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

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

August 2, 2024

I. Prefatory Notes

1. Cube root of 8. Almost every calculator has a square-root button, $\sqrt{\cdot}$. But how would we calculate a *cube* root?

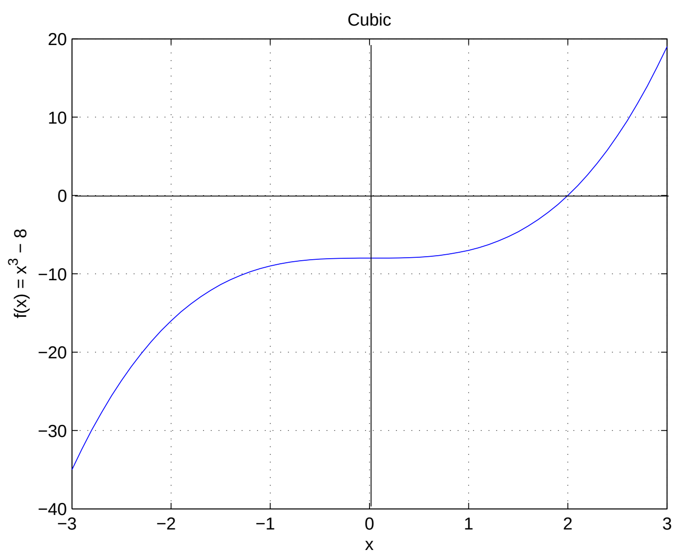
Let's start with one we know: $\sqrt[3]{8} = 2$.

We are going to find this by finding the "zeros" of a "function". For this case

$$0 = f(x) = x^3 - 8$$

will do it: if we can find the value of x which makes this $f(x)$ zero, we will know $\sqrt[3]{8}$.

Here is a plot of this function.

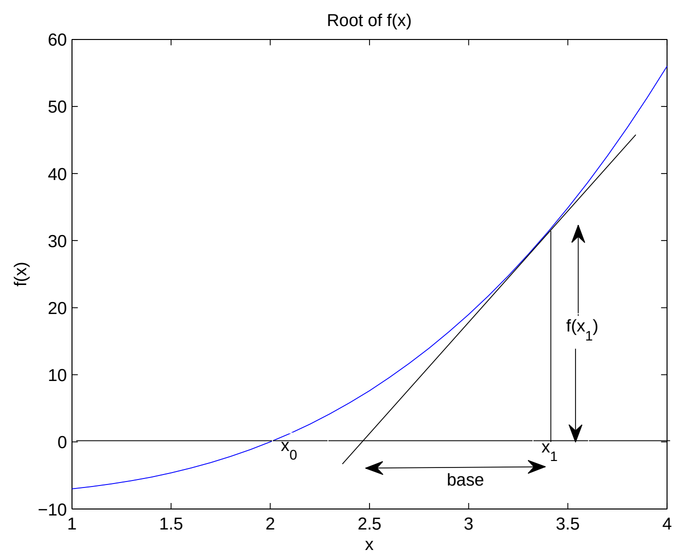


Note that the line crosses the horizontal (x) axis at $x = 2$, which we know to be the cube root of 8.

2. Slope. Now we introduce the idea of a *slope*. Loosely speaking, the slope of a function $f(x)$ at a specific value of x is the tangent to the function curve at that point.

To be precise, we define the slope to be a *number*: the height divided by the base of the triangle generated by that tangent. In the figure

$$\text{slope}_{x=x_1} f(x) \stackrel{\text{def}}{=} \frac{f(x_1)}{\text{base}}$$



3. Approximations. Now we are going to start making approximations. These will be true enough when things get close enough together, so in the end we will have exact results.

(We'll use the symbol $\approx$ to mean "approximately equals". If we use it in a chain of equalities, the last expression in the chain will not exactly equal the first expression—even though there may be many $=$ signs and only one $\approx$ sign—but will only approximately equal it.)

In the above figure, we don't know what the *base* is, so we'll approximate it by $x_1 - x_0$.

We don't know what x_0 is, either, but we want to find it: it is the unknown place where $f(x_0) = 0$. So we'll use it to stand for what we don't yet know and do some algebra until we can eventually pin it down.

Now we can write the slope down approximately.

$$\text{slope}_{x=x_1} f(x) = \frac{f(x_1)}{\text{base}} \approx \frac{f(x_1)}{x_1 - x_0}$$

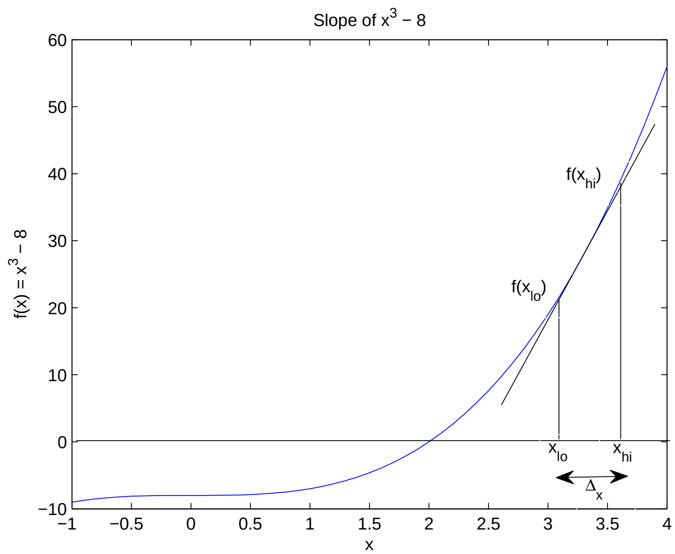
So we can rearrange this to write down x_0, our quarry, approximately.

$$x_0 \approx x_1 - \frac{f(x_1)}{\text{slope}_{x=x_1} f(x)}$$

Since we know $f(x) = x^3 - 8$ we can calculate $f(x_1)$ for any given x_1.

We must now figure out how to calculate $\text{slope}_{x=x_1} f(x)$.

4. Slope of cubic. Here is $f(x) = x^3 - 8$ again, with all the pieces needed to find its slope at x_{lo}.



The definition of slope tells us, if x_{lo} and x_{hi} are close enough together,

$$\mathrm{slope}_{x=x_{\mathrm{lo}}} f(x) = \frac{f(x_{\mathrm{hi}}) - f(x_{\mathrm{lo}})}{x_{\mathrm{hi}} - x_{\mathrm{lo}}} = \frac{f(x_{\mathrm{lo}} + \Delta x) - f(x_{\mathrm{lo}})}{\Delta x}$$

So let's try it for $f(x) = x^3 - 8$.

$$\begin{aligned}
\mathrm{slope}_{x=x_{\mathrm{lo}}} x^3 - 8 &= \frac{(x_{\mathrm{lo}} + \Delta x)^3 - (x_{\mathrm{lo}})^3}{\Delta x} \\
&= \frac{3(x_{\mathrm{lo}})^2 \Delta x + 3x_{\mathrm{lo}}(\Delta x)^2 + (\Delta x)^3}{\Delta x} \\
&\approx \frac{3(x_{\mathrm{lo}})^2 \Delta x}{\Delta x} \\
&= 3(x_{\mathrm{lo}})^2
\end{aligned}$$

That is, slope $x^3 - a = 3x^2$ for any a.

The $\approx$ in the above chain holds more and more exactly the smaller Δx gets, i.e., the closer x_{hi} and x_{lo} get together. But this is what we intend to happen in calling the result the "slope *at*" the specified value. What we really do is take the *limit* as the two points get closer and closer together. This justifies using the exact $=$ in saying above "slope $x^3 - a = 3x^2$".

