WARREN D. MACEVOY

LITTLE-$o()$ CALCULUS

LECTURE NOTES

PUBLISHED BY LECTURE NOTES

WWW.COLORADOMESA.EDU

First printing, May 2014

Contents

List of Figures

List of Tables

Preamble

The point of these notes is to introduce the main concepts of calculus using what is known as little-$o()$ notation, in the hope to provide a perspective that is simple, expressive, intuitive and correct.

Stitz-Zeager [www.stitz-zeager.com] and Boundless [www.boundless.com] are good backgrounds references.

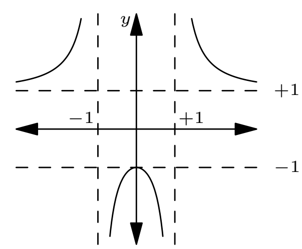

Figure 1: $y = f(x) = \frac{x^2+1}{x^2-1}$

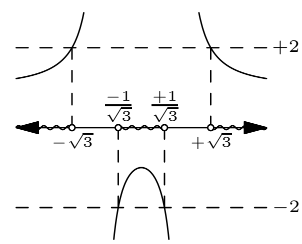

Figure 2: Where $|f(x)|$ is less than 2.

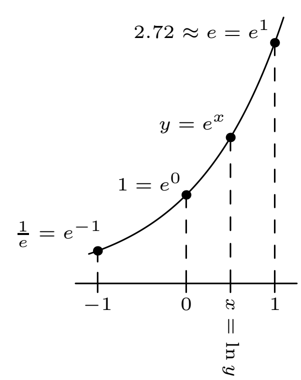

Figure 3: How $y = e^x$ and $x = \ln y$ are related, where $e \approx 2.72$ is Euler's constant.

Algebra is essential

I assume you are comfortable with algebra, including working with rational algebraic expressions, exponents and logarithms, function notation, absolute values, and inequalities.

For example, seeing

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

And being asked to show

$$f(x+h) = \frac{x^2+1}{x^2-1} + \frac{2h \cdot (2x+h)}{(x-1)(x+1)(x+h+1)(x+h-1)}$$

and

$$|f(x)| < 2 \text{ if and only if}$$
$$x \in (-\infty, -\sqrt{3}) \cup (-\frac{1}{\sqrt{3}}, +\frac{1}{\sqrt{3}}) \cup (+\sqrt{3}, +\infty)$$

should seem (at worst) tedious, but not mysterious. Neither should the following[1]:

$$2^{\log_3(x)} = x^{\log_2(3)}.$$

If these are mysterious to you, then you need to learn college algebra.

[1] Recall $b^x = e^{x \ln b}$ and $\log_b x = \ln x / \ln b$

Trigonometry is helpful

It will be useful if you have some basic trigonometry, including measuring angles in radians, which is the only unit of angle we care about here. For example,

$$(\cos a)^2 + (\sin a)^2 = 1\,,$$
$$\cos a = \frac{1}{2} \text{ if and only if } a = \pm\frac{\pi}{3} + 2\pi n, \text{ for some integer } n\,,$$

and

$$\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b)\,,$$
$$\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)\,,$$

should at least be familiar to the point of looking something up to remember the details.

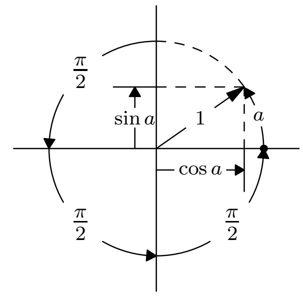

Figure 4: Defining $\cos a$ and $\sin a$ for radian angle a on the unit circle. Notice there are $2\pi \approx 6.28$ radians in a circle.

		6	1		5			
9				4		7		
8			7				2	
4								
		1				4		
				6				
	5		8		2	1		
	6				9			7
	4				6	9		5

Figure 5: A simple Sudoku puzzle. Each row, column and 3×3 sub-grid contains all of the digits from 1 to 9.

