Excursions in Computing Science: Week v Milli-micro-nano-..math Part I

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

May 24, 2022

I. Prefatory Notes

1. The Excursion *Golden ratio* in Week ii gives the relationship

$$\phi^2 - \phi - 1 = 0$$

for ϕ, the "golden ratio". It also suggests that

$$\phi = \frac{1+\sqrt{5}}{2}$$

satisfies this relationship.

How would we *discover* such a value for ϕ? Is there any other value which would also satisfy the relationship?

Let's suppose we have no idea about what sort of number ϕ is. In fact let's rename ϕ to x, which traditionally stands for the unknown. We'll suppose that all we know is

$$x^2 - x - 1 = 0$$

and we'd like to "solve" this for x.

2. It would be nice if we could draw a picture of the relationship. The way we can do that is to see how

$$x^2 - x - 1$$

changes as x changes. From this we might get an idea of what value of x makes it zero,

It's handy to give a simple name to something we are trying to explore, so we'll use y.

$$y = x^2 - x - 1$$

This suggests a game: if I tell you what x is, you can in turn tell me what y must be.

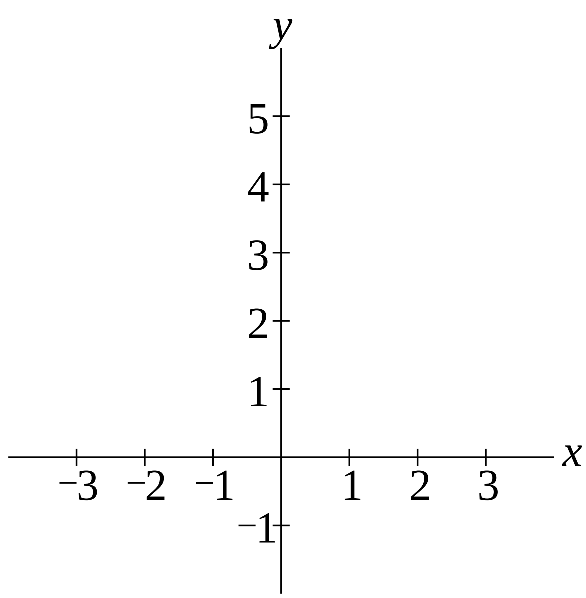
me	you
x	y
0	-1
1	-1
-1	1
2	1
-2	5
3	5
$\vdots$	$\vdots$

These are just some sample values. I could make life harder for you by saying $x = 3/2$, and you'd have to calculate $y = (3/2)^2 - 3/2 - 1 = 9/4 - 3/2 - 1 = -1/4$. And so on: there is no limit to what values I could choose for x.

So you should soon start looking for patterns to make your life easier and also, maybe, to check for mistakes in your calculations. (You can see some patterns already, which we'll come back to.)

3. Cartesian Plane. We mentioned drawing a picture, and that it would be a powerful way to see patterns.

So let's draw two lines: a horizontal one for x and a vertical one for y. Since x and y can be all sorts of different numbers, we'll also measure off distances along these lines corresponding to some of the possible numbers. The me-you, x-y table in the previous Note shows x ranging from -2 to 3 (not in order, but we'll put them in order now), so let's make ticks from -3 to 3 on the horizontal line. Similarly, we'll tick off the vertical line from -1 to 5.



This is called the Cartesian plane, named after Réné Descartes, who first thought of this way of picturing relationships. Note that x is positive to the right and negative to the left. And y is positive going up and negative going down. The point where the two lines meet is called the *origin* and corresponds to $x = 0$ and $y = 0$.

Note something tricky. The relationship $x = 0$, without saying anything about y, is true everywhere on the y-line. (The vertical line: we'll call it the "y-axis" from now on.)

Similarly, the relationship $y = 0$, saying nothing about x, is the x-axis: the horizontal line labelled x.

It is tricky to remember that $x = 0$ means y-axis and $y = 0$ means x-axis, but it makes sense.

If we get the hang of that, we can also say that $x = 1$ is the whole vertical line (not shown but you can imagine it) crossing the x-axis at $x = 1$.

And $y = 1$ is the invisible horizontal line crossing the y-axis at $y = 1$.

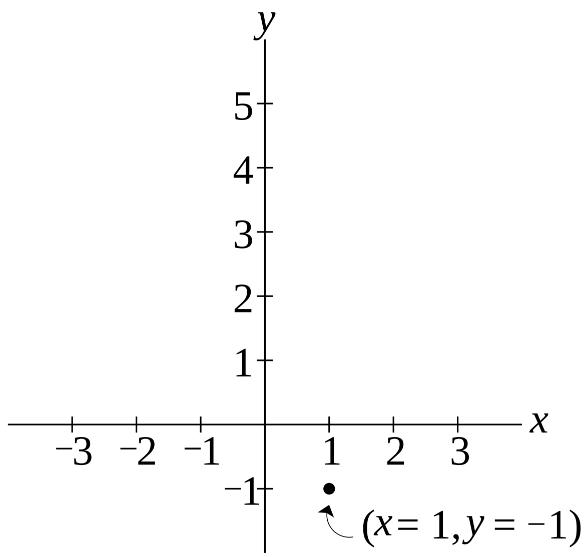
And so on, for any value of x or y.

How do we use the Cartesian plane to draw a picture of

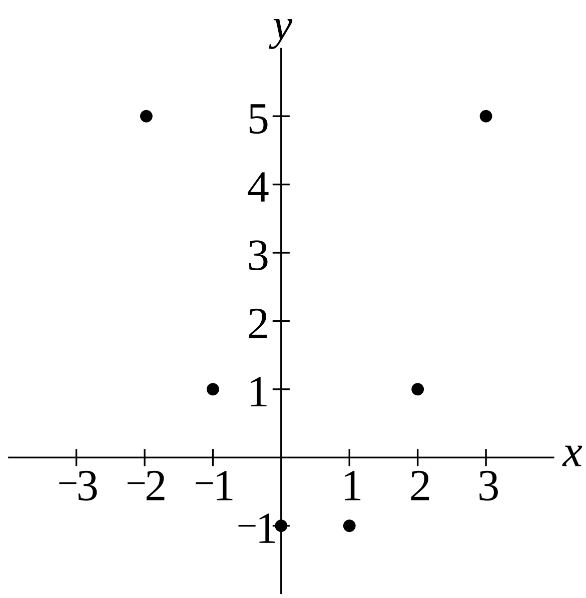
$$x^2 - x - 1?$$

Well, the me-you, x-y table of Note 2 shows *pairs* of values: x and y. The Cartesian plane shows that x-values are horizontal distances and y-values are vertical distances.

So here's what we do with, say, $x = 1 \quad y = -1$: go from the origin rightwards to $x = 1$ on the x-axis; then go downwards on the invisible $x = 1$ (vertical) line until you reach the invisible $y = -1$ (horizontal) line. Mark a point there.



Now do the same for all the other values of x shown and the corresponding values of y you've calculated.

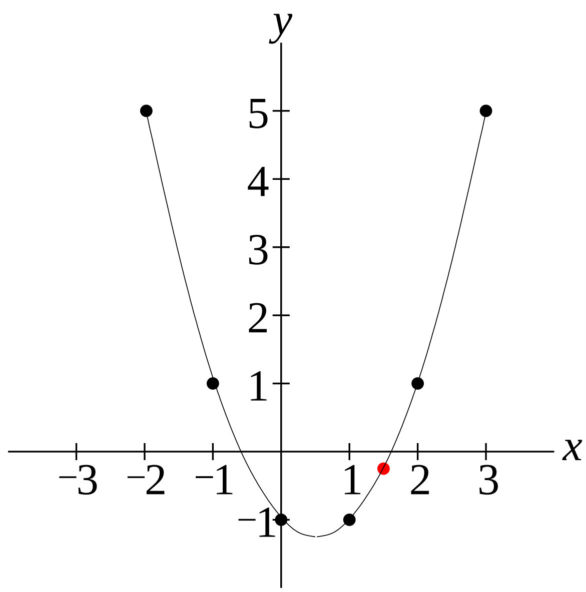


OK, that's the six pairs of numbers in the table. What about all the other possible pairs, such as $x = 2/3, \;\; y = -1/4$?

We could spend forever calculating all these because the possible values of x go on forever—without even going outside the range -2 to 3: $x = 1/16, \;\; x = 17/16, \;\; x = 97/128, \;\; ...$

So we'll see if we can just draw all these extra pairs. They will make an infinite number of dots so we'll just draw them all as one line.

We must be careful, though. We can't just connect the dots we've got with straight lines. To get all the in-between values more or less right we'll have to draw a careful curved line.



Let's see if we got $x = 3/2, \;\; y = -1/4$ right. It's the red dot. Pretty good. A straight line there would have put the point too high: at $x = 3/2, \;\; y = 0$ to be exact.

4. Symmetry. One useful clue about this particular curve is that it is *symmetrical*. We saw that already in the table in Note 2: $x = 0$ and $x = 1$ give the same value for y; so do $x = -1$ and $x = 2$; and so on.

This curve is symmetrical about a particular (invisible) vertical line. That is, it is symmetrical about a particular x-value. You should be able to see that that value is $x = 1/2$: halfway between $x = 0$ and $x = 1$, also halfway between $x = -1$ and $x = 2$, or between $x = -2$ and $x = 3$.

This is going to be helpful for us. We will come back to look at this symmetry in another way.

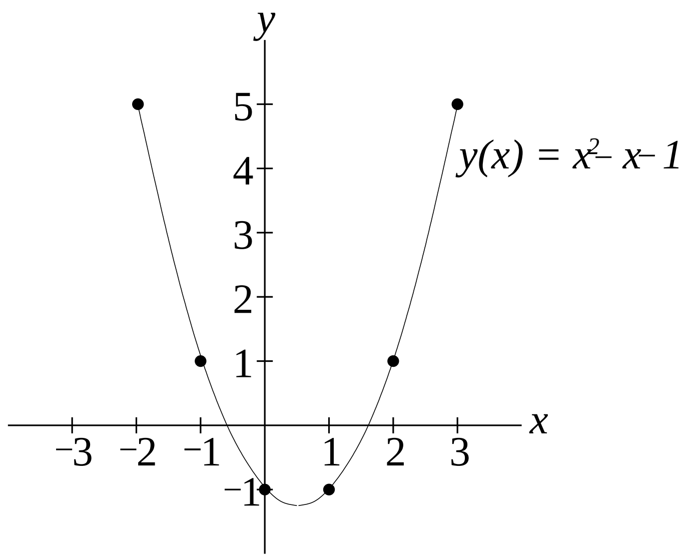
5. Functions. This relationship between x and y is called a *function*. It goes one way, like the game: from the x I give you, you calculate the y. Not the other way around.