5. The root. So now we can go back to Note 3 where we had

$$x_0 \approx x_1 - \frac{f(x_1)}{\mathrm{slope}_{x=x_1} f(x)}$$

and replace it for $f(x) = x^3 - 8$ by

$$x_0 \approx x_1 - \frac{(x_1)^3 - 8}{3(x_1)^2}$$

Let's try it. We'll need a value for x_1. We'll just guess. We know for this case that the final cube root will be 2, but it is a special case we chose to make it easy to check the answer. Usually we can only guess the value of the final cube root, maybe from plotting the function. So we'll pretend we don't know and try $x_1 = 3$ as a guess.

$$x_2 = x_1 - \frac{(x_1)^3 - 8}{3(x_1)^2} = 1.9444$$

This is pretty close to 2, but let's try the result as the *next guess* (call it x_2).

$$x_3 = x_2 - \frac{(x_2)^3 - 8}{3(x_2)^2} = 2.0302$$

Better. Try again.

$$x_4 = x_3 - \frac{(x_3)^3 - 8}{3(x_3)^2} = 1.9856$$

Keep trying.

$$\begin{aligned}
x_5 &= x_4 - \frac{(x_4)^3 - 8}{3(x_4)^2} = 2.0074 \\
x_6 &= x_5 - \frac{(x_5)^3 - 8}{3(x_5)^2} = 1.9964 \\
x_7 &= x_6 - \frac{(x_6)^3 - 8}{3(x_6)^2} = 2.0018 \\
x_8 &= x_7 - \frac{(x_7)^3 - 8}{3(x_7)^2} = 1.9991 \\
x_9 &= x_8 - \frac{(x_8)^3 - 8}{3(x_8)^2} = 2.0005 \\
x_{10} &= x_9 - \frac{(x_9)^3 - 8}{3(x_9)^2} = 1.9998 \\
&\vdots
\end{aligned}$$

How many more steps before the result is 2.0000? 2.000000000000000?

So we see $x_n \to x_0$ in the limit as n gets arbitrarily large.

6. Square roots. This process works for square roots, too. In fact, it is the basis for the $\surd$ button on your calculator, and for the `sqrt` function in any programming language.

We just need to work out that slope$(x^2 - a) = 2x$. Now try finding the square root of 4.

$$x_0 \approx x_1 - \frac{(x_1)^2 - 4}{2(x_1)}$$

7. Antislopes. Now we know about slopes we might wonder about going the other way. Just as division undoes multiplication, *antislopes* undo slopes[1]. Here are some examples.

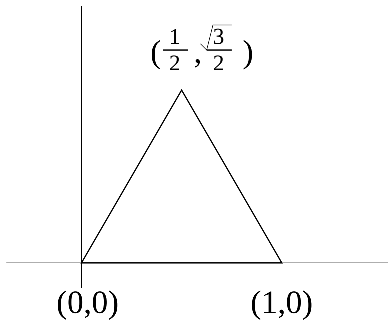
slope$_x$ f	$f(x)$	antislope$_x$ f
0	1	$x + C$
1	x	$x^2/2 + C$
$2x$	x^2	$x^3/3 + C$
$3x^2$	x^3	$x^4/4 + C$
$4x^3$	x^4	$x^5/5 + C$
$\vdots$	$\vdots$	$\vdots$

[1]We saw that division goes with finding slopes and we'll see that multiplication goes with finding antislopes, so it would be more appropriate to reverse this phrase and say "Just as multiplication undoes division, antislopes undo slopes."

where C in each case is some "constant", that is, some number or expression which does not depend on x.

Check these examples by finding the slopes of the antislope$_x$ f column. Then see if you can figure out the antislopes of the slopes$_x$ f column.

8. Areas. In addition to slopes, the other major use of arbitrarily shrinking interval, Δx, is to find areas. Let's find the area of an equilateral triangle.

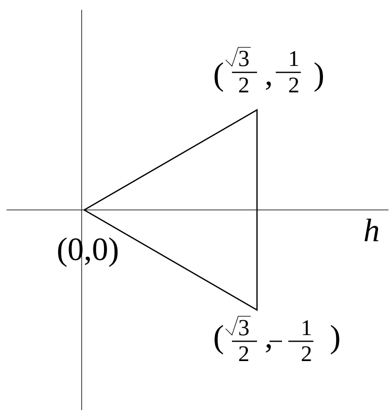


You probably know that the area of a triangle is $bh/2$ where b is the *base* and h is the *height*. For the triangle of the figure, $b = 1$ and $h = \sqrt{3}/4$ so

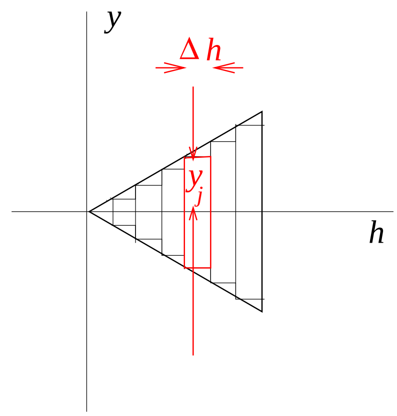
$$\text{area} = \frac{\sqrt{3}}{4}$$

Let's see if we can work out *why* this is the answer.

We'll redraw the triangle rotated so that its "height" is horizontal.



Now slice it into vertical strips of width Δh, and make them rectangles just touching the sides of the triangle at two of the four corners of each rectangle.



The area of the triangle will be

$$\text{area} \;\approx\; \sum_{j=0}^{(\sqrt{3}/2)/\Delta h-1} 2y_j\Delta h$$

where we must figure out what each y_j is.

(Do you see why the number of steps of width Δh is $(\sqrt{3}/2)/\Delta h$?)

I used the approximation sign, $\approx$, because of the missing bits of area between the steps and the sides of the triangle. The important thing about this approximation is that it gets better and better the smaller Δh is: this is where the milli-micro-nano aspect of this Week comes in.

Let's think of y as a *function* of h

$$y(h) = sh$$

where s is the slope of the upper side of the triangle.

We can see

$$s = \frac{1/2}{\sqrt{3}/2} = \frac{1}{\sqrt{3}}$$

and we can check

$$y(0) = 0 \qquad\qquad y(\frac{\sqrt{3}}{2}) = \frac{1}{\sqrt{3}}\frac{\sqrt{3}}{2} = \frac{1}{2}$$

So

$$h_j = j\Delta h$$

between

$$j = 0 \quad \text{and} \quad j = \frac{\sqrt{3}/2}{\Delta h}$$

and we have

$$y_j = y(h_j) = \frac{1}{\sqrt{3}}h_j = \frac{j\Delta h}{\sqrt{3}}$$

So

$$\begin{aligned} \text{area} \;&\approx\; \sum_{j=0}^{(\sqrt{3}/2)/\Delta h-1} 2y(h_j)\Delta h \\ &=\; \sum_{j=0}^{(\sqrt{3}/2)/\Delta h-1} \frac{2\Delta h}{\sqrt{3}}j\Delta h \end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{\sqrt{3}}(\Delta h)^2 \sum_{j=0}^{(\sqrt{3}/2)/\Delta h-1} j \\
&= \frac{2}{\sqrt{3}}(\Delta h)^2 \frac{(\sqrt{3}/2)/\Delta h)((\sqrt{3}/2)/\Delta h-1)}{2} \\
&= \frac{1}{\sqrt{3}}(\Delta h)^2 \frac{\sqrt{3}/2}{\Delta h}(\frac{\sqrt{3}/2}{\Delta h}-1) \\
&\approx \frac{1}{\sqrt{3}}(\Delta h)^2 (\frac{\sqrt{3}/2}{\Delta h})^2 \\
&= \frac{1}{\sqrt{3}}\frac{3}{4} \\
&= \frac{\sqrt{3}}{4}
\end{aligned}
$$

which is the same result as the $bh/2$ calculation at the beginning of this Note.