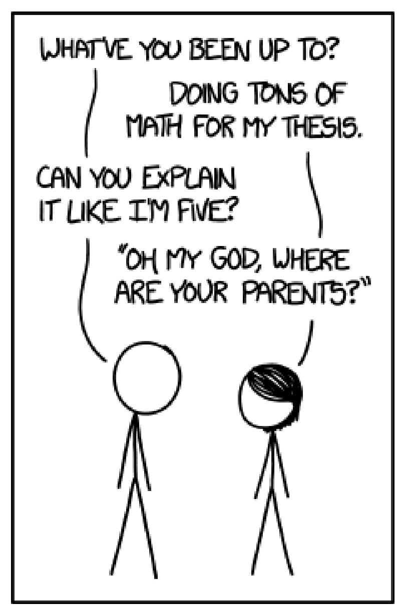

Figure 6: http://xkcd.com/1364

Reading Tips

- Focus. The best environment for learning something mathematical is in a small group willing to help you work out a Sudoku puzzle. If you can work out a Sudoku puzzle in the place and with the people you study with, it is a good sign you can learn math as well. Learning, any learning, is for mono-taskers: you might multitask (usually poorly) on things you already understand. This, however, is something you are trying to understand.

- Drive, don't walk. Learning to use a computer algebra system (like Wolfram Alpha, wxMaxima or Maple), or at least a spreadsheet (like Google Docs, Libre Office or Microsoft Excel) can spare you from a lot of tedium compared to a crummy calculator.

- Understand first. Proofs are less important than understanding. If you are mostly interested in applications, don't worry as much about understanding every proof. Concentrate instead on why the ideas make sense and how you might use these ideas. If you are interested in mathematics itself, understanding the proofs is important, but useless without an intuition of why they make sense and how to use them.

Basics

Little-$o()$ notation is used to denote that some expression is small relative to some other value.

For example, if we say:

$$(1+h)^3 = 1 + 3h + o(h) \text{ as } h \to 0\,.$$

We mean that, as the magnitude of h gets smaller, the magnitude of the difference between the left-hand side $L = (1+h)^3$ and the right-hand side $R = 1 + 3h$ is so small that, even when divided by h, it is still small.

We can check this ratio empirically:

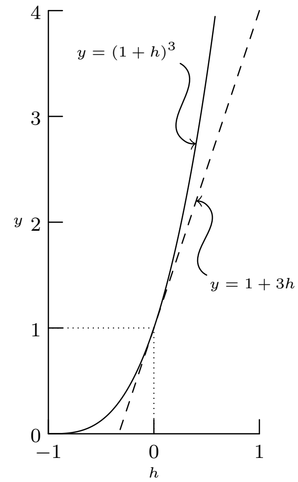

Figure 7: $(1+h)^3 = 1 + 3h + o(h)$ as $h \to 0$. Notice how the curves are *tangent* (touch and align) at $h = 0$.

Table 1: $(1+h)^3 = 1 + 3h + o(h)$ as $h \to 0$.

h	$L = (1+h)^3$	$R = 1+3h$	$E = L - R$	$r = \left\lvert\frac{E}{h}\right\rvert$
0.1	1.331	1.3	0.031	0.31
−0.01	0.970 299	0.97	0.000 299	0.0299
0.001	1.003 003 001	1.003	0.000 000 300 01	0.003 001

Notice how the ratio $r = |E/h|$ shrinks as the magnitude of h shrinks. When we say an expression E is $o(h)$ as $h \to 0$, we mean the magnitude of $|E/h|$ gets small as the magnitude of h gets small.

Graphically, smooth expressions that differ by $o(h)$ are *tangent* (touch and align) at $h = 0$. This is illustrated for this example in figure 7.

The $o()$ notation allows for more general expressions for the relative comparison of smallness, such as $o(h^2)$ or $o(h \ln(h))$. In each case, we are saying that the expression in question is small in magnitude when divided by the expression in the $o()$ as the parameter (h in this case) gets small in magnitude.

The notation works for large parameters as well. For example:

$$2k^3 + 5k^2 - 7k + 3 = 2k^3 + o(k^3) \text{ as } k \to \infty\,. \tag{1}$$

This means, as the magnitude of k gets larger and larger, the cubic polynomial on the left is approximately the leading order term (term with the highest power of k) plus an error small relative to the size of that term. Comparing figure 8 and figure 9 illustrates this trend, and table 2 gives example values as $k \to \infty$.

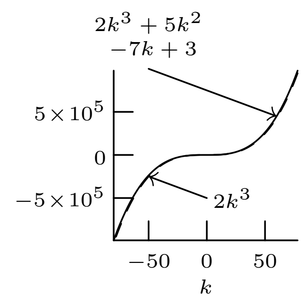

Figure 8: $2k^3 + 5k^2 - 7k + 3 = 2k^3 + o(k^3)$ as $k \to \infty$. Note how similar the curves are for large magnitude k.