We can write the relationship in a way which emphasizes the directional nature of this calculation:

$$y(x) = x^2 - x - 1$$

It says that y depends on x, not the other way around.

We can also label curves with this notation.



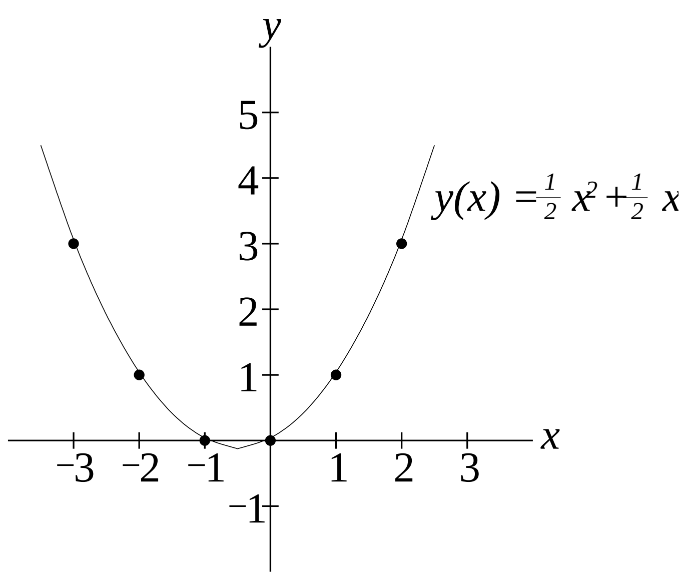
You see that we cannot play the game the other way around. Even using the table of Note 2, if I told you $y = 1$ you would have to give me *two* values for x, $x = -1$ and $x = 2$, not a unique one. And of course without the table, but with only $y = x^2 - x - 1$ we are really stuck for now.

But the problem of going backwards is what we're going to have to deal with if we want to find the value of x at which $y(x) = 0$, i.e., $x^2 - x - 1 = 0$.

This labelled curve is what we can call a *function*—a "function of x" to be more explicit.

What we've drawn is not the only function of x. Here's another one, based on the triangular numbers of Week i Note 1

$$y(x) = \frac{x(x+1)}{2} = \frac{1}{2}x^2 + \frac{1}{2}x$$



You should make the table for the points I've shown. You can also check that at $x = -1/2$, $y = -1/8$, so the curve really does go a little below the x-axis as shown.

(It could be confusing to call *both* of these functions $y(x)$, and other names are often invented, such as $f(x)$ and $g(x)$. But we'll be considering only one function at a time, so we'll stick with $y(x)$.)

6. Zeros of a function. Now let's go backwards. Instead of finding y given x, let's try to find (an) x given y. In particular, suppose we are given $y = 0$.

Solving

$$y(x) = 0$$

is called "finding the zeros of the function". (The other way around, going forwards to find $y(0)$ is too easy to rate a special name.)

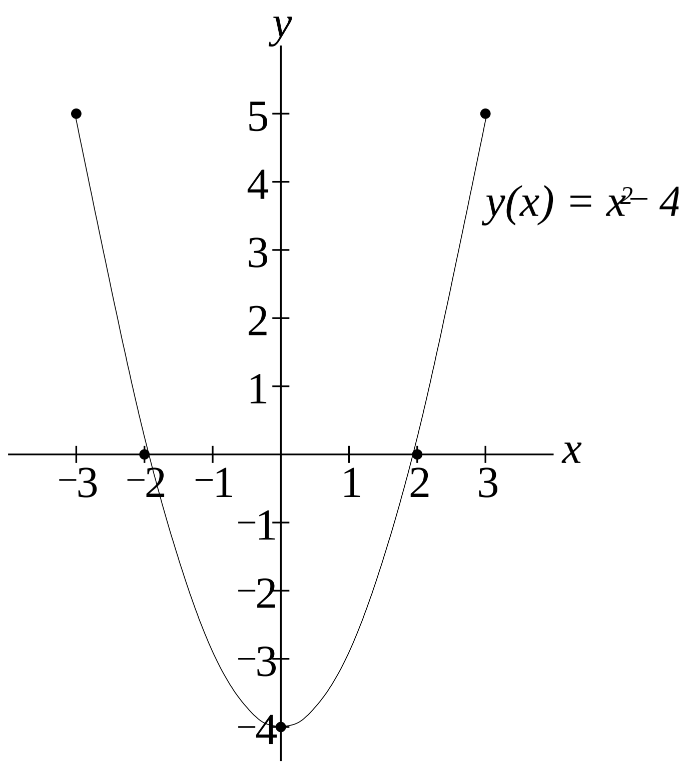
For the second function in Note 5, $y(x) = x(x + 1)/2$, this is easy. We can see that $x = 0$ gives $y = 0$ and so does $x = -1$. This is also on the drawing.

The zeros of the original function, $y(x) = x^2 - x - 1$, are not so obvious, although at least the drawing tells us that they are somewhere between $x = -1$ and $x = 0$, and between $x = 1$ and $x = 2$, where we see the curve crossing the x-axis.

Maybe we can use the symmetry of the curve to help.

7. Symmetry, again. Both functions above are symmetrical. The first is clearly symmetrical about $x = 1/2$ and the second is symmetrical about $x = -1/2$.

Let's look at a third function, this one symmetrical about $x = 0$.



This one clearly has zeros at $x = \pm 2$. The exact value of x is not the point. The point is that the two zeros are symmetrical about the line of symmetry, and when that line is $x = 0$, the function has only x^2 and constant terms in it: no x term.

When we note, for this function

$$x^2 - 4 = 0,$$

it follows immediately that

$$x = \pm 2$$

because 2 is the square root of 4, and there are always two square roots: $(+2)^2 = 4 = (-2)^2$.

So this encourages us to think that if we can make $x = 0$ the line of symmetry of any of our functions, say the original one, then finding the zeros will be easy.

Of course, we cannot, because the line of symmetry of $x^2 - x - 1$ is $x = 1/2$. But we can *change* x.

How about $x' = x - 1/2$? Then the line $x = 1/2$ is also the line $x' = 0$ and the line of symmetry is $x' = 0$.

So, if we convert $x^2 - x - 1$ into a function of x', will it leave us with an x'^2 term and a constant term but no x' term? If so, we will get x' as a simple pair of $\pm$ square roots.

Let's see. We need to turn $x' = x - 1/2$ around

$$x = x' + 1/2$$

Then

$$\begin{aligned} x^2 - x - 1 &= (x' + \frac{1}{2})^2 - (x' + \frac{1}{2}) - 1 \\ &= (x'^2 + x' + \frac{1}{4}) - (x' + \frac{1}{2}) - 1 \\ &= x'^2 - \frac{5}{4} \end{aligned}$$

Bingo! No x' term. So

$$x'^2 - \frac{5}{4} = 0$$

means

$$x' = \frac{\pm\sqrt{5}}{2}$$

and then

$$\begin{aligned} x &= x' + \frac{1}{2} \\ &= \frac{1}{2} \pm \frac{\sqrt{5}}{2} \\ &= \frac{1 \pm \sqrt{5}}{2} \end{aligned}$$

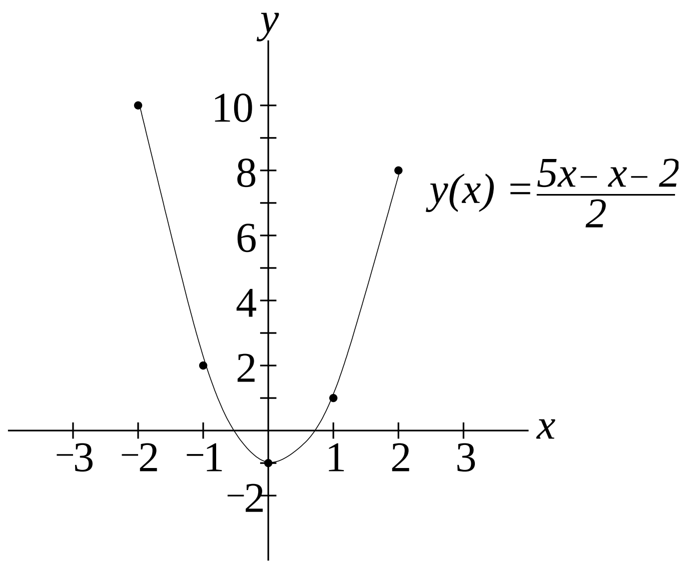
This is two solutions. One of them is the $(1+\sqrt{5})/2$ that the Excursion inWeek ii called the Golden Ratio, ϕ.

The other one is a second solution we didn't know about before: $(1 - \sqrt{5})/2$. if you put these two into a calculator, you'll find that $(1 - \sqrt{5})/2$ indeed lies between -1 and 0, while $(1 + \sqrt{5})/2$ lies between 1 and 2, as the curve for $y(x) = x^2 - x - 1$ told us..

So symmetry helps us a lot in finding the zeros of this kind of function.

8. Slopes. At last we come to the topic for this Week. What if we have a function whose symmetry is not obvious?

$$y(x) = \frac{5x^2 - x - 2}{2}$$



How do we find the line of symmetry?

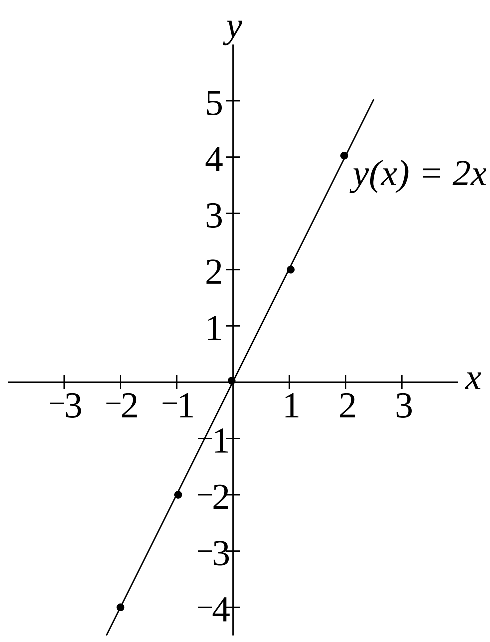
We use the fact that the curve is *horizontal* at the line of symmetry: it has to be, or it wouldn't be symmetrical.

A horizontal line has *slope* 0. What is "slope"?

Slope is defined as

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

Let's start with straight lines.

