Calculus Fundamentals using Little-o()

Why I want to read this.

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.

Why I don't want to read this.

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 asking to show

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

and

$|f(x)| < 2$ 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:

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

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

Trigonometry is helpful.

Additionally, it would be nice if you have some basic trigonometry, including measuring angles in radians, which is the only unit of angle we care about here. For example,

$cos(x) = \pm\sqrt{1-(sin(x))^2}$

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

and

$$sin(x+y) = sin(x)cos(y) + cos(x)sin(y)$$

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

Reading Tips

Distractions make this difficult.

The best environment for learning something mathematical is in a small group willing to help you work out a sudoku puzzle. If you can't work out a sudoku puzzle in the place and with the people you study with, you will not learn mathematics. You are also not a multi-tasker: you might multitask (usually poorly) on things you already understand; this is something you are trying to understand.

Drive, don't walk.

Learning to use a computer algebra system can spare you from a lot of tedium compared to a crummy calculator. Wolfram Alpha, Maxima, Maple, and Mathematica are all good choices.

Understand first.

Proofs are less important than understanding. If you are mostly interested in the applications of calculus, don't worry as much about understanding every proof. Concentrate instead on understand why the facts that are shown make sense and how you mights use these facts. If you are interested in mathematics itself, understanding the proofs are important, but pretty useless without an understanding of why they make sense and how to use them.

References

Stitz Zeager [www.stitz-zeager.com]: algebra, trigonometry, and precalculus
Boundless [www.boundless.com]: on
Algebra - https://www.boundless.com/algebra
Trigonometric Animation - http://www.analyzemath.com/unitcircle/unit_circle_applet.html
Calculus - https://www.boundless.com/calculus

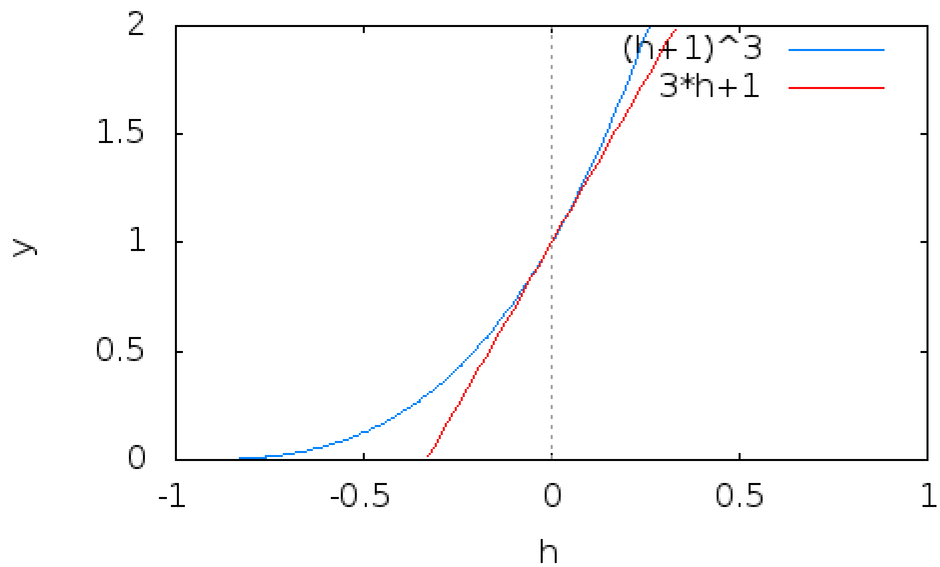
Little-o() notation

Little-o notation is used to denote that some expression is small relative to some other value. For example, if we say, as $h \to 0$ (as h approaches zero),

$$(1+h)^3 = 1 + 3h + o(h).$$

We mean that 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.

A plot $L = (1+h)^3$ vs. $R = 1 + 3h$ for small values of h :



We can check this ratio empirically:

h	$L = (1+h)^3$	$R = 1 + 3h$	$E = L - R$	$r = \left\|\frac{E}{h}\right\|$
0.1	1.331	1.3	0.031	0.31
-0.01	0.970299	0.97	0.000299	0.0299
0.001	1.003003001	1.003	0.00000030001	0.003001

Notice how the ratio r 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.