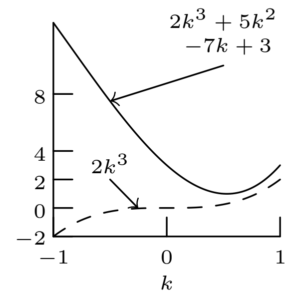

Figure 9: $2k^3 + 5k^2 - 7k + 3 = 2k^3 + o(k^3)$ as $k \to \infty$. Note how dissimilar the curves are for small magnitude k.

Table 2: $2k^3 + 5k^2 - 7k + 3 = 2k^3 + o(k^3)$. Notice how the error E grows, but it is still small when compared to $\varepsilon(k) = k^3$.

k	$L = 2k^3 + 5k^2 - 7k + 3$	$R = 2k^3$	$E = L - R$	$\varepsilon(k) = k^3$	$r = \left\lvert \frac{E}{\varepsilon(k)} \right\rvert$
10	2433	2000	433	1000	0.433 00
−100	−1 949 297	−2 000 000	50 703	−1 000 000	0.050 70
1000	2 004 993 003	2 000 000 000	4 993 003	1 000 000 000	0.004 99

Euler's constant, e.

Euler's constant $e \approx 2.718\,281\,828\,459\,045$ is fundamental for exponents and logarithms, just as $\pi \approx 3.141\,592\,653\,589\,793$ is fundamental in trigonometry.

The value e can be thought of as the value of the expression $(1 + h)^{1/h}$ as $h \to 0$. In $o()$ notation[2]:

$$(1+h)^{1/h} = e + o(h^0) \text{ as } h \to 0\,. \tag{2}$$

[2] Writing 1 as h^0 in $o(h^0)$ may seem surprising, but it is a way of noting what parameter is getting small.

This means, as the magnitude of h gets smaller and smaller, the expression on the left is approximately e, plus an error small relative to $h^0 = 1$. Figure 10 and table 3 demonstrate this. Notice $(1+h)^{1/h}$ is not defined for $h = 0$, but we only care about small *nonzero* values of h.

Curves that differ by $o(h^0)$ need only try to touch at $h = 0$ (unlike $o(h)$ errors, where they need to be tangent).

Table 3: $(1+h)^{1/h} = e + o(h^0)$.

h	$(1+h)^{1/h}$	$r = \left\lvert \frac{(1+h)^{1/h} - e}{1} \right\rvert$
0.01	2.704 813 829	0.0135
−0.0001	2.718 417 755	0.000 136
0.000 001	2.718 280 469	0.000 001 36

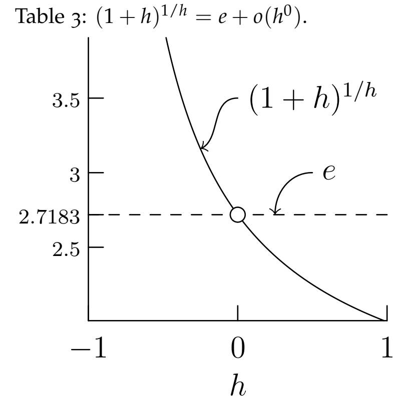

Figure 10: $(1+h)^{1/h} = e + o(h^0)$. Notice the curves only try to touch (instead of be tangent) at $h = 0$.

Sine.

As $h \to 0$, $\sin(x+h) = \sin(x) + \cos(x)h + o(h)$.

We will show this is true later, but, geometrically, it means that evaluating $y = \sin(x)$ near x is approximately a line[3] going through the point $(x, \sin(x))$ with slope $\cos(x)$.

This last example is really important. The idea that, near a given point x, many functions are well approximated by a line is a foundational idea of differential calculus. We call such a function differentiable at x, and the slope of its tangent line $f'(x)$, so that

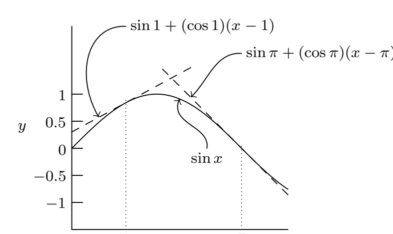

Figure 11: $\sin(x+h) = \sin(x) + \cos(x)h + o(h)$ for $x = 1$ and π.

[3] A line going through the point (x_0, y_0) with slope m is $y = y_0 + m \cdot (x - x_0)$.