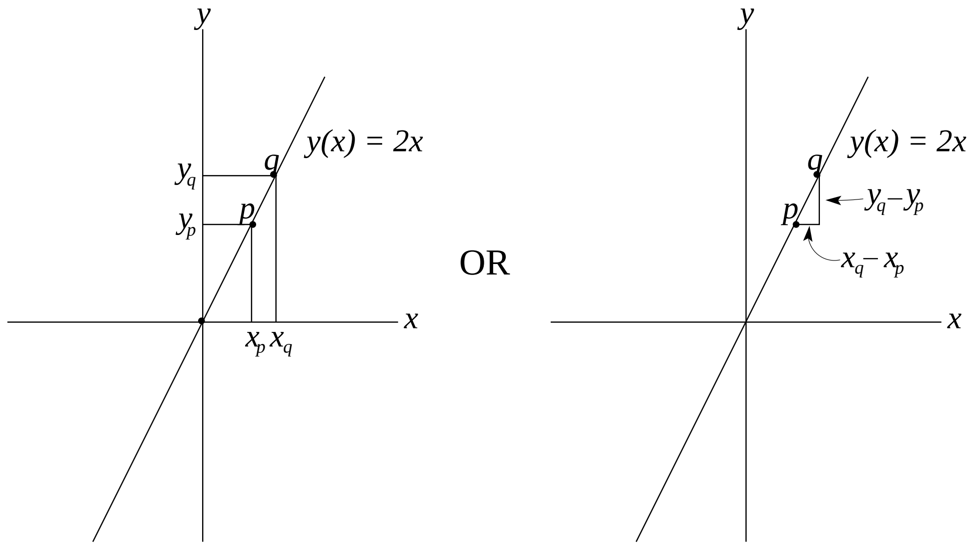


This is also a function, and you can make a table for it to show that $y(x) = 2x$ means $y = -2$ when $x = -1$, $y = 0$ when $x = 0$, $y = 2$ when $x = 1$, and so on.

Its slope is 2 because every time x increases by 1, y increases by 2. Here's the calculation.

$$\text{if} \quad x_q = 2 \quad \text{then } y_q = 4$$

$$\begin{aligned} \text{if} \quad x_p &= 1 \quad \text{then } y_p = 2 \\ x_q - x_p &= 1 \quad y_q - y_p = 2 \\ \frac{\text{rise}}{\text{run}} &= \frac{y_q - y_p}{x_q - x_p} = \frac{4}{2} = 2 \end{aligned}$$



What happens if $x_q = 3, \;\; x_p = 1$? If $x_q = 1, \;\; x_p = -1$? If $x_q = -1, \;\; x_p = -2$?

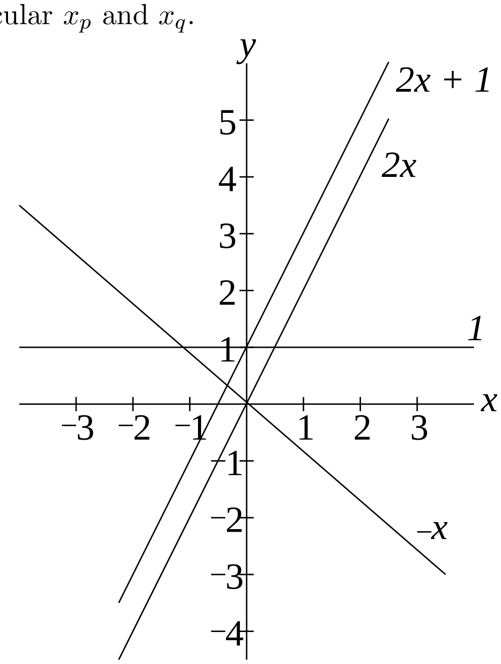
You should always find the slope is 2 for $y(x) = 2x$.

We write this

$$\text{slope}(2x) = 2$$

What about straight lines with other slopes? Here is a table, each line of which requires a calculation of $(y_q - y_p)/(x_q - x_p)$ having chosen some particular x_p and x_q.

$y(x)$	slope(y)
$2x$	2
$3x$	3
$0x$	0
$-1x$	-1
$2x + 1$	2
$2x - 1$	2
$3x + 1$	3
1	0



For example, $y(x) = 3x$: say $x_q = 1, \;\; x_p = 0$. Then $y_q = 3, \;\; y_p = 0$ and

$$\text{slope}(3x) = \frac{y_q - y_p}{x_q - x_p} = \frac{3}{1} = 3$$

Note that parallel lines have the same slope.

Note that horizontal lines have slope 0. This makes sense: they go neither up nor down.

Note that lines going downwards have negative slope.

These are all quite intuitive: positive slope goes up; negative slope goes down; zero slope is flat.

Note that it makes no difference to the slope if you add or subtract a constant to or from the function.

Let's show that slope$(ax + b) = a$ no matter what numbers a and b are.

We won't use particular numbers for x_p and x_q but instead use

$$
\begin{aligned}
y_q &= ax_q + b \\
y_p &= ax_p + b
\end{aligned}
$$

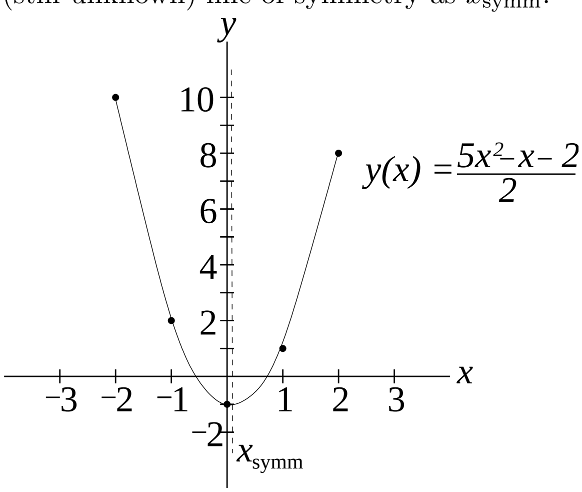
So

$$
\begin{aligned}
\text{slope}(ax + b) &= \frac{y_q - y_p}{x_q - x_p} \\
&= \frac{(ax_q + b) - (ax_p + b)}{x_q - x_p} \\
&= \frac{a(x_q - x_p) + b - b}{x_q - x_p} \\
&= a
\end{aligned}
$$

9. Slopes of curves. To find the line of symmetry of our U-shaped function we must go beyond straight lines to curves. The slope of a curve is not the same for every x but itself depends on x.

So the slope of a (curvy) function is itself a function.

Let's try to go directly from drawing $y(x) = (5x^2 - x - 2)/2$ to drawing its slope. We'll mark the (still unknown) line of symmetry as x_{symm}.



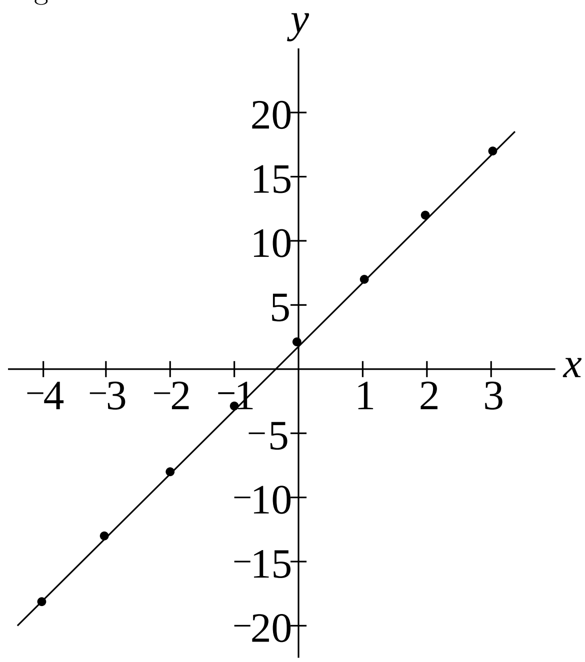
x	slope$\frac{5x^2-x-2}{2}$
$= x_{\text{symm}}$	0
$> x_{\text{symm}}$	$+$
$\gg x_{\text{symm}}$	$++$
$< x_{\text{symm}}$	$-$
$\ll x_{\text{symm}}$	$--$

You see we can't use exact numbers in this sketch. But we can say some things *qualitatively*:

- at the line of symmetry, $x = x_{\text{symm}}$, the slope is 0
- at $x > x_{\text{symm}}$ the slope is positive and increasing as x increases
- at $x < x_{\text{symm}}$ the slope is negative and decreasing as x decreases.

We might guess from this that the slope is a straight line. To *show* that it is we need some numbers.

x	$y = \frac{5x^2-x-2}{2}$	diff $y - y$	diff diff
4	37		
3	20	17	
2	8	12	5
1	1	7	5
0	-1	2	5
-1	2	-3	5
-2	10	-8	5
-3	23	-13	5
-4	41	-18	5



The third column subtracts adjacent values of y to get the rise at each value of x. The run is always 1 because I chose the x values to differ by 1. So plotting `diff` against x gives us the slope. As a function of x this slope is indeed a straight line.

Can we figure out what the equation of this line is? Call it

$$Ax + B$$

and note that when $x = 0 \quad Ax + B = 2$ so $B = 2$. (We're going to find a small mistake here, but let's not worry about it yet.)

To find A, we remember that A is the slope of $Ax + B$. So if we calculate, say

$$\frac{17 - 12}{3 - 2} = 5$$

we see that the slope of this slope at $x = 2$ is 5. But it's a straight line, so its slope everywhere = 5. So $A = 5$.

$$\text{slope}\frac{5x^2 - x - 2}{2} = 5x + 2 \qquad \text{[not quite right]}$$

Check it

x	$5x + 2$
3	17
2	12
1	7
0	2
-1	-3
-2	-8

The "diff diff" column in the above table is interesting: the numbers are all the same and they are the slope of the slope.

That's for $y(x) = (5x^2 - x - 2)/5$. We have other curves.

$y(x)$	slope(y)
$x^2 - x - 1$	$2x - 1$
$\frac{1}{2}x^2 - \frac{1}{2}x$	$x - \frac{1}{2}$
$x^2 - 4$	$2x$

Instead of working each one out, let's do it for any "coefficients" a, b, c:

$$\text{slope}(ax^2 + bx + c) = ?$$

Here I'm going to tell you something without proving it. (It can be proved directly from the definition, slope = rise/run.) Here it is:

$$\begin{aligned}\text{slope}(ax^2 + bx + c) &= a \times \text{slope}(x^2) + b \times \text{slope}(x) + c \times \text{slope}(1) \\ &= a \times \text{slope}(x^2) + b\end{aligned}$$

So we need only focus on slope(x^2).

$$\begin{aligned}\text{slope}(x^2) &= \frac{x_q^2 - x_p^2}{x_q - x_p} \\ &= \frac{(x_q + x_p)(x_q - x_p)}{x_q - x_p} \\ &= x_q + x_p\end{aligned}$$

What does this mean? We're going to have to do something we didn't think of yet—and incidentally find the mistake we made by saying above that $B = 2$.