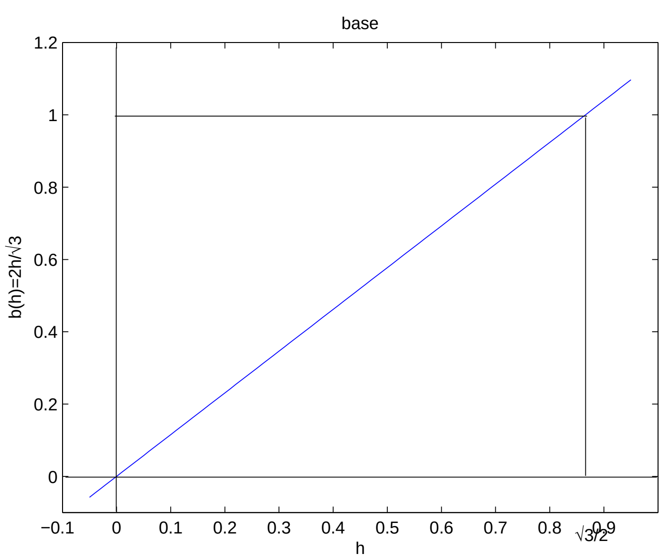
Notice that I made *two* approximations, which happened, this time, to cancel each other out. The second approximation was that $\frac{\sqrt{3}/2}{\Delta h}$ is big enough that $\frac{\sqrt{3}/2}{\Delta h} - 1$ essentially equals $\frac{\sqrt{3}/2}{\Delta h}$. This approximation also gets better and better as Δh gets arbitrarily small: more milli-micro-nano-.. math.

Note also that the sum, $\sum$, disappears in the fourth line above and is replaced by $n(n-1)/2$ (letting n be $\frac{\sqrt{3}/2}{\Delta h}$): the sum is just a triangular number (Week i Note 1). (Please do not confuse the Δ in Δh (this Note) with the Δ in Δ_n etc. (Week i Notes). Here Δ modifies the variable following it to mean "a little bit of". There the Δ means "triangle" and has a *subscript* n to say how big the triangle is.)

To summarize this procedure a little differently, we summed the base, b, as a function of the height, h

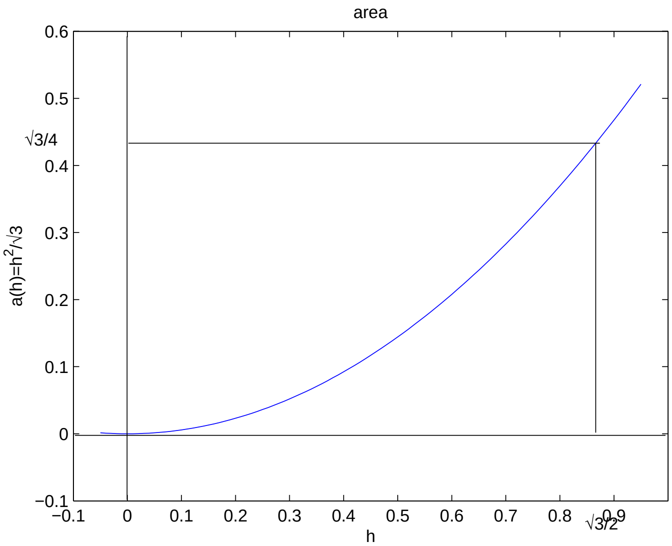
$$b(h) = \frac{2}{\sqrt{3}}h$$

over a discrete set of values, h_j, separated by steps of size Δh from 0 to $\sqrt{3}/2$.



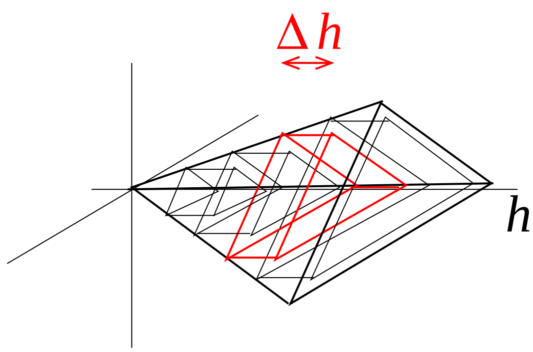
We can use $b = 2h/\sqrt{3}$ to generalize the area from the height $h = \sqrt{3}/2$ triangle to a triangle with *any* height h:

$$\text{area} \;=\; \frac{1}{2}bh = \frac{1}{\sqrt{3}}h^2$$



9. Volumes. We can use the same technique to find the volume of a regular tetrahedron. This has all six edges of the same length, say 1, and height $\sqrt{2/3}$ in the third dimension. (Each of the four faces is an equilateral triangle of sides 1 and "height" on the face—we won't call this the height any more—$\sqrt{3}/2$ as in the previous Note.)

We'll turn the tetrahedron sideways, so its 3-D height is shown horizontally, and slice it into triangular slabs of thickness Δh and sizes from 0 at the apex to sides=1 at the base.



$$area_j = \frac{1}{\sqrt{3}}\, h_j^2$$

$$\Delta\, vol_j = \frac{1}{\sqrt{3}}\, h_j^2 \Delta\, h$$

So the sum we need this time is

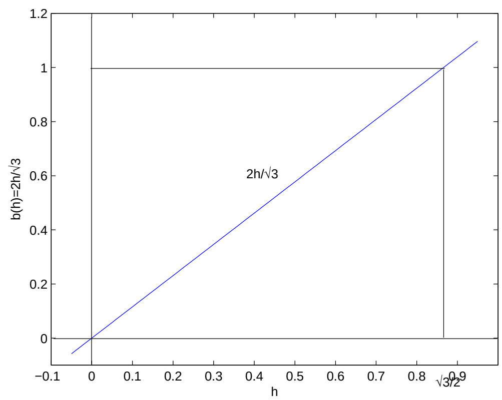
$$\text{vol} \;\approx\; \sum_{j=0}^{\sqrt{2/3}/\Delta h - 1} \frac{1}{\sqrt{3}} h_j^2 \Delta h$$

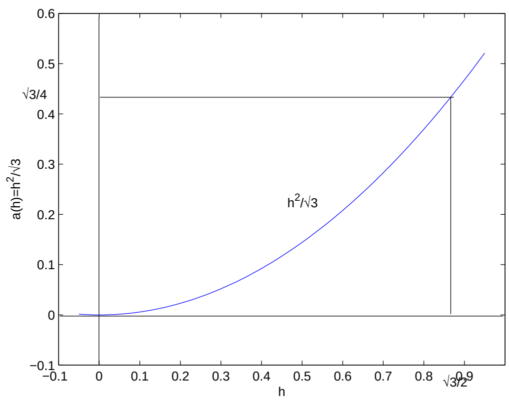
with $h_j = j\Delta h$ as before, and with the same reasons for noting that this is an approximation (which will improve as Δh gets small).

$$\begin{aligned}
\text{vol} &\approx \sum_{j=0}^{\sqrt{2/3}/\Delta h-1} \frac{1}{\sqrt{3}} j^2 (\Delta h)^3 \\
&= \frac{(\Delta h)^3}{\sqrt{3}} \sum_{j=0}^{\sqrt{2/3}/\Delta h-1} j^2 \\
&= \frac{(\Delta h)^3}{\sqrt{3}} \frac{1}{3} \frac{\sqrt{2/3}}{\Delta h} \left(\frac{\sqrt{2/3}}{\Delta h} - \frac{1}{2}\right)\left(\frac{\sqrt{2/3}}{\Delta h} - 1\right) \\
&\approx \frac{1}{3\sqrt{3}} \left(\frac{\sqrt{2/3}}{\Delta h}\right)^3 \\
&= \frac{2\sqrt{2}}{27}
\end{aligned}$$

In the second line I summed j^2 from 0 to $n-1 = \sqrt{\frac{2}{3}}\frac{1}{\Delta h} - 1$ which we can work out from Week i Note 1 as $\frac{1}{3}n(n-\frac{1}{2})(n-1)$. I approximated this by $n^3/3$, once again an approximation which improves as Δh gets small, and an approximation which happens to counteract the first approximation above exactly, giving the exact right answer.

10. Antislopes and areas. In Notes 8 and 9 we found the areas under the following two functions.





Let's just find the antislopes of these functions (Note.7).

$$\begin{aligned}
\text{antislope}_h \frac{2}{\sqrt{3}} h &= \frac{2}{\sqrt{3}} \text{ antislope}_h\ h = \frac{1}{\sqrt{3}} h^2 + C_1 \\
\text{antislope}_h \frac{1}{\sqrt{3}} h^2 &= \frac{1}{\sqrt{3}} \text{ antislope}_h\ h^2 = \frac{1}{3\sqrt{3}} h^3 + C_2
\end{aligned}$$

where the constants C_1 and C_2 can be any numbers not depending on h.

Let's evaluate the first at $h = \sqrt{3}/2$ and at $h = 0$ and subtract.

$$\text{antislope}_h \left. \frac{2}{\sqrt{3}} h \right|_0^{\sqrt{3}/2} = \left. \frac{1}{\sqrt{3}} h^2 \right|_0^{\sqrt{3}/2} = \frac{\sqrt{3}}{4} - 0 = \frac{\sqrt{3}}{4}$$

and the second at $h = \sqrt{2/3}$ and at $h = 0$ and subtract.

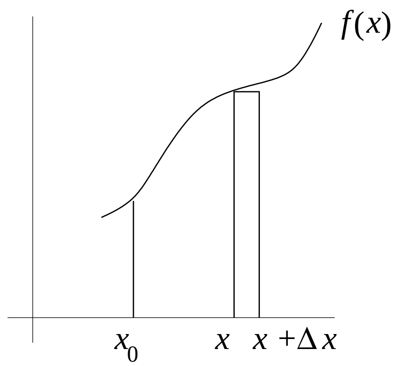
$$\text{antislope}_h \left. \frac{1}{\sqrt{3}} h^2 \right|_0^{\sqrt{2/3}} = \left. \frac{1}{3\sqrt{3}} h^3 \right|_0^{\sqrt{2/3}} = \frac{2\sqrt{2}}{27} - 0 = \frac{2\sqrt{2}}{27}$$

What is the connection with the triangular area in Note 8 and the tetrahedral volume in Note 9?

It turns out that antislopes give *areas* when suitably evaluated.

11. The Fundamental Theorem of Calculus. Areas between two points under a function are linked to the antislope of the function, i.e., to the function whose slope is that function. Let's see why.

Here is a function $f(x)$ and a small section of $f(x)\Delta x$ of the area under $f(x)$.



We can call the area $A(x_0, x)$, supposing that we are in the process of finding the area under $f(x)$ starting at x_0, and that we have so far reached x. If x_0 is fixed, we are interested only in the area as a function of x and we can write, simply, $A(x)$. Then

$$\Delta A(x) = f(x)\Delta x$$

for the small increment, $\Delta A(x)$, of the area.

This is very familiar: just rearrange

$$f(x) = \frac{\Delta A(x)}{\Delta x}$$

and we have that $f(x)$ is the slope of $A(x)$.

So $A(x)$ is the antislope of $f(x)$.

And if we want to find the area from say x_0 to say x_1,we just evaluate

$$A(x_1) - A(x_0)$$

We can write this

$$A(x) \mid_{x_0}^{x_1}$$

or

$$\text{antislope}_x\ f(x) \mid_{x_0}^{x_1}$$

Since these "definite integrals" are the same antislope evaluated at two different places and subtracted, the arbitrary constant that we introduced with antislopes in Note 7 disappears and we can forget about it.

12. Summary

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

The reason I called these Notes "Milli-micro-nano-..maths" is because of the approximations. To make slopes exact we must apply the definition to ever smaller intervals.

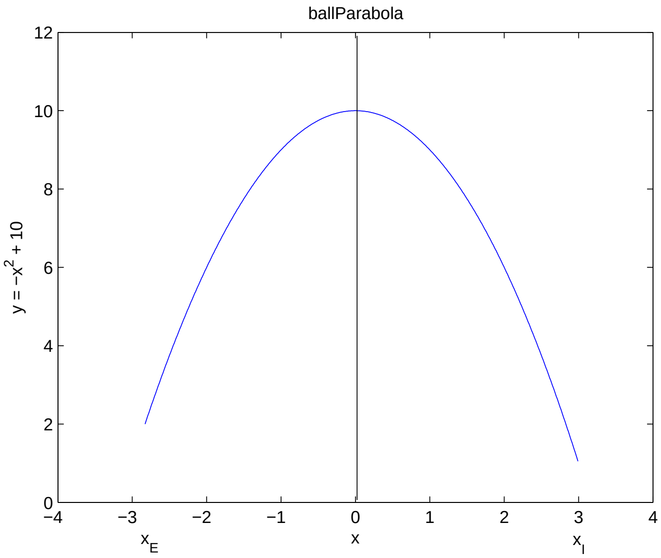
The method I have presented to find cube and square roots is "Newton's method". It is an application of differential calculus, also invented by Isaac Newton, and independently by Gottfried

Leibniz whose notation is now used, although I have not used it. Newton's method works for finding roots (zeros) of almost any function and it converges very quickly.

II. **The Excursions**

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

1. **Throwing balls: parabola and square roots.** (This Excursion is meant to help if you found this topic tough going right from the start. It is about functions and their roots ("zeros").)
   Eric is throwing a ball for Iris to catch. Here is how the ball goes.



The ball leaves Eric's hand 2 meters from the ground, rises to a maximum height of 10 meters, then drops to 1 meter where Iris catches it. Gravity causes the ball to follow a *parabola* in between Eric and Iris. A parabola is an example of a *function*, which gives a unique value for y (which is what I've called the height in the graph) for each possible value of x (which is what I call the horizontal position of the ball).
The relationship between x and y is given by the *equation*

$$y = -x^2 + 10$$

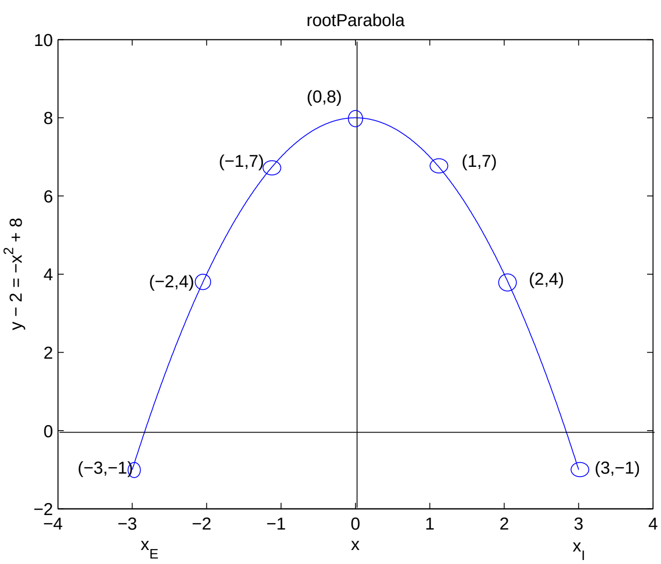
That is the simplest equation to describe the particular path I've plotted. To make it this simple, I chose the axes so that the maximum height is attained when $x = 0$. I also chose the horizontal scale so that the ball is moving from somewhere around $x = -3$ to somewhere around $x = 3$. We'll see where exactly these endpoints are. In fact, that is the purpose of this Excursion.
Let's make a table showing some of the values for y that the rule $y = -x^2 + 10$ calculates for different values of x. The table spells this out in the second column. (We'll come to the meanings of the third and fourth columns shortly.)

x	$-x^2 + 10$	$= y$	$y - 1$	$y - 2$
-3	$-9 + 10$	$= 1$	0	-1
-2	$-4 + 10$	$= 6$		4
-1	$-1 + 10$	$= 9$		7
0	$0 + 10$	$=10$		8
1	$1 + 10$	$= 9$		7
2	$4 + 10$	$= 6$		4
3	$9 + 10$	$= 1$	0	-1

Check these calculations. What do you notice about the value of y for positive and negative values of x?
Now, where must Iris be to catch the ball 1 meter off the ground? The table says the ball is 1 meter off the ground at $x = -3$ and at $x = 3$. Since I've shown Iris on the *x-positive* side, we conclude that Iris is at $x_{\mathrm{I}} = 3$.
The two possible positions, from which we've chosen the positive one, also correspond to $y - 1 = 0$, which I've marked with 0 in the $y - 1$ column.
Finding out where Eric must stand to throw the ball is trickier. We don't see any zeros in the $y - 2$ column. But we do see some negative numbers and some positive numbers in that column. $y - 2 = 0$, or $y = 2$, must occur:
a) between $x = -3$, where $y - 2 = -1$, and $x = -2$, where $y - 2 = 4$—and it plainly must be somewhat closer to $x = -3$ than to $x = -2$; or
b) between $x = 2$, where $y - 2 = 4$, and $x = 3$, where $y - 2 = -1$—and it plainly must be somewhat closer to $x = 3$ than to $x = 2$.
Let's do some algebra.

$$\begin{aligned} 0 &= y - 2 \\ &= -x^2 + 10 - 2 \\ &= -x^2 + 8 \end{aligned}$$

and here is the graph of this $y - 2 = -x^2 + 8$. (I could have given $y - 2$ a new name, say y' or z, but why burden ourselves with additional symbols?)



We see it is another parabola, this time one which *crosses the x-axis* at the "zeros" of $y - 2$, i.e., where $y - 2 = 0$, or $y = 2$, which is just the height of the ball as it leaves Eric's hand. These zeros are the *roots* of the parabola.
(I've also shown the pairs (x, y) for the seven values of x that are in the table. Clearly these are only *some* of the possible pairs. How many could there be?)
To find where Eric is, we need the roots of this second parabola.
We must do some more algebra. Continuing from before

$$\begin{aligned} 0 &= -x^2 + 8 \\ x^2 &= 8 \\ x &= \sqrt{8} \\ &= 2.828\cdots \end{aligned}$$

Thus Eric is at $x_{\mathrm{E}} = -\sqrt{8} = -2.828\cdots$.
In the last line, I used the "square root" button on my calculator to find the number. Check this and try it on your own calculator. But why have I used the minus sign? Does it satisfy the requirement (b) above? What about requirement (a)?
Why does a square root almost always have two possible values? How do they relate to one another? What square root has only one value?

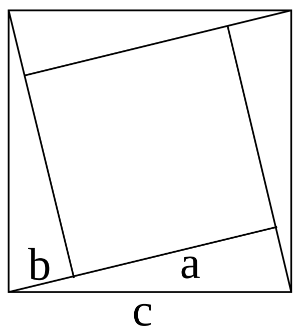
2. Write a program, say `cubeRT(a)`, which finds the cube root of a correctly to some number of decimal places which you can fix inside the program or else control by a second parameter, say `decpl`, for any number, a.

3. a) Show that slope $x^n = nx^{n-1}$ for any positive integer n. Hint: use the binomial coefficients of Week ii Note 6 in a way similar to the way they are used to expand $(1+x)^n$ in Week ii Note 7.
b) Show that slope $x^n = nx^{n-1}$ for any rational number n.

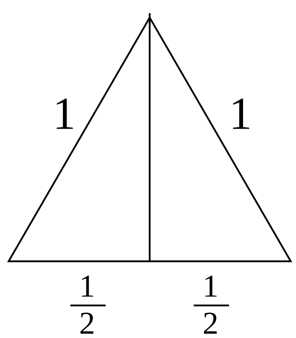
4. **Pythagoras.** In the next three excursions we are going to show that the height of an equilateral triangle of base b is indeed $(\sqrt{3}/2)b$ (i.e., $\sqrt{3}/2$ when $b = 1$ as in the example of Note 8). For this, we need to know *Pythagoras' theorem* that for a right-angled triangle of sides a, b and c (where c is the long side, the *hypotenuse*)

$$c^2 = a^2 + b^2$$

Show this using the following diagram

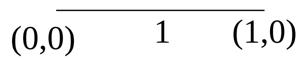


5. Using Pythagoras and the following diagram, show that the height of an equilateral triangle of base $b = 1$ is $\sqrt{3}/2$.

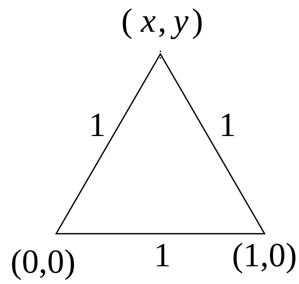


Why is the height $(\sqrt{3}/2)b$ for any base b?

6. Another way to find the height of an equilateral triangle is to start with its base, say of length 1



and find the point (x, y) which is a distance 1 from both (0,0) and (1,0).



a) Show that Pythagoras says the distance between two points (x, y) and (x', y') is

$$(x - x')^2 + (y - y')^2$$

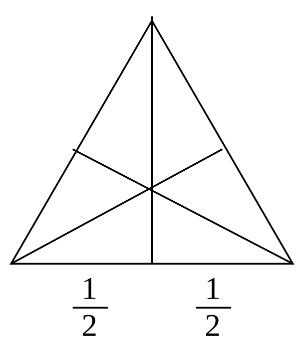
b) Hence show that the two equations

$$\begin{aligned} x^2 + y^2 &= 1 \\ (x-1)^2 + y^2 &= 1 \end{aligned}$$

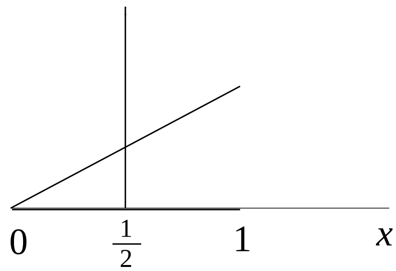
must be solved.
c) Finally, use this to show that $x = 1/2$ and $y = \sqrt{3}/2$

7. The sum of j^2, $\sum_{j=0}^{n} j^2$ where $n = \sqrt{\frac{2}{3}} \frac{1}{\Delta h}$ in Note 9, could be approximated as $\frac{1}{2} \sum j(j+1)$ for large n (small Δh). This is the sum of triangular numbers in Week i Note 1, and produces the nth tetrahedral number. Show that it gives the same answer as the volume calculation in Note 9.