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 relative to the expression in the $o()$ as the parameter (h in this case) gets small in magnitude.

Here are some other examples to help get the idea. You should check these with a calculator or spreadsheet. We give a formal definition of the notation at the end of this section.

As $k \to \infty$, $2k^3 + 5k^2 - 7k + 3 = 2k^3 + o(k^3)$.

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.

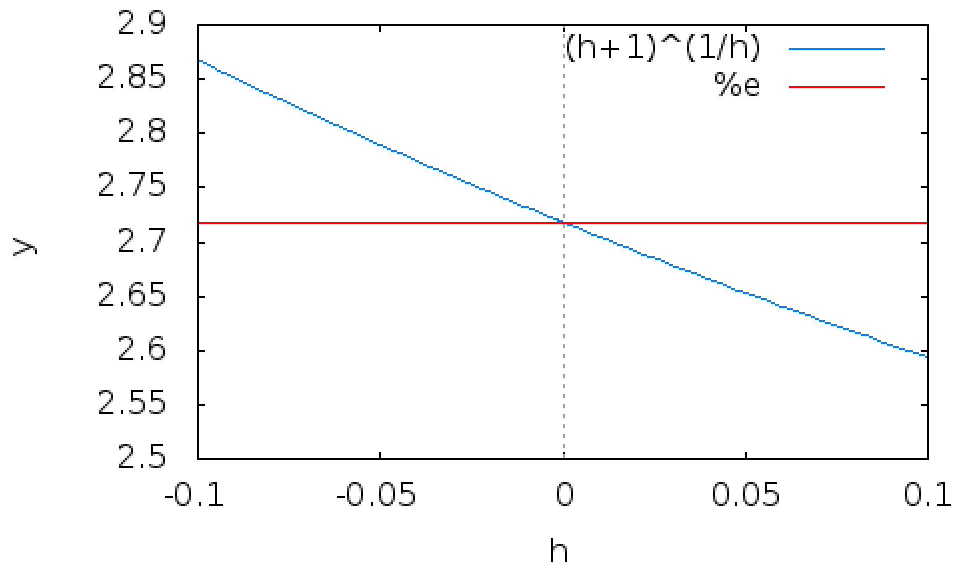
k	$L = 2k^3 + 5k^2 - 7k + 3$	$R = 2k^3$	$E = L - R$	$\varepsilon(k) = k^3$	$r = \left\lvert \frac{E}{\varepsilon(k)} \right\rvert$
10	2,433	2,000	433	1,000	0.43300
-100	-1,949,297	-2,000,000	50,703	-1,000,000	0.05070
1,000	2,004,993,003	2,000,000,000	4,993,003	1,000,000,000	0.00499

Notice how the error E grows, but it is still small *when compared to* $\varepsilon(k) = k^3$.

As $h \to 0$, $(1+h)^{(1/h)} = e + o(h^0)$.

This means, as the magnitude of h gets smaller and smaller, the expression on the left is approximately Euler's constant $e = 2.718\ 281\ 828\ 459\ 045..$, plus an error small relative to $h^0 = 1$. *Just as π plays a fundamental role in trigonometry, e is fundamental for exponents and logarithms.*

A plot of $(1+h)^{1/h}$ vs $e \approx 2.718$ for a range of small (but nonzero) values of h :



Here is a tabular comparison for some small values of h :