When we got 17 for the slope of $(5x^2 - x - 2)/2$ at $x = 3$ we were a little sloppy: 17 is the slope but we are not sure if we should say at $x = 3$ or at $x = 4$ or somewhere in between.

Let's try halfway between: suppose that at $x = 3\frac{1}{2}$ the slope of $(5x^2 - x - 2)/2$ is 17. We'll keep the $A = 5$ but change B.

$$17 = Ax + B = 5 \times 3\frac{1}{2} + B = 17\frac{1}{2} + B$$

so $B = -\frac{1}{2}$.

10. Centering. This correction agrees with the rule I just gave you

$$\text{slope}(\frac{5}{2}x^2 - \frac{1}{2}x - 1) = \frac{5}{2}\,\text{slope}(x^2) - \frac{1}{2}$$

and if we stick with $A = 5$ we see that

$$\text{slope}(x^2) = 2x$$

But we got

$$\text{slope}(x^2) = x_q + x_p$$

just now, from the definition.

To reconcile these, we'll have to improve the definition of slope.

Here is the problem. For a curve, the slope depends exactly on x, not approximately. We see that if we shifted x from 3 to $3\frac{1}{2}$ we changed the slope of $(5x^2 - x - 2)/2$.

So what we must do is note that x_p and x_q are on opposite sides of and equidistant from their average value. This average value we'll call x, the location of the slope we're looking for.

So

$$x_p + x_q = 2x$$

And

$$\text{slope}(x^2) = 2x$$

Hence our table of slopes

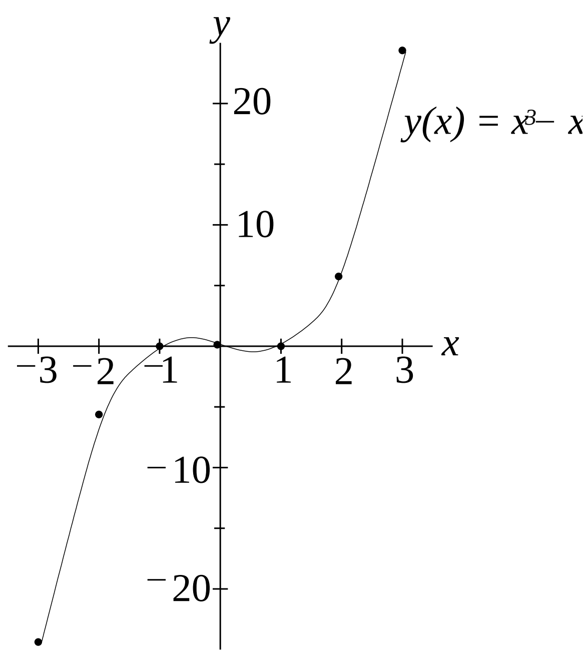
$y(x)$	slope(y)
$\frac{5}{2}x^2 - \frac{1}{2}x - 1$	$5x - \frac{1}{2}$
$x^2 - x - 1$	$2x - 1$
$\frac{1}{2}x^2 - \frac{1}{2}x$	$x - \frac{1}{2}$
$x^2 - 4$	$2x$
$ax^2 + bx + c$	$2ax + b$

11. Limits. If we look at powers of x, we can begin to see a pattern in their slopes.

n	x	slope(x^n)
0	1	0
1	x	1
2	x^2	$2x$
3	x^3	$3x^2$
4	x^4	$4x^3$
:	:	:
n	x^n	nx^{n-1}

We've already shown $n = 0$, 1 and 2. For $n = 3$ we can go back to drawings for a qualitative glimpse.

x	x^3	$x^3 - x$
3	27	24
2	8	6
1	1	0
0	0	0
−1	−1	0
−2	−8	−6
−3	−27	−24



I've drawn $y(x) = x^3 - x$ because it gives a better picture than $Y(x) = x^3$. If you calculate it for a few more numbers, especially $x = -1/2$ and $x = 1/2$, you'll see that it does indeed cross $y = 0$ (the x-axis) at the three points shown.

On this drawing, where is the slope = 0? You should find two places, between $x = -1$ and $x = 0$, and between $x = 0$ and $x = 1$. We won't say $x = -1/2$ and $x = 1/2$ because that's not right.

Instead we will call them $x_{\max}$ and $x_{\min}$—not because they are where $y(x)$ is the highest or lowest but because they do give a *local* maximum and a *local* minimum for $y(x)$.

Here's the qualitative table

x	slope$(x^3 - x)$
$x \ll x_{\max}$	$++$
$x < x_{\max}$	$+$
$x = x_{\max}$	0
$x_{\max} < x < x_{\min}$	$-$
$x = x_{\min}$	0
$x > x_{\min}$	$+$
$x \gg x_{\min}$	$++$

This could be one of our U-shaped curves. Let's use the definition—back to just x^3.

$$\begin{aligned}\text{slope}_{x=\frac{x_q+x_p}{2}}(x^3) &= \frac{x_q^3 - x_p^3}{x_q - x_p}\\ &= x_q^2 + x_q x_p + x_p^2\end{aligned}$$

Hint: $(x_q - x_p)(x_q^2 + x_q x_p + x_p^2) = x_q^3 - x_p^3$ using the symbol multiplication table

x	x_q^2	$x_q x_p$	x_p^2
x_q	x_q^3	$x_q^2 x_p$	$x_q x_p^2$
$-x_p$	$-x_q^2 x_p$	$-x_q x_p^2$	$-x_p^3$

But this time, just looking at the average does not work: we cannot simplify $x_q^2 + x_q x_p + x_p^2$ down to just x using $x = (x_q + x_p)/2$.

So we must improve the definition once again. (This improvement can replace the previous improvement of simply taking the average of x_q and x_p.)

We'll now take the definition *in the limit* as x_q and x_p both *approach* x.

$$\text{slope}_x(y(x)) = \lim_{x_q \to x, x_p \to x} \frac{y(x_q) - y(x_p)}{x_q - x_p}$$

We can see that this is tricky: the denominator approaches zero when both x_q and x_p approach the common value x.

Fortunately so far the denominator has cancelled out.

$$\text{slope}_x(x^3) = \lim_{x_q \to x, x_p \to x} (x_q^2 + x_q x_p + x_p^2) = 3x^2$$

(And also

$$\text{slope}_x(x^2) = \lim_{x_q \to x, x_p \to x} (x_q + x_p) = 2x.)$$

(Once we have x_q and x_p approaching the same thing, we will get the same result if we just, say, let x_q approach x_p: lopsided but it won't matter.

We can write this directly in terms of x if we say

$$\begin{aligned}x_p &= x\\ x_q &= x + \Delta x\end{aligned}$$

where Δx is a small increase in x which we'll take to the limit $\Delta x \to 0$

$$\text{slope}(y(x)) = \lim_{\Delta x \to 0} \frac{y(x+\Delta x) - y(x)}{x + \Delta x - x}$$

This is the standard definition of the slope.)

12. Back to the zeros of $\frac{5}{2}x^2 - \frac{1}{2}x - 2$. For cubic curves (involving up to x^3) and higher we must worry about limits. (Limits work also just as well for powers 0, 1 and 2.) For the quadratic curve (involving up to x^2), averages will do.

Anyhow, we know

$$\text{slope}_x(\frac{5}{2}x^2 - \frac{1}{2}x - 1) = 5x - \frac{1}{2}$$

So we can find out where it is zero.

$$\begin{aligned} 5x - \frac{1}{2} &= 0 \\ x &= \frac{1}{10} \end{aligned}$$

This is the line of symmetry of $\frac{5}{2}x^2 - \frac{1}{2}x - 1$.

So we can find its zeros by shifting x to $x' = x - \frac{1}{10}$ and later shifting back again.

$$\begin{aligned} 0 &= \frac{5}{2}x^2 - \frac{1}{2}x - 1 \\ &= \frac{5}{2}(x' + \frac{1}{10})^2 - \frac{1}{2}(x' + \frac{1}{10}) - 1 \\ &= \frac{5}{2}(x'^2 + \frac{1}{5}x' + \frac{1}{100}) - \frac{1}{2}(x' + \frac{1}{10}) - 1 \\ &= \frac{5}{2}x'^2 + \frac{1}{40} - \frac{1}{20} - 1 \\ &= \frac{5}{2}(x'^2 - \frac{2}{5}(1 + \frac{1}{40})) \\ x'^2 &= \frac{2}{5}(\frac{40+1}{40}) \\ x' &= \pm\sqrt{\frac{2}{5}\frac{41}{40}} \end{aligned}$$

So

$$\begin{aligned} x &= x' + \frac{1}{10} \\ &= \frac{1}{10} \pm \sqrt{\frac{2}{5}\frac{41}{40}} \end{aligned}$$

You can check with a calculator what these two numbers are and how they fit the plot back in Notes 8 and 9.

We've found the zeros of two of our U-shaped quadratic curves, here and in Note 7. Let's do it once more for *any* coefficients a, b and c.

1) $\qquad 0 = \text{slope}_x(ax^2 + bx + c) = 2ax + b$
gives

$$x = -\frac{b}{2a}$$

as the line of symmetry.
2) So shifting x to $x' = x + \frac{b}{2a}$

$$
\begin{aligned}
0 &= ax^2 + bx + c \\
&= a(x' - \frac{b}{2a})^2 + b(x' - \frac{b}{2a}) + c \\
&= a(x'^2 - \frac{b}{a}x' - \frac{b^2}{4a^2}) + b(x' - \frac{b}{2a}) + c \\
&= a(x'^2 + \frac{b^2}{4a^2} - \frac{b^2}{2a^2} + \frac{c}{a}) \\
&= a(x'^2 - (\frac{b^2}{4a^2} - \frac{4ac}{4a^2}))
\end{aligned}
$$

So

$$
\begin{aligned}
x' &= \pm\sqrt{\frac{b^2 - 4ac}{4a^2}} \\
&= \frac{\pm\sqrt{b^2 - 4ac}}{2a}
\end{aligned}
$$

3) Shift back again

$$
\begin{aligned}
x &= x' - \frac{b}{2a} \\
&= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\end{aligned}
$$

This is the *quadratic formula* which will find the zeros of any quadratic $ax^2 + bx + c$. You'll use this a lot in high school, but they never will have told you *why* it works.

13. Square roots. Another thing you will probably not learn in school is how to find a square root. Even if you only ever do it with a calculator, you might wonder how the calculator does it. We need slopes for this, too.