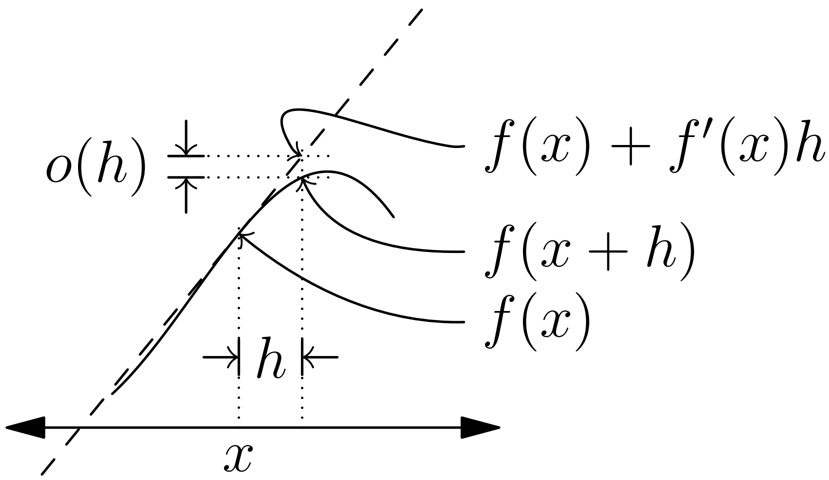

Figure 12: Differential calculus in a nutshell: often, the value of $f(x)$ near x is approximately a line: $f(x+h) = f(x) + f'(x)h + o(h)$.

$f(x+h) = f(x) + f'(x)h + o(h)$. For the example of figure 11, we are saying that, if $f(x) = \sin(x)$, then $f'(x) = \cos(x)$.

Always writing "as $h \to 0$" is tedious. It will be clear from the $o()$ notation which parameter we consider large or small.

In general, $o(h^0)$ approximations touch at $h = 0$, $o(h^1)$ approximations have the same tangent (best fitting line), and $o(h^2)$ approximations have the same curvature (best fitting circle). Figure 13 illustrates this with approximations of e^h.

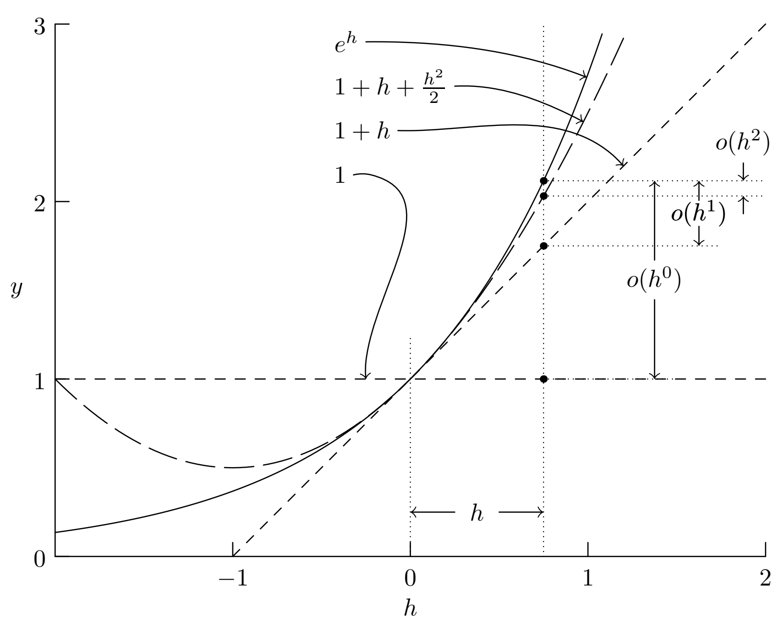

Figure 13: Three progressively better approximations of e^h as $h \to 0$:
$e^h = 1 + o(h^0)$,
$e^h = 1 + h + o(h^1)$, and
$e^h = 1 + h + \frac{h^2}{2} + o(h^2)$.
Notice the $o(h^0) = o(1)$ approximation touches, the $o(h^1) = o(h)$ approximation is tangent (best fitting line), and the $o(h^2)$ approximation has the same curvature (best fitting circle) at $h = 0$.