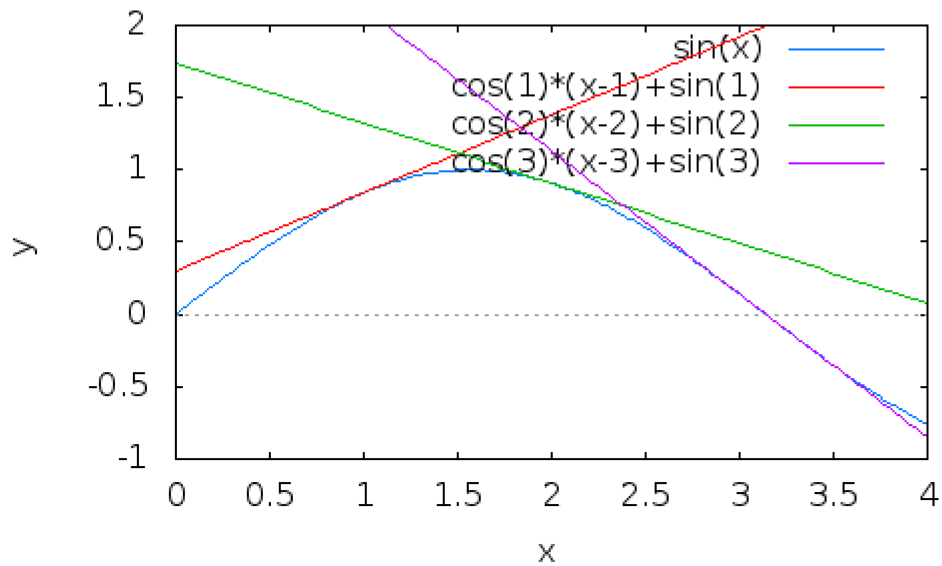
h	$L = (1+h)^{1/h}$	$R = e$	$E = L - R$	$\varepsilon(h) = h^0$	$r = \left\lvert \frac{E}{\varepsilon(k)} \right\rvert$
0.01	2.704813829	2.7182818...	-0.0135	1	0.0135
-0.0001	2.718417755	2.7182818...	0.000136	1	0.000136
0.000001	2.718280469	2.7182818...	-0.00000136	1	0.00000136

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

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 $sin(x)$ near x is approximately a line going through the point $(x, sin(x))$ with slope $cos(x)$.

Here is a plot of these approximations for $sin(x)$ at $x = 1, 2$, and 3 :



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 the slope of that approximating line $f'(x)$, so that $f(x+h) = f(x) + f'(x)h + o(h)$. For this example, we are saying that, if $f(x) = sin(x)$, then $f'(x) = cos(x)$.