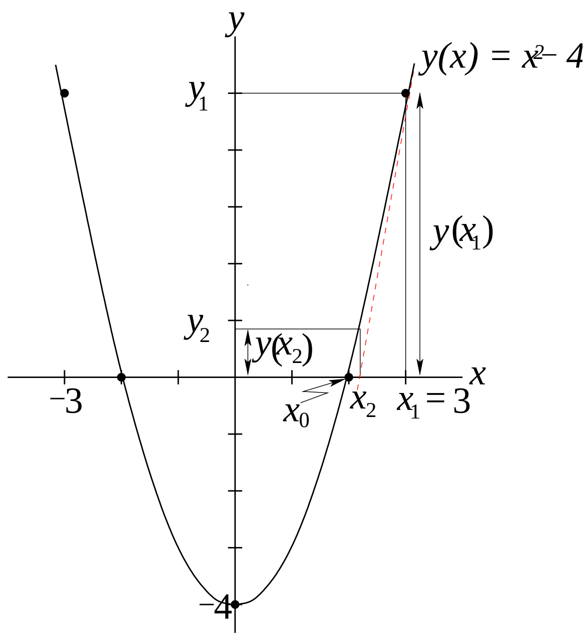
Let's find one we know already, just so we can see that the method works: $\sqrt{4}$.

We can make use of what we've just learned by asking: is there a function whose zeros give $x = \pm\sqrt{4}$? Yes:

$$y(x) = x^2 - 4$$

Now we're pretending that we don't know that $\sqrt{4} = 2$, but we have a picture of this function (Note 7) and we can take a *guess*.

Let's try $x = 3$. We'll call it x_1. And we'll call x_0 the zero we're looking for, $x_0^2 = 4$.



Since we know $y(x) = x^2 - 4$, we also know its slope.

$$\text{slope}_x y = 2x$$

So we know two more things about x_1 (the first was its value, e.g., 3),

- $y(x_1) = x_1^2 - 4$
- $\text{slope}_{x=x_1} y = 2x_1$

The slope of y at x_1 is also the slope (everywhere) of the straight line that is *tangent* to $y(x)$ at $x = x_1$. That is, the tangent *touches* $y(x)$ at $x = x_1$ ("tangere" is Latin for "to touch") but does not cross it: both lines are pointing in exactly the same direction at $x = x_1$.

So we draw that tangent, and call $x = x_2$ where it crosses the x-axis (red dashed line).

Clearly x_2 is nearer to x_0 than x_1 was: it is an improved guess. Can we find x_2 from the three things we know about x_1?

From the definition of slope, at $x = x_1$ the slope of $y(x) = x^2 - 4$ is

$$\text{slope}_{x=x_1} y(x) = \frac{y_1 - y_2}{x_1 - x_2} \approx \frac{y_1}{x_1 - x_2}$$

because y_2 is pretty close to 0, and would be exactly zero if the x_2 guess happened to be right. So the $\approx$ is an optimistic (hopeful) approximation.

Assuming this approximation, we can do some rearranging. We'll call $\text{slope}_x y = y'(x)$ to make things shorter.

$$\begin{aligned}
y'(x) &= \frac{y(x_1)}{x_1 - x_2} \\
x_1 - x_2 &= \frac{y(x_1)}{y'(x)} \\
x_2 &= x_1 - \frac{y(x_1)}{y'(x)} \\
&= x_1 - \frac{x_1^2 - 4}{2x_1}
\end{aligned}$$

There we go: if we know x_1 we can calculate x_2. For example, x_1 is our guess of 3:

$$x_2 = 3 - \frac{9-4}{2\times 3} = 3 - \frac{5}{6} = 2\frac{1}{6} = 2.1\overline{6}$$

This is not exactly $\sqrt{4}$ but it is closer than 3 was.

So we do it all over again, stepping from x_2 to an even better guess x_3:

$$\begin{aligned}
x_3 &= x_2 - \frac{x_2^2 - 4}{2x_2} \\
&= \frac{13}{6} - \frac{\frac{169}{36} - \frac{144}{36}}{\frac{13}{3}} \\
&= \frac{169\times 6}{36\times 13} - \frac{25}{36}\frac{3}{13} \\
&= \frac{939}{36\times 13} \\
&= 2.0064
\end{aligned}$$

This is a lot closer.

We could actually write a little calculator program to do the arithmetic for us. Suppose we've been given a value for X; let's use Y, and Z for y'.

$$\begin{aligned}
Y &= X^2 - 4 \\
Z &= 2\times X \\
X &= X - Y/Z
\end{aligned}$$

This changes X to the next approximation. We can even take an important step and replace 4 by A as the number we want to find the square root of.

```
PROGRAM:NEWT
   X^2− A → Y
   2*X → Z
   X − Y/Z → X
```

If we initialize by storing 4 in A

```
4 → A
```

and our guess, 3, in X

```
3 → X
```

and run

```
Prgm:NEWT
```

we get

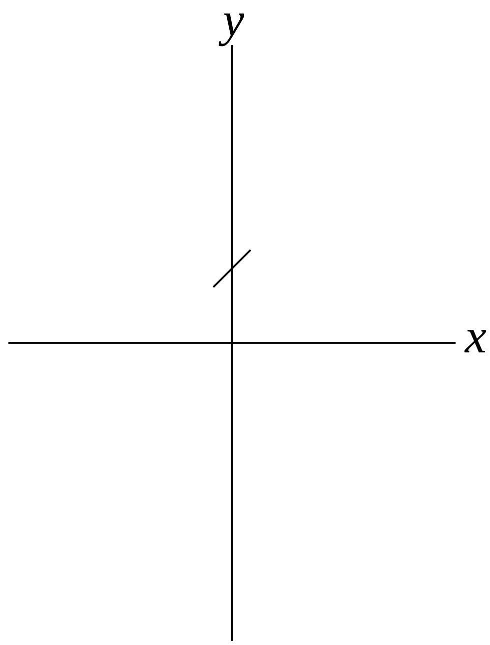
```
2.166666667
2.006410256
2.00001024
2
```

So it takes the calculator only four tries before it cannot distinguish the approximation from the exact answer $\sqrt{4} = 2$. That's to ten significant figures. (Actually, if the calculator had infinite precision, it would never get to exactly 2, but, as engineers know, it would soon get close enough.

14. Self-slope. So slopes are useful. They are also fun. Let's ask about the function that is its own slope.

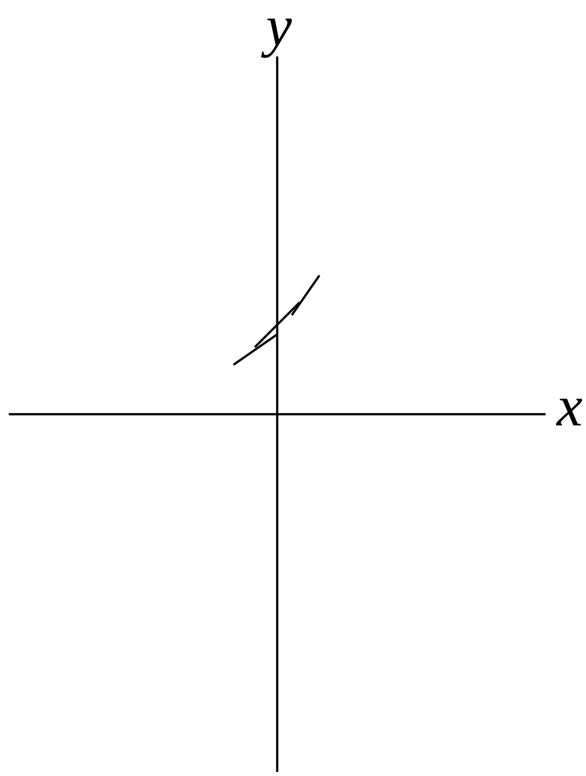
This question does not give you very much to go on, so I'll give us a start: suppose $y(0) = 1$.

So when $x = 0$, $y = 1$. What is slope(y)? Why, the same: 1. We draw a little line of slope 1 crossing the y-axis ($x = 0$) at $y = 1$.

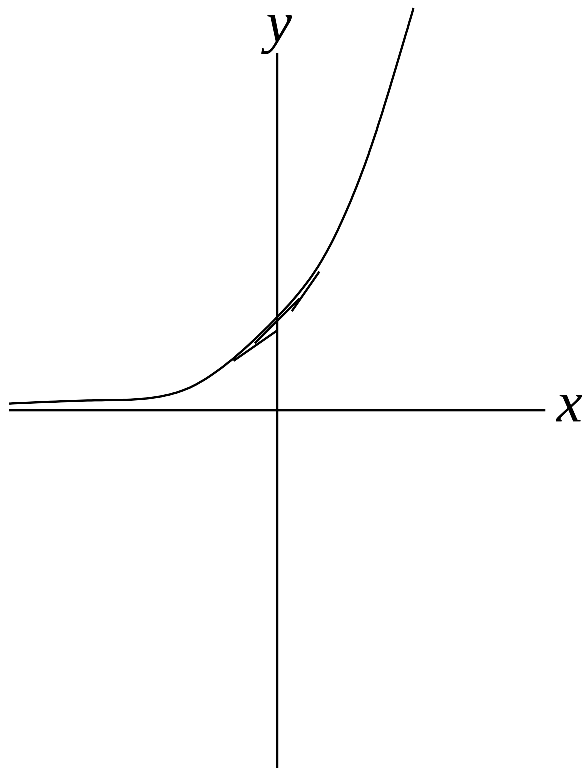


When x is just a little bigger than 0, y will be on this line and so just a little bigger than 1. So will slope(y) here be a little bigger than 1. We draw another line, sloping up a little more, just above $x = 0, y = 1$.

Similarly, just below $x = 0$, y and its slope are just a little smaller than 1, se we draw a slightly less steep line.



If we keep doing this, moving x alternately more positive and more negative, we'll get a curve.



This can never have slope 0 or else it will itself be 0: it can neither cross nor touch the x-axis. It can never have slope ∞ either, for similar reasons: it never results in a vertical line.

15. Infinite series. Can we do this mathematically? In Note 11 we have a table of slopes of powers of x. These slopes are themselves almost powers of x. There may be a clue in the way that slopes of powers are themselves powers—lower powers—albeit multiplied by a constant.

Let's try

$$y(x) = 1 + x + x^2 + x^3 + ...$$

(I won't stop: you'll see why.) Then

$$y'(x) = 0 + 1 + 2x + 3x^3 + ...$$

(Remember that slopes are transparent to addition.)

Dropping the 0, we can match up each term in the slope to the previous term in y. It only goes wrong when we get to the x terms: the slope term is twice too big.

Well, we can fix that: slope is transparent to multiplication by a constant, so let's just divide x^2 by 2.

Now we are OK up to the x^2 terms. We'll have to divide x^3 by 3 to get x^2. But we've now changed x^2 to $x^2/2$, so we must divide x^3 by 3×2.

$$\begin{aligned} y(x) &= 1 + x + \frac{x^2}{2} + \frac{x^3}{3 \times 2} + .. \\ y'(x) &= 0 + 1 + x + \frac{x^2}{2} + .. \end{aligned}$$

Pretty good. But we're always missing the last term in y. So this series has to go on forever.

