

How to Think About Exponentials

www.3blue1brown.com

March 4, 2015

This article is written for anyone who watched my video on why $e^{\pi i} = -1$ and wishes for some more details on exponential functions.

1 Recap of the $e^{\pi i} = -1$ Video

In this video, I aimed to provide an intuition that is understandable even to those who have not studied a lot of math. In particular, I wanted it to be calculus-free. My approach requires viewers to reframe how they view numbers, addition, multiplication and exponentiation as follows:

1. Every real number is simultaneously three things: A point on the infinitely extending line, an “adder”, and a “multiplier”. An adder is an action which slides that line along itself, and a multiplier is an action which stretches or squishes that line¹. In the case of complex numbers, just replace the word “line” with “plane”. Note, this is *not* to say that adders and multipliers are just a way to visualize what addition and multiplication do by moving around numbers on the line. If anything you should think of numbers as *primarily* being movements of the pure unlabeled line (or plane), and associating them with points on the line is just a convenient way to record the specific movement of each adder/multiplier².

This allows us to think of operations such as addition, multiplication and exponentiation as applying specifically to adders or multipliers, not just numbers.

2. Addition is an operation which takes in two adders and outputs an adder. Specifically, if you successively apply two adders, the composite action is the same as that of some other adder. This other adder is, by definition, their sum.
3. Likewise multiplication take in two multipliers and outputs a multiplier, defined to be the composite action of the two input multipliers.
4. Exponentiation takes in an adder, and outputs a multiplier, and it does this in such a way that adding inputs is the same as multiplying outputs. That is to say, it is a function $f(x)$ that satisfies

¹If we wanted to be both more pure and more confusing, we could say multipliers don't really act on the line (or plane), but they instead transform adders to adders, preserving addition in the sense that $a(b+c) = ab+ac$. In fact, with real and complex numbers they encompass all the possible ways to map adders to adders while preserving addition in this way. So in short we have points, functions of points (adders), and functions of functions of points (multipliers), and each number is an over-loaded term referring to any one of these three things.

²Choose an arbitrary point on the line and call it zero, then each adder is associated with the point where it takes zero. If you think of multipliers as acting on adders, as in the previous footnote, then you choose an arbitrary non-zero adder to be called “one”, and each multiplier is associated with the adder where it takes one. So in some sense the “purest” form of a number is its capacity as a multiplier, since associating it with either an adder or a point requires making arbitrary choices.

$f(x + y) = f(x)f(y)$ where x and y can be any numbers, or really any object where addition and multiplication make sense³.

- (a) For any function satisfying this property, we write it as a^x where $a = f(1)$. This terrible notation⁴ stems from the relationship these function have with repeated multiplication, which has traditionally defined exponentiation. If x is a counting number, i.e. $1, 2, 3, \dots$, we can write it as $\underbrace{1 + \dots + 1}_{x \text{ times}}$, hence the property that $f(x + y) = f(x)f(y)$ implies

$$f(x) = f\left(\underbrace{1 + \dots + 1}_{x \text{ times}}\right) = \underbrace{f(1) \cdot \dots \cdot f(1)}_{x \text{ times}}$$

- (b) One of these functions is in some sense the “most natural”, and we write it either as $\exp(x)$ or e^x . Note, this just means the number e is, by definition, the value of this special function at 1, not that e is some magic number which gives the function its specialness. It is defined explicitly with the following infinite sum:

$$e^x := 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots + \frac{1}{n!}x^n + \dots$$

This series is almost always derived for students using calculus, with Taylor’s formula, the fact that e^x is its own derivative, and in many cases a fair bit of handwaving⁵. In the following section I would like to go over a way one might derive it simply by seeking a function which satisfies $f(x + y) = f(x)f(y)$ and trying to keep things natural.

2 Constructing the series for e^x

Imagine you didn’t know the series defining e^x , and you were asked to define a function $f(x)$ with the property that $f(x + y) = f(x)f(y)$ for all number x and y . How would you go about constructing such a thing? One silly function satisfies this property is $f(x) = 0$, (which corresponds to 0^x), so to ensure our answer is interesting let’s state from the get go that we want a *non-zero* function. If you only know how to add and multiply numbers, the only functions you can define explicitly are polynomials, so we will begin by assuming f is of the form⁶.

$$f(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n$$

For some constants c_0, \dots, c_n . To start, adding 0 to the input should not change the output, which by our property means $f(x) = f(x+0) = f(x)f(0)$, so presuming $f(x)$ is not always 0, this implies $f(0) = 1$. Plugging 0 in for x above, this implies $c_0 = 1$.

³For those of you into fancy terminology, we could say “exponential” is synonymous with “homomorphism from a ring’s additive group to its multiplicative group”

⁴I say this is terrible notation for a couple reasons. First of all, it confuses people into thinking that repeated multiplication defines exponentiation, when in reality it arises as a consequence in a specific case (though historically we did, of course, start thinking about exponentiation as repeated multiplication). Secondly, it leaves certain functions ambiguous, such as i^x . Have you ever wondered what i^i is? Most people say it is $e^{2\pi n + \pi/2}$ for any integer n . What the heck is that supposed to mean? It has infinitely many values? I think a much better understanding of what is going on is to say there are infinitely many function f satisfying $f(x + y) = f(x)f(y)$ and also $f(1) = i$, and the choice for which of these functions i^x is supposed to represent determines what i^i should be.

⁵“Handwaving”, for those of you who don’t know, is the playful term to describe the act of skipping over crucial details in an explanation while trying to play them off as obvious or unimportant.

⁶Of course, we know that our answer will ultimately look like an infinite polynomial, or “power series”, but for now let’s avoid having to make sense of infinite sums.

Next, let's write $f(x+y) = f(x)f(y)$ in terms of the polynomial, but for now only look at terms with exponents ≤ 1 .

$$1 + c_1(x+y) + \dots = (1 + c_1x + \dots)(1 + c_1y + \dots)$$

No matter what c_1 is, each side will look like $1 + c_1x + c_1y +$ (some higher order terms). This means we can freely choose what c_1 will be! Don't let this fool you into thinking our function is not special, this will be the only choice we make, and from here on out our hand is forced. This choice completely determines which exponential our function $f(x)$ will be⁷. The only numbers which don't somehow feel arbitrary are 0 and 1. Choosing 0 would result in the function $f(x) = 1$, which corresponds with 1^x , so to get a function which is both natural and interesting we choose $c_1 = 1$.

Next, we expand the expression $f(x+y) = f(x)f(y)$ a little further to look at the quadratic terms, i.e. x^2, xy and, y^2

$$\begin{aligned} \dots + c_2(x+y)^2 + \dots &= (1 + x + c_2x^2 + \dots)(1 + y + c_2y^2 + \dots) \\ \dots + c_2x^2 + 2c_2xy + c_2y^2 + \dots &= \dots + c_2y^2 + xy + c_2x^2 + \dots \end{aligned}$$

The x^2 and y^2 terms are guaranteed to match up, since on either side they each have a single c_2 coefficient, but for the xy terms to match up on either side we must have $2c_2 = 1$. This implies $c_2 = \frac{1}{2}$.

To help build the intuition for the ensuing pattern, let's do one more term directly. When we only look at the cubic terms, x^3, x^2y, xy^2, y^3 , we get the following:

$$\begin{aligned} \dots + c_3(x+y)^3 + \dots &= (1 + x + \frac{1}{2}x^2 + c_3x^3 + \dots)(1 + y + \frac{1}{2}y^2 + c_3y^3 + \dots) \\ \dots + c_3x^3 + 3c_3x^2y + 3c_3xy^2 + c_3y^3 + \dots &= \dots + c_3y^3 + \frac{1}{2}xy^2 + \frac{1}{2}x^2y + c_3y^3 + \dots \end{aligned}$$

All these terms will match up so long as $3c_3 = \frac{1}{2}$, so we must have $c_3 = \frac{1}{2 \cdot 3} = \frac{1}{6}$.

Continuing in this fashion, one will quickly find that at each step we require $nc_n = c_{n-1}$, and hence that each term must be $\frac{1}{n(n-1)(n-2)\dots(2)(1)} = \frac{1}{n!}$. Of course, one would still have to verify that once you are forced into these coefficients, every term will indeed match up. That is to say, why will the $x^m y^k$ term of $f(x+y)$ always be the same as that of $f(x)f(y)$ no matter what m and k are? Expanding the polynomials for $f(x)f(y)$, this term will simply be $\frac{1}{m!} \frac{1}{k!} x^m y^k$, and for $f(x+y)$ this term will come from the expansion of the expression $(x+y)^{m+k}$, where those familiar with binomial expansion will know that it is $\frac{1}{(m+k)!} \binom{m+k}{m} x^m y^k = \frac{1}{m!k!} x^m y^k$.

I would like to stress one more time that we should think of exponentials purely in terms of the property $f(x+y) = f(x)f(y)$, not in terms of the infinite sum we use to define them explicitly. Such formula should be used only for explicit computation (this is, for instance, what your calculator does), and to prove that functions satisfying $f(x+y) = f(x)f(y)$ for all x and y exist in the first place.