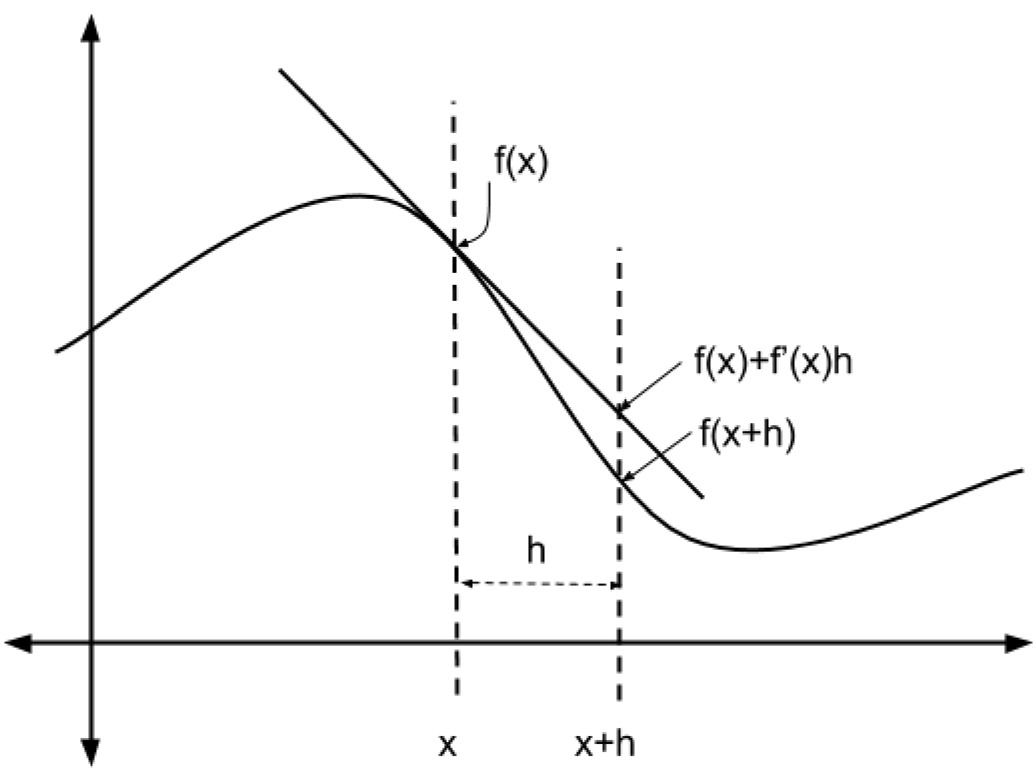


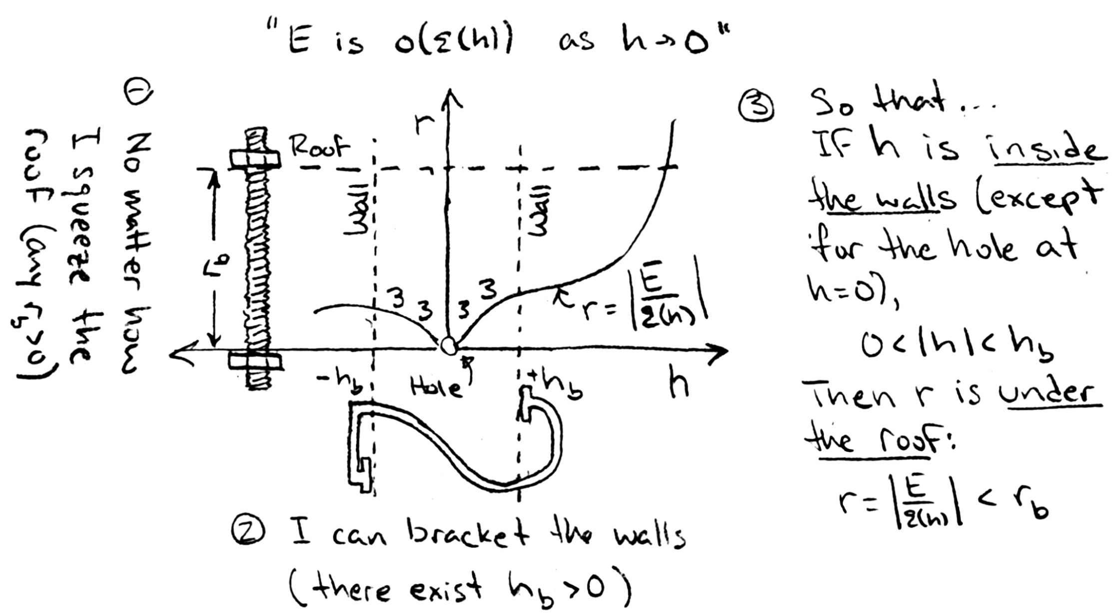
Figure 1. The most important figure in calculus. For many functions, the value near a given

point x is well approximated by a line called *the tangent line of* $f(x)$ *at* x *with slope* $f'(x)$. *As an equation: as* $h \to 0$, $f(x+h) = f(x) + f'(x)h + o(h)$. *Such functions are called differentiable at* x.

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

Formalities

An expression E is $o(\varepsilon(h))$ as $h \to 0$, means the ratio $r = \left|\frac{E}{\varepsilon(h)}\right|$ is as small as we choose, provided we force the magnitude of h to be small enough (but not zero). Specifically, for any bound $r_b > 0$, there exists an $h_b > 0$, so that, if $0 < |h| < h_b$, then $\left|\frac{E}{\varepsilon(h)}\right| < r_b$.



"I can make $|E/\varepsilon(h)|$ as small as you wish, so long as I can keep the magnitude of h small enough (but not zero)"

Redraw this in your notes so you remember the definition of little-o.