Let's find the general term before dealing with the mind-boggling possibilities of an infinite series.

The last denominator was 3×2, which is actually $3 \times 2 \times 1$. The second-last denominator is thus 2×1. Pattern? Factorial: $n! = n(n-1)(n-2)...3 \times 2 \times 1$.

If we use our slope rules for powers (first table in Note 11)

$$\text{slope}\left(\frac{x^n}{n!}\right) = n\frac{x^{n-1}}{n!} = \frac{x^{n-1}}{(n-1)!}$$

and that's the previous term. Just what we need.

So what about this sum which goes on forever? How can it help but be infinite?

This question would have troubled the ancient Greek thinkers: Zeno's "paradox" goes:

> I want to get to the window.
> But first I must get halfway there.
> Then I must get halfway of the remaining half.
> Then I must get halfway along that last quarter.
> :

This goes on forever, too, but I do know that I can get to the window just by walking over. So this infinite sum must *converge*.

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + ... + \frac{1}{2^n} + .. = 1$$

Similarly our series for the self-slope function converges. Indeed it converges for *any* value of x: even though it has arbitrarily large powers, x^n, for n arbitrarily large, these numerators are essentially wiped out by the factorial denominators:

$$n! \gg x^n$$

for any n bigger than some threshold value, no matter what x is.

Not all infinite sums converge:

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + .. + \frac{1}{n} + ..$$

does not. Nor do all infinite series in powers of x converge for all values of x:

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^5}{5} + ...$$

even though the sign alternates, does not converge if $\mid x \mid \geq 1$ (i.e., $-x \leq -1$ or $x \geq 1$). But we won't get into these subtleties here. Our self-slope series does converge everywhere.

16. Programming the infinite series. Even if we accept that the series will converge, there's still a lot of calculation to do. So let's see if we can get a machine to do it,

Before choosing a machine, we must think out what we'll tell it to do.

The series is a *sum* of *terms* depending on n. The sum is what it was before, plus the term:

$sum = sum + term$ `S+T→ S`

For each subsequent term we must increment n by 1:

$n = n + 1$ `N+1→ N`

The next term is the previous term times x divided by n:

$term = term \times x/n$ `T*X/N→T`

If that's our program, we must start it off:

$sum = 0$	`0→S`
$n = 0$	`0→N`
$term = 1$	`1→T`

And, always before committing a program into a machine, we had better do a hand check. Let's suppose $x = 1$ for this check.

sum	n	*term*
0	0	1
1	1	1
2	2	1/2
5/2	3	1/6
8/3	4	1/24
65/24	5	1/120

Let's check this against the series for $x = 1$

$$1 + 1 + \frac{1}{2} + \frac{1}{3 \times 2} + \frac{1}{4 \times 3 \times 2} + \frac{1}{5 \times 4 \times 3 \times 2} + ..$$

it seems to be good. This is good because getting the order of the program right can be tricky, as is getting the right initialization for the program.

From what we've seen above, the machine can be a programmable calculator. In fact I've given TI81 code on the right of each step. Here's our program for the TI81.

```
Prgm EXP
S+T→S
N+1→N
T*X/N→T
Disp S
```

And we might as well put the initialization into a single package, too, and call it a program, so that we can restart the whole calculation easily.

```
Prgm RESTART
0→S
0→N
1→T
Prompt X
```

To run this, we restart then run `EXP` repeatedly until the answers stop changing.

`Prgm RESTART`	
`X=?1`	
`Prgm EXP`	1
<ENTER>	2
<ENTER>	2.5
<ENTER>	2.666666667
<ENTER>	2.708333333
<ENTER>	2.716666667
<ENTER>	2.718055556
	:
13th time	2.718281828

Compare this with

$$\texttt{e}^\wedge\texttt{(1)}\quad 2.718281828$$

Try again with `x=2`: after 17 iterations I got 7.38905099, which is exactly what the calculator also gives for `e^(2)`.

What we seem to have is a way of calculating the *exponential* function, $\exp(x)$, or

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ..$$

where *Euler's number* (L. Euler, pronounced "Oiler") $e = 2.718281828...$

Looking back at the self-slope curve we plotted in Note 14 we see that it could indeed be an exponential. Especially, $e^0 = 1$, where we started. You can also use the TI81 calculator (or a successor) to plot e^x to see the resemblance.

17. Slope equations. What we've been discussing is the simplest of the interesting *slope equations*: the self-slope curve satsifies

$$\text{slope}_x y = y$$

and we've solved this equation:

$$y(x) = e^x$$

Actually, we could add a constant to x (or multiply e^x by a different but related constant) and still have a solution.

$$y(x) = e^{x+c} = e^c e^x = ae^x$$

taking advantage of the transparency of slope to multiplying by a constant (Note 9)

$$\text{slope}_x y = \text{slope}_x(ae^x) = a \times \text{slope}_x e^x = ae^x = y$$

But what happens if we *multiply* x by a constant?

$$\text{slope}_x e^{cx} =?$$

We'll have to go back to the series

$$y(x) = 1 + cx + \frac{(cx)^2}{2!} + \frac{(cx)^3}{3!} + \frac{(cx)^4}{4!} + ..$$

The slope of this, using the table of slopes in Note 11 and the transparency rules of Note 9, is

$$\begin{aligned}
\text{slope}_x e^{cx} &= 0 + c + c^2 x + c^3\frac{x^2}{2!} + c^4\frac{x^3}{3!} + .. \\
&= c(1 + cx + \frac{(cx)^2}{2!} + \frac{(cx)^3}{3!} + ..) \\
&= ce^{cx}
\end{aligned}$$

So we have a second slope equation

$$\text{slope}_x y = cy$$

which has the solution

$$y(x) = e^{cx}$$

And, more generally,

$$y(x) = ae^{cx}$$

Let's go on to *second order* slope equations. These involve *double* slopes

$$\text{slope}_x^2 y = \text{slope}_x \text{slope}_x y = y$$

for instance. This one is easy: $y(x) = ae^x$ solves $\text{slope}_x y = y$ and, for this y

$$\text{slope}_x \text{slope}_x y = \text{slope}_x y = y$$

and we're done.

What about

$$\text{slope}_x^2 y = qy?$$

We'll do what everybody must do when confronted by a new slope equation: guess and try. Let's try $y(x) = ae^{cx}$.

$$\text{slope}_x \text{slope}_x y = \text{slope}_x ace^{cx} = ac^2 e^{cx} = c^2 y$$

So this works if $c = \pm\sqrt{q}$.

But what if q is negative? Or, let's say q is positive but now we want to solve

$$\text{slope}_x^2 y = -qy$$

We need to return to rotation matrices (Week iv Note 8) for advanced treatment.

18. Ninety-degree rotations. We saw a few rotation matrices in Week iv and its Excursions. They all followed the pattern

$$R = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

with the special restriction $c^2 + s^2 = 1$ to make them pure rotations and so not change the sizes of the rotated objects.

For example, in Note 8 of Week iv

$$\begin{array}{rcl} c & = & 4/5 \\ s & = & 3/5 \end{array}$$

and you can check both the pattern and the restriction.

The values of c and s are determined by the *angle* of the rotation. What are c and s for a 90-degree rotation?

We'll use an important fact about the two special vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}:$$

they pick out, respectively, the first column and the second column of any matrix which multiplies them.

$$\begin{pmatrix} a & d \\ b & c \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \qquad\qquad \begin{pmatrix} a & d \\ b & c \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}$$

With a 90-degree rotation we know what is supposed to happen to these two vectors:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad\qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

so the rotation matrix must be

$$R_{90} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Check that this obeys the pattern and restriction for rotation matrices.

Now let's consider a general rotation again, and the two special matrices, the identity matrix, I, which changes nothing (0-degree rotation: see the Excursions), and this 90-degree rotation.

$$R = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

We need a less clunky name than R_{90} for the last: we'll call it i—because, if you like, we already have I, but mainly because that is what mathematicians conventionally call it.

Notice what follows from this:

$$R = cI + si$$

where R, I, i are matrices (R is the rotation matrix at the start of this Note) and c, s are just numbers. But each of R, I and i could also be thought of as numbers. They obey all of the familiar rules of numbers, such as commutativity

$$m + n = n + m \qquad m \times n = n \times m$$

associativity

$$k + (m + n) = (k + m) + n \qquad k \times (m \times n) = (k \times m) \times n$$

having an identity element (I plays the role of 1)

$$m \times 1 = m = 1 \times m$$

and having an inverse (a rotation can always be reversed)

$$m \times m^{-1} = 1 = m^{-1} \times m$$

So, treating the identity matrix as 1, I can rewrite the above as

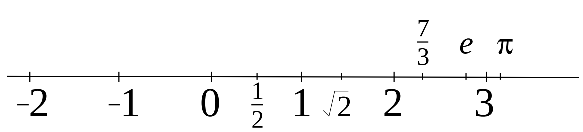
$$R = c + is$$

What is i^2? Back to matrices

$$i \times i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I = -1$$

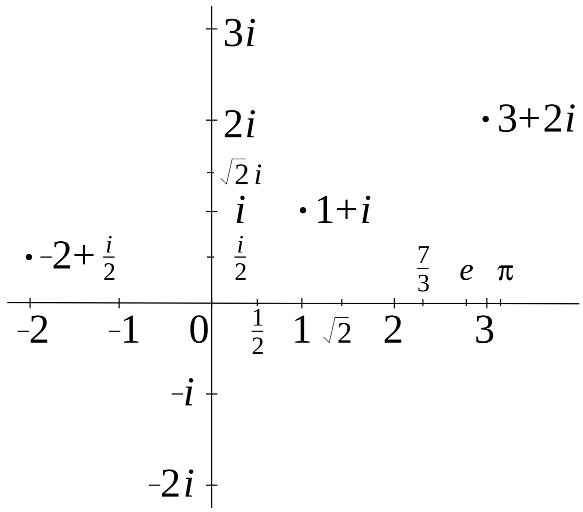
Well: it is the square root of -1! (Another reason for the name i is that it could stand for "imagine that!") But is is unproductive to think of i as $\sqrt{-1}$. Much better to think of it as it is: a 90-degree rotation.

What does this do to the usual number line?



Regular numbers are one-dimensional. Including i makes them two-dimensional: it rotates this line 90 degrees about 0.

19. Two-dimensional numbers. We thus have



The complete set of numbers is 2-dimensional. We'll call them 2-dimensional numbers, or *2-numbers* for short. (If we need to, we could refer to *1-numbers* for what we used to think of as just numbers.)

