r$ to gravitational force GM/r^2

$$\partial(-r^{-1}) = r^{-2}$$