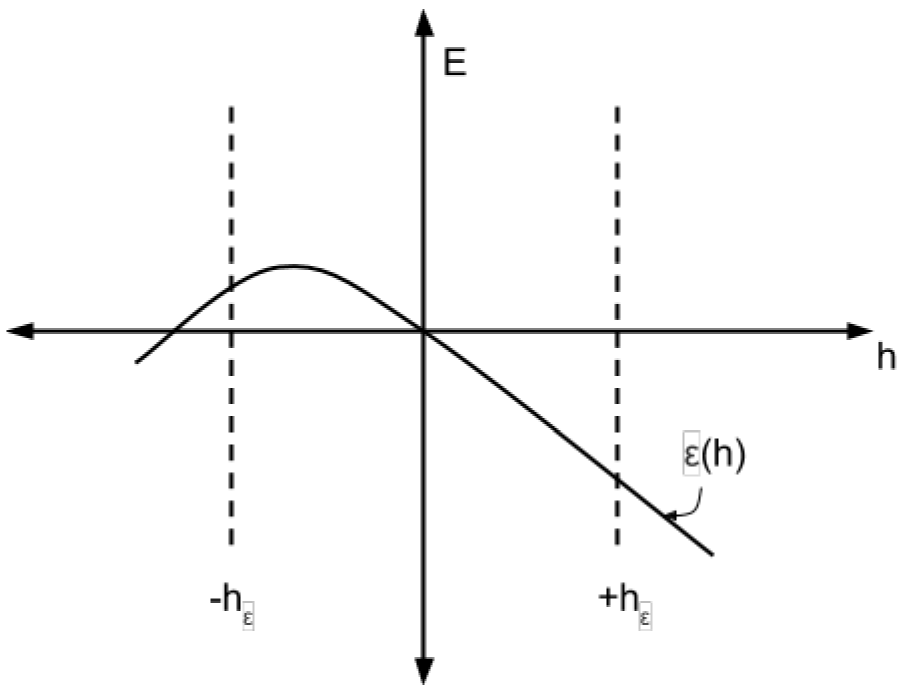
If you are familiar with limit notation, this can be written as $\lim_{h \to 0} \left|\frac{E}{\varepsilon(h)}\right| = 0$.

In terms of large parameters, $E = o(\varepsilon(k))$ as $k \to \infty$ is the same as $E = o(\varepsilon(1/h))$ as $h \to 0$ by means of the substitution $h = 1/k$.

If you are familiar with limit notation, this can be written as $\lim_{k\to\infty} \left|\frac{E}{\varepsilon(k)}\right| = 0$.

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 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 $h_\varepsilon > 0$ so that $\varepsilon(h)$ is defined (finite) and nonzero for $0 < |h| < h_\varepsilon$.



> All we ask of $\varepsilon(h)$ to be admissible, is that there is a range near zero (but we don't care about at zero), where it is defined (finite) and nonzero. This way we can divide E by $\varepsilon(h)$ in this range to compute the ratio $r = |E/\varepsilon(h)|$.

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.

Summary

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

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

  For any $r_b > 0$, there exists $h_b > 0$, so that:
  If $0 < |h| < h_b$, then $\left|\frac{E}{\varepsilon(h)}\right| < r_b$.

- E is $o(\varepsilon(k))$ as $k \to \infty$ means E is $o(\varepsilon(1/h))$ as $h = 1/k \to 0$.

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

- As $h \to 0$, $(1+h)^{(1/h)} = e + o(h^0)$, 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 $o(\varepsilon(h))$ only if it is defined and nonzero for small enough nonzero h: there must exist $h_\varepsilon > 0$ so that $\varepsilon(h)$ is defined (finite) and nonzero for $0 < |h| < h_\varepsilon$.

- 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.

Useful General Facts about o() notation

Zero is small.

For any admissible $\varepsilon(h)$, 0 is $o(\varepsilon(h))$.

It is sometimes handy to think of an exact relationship as an approximation.