⁷In fact, we will have $f(x) = a^x$ for $a = \ln(c_1)$. The fact that exponentials like i^x require a choice for what function they represent reflects that fact that in the complex plane, one must make a choice for how to define $\ln(x)$

3 Infinite Sum

The only problem with the derivation process in the previous section is that you never stop. At no point can we say c_n is 0 for n larger than some amount. We end up saying that e^x is defined as an *infinite* sum:

$$e^x := 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots + \frac{1}{n!}x^n + \cdots$$

But what does that mean? How could you add infinitely many things? At some point I would like to make a video on how to think about infinite sums, but for now I will say two things:

1. Every irrational number you write down is secretly an infinite sum, so I think most people will have some intuitive comfort with what such a sum has to be. For instance, when you write π as 3.14159265357... with the understanding that there are infinitely many more digits, what you are really saying is that

$$\pi = 3 + \frac{1}{10} + \frac{4}{10^2} + \frac{1}{10^3} + \frac{5}{10^4} + \frac{9}{10^5} + \cdots$$

where the sum continues infinitely. That said, many people are also very uncomfortable with the fact that $0.9999\cdots = 1$, so this intuition often needs to be sharpened.

2. Instead of treating expressions of the form

$$c_0 + c_1x + c_2x^2 + \cdots + c_nx^n + \cdots$$

as something you actually have to add up, you could instead view them as a way to treat infinite sequences of numbers such as (c_0, c_1, \dots) as objects in and of themselves which can still be added and multiplied together. In this sense, the terms x^n act as indices for this collection rather than as components of a function. Addition and multiplication work exactly the way you would expect them to. So addition looks like this,

$$\begin{aligned} f(x) &= a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots \\ + \quad g(x) &= b_0 + b_1x + b_2x^2 + b_3x^3 + \cdots + b_nx^n + \cdots \end{aligned}$$

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n + \cdots$$

and multiply works like this⁸:

$$\begin{aligned} f(x) &= a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots \\ \times \quad g(x) &= b_0 + b_1x + b_2x^2 + b_3x^3 + \cdots + b_nx^n + \cdots \end{aligned}$$

$$f(x)g(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + \cdots + \left(\sum_{\substack{i,j \\ i+j=n}} a_ib_j \right) x^n + \cdots$$

These objects are called “power series”. We could also have such infinite collections of coefficients which are indexed by multiple variables, for instance with $x^m y^k$ terms, rather than just x^n terms. Treating infinite sums like this is a way to abstractly discuss what properties such sums must have

⁸Notice, each sum in this expression will only be finite, so they all make sense

before (or perhaps without) ever making sense of evaluating such sums.

What we did in the previous section does indeed show that $f(x) = 1 + x + \frac{1}{2}x^2 + \dots + \frac{1}{n!}x^n + \dots$ will satisfy $f(x+y) = f(x)f(y)$ when we view it as a power series. In this sense, you should believe that if there is *any* way to make sense out actually adding up infinitely many terms, our expression for e^x will have to satisfy the desired property.

4 Other Inputs

Those of you with some background in linear algebra might enjoy seeing what other “adders” and “multipliers” exponential functions can deal with. For instance, matrices can also be added and multiplied, and indeed we can apply e^x , perhaps better written as $\exp(x)$. For instance

$$\begin{aligned} \exp\left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 + \dots + \frac{1}{n!} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^n + \dots \\ &\approx \begin{pmatrix} 51.9689562 & 74.73656456 \\ 112.10484685 & 164.07380304 \end{pmatrix} \end{aligned}$$

As with complex and real numbers, the infinite sum should not govern our intuition for what is going on, we again can think of turning adders to multipliers. In the case of $n \times n$ matrices, “adders” are vectors in the n^2 -dimensional vector space of $n \times n$ matrices, and “multipliers” are linear transformations on an n -dimensional vector space.

Just as matrices are linear operators of finite dimensional vector spaces, we can do the same thing to operators on function spaces. For instance, the derivative $\frac{d}{dx}$ is a linear operator on the vector space of all infinitely differentiable functions, and we can put it through \exp to get another linear operator:

$$\exp\left(\frac{d}{dx}\right)f(x) = 1f(x) + \frac{d}{dx}f(x) + \frac{1}{2}\frac{d^2}{dx^2}f(x) + \dots + \frac{1}{n!}\frac{d^n}{dx^n}f(x) + \dots$$

In fact, Taylor’s formula tells us that the above expression is precisely $f(x+1)$, so $\exp\left(\frac{d}{dx}\right)$ is the operator which transforms $f(x)$ into $f(x+1)$. Who would have ever thought we could make sense out of the sentence “ e to the derivative is a left-shift”?