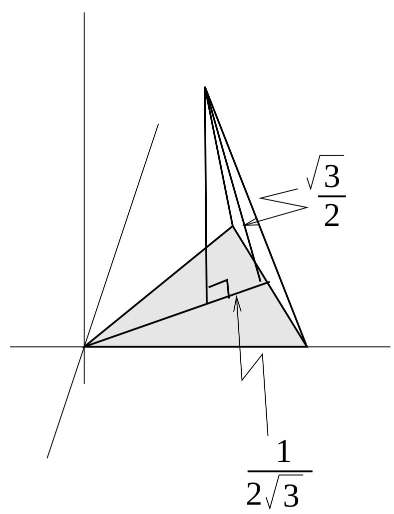
8. Pythagoras can show that the height of unit-sided tetrahedron is $\sqrt{2/3}$ as said in Note 9. Here is the geometrical approach.
The fourth point, or apex, of the tetrahedron will lie directly above the center of the equilateral triangle that is its base (why?)



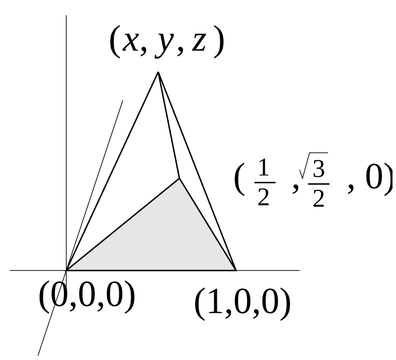
a) This can be found by bisecting the angle at each corner and noting where the bisectors meet. By symmetry, all three bisectors meet at the same point, and each is perpendicular to the edge of the triangle opposite to the angle it bisects.



Use the slope of the bisector (it's the same as the slope of the edge of the rotated triangle in Note 8) and Pythagoras to show that its length from $x = 0$ to $x = 1$ is $2/\sqrt{3}$. Since the *vertical* bisector cuts this bisector in half, show that the common point where all bisectors meet is $1/\sqrt{3}$ from each vetex and $1/(2\sqrt{3})$ from each opposite edge.
b) Now rise straight up in the third dimension from this common point to the apex of the tetrahedron, and note where it meets the 2-D "height" of the triangular face shown and use Pythagoras to show that the 3-D height is $\sqrt{2/3}$.



9. The algebraic approach finds the height of the tetrahedron much more quickly than the geometrical approach of the last excursion.



Requiring the apex, (x, y, z), to be unit distance from the three corners (0,0,0), (1,0,0) and $(1/2,\ \sqrt{3}/2,\ 0)$ of the base means

$$x^2 + y^2 + z^2 \quad = \quad 1$$

$$\begin{aligned}(x-1)^2+y^2+z^2 &= 1\\ (x-\frac{1}{2})^2+(y-\frac{\sqrt{3}}{2})^2+z^2 &= 1\end{aligned}$$

Show $(x,y,z) = \pm(\frac{1}{2},\ \frac{1}{2\sqrt{3}},\ \sqrt{\frac{2}{3}})$

10. Go on from the previous excursion to find the "height" in "the 4th dimension" of the 4-D regular simplex with tetrahedral "base"

$$(0,0,0,0), (1,0,0,0), (\frac{1}{2},\ \frac{\sqrt{3}}{2},\ 0,0) \text{ and } (\frac{1}{2},\ \frac{1}{2\sqrt{3}},\ \sqrt{\frac{2}{3}}, 0)$$

11. **Gini and IGE.** Why did the U.S.A. become so desperate as to elect a child as President? We can use this question to motivate a discussion of slopes and matrices.
A reason might be a disappointed American Dream. The Dream is the hope that we are not conditioned by birth or class but can do and become whatever we want despite our origins. It is the dream of equal opportunity.
These characterizations lead to two quantities which can be calculated and related to each other. The Gini coefficient of inequality (Corrado Gini, 1912 and 1936) is 0.34 for the U.S.: 0 means everybody is equal; 1 means extreme inequality. The IGE coefficient of social rigidity (intergenerational earnings rigidity) is 0.47 for the U.S.: 0 means freedom from birth or class; 1 means being completely locked in by your origins.
Economists have established that these two measures are strongly correlated. The American Dream is powerfully hampered by inequality.
a) Let's start with social classes. Matthew Stewart [Ste18] divides the U.S. population into the 90% the 9.9% and the 0.1% and characterizes their respective wealths as the following percentages of the total wealth of the country.

Class	Wealth
90%	20%
9.9%	60%
0.1%	20%

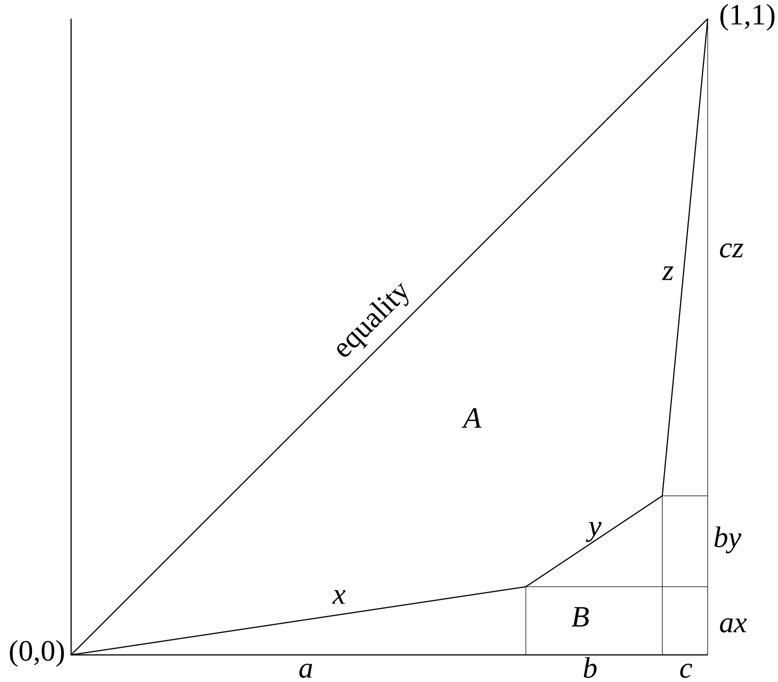
If the population was 1000 people and the wealthiest was worth 1000$ what would the others each be worth?
This data shows that the U.S. is already extremely unequal. Since wealth can be inherited (especially real estate), as can debts, this tally of wealth exaggerates the inequality. In the following we will discuss income rather than wealth.
b) Income is also unequal. Here is a table analogous to the first. I have kept the three classes but invented income figures to give the Gini coefficient measured for the U.S. at 0.34. That coefficient would normally be calculated on much more extensive income data, but my model will do two things. It will enable us to show very simply how the Gini coefficient is calculated. And it will illustrate very simply what the Gini coefficient means.

	Class	Income	
a	90%	1$	x
b	9.9%	4$	y
c	0.1%	256$	z

I have added symbols a, b, c for the population of each class and x, y, z for the income of each person in that class. (This income is in relative dollars. It might mean how much each earns every four or five minutes of the working day. The absolute values are not important in this

comparison of classes.)
The figure helps us calculate the Gini coefficient of inequality

$$\Gamma = \frac{A}{A+B} = 2A = 1 - 2B$$



A + B = 1/2: ax+by+cz = 1 a+b+c = 1
B = (a/2 + b + c) ax + (b/2 + c)by + (c/2)cz

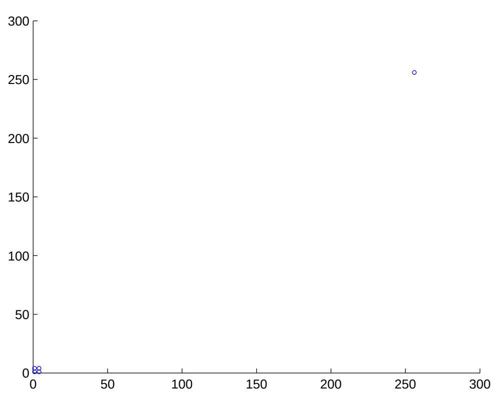
The line labelled xyz is the *Lorenz curve* (Max Lorenz, 1905). The labels x, y, z are the *slopes* of the lines they label, while the labels a, b, c are the *lengths* of the segments of the horizontal axis that they label. So, from what we know about slopes, the x line rises an amount ax over the length a, the y line rises an amount by over the length b, and the z line rises an amount cz over the length c. These rises are shown. What they mean is the total amounts earned in each class: ax for class a, by for class b and cz for class c.
Relating the figure to the above table, $a = 90$, $b = 9.9$, $c = 0.1$, $x = 1$, $y = 4$, $z = 256$. The figure is not to scale.
And finally I have forced the sum of these totals $ax + by + cz = 1$ so that the values of x, y and z must be scaled down by the amount $90 \times 1 + 9.9 \times 4 + 0.1 \times 256$. Again, only the relative amounts of x, y and z count.
The Gini coefficient of inequality is the ratio of the area A (between the line "equality" of slope 1 and the Lorenz curve) to the total area $A + B$ of the triangle (so $A + B = 1/2$) where B is the area under the Lorenz curve.
To start to understand this, note that if everybody were equal their earnings would be equal, $x = y = z$, and the requirements $a + b + c = 1$ and $ax + by + cz = 1$ then force each of these slopes to be $x = y = z = 1$. So the Gini coefficient of inequality $\Gamma = 0$, its minimum value. Thus the xyz line would coincide with the "equality" line.
On the other hand, maximum inequality would happen if $ax = 0$, $by = 0$ and $cz = 1$ so everybody gets nothing except the richest person who gets it all.
Show that the values in the table give $\Gamma = 0.34$.
What are the contributions of each class to the total income of the country (in the same way

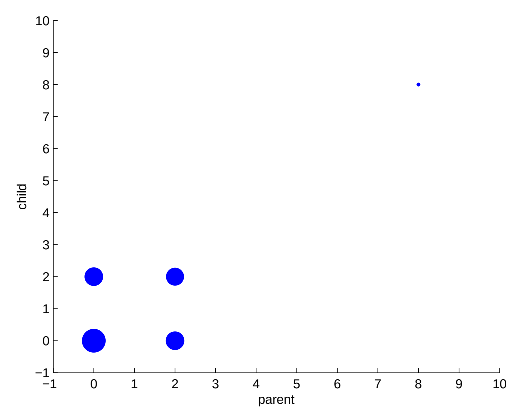
that 20%, 60% and 20$ were the respective contributions to wealth)?
c) The rigidity of social stratification can be measured, at least for incomes, by comparing a child's income with the parent's (say both at the same age). If, over a population, the child's income is unrelated to the parent's, then we have the kind of equality of opportunity that constitutes the American Dream. If, on the other hand, every child's income is identical to the parent's then the earnings classes are rigid because the child has not escaped the class.
The relatedness of child's to parent's income is calculated as the IGE (intergenerational earnings) rigidity coefficient. We look at that now.
Again I will keep the model of three income classes but I will invent figures so as to arrive at the U.S. IGE of 0.47. This simplification will again make it easy to show how to do the calculation and make it easy to understand the meaning of the coefficient.
The next table shows transitions from the parent's income class to the child's income class for a population of 1000. The distributions remain fixed: 90%, 9.9% and 0.1% of the population in each class for both generations. (And the table has an equivalent matrix. Note the row and column sums.)

p	c	#
1$	1$	845
1$	4$	55
4$	1$	55
4$	4$	44
256$	256$	1

$$\begin{matrix} 845 & 55 & \\ 55 & 44 & \\ & & 1 \end{matrix}$$

What the table means is that, of the 900 parents in salary bracket 1$, 845 children remained in salary bracket 1$ while 55 moved up to salary bracket 4$. At the same time, of the 99 parents in salary bracket 4$, 44 children remained and 55 moved down. The child of the one person in bracket 256$ stayed there.
To visualize the relationship between parent and child we should plot this data—a third representation of the same data. Here are two versions. The second is clearer.



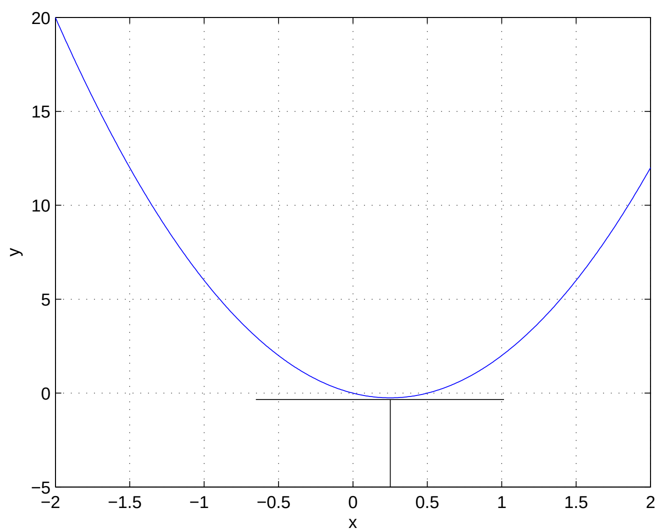


The lefthand figure plots the table as given but without any indication of the numbers of people involved. Because of the extreme range of incomes this plot is very empty.
The righthand figure plots, instead of the salary brackets, the power that 2 must be raised to in order to get that salary bracket. Thus $1 = 2^0$, $4 = 2^2$ and $256 = 2^8$. It also enlarges the plot circles to indicate the numbers of children who stay or move.
d) **Least squares fit.** One way to find out if there is a relationdship between two sets of numbers as in the table and plots above is to see what the straight line $y = ax + b$ looks like that most closely fits all the (x,y) pairs. We'll use the 0-2-8 plot and make a table with subscripts j. We'll rename to x the values corresponding with p, and those for c to y. That is, $p = 2^x$ and $c = 2^y$.

x_j	y_j	j
0	0	1:845
0	2	846:900
2	0	901:955
2	2	956:999
8	8	1000

What we must do is to minimize the differences between y_j and $ax_j + b$. To do this just once instead of for 1000 differences we'll have to sum and minimize the sum. It is not good enough to sum just $y_j - ax_j + b$ because typically some of the differences will be positive and some negative and so will tend to cancel each other out. This would reduce the effectiveness of the method. So instead we'll sum the *squares* of those differences and minimize that sum. Each square is a positive (well, non-negative) number and so minimizing the sum *of squares* will be the best way to minimize *each* of the differences. Hence the name "least squares".

e) **Minimizing.** We need to use slopes to minimize. For instance, suppose we wanted to find the minimum of a function $4x^2 - 2x$.



The picture shows that the minimum is where the slope goes to zero. Knowing how to find slopes we can calculate this exactly.

$$\begin{aligned} 0 &= \text{slope } (4x^2 - 2x) \\ &= 8x - 2 \\ x &= 2/8 = 0.25 \end{aligned}$$

f) So that is what we are going to do with the sum of the squares of the differences. First we must straighten out some possible confusion. In this least-squares minimization, the x_j and y_j are *given* numbers: they are constants. It is the a and b which are the variables. So we must minimize—and use slopes— with respect to a and b.

Worse, there are *two* of these variables to minimize with respect to. So we'll have to do two minimizations.

I'll use the Σ symbol, which we saw in Week_i and Week_ii just means sum. So $\Sigma_j x_j = x_1 + x_2 + ..$ for all the x_js.

Here's the problem: minimize $\Sigma_j (ax_j + b - y_j)^2$ with respect to a and then again with respect to b. That means setting slope$_a$ of this sum to zero and then setting slope$_b$ of the sum to zero. Check that the following correctly uses the rules for finding slopes of squares and of constant multiples.

$$\begin{aligned}
0 &= \text{slope}_a((ax_1+b-y_1)^2+(ax_2+b-y_2)^2+..) \\
&= 2(ax_1+b-y_1)x_1+2(ax_2+b-y_2)x_2+.. \\
&= 2a\Sigma_j x_j^2+2b\Sigma_j x_j-2\Sigma_j x_j y_j \\
0 &= \text{slope}_b((ax_1+b-y_1)^2+(ax_2+b-y_2)^2+..) \\
&= 2(ax_1+b-y_1)+2(ax_2+b-y_2)+.. \\
&= 2a\Sigma_j x_j+2b\Sigma_j 1-2\Sigma_j y_j
\end{aligned}$$

Remember that $\Sigma_j x_j^2$, $\Sigma_j x_j$, $\Sigma_j y_j$ and $\Sigma_j x_j y_j$ are all just constant numbers which can be calculated from the table:

$$\begin{aligned}
\Sigma_j x_j &= 900*0+99*2+1*8=206 \\
\Sigma_j y_j &= 900*0+99*2+1*8=206 \\
\Sigma_j x_j^2 &= 900*0+99*4+1*64=460 \\
\Sigma_j x_j y_j &= 955*0+44*4+1*64=240
\end{aligned}$$

And note finally that $\Sigma_j 1$ is just the count of the population: $\Sigma_j 1=1000$.
Back to the two equations.

$$\begin{aligned}
a\Sigma_j x_j^2+b\Sigma_j x_j &= \Sigma_j x_j y_j \\
a\Sigma_j x_j+b\Sigma_j 1 &= \Sigma_j y_j
\end{aligned}$$

We can rewrite them as a matrix equation.

$$\begin{pmatrix} \Sigma_j x_j^2 & \Sigma_j x_j \\ \Sigma_j x_j & \Sigma_j 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} \Sigma_j x_j y_j \\ \Sigma_j y_j \end{pmatrix}$$

and solve

$$\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} \Sigma_j x_j^2 & \Sigma_j x_j \\ \Sigma_j x_j & \Sigma_j 1 \end{pmatrix}^{-1}\begin{pmatrix} \Sigma_j x_j y_j \\ \Sigma_j y_j \end{pmatrix}$$

That's the "inverse matrix" I've written with the -1 exponent. For a 2-by-2 matrix it is easy to find.

$$\begin{pmatrix} p & r \\ q & s \end{pmatrix}^{-1}=\frac{1}{ps-qr}\begin{pmatrix} s & -r \\ -q & p \end{pmatrix}$$

Use your matrix-multiplication skill to multiply these two matrices—without the -1 on the left-hand matrix.
So now you can put in the numbers and show that

$$\begin{pmatrix} 460 & 206 \\ 206 & 1000 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} 240 \\ 206 \end{pmatrix}$$

tells us that $a=0.47$ or close enough. This is the IGE rigidity coefficient for the U.S.A.
Try plotting the line $y=0.47313x+0.10569$ on the -1 to 10 parent-child graph above. What is strange and how does it make sense in the end?
The simple, 3-class, model for the U.S. shows (i) that the ratio of worker to ChiefEO salaries is less than 1/250 (Ben Cohen and Jerry Greenfield, when founding their ice-cream company, originally idealistically decreed that this ratio should not be less than 1/7: what would the Gini inequality coefficient be in that case?) and (ii) that lower-class upward mobility is just 6% of the population.

g) For Canada, the Gini inequality coefficient is 0.29 and the IGE rigidity coefficient is 0.19. Show that the following data for our 3-class model leads to these coefficients.

	Class	Income	
a	90%	1$	x
b	9.9%	4$	y
c	0.1%	128$	z

p	c	#
1$	1$	818
1$	4$	82
4$	1$	82
4$	4$	17
128$	128$	1

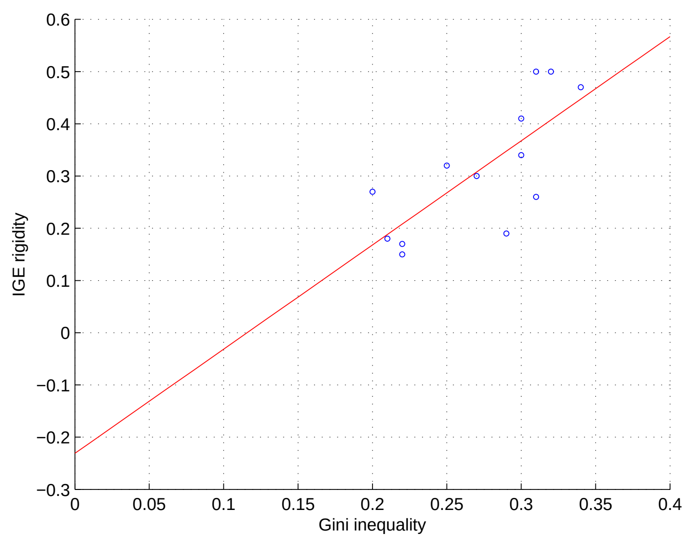
This is not anything to be proud of either.
Write a paragraph summarizing the differences between the U.S.A. and Canada.
h) To spot the relationship between inequality and social rigidity we need more countries. From [Ste18, Cor06] I extracted the following (there was one disagreement about Canada's IGE and I followed Corak, who gives numbers, rather than my interpolation on Stewart's graph: this graph appears in [Cor13] and Stewart credits Alan Krueger (2012) with pointing out the relationship he called the Great Gatsby curve).

	UK	It	US	Fr	Jp	Ger	NZ	Oz	Swe	Can	Fin	Nor	Dk
Gini	.32	.31	.34	.30	.30	.25	.27	.31	.20	.29	.21	.22	.22
IGE	.50	.50	.47	.41	.34	.32	.30	.26	.27	.19	.18	.17	.15

Using the least-squares method we developed above, with x_j now standing for the Gini data and y_j for the IGE data, show that the matrix equation

$$\begin{pmatrix} 0.9906 & 3.54 \\ 3.54 & 13 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1.1587 \\ 4.06 \end{pmatrix}$$

solves to give $ax + b = 1.9951x - 0.2310$. Here's the plot.



Where would the Ben-and-Jerry Gini fall on this line, and what model parameters would fit the resulting IGE rigidity?
i) This relationship is not necessarily causal. Inequality could cause class rigidity. Rigidity could cause inequality. Both could be caused by something else. Stewart seems to argue that inequality causes rigidity and he discusses five "goods" which characterize the 9.9% and keep it inaccessible to others: good family, good health, good education, good neighbourhood and

good jobs.
j) Is the Dream collapsing in other countries? Is there an explanation here for political developments such as the U.K. leaving the European Union or widespread election of rightist or populist governments?

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

References

[Cor06] Miles Corak. Do poor children become poor adults? Technical Report DP 1993,, The Institute for the Study of Labor (IZA), Bonn, March 2006. URL http://ftp.iza.org/dp1993.pdf (last accessed 18/10/3).

[Cor13] Miles Corak. Income inequality, equality of opportunity, and intergenerational mobility. Technical Report DP 7520,, The Institute for the Study of Labor (IZA), Bonn, July 2013. URL http://ftp.iza.org/dp7520.pdf (last accessed 18/10/3).

[Ste18] Matthew Stewart. The birth of the new American aristocracy. *The Atlantic*, 321(5):48–63, June 2018.