Two useful examples of this we will use later are:

For constant functions, $f(x) = c$, then

$$\begin{aligned} f(x+h) &= c \\ &= f(x) + 0h + o(h). \end{aligned}$$

For linear functions, $f(x) = mx + b$, then

$$\begin{aligned} f(x+h) &= m(x+h)+b \\ &= mx+b+mh \\ &= f(x)+mh \\ &= f(x)+mh+o(h). \end{aligned}$$

Proof. Since $\varepsilon(h)$ is admissible, there is an $h_\varepsilon > 0$ so that $\varepsilon(h) \neq 0$ so long as $0 < |h| < h_\varepsilon$. For any $r_b > 0$, by choosing $h_b = h_\varepsilon$, the ratio $r = \left|\frac{0}{\varepsilon(h)}\right| = 0 < r_b$ whenever $0 < |h| < h_b$.

Higher powers of h are small compared to lower powers of h.

Polynomials and polynomial approximations are common, which makes integer powers of h the most important kinds of terms to approximate. So this is the most important relative comparison to use and understand.

$|h|^q$ is $o(|h|^p)$ for any $p < q$. The $|\cdot|$ are optional for integer values of p or q.

An example of this would be $(x+h)^2 = x^2 + 2xh + h^2 = x^2 + 2xh + o(h)$.

Proof. For any $r_b > 0$, choose $h_b = r_b^{1/(q-p)}$, then
whenever $0 < |h| < h_b$, $r = \left|\frac{|h|^q}{|h|^p}\right| = |h|^{q-p} < h_b^{q-p} = r_b$.

Notice that $q-p$ must be non-zero for h_b to exist and positive for $|h|^{q-p} < h_b^{q-p}$ to be true whenever $0 < |h| < h_b$.

The product of a bounded thing and a small thing is small.

If $E = ab$, where $|a|$ is bounded for small enough non-zero h, and b is $o(\varepsilon(h))$, then E is $o(\varepsilon(h))$.

For example, xh^3 is $o(h^2)$, $\frac{h}{x} = \frac{1}{x}h$ is $o(h^0)$ when $x \neq 0$, and $(x+h)h^2$ is $o(h)$.

Proof.
First, "$|a|$ is bounded for small enough non-zero h" means
there exists an $A > 0$ and $h_0 > 0$, for which $|a| < A$ whenever $0 < |h| < h_0$.

For any Loading... since Loading... is Loading... we can choose Loading...so that Loading... whenever Loading... By choosing Loading...then Loading... whenever Loading...

If Loading... are Loading..., and Loading... is independent of Loading... then Loading...is Loading....

For example,

Loading...

Loading...

Loading...

Loading...

Proof. For any Loading... “Loading... is Loading...” means there is an Loading... so that Loading... whenever Loading... By choosing Loading..., then Loading... whenever Loading...

Something smaller is something small, version 1

If Loading... for small enough non-zero Loading... and Loading...is Loading... then Loading... is Loading...

For example,

Loading... and Loading... is Loading... so Loading...is Loading...

Proof. Loading... for small enough non-zero Loading...” means there is an Loading... so that Loading... whenever Loading...

For any Loading... “Loading... is Loading...” means there is an Loading... so that Loading... whenever Loading... By choosing Loading... then Loading... whenever Loading...

Something smaller is something small, version 2

If Loading... is Loading...and Loading...is Loading... then Loading... is Loading...

For example,

Loading...Loading...

Loading...

Loading...

Loading...

Proof. Loading... is Loading...means there is an Loading... so that Loading... or Loading... whenever Loading... By version 1 of something smaller is something small, Loading... is Loading...

Products of small things are small, version 1.

Saying Loading... is the same as Loading...provided Loading... for small enough non-zero Loading...

For example

Loading...
Loading...
Loading...

Proof. Write Loading... where Loading... is Loading... "Loading... for small enough non-zero Loading..." means there is an Loading... so that Loading...whenever Loading... For any Loading... there is an Loading... so that Loading... By choosing Loading... then Loading...

Products of small things are small, version 2.

If Loading... is Loading... then Loading... is Loading...

For example,

Loading...
Loading...
Loading...