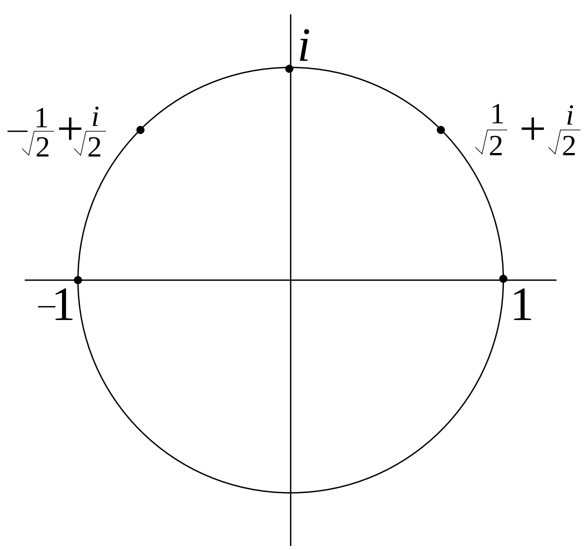
In these 2-numbers, what does

$$c + is$$

mean? Remember, $c^2 + s^2 = 1$. Some examples are

c	s
1	0
$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
0	1
$-\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
-1	0

Let's draw these as 2-numbers.

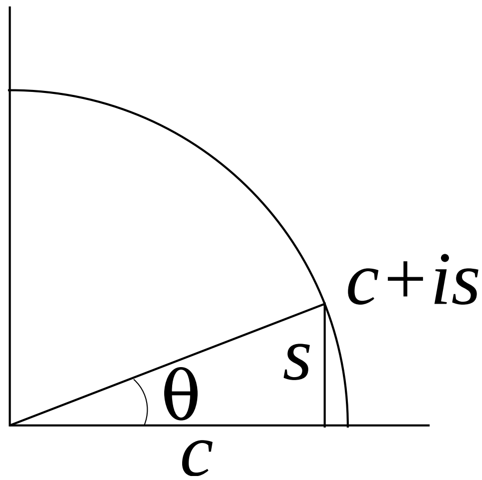


It seems we're going in *circles*. In fact,

$$c + is$$

describes the "unit circle"—the circle of radius 1 centred at 0—if we take all possible c and all possible s subject to $c^2 + s^2 = 1$.

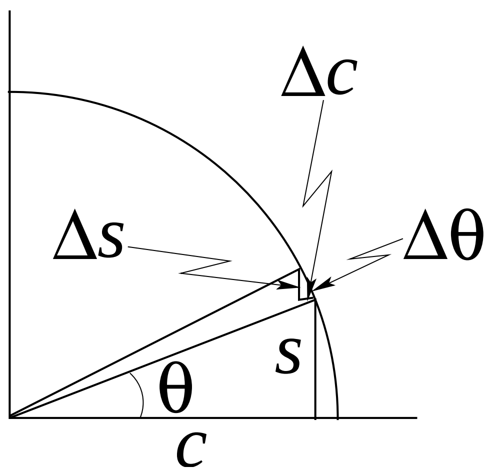
20. Slope of c and s. If we think about it, we see that c and s each depend on only one thing: the *angle* to the point labelled $c + is$. We can use the traditional Greek letter θ (theta) as the angle.



Since c and s each depend on θ, they are *functions* of θ, $c(\theta)$ and $s(\theta)$.

So we can ask about their *slopes* with respect to θ.

Here's what happens when we make a small change, $\Delta\theta$, in θ.



Note the two similar triangles:

$$\begin{aligned} \Delta c &\propto -s \\ \frac{\Delta c}{\Delta\theta} &\approx -s \\ \text{slope}_\theta c &= -s \end{aligned} \qquad\qquad \begin{aligned} \Delta s &\propto c \\ \frac{\Delta s}{\Delta\theta} &\approx c \\ \text{slope}_\theta s &= c \end{aligned}$$

We want to know how c and s each change as θ changes to $\theta + \Delta\theta$. The following is an intuitive argument, glazing over some important steps.

First look at the little triangle with sides labelled Δs and Δc. It is similar to the big triangle, given by angle θ, with sides c and s. So

$$\Delta c \propto -s \qquad\qquad \Delta s \propto c$$

where $\propto$ means"proportional to" and the negative sign for Δc captures the fact that c is getting smaller as θ gets bigger: $c + \Delta c < c$.

Second, we can say that the arc of the unit circle cut out by angle $\Delta\theta$ is itself $\Delta\theta$. This depends on how we measure angles: for this, not in degrees but in "radians". Let's just take it as given.

Third, the arc is just the hypotenuse of the small triangle, if $\Delta\theta$ is a small angle. So we can improve our "proportional to", above, to close to exact equality.

$$\frac{\Delta c}{\Delta\theta} \approx -s \qquad\qquad \frac{\Delta s}{\Delta\theta} \approx c$$

Finally, that is just the slope (if $\Delta\theta$ is small enough)

$$\text{slope}_\theta c = -s \qquad\qquad \text{slope}_\theta s = c$$

Now what about $c + is$?

$$\text{slope}_\theta(c + is) = -s + ic$$

This may or may not be interesting. But if we do it again it is very interesting.

$$\text{slope}^2_\theta(c + is) = -c - is = -(c + is)$$

So $\text{slope}^2(c + is)$ is the negative of $c + is$. Let's use a single letter for $c + is$ and show explicitly that it is a function of θ.

$$Y(\theta) = c + is$$

What we've just found is

$$\text{slope}^2_\theta Y = -Y$$

21. Connecting with slope equations. Looking back at the second-order slope equation

$$\text{slope}^2_x y = -qy$$

in Note 17, we see that 2-numbers are necessary.

Since the solution to

$$\text{slope}^2_x y = qy$$

is

$$y = ae^{\sqrt{q}x}$$

now we see that

$$y = ae^{i\sqrt{q}x}$$

solves

$$\text{slope}^2_x y = -qy$$

Let's set $a = 1$ and $q = 1$, so $y = e^{ix}$:

$$\text{slope}^2_x y = -y$$

Change the y to Y and the x to θ and we have

$$\text{slope}^2_\theta Y = -Y$$

and we've just found $Y = c + is$.

Here's the connection

$$e^{i\theta} = c + is$$

A special case: what happens if θ is half a turn (that is, 180 degrees in our old way of measuring angles)? Then $c = -1$ and $s = 0$ from our table of examples in Note 19.

We could say θ is 180 degrees, but we said in Note 20 that we can't measure angles in degrees and get the arc to equal $\Delta\theta$. We must measure in the length of the arc. For a half turn, this is half of the 2π circumference of a circle whose radius $r = 1$.

So, at half a turn

$$\theta = \pi \qquad\qquad c = -1 \qquad\qquad s = 0$$

and we have

$$e^{i\pi} = -1$$

or, to bring in all five of the most important numbers in math

$$e^{i\pi} + 1 = 0$$

22. Summary

(These notes show the trees. Try to see the forest!)

II. **The Excursions**

You've seen lots of ideas. Now *do* something with them!

1. Show that $(1 - \sqrt{5})/2 = 1 - \phi$ as calculated in Note 7. Hence confirm that $\phi^2 - \phi - 1 = 0$.

2. Show that slope$(f(x) + g(x))$ = slope$(f(x))$ + slope$(g(x))$ directly from the definition, slope = rise/run. Hint:
$$\text{slope}(f(x)) = \frac{f(x_q) - f(x_p)}{x_q - x_p}$$

3. Show that slope$(a \times f(x)) = a\times$slope$(f(x))$ for any "constant" a. (A *constant* in this context is something which does not depend on x, e.g., a number.)

4. Enter and run the "Newton's method" calculator program of Note 13. What does it give you for $\sqrt{2}$ (try starting guess 1)? Square this answer.

5. Write a Newton"s method program which finds cube roots. Write one which finds Nth roots.

6. **Mensuration** a) The area of a circle of radius r is πr^2. What is the slope of this? What is the length of its circumference? Why are these equal? (Hint. If the circle were to grow by a small amount Δr its additional area would be $\Delta r\times$ its circumference.
b) The volume of a sphere of radius r is $(4/3)\pi r^3$. What is the slope of this? What is the area of its surface?
c) The volume of a cube of "radius" a is $(2a)^3 = 8a^3$. What is the slope of this? What is the area of all its faces?
d) Repeat this for a cube of side a: why must we take the slope with respect to $a/2$ and how do we do this?
e) Excursion *Hyperspheres* in Book 9c Part I shows that the "volume" of a 4-dimensional sphere of radius r is $\pi^2 R^4/2$. What is its "surface area"?

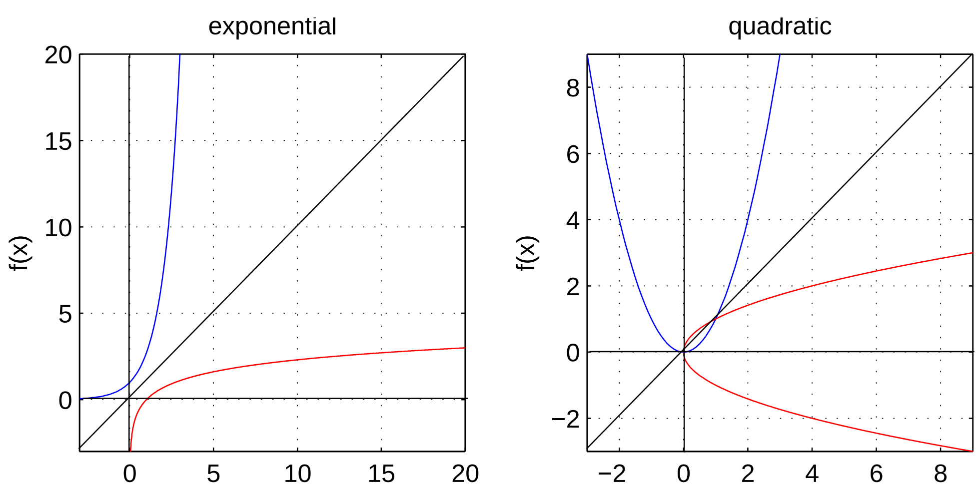
7. In Note 18, c and s are both functions of the angle of rotation, $c(\theta)$ and $s(\theta)$. We saw there that $c(90) = 0$ and $s(90) = 1$.
a) For 0 degrees show that $c(0) = 1$ and $s(0) = 0$. Hint. For 0 degrees $\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$ must be the matrtix that does nothing, the *identity* matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Or use the reasoning of Note 18.

b) What about 180 degrees and 270 degrees? Hint: these just change the signs of 0 degrees and 90 degrees, respectively. Or, a 180-degree rotation is two 90-degree rotations and 270 is three.
c) For 45 degrees, solve

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix}\begin{pmatrix} c & -s \\ s & c \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

because two 45-degree rotations makes a 90-degree rotation. (There are more than one solution: stick to both c and s being positive.) Or use the reasoning of Note 18.
d) What about 135, 225 and 315 degrees?
e) Draw a plot of $c(\theta)$ for $\theta = 0, 45, 90, 135, 180, 225, 270, 315$ and 360 degrees,and the same for $s(\theta)$.

8. **Inverses.** a) The *inverse* of an ordinary number is its *reciprocal*: the inverse of 2 is $1/2$; the inverse of $2/3$ is $3/2$. What is the connection between turning a fraction upside-down and dividing it into 1?
The common property of both concepts is that either, when multiplied by the original number, gives 1. This is the definition of "inverse", which is a more general concept than "reciprocal". (See Excursion *Modular arithmetic* in Week ii for inverses that are not reciprocals.) The inverse can be thought of as combining with the number to produce the number that "does nothing", since 1 multiplied by any further number has no effect on it.
The inverse of a *function* is another function which, when applied to it, gives the function that"does nothing", i.e., the identity function $id(x) = x$. So, for example, the inverse of squaring is the square root: $\sqrt{x^2} = x$. Note that it works in reverse, too: $\sqrt{x}^2 = x$.
Similarly the inverse of the exponential is the *natural logarithm* (see Note 5 of Week ii): $\ln e^x = x$ and $e^{\ln x} = x$.
b) What is a function which is its own inverse? Is there more than one such function?
A hint for this last question comes from picturing inverse functions. Here are the exponential and logarithmic functions, e^x and $\ln x$, as well as the quadratic and square root functions, x^2 and $\sqrt{x}$.



As well as the x and y axes, I've shown the line $y = x$ which acts as a mirror.
You should note that the curve I've shown for $\sqrt{x}$ is *not* a function: it has *two* values for any positive value for x. But functions are supposed to be uniquely determined by their argument. I had to draw it by plotting both `sqrt(x)` and `−sqrt(x)`. (The fact that $\sqrt{x}$ has *no* values for some x, namely negative x, is not an issue: $\ln x$ also has no values for negative

x, but that does not stop it from being a function.)
c) Why does a function become its inverse on exchanging $x \leftrightarrow y$? Why does this exchange act like a mirror along the line $y = x$?
d) Use this visualization to show that the slope of the inverse of a function is the reciprocal of the slope of the function.

9. **Radians.** There are various ways of measuring angles. Commonly we use degrees, with 360 degrees being a full turn. (This number is a legacy from the Babylonians and has the advantage that it is an integer which has many factors—2, 3, 5 and their multiples: $360 = 2^3 3^2 5$—so that it is easy to get integer measures for many important fractions of a turn.) We could alternately use a much more natural measure, namely fractions of a *turn*, so that a right angle is a quarter-turn, and so on.
Because of the need in Note 20, where we are measuring triangles, to express an angle as a *length*, we must learn a third way of measuring angles. A *radian* is the angle subtended by an arc of the unit circle which has length 1.
How many radians in a full turn (360 degrees)? In a half turn (180 degrees)? In a quarter turn (a right angle, 90 degrees)?

10. **Rotation generators.** In Note 18 we decomposed
$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$
into
$$R = cI + si$$
by introducing the 90-degree rotation
$$i = \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$$
In Note 21 we wrote the rotation as 2-numbers in the form
$$e^{i\theta} = c + is$$
In Note 16 we equated
$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ..$$
so that e^x is a function which is its own slope.
Here we combine matrices and 2-numbers to explore
$$\begin{aligned} e^{i\theta} &= I + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + .. \\ &= \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}\theta + \frac{1}{2!}\begin{pmatrix} & -1 \\ 1 & \end{pmatrix}^2\theta^2 + \frac{1}{3!}\begin{pmatrix} & -1 \\ 1 & \end{pmatrix}^3\theta^3 + .. \end{aligned}$$
where i goes back to being a matrix, and the meaning of $e^{i\theta}$ is given by the infinite series.
Since we found $i^2 = -I$ (as matrices: Note 18)
$$\begin{aligned} e^{i\theta} &= I + i\theta - \frac{\theta^2}{2!}I - \frac{\theta^3}{3!}i + .. \\ &= I\left(1 - \frac{\theta^2}{2!} + ..\right) + i\left(\theta - \frac{\theta^3}{3!} + ..\right) \\ &= I\cos\theta + i\sin\theta \\ &= \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \end{aligned}$$

So $e^{i\theta}$, with i considered as a matrix, is just the rotation matrix.
The matrix i is thus the *generator* of 2D rotations.
Complete the above argument by showing that the series expansions for cos() and sin() are

$$\begin{aligned}\cos\theta &= 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!}.. \\ \sin\theta &= \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!}..\end{aligned}$$

11. **3D rotation generators.** In Excursion *Three-dimensional rotations* of Week iv we wrote three matrices for 3D rotations. They can also be expressed with generators, J'_z, J'_y, J'_x, following the previous Excursion. (I've used primes because we are going to modify the definitions at the end of this Excursion.)

$$R_{xy} = \begin{pmatrix} c_{xy} & -s_{xy} & \\ s_{xy} & c_{xy} & \\ & & 1 \end{pmatrix} = e^{\theta_{xy}J'_z} \qquad J'_z = \begin{pmatrix} & -1 & \\ 1 & & \\ & & \end{pmatrix}$$

$$R_{zx} = \begin{pmatrix} c_{zx} & & s_{zx} \\ & 1 & \\ -s_{zx} & & c_{zx} \end{pmatrix} = e^{\theta_{zx}J'_y} \qquad J'_y = \begin{pmatrix} & & 1 \\ & & \\ -1 & & \end{pmatrix}$$

$$R_{yz} = \begin{pmatrix} 1 & & \\ & c_{yz} & -s_{yz} \\ & s_{yz} & c_{yz} \end{pmatrix} = e^{\theta_{yz}J'_x} \qquad J'_x = \begin{pmatrix} & & \\ & & -1 \\ & 1 & \end{pmatrix}$$

Note that I've tweaked the notation to describe, e.g., the angle θ_{xy} in the x-y plane as applied to the generator J'_z for rotations about the z-axis. NB. $c_{xy} \stackrel{\text{def}}{=} \cos\theta_{xy}$, etc.
a) Show that $e^{\theta_{xy}J'_z} = R_{xy}$ and so on.
b) Since Excursion *Three-dimensional rotations* (Week iv) showed that 3D rotations do not commute, we can explore the commutators of these generators

$$[J'_x, J'_y]_- \stackrel{\text{def}}{=} J'_xJ'_y - J'_yJ'_x$$

Show that

$$[J'_x, J'_y]_- = J'_z \qquad [J'_z, J'_x]_- = J'_y \qquad [J'_y, J'_z]_- = J'_x$$

c) We could write a combination of rotations

$$R_{xy}R_{zx} = e^{\theta_{xy}J'_z}e^{\theta_{zx}J'_y}$$

as

$$e^{\theta_{xy}J'_z + \theta_{zx}J'_y}$$

but why should we be careful? Hint: does matrix addition commute?
d) The J's are 3-by-3 matrices. Can we find 2-by-2 matrices which also describe 3D rotations? Let's look for 2-by-2 generators that satisfy the same commutation properties as in (b) above. We must use 2D numbers in these matrices: i is a 2D number, not a matrix, in the following.

$$\sigma'_z = \frac{i}{2}\begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \qquad \sigma'_y = \frac{i}{2}\begin{pmatrix} & -i \\ i & \end{pmatrix} \qquad \sigma'_x = \frac{i}{2}\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$$

Show that

$$[\sigma'_x, \sigma'_y]_- = \sigma'_z \qquad [\sigma'_z, \sigma'_x]_- = \sigma'_y \qquad [\sigma'_y, \sigma'_z]_- = \sigma'_x$$

(These matrices—well, without the $i/2$ factors—are called *Pauli* matrices and are very important in quantum physics: see Week 6 Note 3.)
e) Show that the Pauli matrices also satisfy

$$[\sigma'_x, \sigma'_y]_+ = 0 \qquad [\sigma'_z, \sigma'_x]_+ = 0 \qquad [\sigma'_y, \sigma'_z]_+ = 0$$

where the *anticommutator* is, e.g.,

$$[\sigma'_x, \sigma'_y]_+ \stackrel{\text{def}}{=} \sigma'_x\sigma'_y + \sigma'_y\sigma'_x$$

In other words

$$\sigma'_x\sigma'_y = -\sigma'_y\sigma'_x$$

etc. Compare these to the Clifford algebra ("interval algebra") of Week 7c Part B.
f) Note that the Js are all *antisymmetric* matrices, e.g., $J_z'^T = -J_z'$ where ^{T} is the *transpose* operator, e.g.,

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}^T \stackrel{\text{def}}{=} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

The rotation matrices are *orthogonal*, e.g.,

$$R_{xy}R_{xy}^T = I$$

Show that e^A is orthogonal if A is antisymmetric. Hint: show that $(e^A)^T = e^{A^T}$ so $(e^A)(e^A)^T = e^Ae^{-A} = e^0 = 1$ with matrix versions of 0 and 1.
g) The above ideas of antisymmetric and orthogonal go respectively over to *antihermitian* and *unitary* for matrices involving 2-numbers such as the Pauli matrices. The only change is that the transpose operator goes over to an operator which also changes the sign on every i: transpose becomes *hermitian conjugate*, designated †.
Show that iH is antihermitian $(iH)^\dagger = -iH^\dagger$ if H is hermitian $H^\dagger = H$,
Show that the Pauli matrices are antihermitian.
Show that $U = e^{iH}$ is unitary if H is hermitian (or iH is antihermitian).
h) The significance of orthogonal or unitary matrices is that they preserve distances when used to transform space, e.g.,

$$x'^2 + y'^2 = r^2 = x^2 + y^2$$

in

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R\begin{pmatrix} x \\ y \end{pmatrix}$$

for a rotation R—which is an orthogonal matrix.
Show this.
The significance of symmetric or hermitian matrices (or their antis) is that these properties are very easy to check. Symmetric matrices have 1-number ("real") eigenvalues and orthogonal eigenvectors (see Week iv, Excursions *Diagonalizing matrices* and following); hermitian matrices have the same, in 2-number terms.
i) Generators are thus often written as symmetric or hermitian matrices J, and the rotations (say) generated are, e.g.,

$$R = e^{i\theta J}$$

Work out symmetric Js and hermitian σs for the 3-by-3 and 2-by-2 cases, respectively, above, and find their commutation relationships. What are the anticommutators of the σs?

12. Any part of the Preliminary Notes that needs working through.