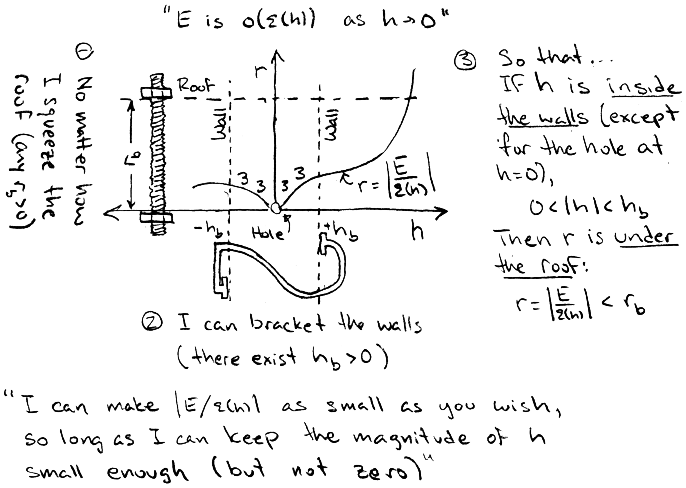

Figure 14: Sketch of the meaning of $o()$. Redraw this in your notes so you remember the definition of little-$o()$.

Formalities

An expression E is $o(\varepsilon(h))$ as $h \to 0$ when the ratio $r = \left|\frac{E}{\varepsilon(h)}\right|$ is smaller than any positive bound r_b, provided we force the magnitude of h to be small enough (but not zero).

Specifically[4] E is $o(\varepsilon(h))$ if, and only if, for any bound $r_b > 0$, there exists a bound $h_b > 0$, so that $r = |E/\varepsilon(h)| < r_b$ whenever $0 < |h| < h_b$.

We define[5] large k in terms of small h by saying E is $o(\varepsilon(k))$ as $k \to \infty$ means E is $o(\varepsilon(1/h)$ as $h \to 0$. In other words, for any bound $r_b > 0$, there exists a bound $k_b > 0$, so that $r = |E/\varepsilon(h)| < r_b$ whenever $|k| > k_b$.

Note that we divide E by $\varepsilon(h)$ to get r in the above definition for some range of nonzero values of h. So, for any expression to be $o(\varepsilon(h))$, $\varepsilon(h)$ must be defined and nonzero for some range of nonzero values of h. We therefore restrict $\varepsilon(h)$ to such admissible functions:

> For $\varepsilon(h)$ to be admissible in $o(\varepsilon(h))$ notation, there must exist some $h_0 > 0$ so that $\varepsilon(h)$ is defined (finite) and nonzero for $0 < |h| < h_0$. This way we can form the ratio $r = |E/\varepsilon(h)|$ without dividing by zero for at least some range of h.

We only use admissible $\varepsilon(h)$ in these notes. In particular, $\varepsilon(h) = |h|^p$ is admissible for any value of p , and $\varepsilon(h) = h^p$ is admissible for any integer value of p.

[4] If you are familiar with limit notation, E is $o(\varepsilon(h))$ as $h \to 0$ means $\lim_{h\to 0}\left|\frac{E}{\varepsilon(h)}\right| = 0$.

[5] If you are familiar with limit notation, E is $o(\varepsilon(k))$ as $k \to \infty$ means $\lim_{k\to\infty}\left|\frac{E}{\varepsilon(k)}\right| = 0$.

Summary

- Little-$o()$ notation is a way of describing an expression as small in comparison to some other value:

  E is $o(\varepsilon(h))$ as $h \to 0$ means the ratio $r = |E/\varepsilon(h)|$ can be made as small as desired provided the magnitude of h is small enough (but not zero).

- Graphically, expressions that differ by $o(h^0)$ need to touch at $h = 0$, but expressions that differ by $o(h^1)$ also need to be tangent at $h = 0$.

- Specifically, E is $o(\varepsilon(h))$ as $h \to 0$ means:

  For any $r_b > 0$, there exists $h_b > 0$, so that:

  $$\text{if } 0 < |h| < h_b\,, \text{ then } r = \left|\frac{E}{\varepsilon(h)}\right| < r_b\,.$$

- E is $o(\varepsilon(k))$ as $k \to \infty$ means

  E is $o(\varepsilon(1/h)$ as $h \to 0$.

  or

  For any $r_b > 0$, there exists $k_b > 0$, so that:

  $$\text{if } |k| > k_b\,, \text{ then } r = \left|\frac{E}{\varepsilon(h)}\right| < r_b\,.$$

- For differentiable functions, the value $f(x+h)$ near a given point x is well approximated by a line called the tangent line of $y = f(x)$ at x with slope $f'(x)$.

- $(1+h)^{(1/h)} = e + o(h^0)$ as $h \to 0$, where $e \approx 2.718$ is Euler's constant.

- Because an expression E is divided by $\varepsilon(h)$ in the definition of $o(\varepsilon(h))$, $\varepsilon(h)$ is admissible in $o(\varepsilon(h))$ notation only if it is defined and nonzero for small enough nonzero h.

  For $\varepsilon(h)$ to be admissible in $o(\varepsilon(h))$, there must exist $h_0 > 0$ so that $\varepsilon(h)$ is defined (finite) and nonzero for $0 < |h| < h_0$.

- We only use admissible $\varepsilon(h)$ in these notes. In particular, $\varepsilon(h) = |h|^p$ is admissible for any value of p, and $\varepsilon(h) = h^p$ is admissible for any integer value of p.

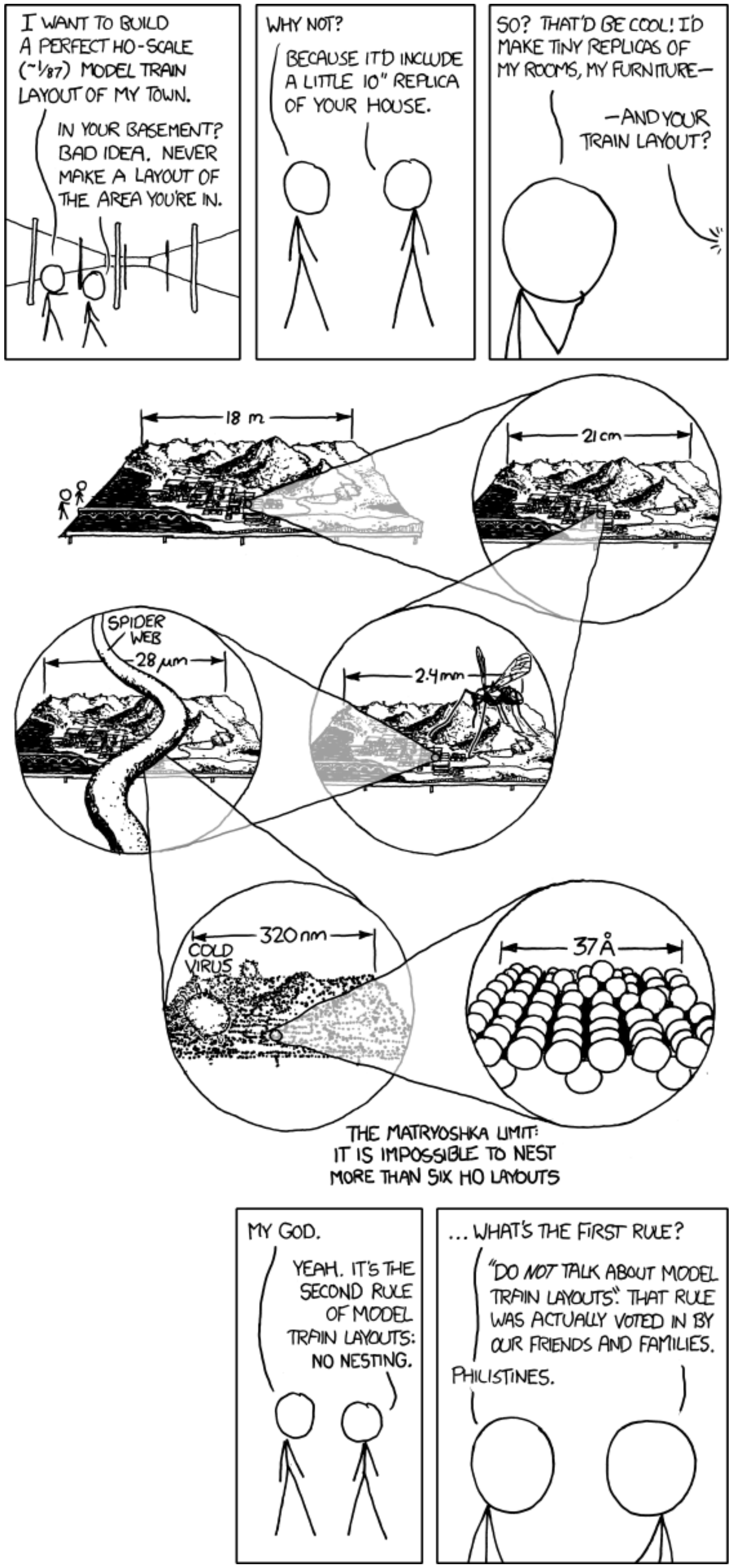

Figure 15: http://xkcd.com/878 Math is not limited by physics, however.

Index