Proof. Saying **Loading...** means Loading... where Loading... and Loading... For any Loading..., we can choose Loading... small enough to make Loading...for all Loading... and can choose Loading... small enough to make Loading... for all Loading... By choosing Loading... then Loading... for all Loading...

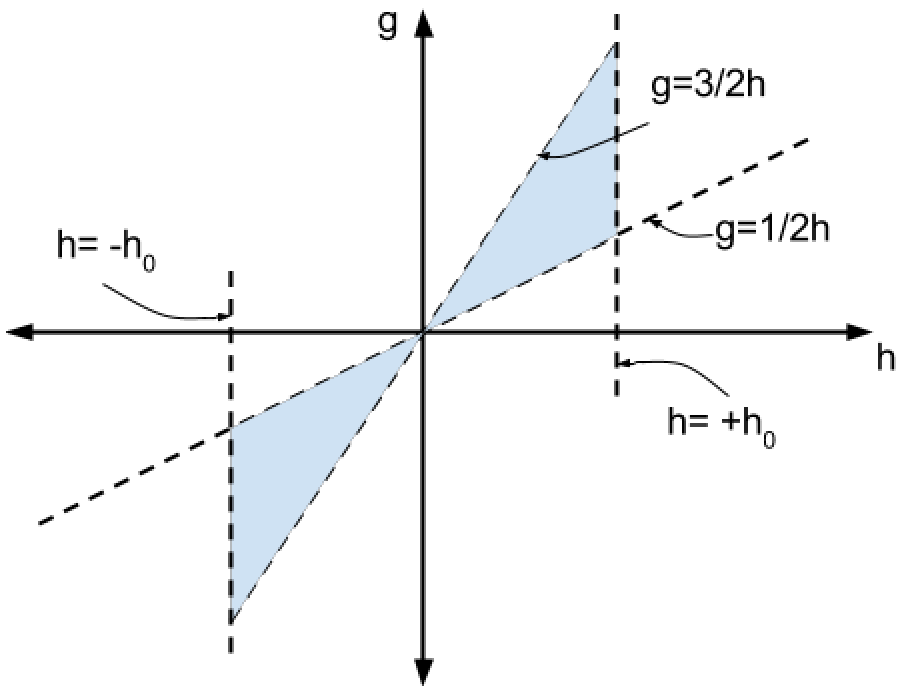
For Loading... the small parameter can change a small amount.

If Loading...and Loading...then Loading...

For example,

If Loading...then Loading...

Proof. Loading... means there is an Loading... for which Loading... for Loading... Solving this inequality gives us bounds on Loading...in terms of Loading...as in the following diagram (Loading...must be in the blue region when Loading...



For any Loading... there exists an Loading... so that Loading.... Choose Loading... then Loading...

For Loading... the small parameter can be scaled.

If Loading...and Loading...for some Loading... then Loading...

For example,
Loading... provided Loading...

Proof. For any Loading... there is an Loading... so that Loading.... Choose Loading... then Loading...

Summary

- Zero is small:

  For any admissible Loading... Loading... is Loading...

- Higher powers are smaller than lower powers:

If Loading... then Loading... is Loading...
where Loading... are optional for integer values of Loading... or Loading...

- Finite multiples of small things are small:

  If Loading... whenever Loading... for some positive bounds Loading... and Loading...
  then Loading... is Loading...

- Finite sums of small things are small:

  Loading... is Loading...

- Smaller things are small, version 1:

  If Loading..., whenever Loading... for some positive bound Loading...
  and Loading...is Loading... then Loading... is Loading...

- Smaller things are small, version 2:

  If Loading... is Loading...,
  and Loading...is Loading... then Loading... is Loading...

- Products of small things are small, version 1:

  If there exists Loading... so that Loading... whenever Loading...then
  Loading... is Loading... is equivalent to Loading...is Loading...

- Products of small things are small, version 2:

  If Loading... is Loading... then Loading...isLoading...

- For Loading... the small parameter can change a small amount:

  If Loading...and Loading...then Loading...
  where the Loading... are optional for integer values of Loading...

- For Loading... the small parameter can be scaled: