Logarithms and Exponentiation

In these notes, we use our results on integration to define the natural logarithm function and derive its properties. We then use this function to define a^r for all $a > 0$ and $r \in \mathbf{R}$, in a unified way that does not depend on whether r is positive or negative, is an integer, is rational, or is irrational. We see that this elegant (albeit nonintuitive) definition is consistent with the usual definition for rational r, and implies that for irrational exponents, the "intuitive" definition of a^r actually works—i.e. that a^r can be defined *unambiguously* (if inelegantly) as a limit obtained by approaching r through rational exponents. (It's very unlikely that the student was shown in high school, or wherever he/she first encountered irrational exponents, that this definition is *unambiguous*, i.e. that the value of the limit does not depend on *which* of the uncountably many rational sequences approaching r is used.) We also show that all the usual algebraic "rules of exponents" follow, and that the functions $x \mapsto a^x$ (for any $a > 0$) and $x \mapsto x^r$ (for any $r \in \mathbf{R}$) are differentiable and have the "expected" formulas for their derivatives. For the function $x \mapsto x^r$ with r irrational, this would be extraordinary difficult using only the "intuitive" definition of x^r, but with our unified definition the derivative computations are identical for *all* r.

These notes are for my MAA 4211 class, Fall 2020. A reference such as "LN Corollary 35.9" means "Corollary 35.9 in my posted lecture notes" (available only to this class).

1 Logarithmic and exponential functions

Since the function $t \mapsto \frac{1}{t}$ is continuous on $(0, \infty)$, its integral over any closed interval with endpoints in $(0, \infty)$ exists. This allows us to make the following defintion:

Definition 1.1 The function $\log : (0, \infty) \to \mathbf{R}$ is defined by

$$\log x := \log(x) = \int_1^x \frac{1}{t}\, dt \quad \text{(for each } x > 0). \tag{1.1}$$

Remark 1.2 This function "log" is the *natural logarithm* function that you are probably used to denoting "ln". Mathematicians tend to call this function "log" rather than "ln" (except when teaching lower-level calculus courses and other courses populated largely by engineering majors), and write "$\log_{10}$" for the function you may be used to denoting simply as "log", because $\log_{10}$ has no special mathematical significance. Prior to the age of pocket calculators, $\log_{10}$ had much greater *practical* significance than it has now; bankers, scientists, and other people who used to need to do a lot of multiplication by hand would use tables of values of the $\log_{10}$ function in order to reduce multiplication to the addition of logs. Nowadays, these tasks are done by computers and pocket calculators. The $\log_{10}$ function still survives in a few $\log_{10}$-based scales in the sciences, such as the pH scale in chemistry and the decibel scale for sound-intensity. We also still use the phrase "order of magnitude" in a sense coming from the $\log_{10}$ function, since human beings brought up with base-10 arithmetic naturally find it easy to think in terms of how many powers of 10 are involved.

Proposition 1.3 *The function* $\log : (0,\infty) \to \mathbf{R}$ *is differentiable, strictly increasing, bijective, and satisfies the following for all* $x, y \in (0,\infty)$ *and all integers* n*:*

(i) $\log'(x) = \frac{1}{x}$ *(where "*$\log'$*" denotes the derivative of* $\log$*.)*

(ii) $\log(1) = 0$.

(iii) $\log(xy) = \log x + \log y$.

(iv) $\log(\frac{1}{x}) = -\log x$.

(v) $\log \frac{x}{y} = \log x - \log y$.

(vi) $\log(x^n) = n \log x$.

Proof: Since $t \mapsto \frac{1}{t}$ is continuous, the Fundamental Theorem of Calculus implies that log is differentiable and that $\log'(x) = \frac{1}{x}$. Since $\frac{1}{x} > 0$ for all $x > 0$, log is a strictly increasing (hence injective) function; we will show later that its range is all of $\mathbf{R}$.

Property (ii) is immediate from the defining equation (1.1). To establish property (iii), let $x, y > 0$. By LN Corollary 35.9 (a corollary of "additivity of the integral)", we have

$$\log(xy) = \int_1^{xy} \frac{1}{t}\,dt = \int_1^{x} \frac{1}{t}\,dt + \int_x^{xy} \frac{1}{t}\,dt = \log x + \int_x^{xy} \frac{1}{t}\,dt. \tag{1.2}$$

(We use LN Corollary 35.9 rather than the our original "additivity of the integral" result, LN Proposition 35.4, to ensure that (1.2) is true regardless of the size-order of x, y, and 1, or whether these numbers are all distinct.) In the last integral in (1.2), we may make the substitution $t = xs$ (More formally, we define the function $\varphi : [\min\{1,y\}, \max\{1,y\}] \to [\min\{x,xy\}, \max\{x,xy\}]$ by $\varphi(s) = xs$. Then $\varphi(1) = x$, $\varphi(y) = xy$, and $\varphi'(s) = x$ for all s.) Applying the change-of-variable theorem (LN Proposition 35.18), we have

$$\int_x^{xy} \frac{1}{t}\,dt = \int_1^{y} \frac{1}{xs}\,x\,ds = \int_1^{y} \frac{1}{s}\,ds = \log y.$$

Hence (1.2) implies that $\log(xy) = \log x + \log y$, which is property (iii).

Property (iv) now follows from properties (ii) and (iii) (since $x\,\frac{1}{x}=1$), and property (v) then follows from properties (iii) and (iv) (since $\frac{x}{y}=x\,\frac{1}{y}$).

Property (vi) holds for $n=0$ by property (ii), and is true trivially for $n=1$. Using property (iii) and induction, it follows easily that property (vi) holds for all $n\geq 1$ (details are left to the student). From this and property (iv), we then find that property (vi) holds for $n\leq -1$ as well.

We have now proven everything except that the range of log is $\mathbf{R}$. For this, we first note that since $\frac{1}{t}\geq\frac{1}{2}$ for $t\in[1,2]$, the order property of the integral (LN Corollary 33.3) implies

$$\log 2=\int_1^2\frac{1}{t}\,dt\geq\int_1^2\frac{1}{2}dt=\frac{1}{2},$$

so

$$\log 2\geq\frac{1}{2}>0. \tag{1.3}$$

Let $y\in[0,\infty)$, and let n be a positive integer such that $n\log 2\geq y$; such n exists since $\log 2>0$. Then, by property (vi), we have $\log(2^n)\geq y$, so $y\in[\log(1),\log(2^n)]$. Since log is differentiable, log is continuous, so the Intermediate Value Theorem implies that there exists $x\in[1,2^n]$ such that $\log x=y$. Hence log achieves every non-negative real value. Property (iv) then shows that log achieves every non-positive real value as well, hence achieves every real value. Thus the range of log is $\mathbf{R}$, as claimed. ■

Remark 1.4 (A pause to smell the roses) Thanks to our hard work on integration, up through the FTC (Fundamental Theorem of Calculus), the proof of part (i) of Proposition 1.3 was very short, so let us take a moment to reflect on something we've achieved with the formula "$\log'(x)=x^{-1}$".

When we first learn calculus, the first functions we learn how to differentiate are the power-functions $x\mapsto x^n$, where n is a positive integer or zero. Shortly thereafter, we work out the derivative for negative n as well, discovering the beautiful fact that $\frac{d}{dx}x^n=nx^{n-1}$ for *all* integers n. Later, when we start to study antidifferentiation, our first tool is *recognition*: having seen power-functions arise as derivatives of other power functions, we can easily invert the process. Since $3x^2$ is the derivative of x^3, we know that any multiple of x^2 will have some multiple of x^3 as an antiderivative (where "multiple of" means "constant times", and where I'm using Calculus 1 notation and terminology for functions, e.g. "x^3" instead of "$x\mapsto x^3$").

More generally, our derivative formula for integer-exponent power functions tells us that, for almost every integer n, the function x^n is an antiderivative of nx^{n-1}, hence that $\frac{x^n}{n}$ is an antiderivative of x^{n-1}, hence that $\frac{x^{n+1}}{n+1}$ is an antiderivative of x^n—with only one exception, the case $n=-1$. We never see x^{-1} arising as the derivative of a multiple of a power function, so at this early stage of our learning in Calc 1, we have no way to find an antiderivative. But since x^{-1} is continuous on $(0,\infty)$, LN Theorem 35.12 ("part of" the FTC) gives us a *formula* for an antiderivative of x^{-1}. The gap is filled![1]

Since $\log : (0,\infty) \to \mathbf{R}$ is bijective, it has an inverse, so we may make the following definition:

Definition 1.5 We define the function $\exp : \mathbf{R} \to (0,\infty)$ to be the inverse of $\log : (0,\infty) \to \mathbf{R}$.

Proposition 1.6 *The function* $\exp : \mathbf{R} \to (0,\infty)$ *is differentiable, strictly increasing, bijective, and satisfies the following for all* $x, y \in (0,\infty)$ *and all integers* n*:*

(i) $\exp' = \exp$ *(where "*$\exp'$*" denotes the derivative of* $\exp$*.)*

(ii) $\exp(0) = 1$.

(iii) $\exp(x+y) = \exp(x)\exp(y)$

(iv) $\exp(-x) = \frac{1}{\exp(x)}$.

(v) $\exp(x-y) = \frac{\exp(x)}{\exp(y)}$.

(vi) $\exp(nx) = (\exp(x))^n$.

Proof: The fact that exp is bijective and strictly increasing follow from the fact that it is the inverse of a function with these properties.

We next show that $\exp' = \exp$. Let $y_0 \in \mathbf{R}$ and let $x_0 = \exp(y_0)$ (thus $y_0 = \log x_0$). Since $\log'(x_0) = \frac{1}{x_0} \neq 0$, LN Proposition 28.3 ensures us that exp is differentiable at y_0 and that

$$\exp'(y_0) = ((\log)^{-1})'(x_0) = \frac{1}{\log'(x_0)} = \frac{1}{1/x_0} = x_0 = \exp(y_0)$$

Since y_0 was arbitrary, we conclude that exp is differentiable and that $\exp' = \exp$.

Properties, (ii)—(vi) of exp follow from the corresponding properties for log. For example, for (iii), given any $x, y \in \mathbf{R}$, by the bijectivity of log there exist unique $a, b \in (0,\infty)$ such that $x = \log a$ and $y = \log(b)$. We then have

$$x + y = \log a + \log b = \log(ab) = \log(\exp(x)\exp(y)),$$

implying that $\exp(x+y) = \exp(\log(\exp(x)\exp(y))) = \exp(x)\exp(y)$. Derivations of the remaining properties are left to the student. ■

[1] Unfortunately, the all-too-popular "early transcendentals" calculus textbooks rob students of an appreciation of how marvelous this is, and make it difficult for them to fall in love with calculus.

Consider now any positive, real number a. Letting $x = \log a$, property (vi) in Proposition 1.6, read from right to left, says that for any integer n we have $a^n = \exp(n \log a)$. In view of this fact, the following definition does not alter the meaning of a^n for any integer n, but gives meaning to "a^n" for all *real* n:

Definition 1.7 *Let $a, r \in \mathbf{R}$, with $a > 0$. We define the number $a^r \in (0, \infty)$ by*

$$a^r = \exp(r \log a).$$

The next proposition may be summarized as saying that the "usual algebra of exponentiation" for integer exponents holds more generally for real exponents.

Proposition 1.8 *Let $a, b, x, y \in \mathbf{R}$, with $a > 0$ and $b > 0$. Then:*

(i) $a^0 = 1$.

(ii) $a^{x+y} = a^x a^y$.

(iii) $a^{-x} = \frac{1}{a^x}$.

(iv) $(a^x)^y = a^{xy}$.

(v) $(ab)^x = a^x b^x$.

(vi) *If $a > 1$ then the map $x \mapsto a^x$ is strictly increasing; if $a < 1$ then this map is strictly decreasing.*

Proof: All these properties follow quickly from Proposition 1.6 and Definition 1.7. For example, for (ii) we have

$$a^{x+y} = \exp((x+y)\log a) = \exp(x \log a + y \log a) = \exp(x \log a)\exp(y \log a) = a^x a^y.$$

The remaining parts of the proof are left as an exercise to the student. ■

Exercise 1.1 Complete the proof of Proposition 1.8.

Observe that Definition 1.7 says nothing about how to define 0^r. Of course, if r is a positive integer, we already have a purely algebraic definition of 0^r, yielding $0^r = 0$. Using the fact that for positive integers n, the unique n^{th} root of 0 is 0, we can naturally extend the definition "$0^r = 0$" to all positive *rational* r. But irrational r cannot be handled by these purely algebraic means. For these, we need a separate definition, which we will write in a way that applies in both the rational-r and irrational-r cases:

Definition 1.9 For every $r > 0$, we define $0^r = 0$.

We do not define the expression "0^r" for $r \le 0$; in particular, we do not define "0^0".

Exercise 1.2 Determine how Proposition 1.8 generalizes if we allow $a \ge 0$ and/or $b \ge 0$. Are any restrictions on x and/or y needed (possibly different restrictions for different parts of the proposition)? If so, what?

Exercise 1.3 (Just for fun) Parts (a) and (b) of this exercise can be done in either order; neither is likely to help you much with the other. Part (a) *can* be used to help with part (b)—it's obvious that (b) is *somehow* related to (a)—but there's an easier, independent approach to part (b) that doesn't require figuring out this "somehow" precisely.

As you know, $2^4 = 4^2$. You can probably convince yourself quickly, and perhaps even come up with a proof, that there are no other distinct integers n, m for which $n^m = m^n$. (Certainly for fixed n—e.g. $n = 2$—you should be able to understand intuitively why as $m \to \infty$ we have $n^m >> m^n$; "the larger exponent eventually wins".) But what if we remove the "integer" requirement?

(a) Graph the relation $\{(x, y) \in \mathbf{R}^2 \mid x, y > 0 \text{ and } x^y = y^x\}$. (Observe that the graph contains the ray $\{y = x > 0\}$, but also contains the points $(2, 4)$ and $(4, 2)$ that are not on this ray. Surely there must be some other points ...)

(b) *Without using a calculator*, determine which is greater: e^π or π^e? (You are allowed to use the fact that $e < \pi$.) Note that since $2^4 = 4^2$, and both e and π are close to 3, it is not at all clear whether "the larger exponent wins".

Remark 1.10 (Rational exponents) Proposition 1.8 shows that Definition 1.7 is consistent with our prior definitions of a^r for *rational* exponents r. For example, for $a > 0$ and n a positive integer, Proposition 1.8(iv) shows that $(a^{1/n})^n = a^{n/n} = a$. Since the function $x \mapsto x^n$ is strictly increasing on $(0, \infty)$, the number $a^{1/n}$ is therefore the *unique* positive real number c such that $c^n = a$, i.e. the (positive) n^{th} *root of* a. Similarly, if p, q are integers and $q \ne 0$, Proposition 1.8 shows that, consistently with prior definitions of "$a^{p/q}$", we have

$$(a^{1/q})^p = a^{p/q} = (a^p)^{1/q}. \tag{1.4}$$

However, Definition 1.7 gives a much "cleaner", if less intuitive, definition of a^r for $r \in \mathbf{Q}$ than does taking either the first or second equality in (1.4) to be the definition of $a^{p/q}$, because a rational number does not have a *unique* expression as a quotient of integers; e.g. $\frac{2}{3} = \frac{16}{24} = \frac{-42}{-63}$. When we attempt to use (say) the first equality in (1.4) as the definition of $a^{p/q}$ when p, q are positive integers, we must do one of the following in order to ensure that $a^{p/q}$ is well-defined: (1) require that the exponent be expressed in "lowest terms", i.e.

with p and q having no common divisor greater than 1, or (2) show that if $\frac{p}{q} = \frac{p'}{q'}$, where p, q, p', q' are positive integers, then $(a^{1/q'})^{p'} = (a^{1/q})^p$. (Since every rational number *can* be expressed in lowest terms, (2) can be reduced to the case in which $p' = kp$ and $q' = kq$ for some positive integer k.) Approach (1), however, becomes insufficient the moment we try to show that rational exponents obey property (ii) in Proposition 1.8, based only on the algebra of integer exponents and on a definition of $a^{p/q}$ that requires p and q to be relatively prime. For example, $\frac{1}{5} + \frac{3}{10} = \frac{1}{2}$, but you will not likely succeed in showing that $a^{1/5}a^{3/10} = a^{1/2}$ without knowing that $a^{1/5} = (a^{1/10})^2$ and that $(a^{1/10})^5 = a^{1/2}$. Similarly, you will have difficulty showing that $2^{0.6} > 2^{0.5}$ without knowing that $(2^{1/5})^3 = (2^{1/10})^6$ and that $2^{1/2} = (2^{1/10})^5$. Thus, if we attempt to take (1.4) as the definition of a^r for rational non-integer r, then to obtain the results of Proposition 1.8 even just for rational exponents, based on knowing them for integer exponents, we are forced to prove (at least) that $(a^{1/kq})^{kp} = (a^{1/q})^p$ for all positive integers p, q, and k. This is not *difficult* to prove, but the necessity of proving it can't be avoided if we attempt to use one of the equalities in (1.4) as the definition of a^r for non-integer rational r.

Remark 1.11 You may be accustomed to thinking that the algebraic rules in Proposition 1.8 are "obvious", even though you likely have been using one of the equalities in (1.4) as the definition of a^r when r is rational, positive, and expressed in lowest terms, and have likely been taking Proposition 1.8(iii) as the definition of a^r when r is rational and negative. There is nothing incorrect about these prior definitions. However, as Remark 1.10 shows, if we use these prior definitions then the algebraic rules in Proposition 1.8 are not at all obvious once we leave the realm of integer exponents. There's a difference between something being obvious because we *understand why it's true*, and thinking it "obvious" because we *memorized* it and were *told* it was true.

Remark 1.12 (Irrational exponents, part 1) Modern students, having grown up with pocket calculators that have an "x^y" button on them, may be not be conscious of the fact that there is nothing obvious about what an expression like "$2^{\sqrt{2}}$" should mean. We can use equation (1.4) to define what it means to raise a number to a *rational* exponent, but this equation says nothing about irrational exponents.

Let r be an irrational number. We may attempt to define a^r in an *ad hoc* manner, using the decimal expansion of r as a sequence in $\mathbf{Q}$ approaching r from below (e.g. using the fact that $\sqrt{2}$ is the limit of a sequence 1, 1.4, 1.41, 1.414, 1.4142, ...), and tentatively defining a^r to be the limit of this sequence. Why should the limit exist? If we first do the work mentioned in Remark 1.10, we can then show that this sequence is monotone and bounded, hence convergent. But that's not entirely satisfying (nor can it be rigorously justified early in Calculus 1, let alone prior to calculus): should the value of the number a^r depend on the fact that humans have 10 fingers (the reason that we chose the *decimal* expansion of r)? For *rational* exponents there is no such dependence, so we would certainly *hope* that there is none for irrational exponents either. This leads us to want to prove at least the following: if $(r_n)_{n=1}^{\infty}$ is an increasing sequence of rational numbers approaching r, does $\lim_{n\to\infty} a^{r_n}$ is independent of the choice of the sequence (r_n).

Even this is not wholly satisfying. What if we had chosen a sequence (r_n) that *decreases* to r instead of *increasing* to r? (For example, if $r = -\sqrt{2}$, the sequence $-1, -1.4, -1.41, -1.414, -1.4142, \ldots$ is a decreasing sequence.) What if we had chosen a non-monotonic sequence (r_n) with limit r? Do we always get the same limit? Using (1.4) it is, indeed, possible to prove that $\lim_{n\to\infty} a^{r_n}$ exists for every sequence in $\mathbf{Q}$ converging to r, and that the value of this limit is independent of the choice of sequence (r_n).[2] Using this limit to define a^r, we then have a definition of a^r that is valid for all real r and all $a > 0$. Can we then prove Proposition 1.8 based on this approach? Yes, but still more work is involved, and some questions with non-obvious answers have to be addressed. By contrast, when we use Definition 1.7 to define a^r for all real r, we have a definition that works simultaneously for rational and irrational exponents, that looks identical for *all* real exponents, and that renders the proof of Proposition 1.8 essentially trivial.

Definition 1.13 The number $e \in \mathbf{R}$ is defined by $e = \exp(1)$.

From Definitions 1.7 and 1.13, we immediately have

$$e^x = \exp(x)$$

for all $x \in \mathbf{R}$.

Proposition 1.14 *Let $a, r \in \mathbf{R}$, with $a > 0$. The function $\mathbf{R} \to \mathbf{R}$ defined by $x \mapsto a^x$, and the function $(0, \infty) \to \mathbf{R}$ defined by $x \mapsto x^r$, are differentiable, and their derivatives are given by the following formulas:*

(i) $\frac{d}{dx} a^x = a^x \log a$.

(ii) $\frac{d}{dx} x^r = r x^{r-1}$.

Proof: (i) Define $f, g : \mathbf{R} \to \mathbf{R}$ by $f(x) = a^x$ and $g(x) = x \log a$. Then, by definition, $f = \exp \circ g$. Since exp and g are differentiable, the Chain Rule Theorem implies that f is differentiable and that $f'(x) = \exp'(g(x))g'(x) = \exp(g(x)) \log a = a^x \log a$.

(ii) Define $f, g : \mathbf{R} \to \mathbf{R}$ by $f(x) = x^r$ and $g(x) = r \log x$. Then, by definition, $f = \exp \circ g$. Since exp and g are differentiable, the Chain Rule Theorem implies that f is differentiable and that $f'(x) = \exp'(g(x))g'(x) = \exp(g(x)) \frac{r}{x} = x^r \frac{r}{x}$. Using Proposition 1.8(ii)–(iii), this last expression equals $r x^{r-1}$. ■

[2] But the proofs of these statements are far beyond the level at which today's students are generally taught this approach to defining irrational powers. An "it can be shown" statement is required, and it's for something that really *can't* be shown at the level of Calculus 1.

Remark 1.15 (Exponential and logarithm functions) For each $a > 0$, the function $x \mapsto a^x$ may be called the *exponential function with base* a. All such functions are called "exponential functions". In this terminology, the function exp is the exponential function with base e. The function exp is also referred to as "*the* exponential function" (with no base mentioned).

Note that for any $k \in \mathbf{R}$, $a^{kx} = (a^k)^x$, so every function of the form $a \mapsto a^{kx}$ is also an exponential function.

Since $a^x = \exp(x \log a)$, and $\log a$ is positive for $a > 1$ and negative for $a < 1$, Proposition 1.6 implies that the exponential function with base a is strictly increasing if $a > 1$ and strictly decreasing if $a < 1$ (and is the constant function 1 if $a = 1$), and has range $(0, \infty)$ if $a \neq 1$. Hence for $a \neq 1$ this function has an inverse. This inverse function is called *the logarithm function with base* a, and is denoted $\log_a$ (which is best read "log, base a" [3]; the expression $\log_a x := \log_a(x)$ is best read "log, base a, of x"). Any such function is known as a *logarithm function*, or simply "log function". Observe that $\log_e$ is the same as the *natural logarithm* function that we're denoting simply as log, but which you're probably used to writing as "ln".

Exercise 1.4 Show that for $a, b, x > 0$ and $a \neq 1 \neq b$,

$$\log_b x = \frac{\log_a x}{\log_a b}.$$

Thus, any logarithm function is a constant times any other.

Remark 1.16 (Power functions and their derivatives) For each $r \in \mathbf{R}$, Definition 1.7 defines the expression "x^r" for all $x > 0$, and, as we have seen, this definition agrees with definitions we have previously learned for $r \in \mathbf{Q}$. However, for certain r we do not need $x > 0$ for the expression x^r to be defined. (For example, if n is a positive integer, basic algebra produces a definition of x^n for all $x \in \mathbf{R}$, and we then define $x^{-n} = \frac{1}{x^n}$ for all $x \neq 0$. We also define $x^0 = 1$ for all $x \neq 0$, in order that the property in Proposition 1.8(iii) hold for all real $a \neq 0$ and all integer exponents. For odd integers n, every real number has a unique n^{th} root, so we may define $x^{1/n}$ for all real x.) For each $r \in \mathbf{R}$, and any set $U \subseteq \mathbf{R}$ such that $x \mapsto x^r$ is defined for all $x \in U$, the function $U \to \mathbf{R}$ given by $x \mapsto x^r$ is called a *power function*, or the r^{th}-*power function*. (Of course, for certain r, we

[3] This is one of two common ways that the notation "$\log_a$" is read. The other, "log to the base a", appears to be idiomatic—it makes no sense grammatically, unlike alternatives such as "log *from* base a" or "log *with* base a"—but *is* the terminology used in the classic textbook [1] by Thomas, from which many current mathematicians, including the authors of many Calculus 1-2-3 textbooks, learned calculus. Based on the terminology "raising *to* a power", which (at least for integer exponents) is probably much older than any terminology for logarithms with bases other than 10, if $y = a^x$ there is logical justification to say that "x is the logarithm of y *from* the base a," or that "x is the log, *from* the base a, of y." By contrast, if we apply conventional rules of grammar and usage, "log to the base a" is not consistent with the terminology for powers, or with any other uses of "to" in English.

have other names as well, e.g. the *squaring function*, for $r = 2$ and the *cube-root function* for $r = 1/3$.) We still use the name "(r^{th}-)power function" (or these other names) if the codomain $\mathbf{R}$ is replaced by any set containing $\{x^r : x \in U\}$, as in "the squaring function $\mathbf{R} \to [0, \infty)$" or "the cube-root function $(8, \infty) \to (2, \infty)$".

The remainder of this Remark is optional reading; the student may skip to Remark 1.17. In Calculus 1, one of the first things we learn is that for *positive integers* r,

$$\frac{d}{dx}x^r = rx^{r-1}; \tag{1.5}$$

to derive this fact we use the Binomial Theorem. We also learn that the derivative of a constant function is 0, so that (1.5) holds for $r = 0$ as well (on the domain $\mathbf{R} \setminus \{0\}$), not as a consequence of the Binomial Theorem, but because x^0 has been defined to be 1 for $x \neq 0$. In a *good* Calculus 1 course, we learn progressively that (1.5) holds for more and more exponents, until we have shown that it is true for all *rational* exponents:

1. By one of several methods, we learn that

$$\frac{d}{dx}x^{-1} = -x^{-2}, \tag{1.6}$$

which is (1.5) for $r = -1$. Simple methods by which (1.6) can be shown, without using any "laws of exponents" for anything other than integer exponents (the only exponents for which it is *obvious* that rules like Proposition 1.8(ii), (iv), and (v) are valid), are:

(a) Direct calculation: defining $f : \mathbf{R} \setminus \{0\} \to \mathbf{R}$ by $f(x) = x^{-1}$, we compute

$$f'(x) = \lim_{h\to 0} \frac{\frac{1}{x+h} - \frac{1}{x}}{h} = \lim_{h\to 0} \frac{\frac{x-(x+h)}{x(x+h)}}{h} = \lim_{h\to 0} \frac{-1}{x(x+h)} = \frac{-1}{x^2}\ .$$

(b) First learning the quotient rule, then computing

$$\frac{d}{dx}\frac{1}{x} = \frac{x\frac{d}{dx}(1) - 1\frac{d}{dx}(x)}{x} = \frac{x \cdot 0 - 1}{x^2} = -x^{-2}.$$

(Method (a) is really a special case of the proof of the quotient rule, so it is not entirely different from method (b).)

2. By one of several methods, we learn that for *all* positive integers n,

$$\frac{d}{dx}x^{-n} = -nx^{-n-1} \tag{1.7}$$

on $\mathbf{R} \setminus \{0\}$, which is (1.5) for $r =$ [negative integer]. Simple methods by which (1.7) can be shown are, without using any "laws of exponents" for anything other than integer exponents, are:

(a) First learning the quotient rule, then computing

$$\frac{d}{dx}\frac{1}{x^n} = \frac{x^n \frac{d}{dx}(1) - 1\frac{d}{dx}(x^n)}{(x^n)^2} = \frac{x^n \cdot 0 - nx^{n-1}}{x^{2n}} = -nx^{-n-1}.$$

(b) First establishing (1.6), then learning the Chain Rule, and then computing

$$\frac{d}{dx}x^{-n} = \frac{d}{dx}(x^n)^{-1} = -(x^n)^{-2}\frac{d}{dx}(x^n) = -(x^{-2n})nx^{n-1} = -nx^{-n-1}.$$

Combining (1.7) with the previously-established cases of (1.5), we have now learned that (1.5) holds for all *integer* exponents.

3. We show, by one of several methods, that for positive integers n, the function $x \mapsto x^{1/n}$ is differentiable on $(0, \infty)$, and compute that

$$\frac{d}{dx}x^{\frac{1}{n}} = \frac{1}{n}x^{\frac{1}{n}-1}. \tag{1.8}$$

on this interval. Two methods by which (1.7) can be shown, without using any "laws of exponents" for anything other than *rational* exponents, are:

(a) Proving the " 'Baby' Inverse Function Theorem" and then applying it to $h : x \mapsto x^n$. This shows that the function $f : x \mapsto x^{1/n}$ is differentiable and that

$$\frac{d}{dx}x^{1/n} = f'(x) = \frac{1}{h'(f(x))} = \frac{1}{\frac{d}{dy}y^n\big|_{y=x^{1/n}}} = \frac{1}{n(x^{\frac{1}{n}})^{n-1}} = \frac{1}{n(x^{1-\frac{1}{n}})} = \frac{1}{n}x^{\frac{1}{n}-1}.$$

(b) Restricting attention to $n \geq 2$ (sufficient, since (1.8) has been proven already for $n = 1$), showing first that $f : x \mapsto x^{1/n}$ is continuous on $(0, \infty)$), and then using the algebraic identity $b^n - a^n = (b-a)(b^{n-1} + b^{n-2}a + b^{n-3}a^2 + \cdots + a^{n-1})$ (for $a, b \in \mathbf{R}$) to compute

$$
\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{(x+h)^{\frac{1}{n}} - x^{\frac{1}{n}}}{h} \\
&= \lim_{h\to 0} \frac{\left[(x+h)^{\frac{1}{n}} - x^{\frac{1}{n}}\right]\left[\left\{(x+h)^{\frac{1}{n}}\right\}^{n-1} + \left\{(x+h)^{\frac{1}{n}}\right\}^{n-2} x^{\frac{1}{n}} + \cdots + (x^{\frac{1}{n}})^{n-1}\right]}{h\left[\left\{(x+h)^{\frac{1}{n}}\right\}^{n-1} + \left\{(x+h)^{\frac{1}{n}}\right\}^{n-2} x^{\frac{1}{n}} + \cdots + (x^{\frac{1}{n}})^{n-1}\right]} \\
&= \lim_{h\to 0} \frac{\left\{(x+h)^{\frac{1}{n}}\right\}^{n} - (x^{\frac{1}{n}})^{n}}{h\left[\left\{(x+h)^{\frac{1}{n}}\right\}^{n-1} + \left\{(x+h)^{\frac{1}{n}}\right\}^{n-2} x^{\frac{1}{n}} + \cdots + (x^{\frac{1}{n}})^{n-1}\right]} \\
&= \lim_{h\to 0} \frac{(x+h) - x}{h\left[\left\{(x+h)^{\frac{1}{n}}\right\}^{n-1} + \left\{(x+h)^{\frac{1}{n}}\right\}^{n-2} x^{\frac{1}{n}} + \cdots + (x^{\frac{1}{n}})^{n-1}\right]} \\
&= \lim_{h\to 0} \frac{1}{\left\{(x+h)^{\frac{1}{n}}\right\}^{n-1} + \left\{(x+h)^{\frac{1}{n}}\right\}^{n-2} x^{\frac{1}{n}} + \cdots + (x^{\frac{1}{n}})^{n-1}} \\
&= \frac{1}{\underbrace{\left\{x^{\frac{1}{n}}\right\}^{n-1} + \left\{x^{\frac{1}{n}}\right\}^{n-2} x^{\frac{1}{n}} + \cdots + (x^{\frac{1}{n}})^{n-1}}_{n \text{ terms}}} \\
&= \frac{1}{nx^{1-\frac{1}{n}}} \\
&= \frac{1}{n} x^{\frac{1}{n}-1}.
\end{aligned}
$$

If n is odd, so that $x^{1/n}$ is defined for all x, the methods in (a) and (b) extend from the domain $(0, \infty)$ to the domain $\mathbf{R} \setminus \{0\}$.

4. Having established (1.5) for the cases in which n is an integer or the reciprocal of a positive integer, we apply the Chain Rule Theorem to generalize to other rational exponents, as follows[4]: for an arbitrary rational number $r = \frac{m}{n}$, where m, n are

[4] In our good calculus course, we do not investigate other exponents until we have proven the Chain Rule Theorem (CRT). However, once we have established the CRT, an alternative approach to deriving (1.5) for general rational exponents that does not require that we first handle reciprocal-integer exponents separately, is as follows: (1) Introduce implicit differentiation (which depends crucially on the chain rule). (2) *State* the Implicit Function Theorem (for real-valued functions of a single real variable), advising the student that the proof is beyond the scope of a Calculus 1 course. (3) For an arbitrary rational number $r = \frac{m}{n}$, where m, n are integers and $n > 0$, show that the Implicit Function Theorem implies that the equation $y^n = x^m$ (with $(x, y) \in (0, \infty) \times (0, \infty)$) defines y as a differentiable function of x on $(0, \infty)$, hence that the function $x \mapsto x^{m/n}$ is differentiable on $(0, \infty)$. (4) Implicitly differentiate "$y^n = x^m$" with respect to x, and then apply rules of rational exponents, to deduce that $\frac{d}{dx} x^{m/n} = \frac{dy}{dx} = \frac{m}{n} x^{m/n-1}$.

integers and $n > 0$, we have

$$\frac{d}{dx}x^r = \frac{d}{dx}(x^{\frac{1}{n}})^m = m(x^{\frac{1}{n}})^{m-1}\frac{d}{dx}x^{\frac{1}{n}} = m(x^{\frac{1}{n}})^{m-1}\frac{1}{n}x^{\frac{1}{n}-1} = \frac{m}{n}x^{\frac{m-1}{n}+\frac{1}{n}-1} = rx^{r-1}.$$

In our good Calculus 1 course, we have now *shown* that the derivative formula (1.5) is valid for all rational exponents, using no laws of exponents that we did not know how to prove in high school. But our method of proof was different for different types of rational exponents. This strongly suggests that there must be some underlying principle, undiscovered as yet, that would give a unified derivation of (1.5) for all rational exponents. (Our proof of Proposition 1.14(ii) is exactly this unified derivation; moreover, it works equally well for all real exponents, whether rational or irrational.)

Since the rationals are dense in the reals, we could reasonably conjecture now, in our good Calculus 1 course, that formula (1.5) holds for all *real* exponents. It would have been absurdly bold to make such a conjecture based only on knowing that (1.5) holds for nonnegative integer exponents.

Remark 1.17 (Irrational exponents, part 2) Suppose that, instead of using Definition 1.7 to define irrational powers of positive numbers, we have defined them an "elementary" way, using a limit-procedure such as those discussed in Remark 1.12 (assuming we have already defined rational powers by elementary means, the way we would in high school or or that we did earlier this semester, rather than by Definition 1.7). Assume that we have done this in the best possible way, showing that for any $a > 0$ and any rational sequence $(r_n)_{n=1}^{\infty}$ converging to r, (i) $\lim_{n\to\infty} a^{r_n}$ exists and (ii) the value of this limit is independent of the choice of sequence (r_n). Finally, suppose we have shown that the derivative-formula (1.5) holds for rational exponents, just based on these elementary, intuitive definitions. To then show that (1.5) is true (with this value of r) we must do something like the following:

1. Choose a sequence of rational numbers $(r_n)_{n=1}^{\infty}$ converging to r.
2. For fixed $x > 0$, write down the following computation, which we will do in "shoot first and ask questions later" form:

$$\begin{aligned}
\lim_{h\to 0}\frac{(x+h)^r - x^r}{h} &= \lim_{h\to 0}\left(\lim_{n\to\infty}\frac{(x+h)^{r_n} - x^{r_n}}{h}\right) && (1.9)\\
&\stackrel{?}{=} \lim_{n\to\infty}\left(\lim_{h\to 0}\frac{(x+h)^{r_n} - x^{r_n}}{h}\right) && (1.10)\\
&= \lim_{n\to\infty}\left(\frac{d}{dx}x^{r_n}\right)\\
&= \lim_{n\to\infty} r_n x^{r_n-1}\\
&\stackrel{?}{=} rx^{r-1}\,. && (1.11)
\end{aligned}$$

If we can justify equalities (1.10) and (1.11), we will have shown that the r^{th}-power function is differentiable on $(0, \infty)$, and that its derivative is the function $x \mapsto rx^{r-1}$.

Justifying (1.11) is no problem: $(r_n - 1)_{n=1}^{\infty}$ is a rational sequence converging to $r-1$, so, by what we're assuming we've already proven about our elementary definition of irrational powers, x^{r_n-1} converges to x^{r-1}. Since (r_n) converges to r, one of our basic results about real-valued sequences shows that the sequence $(r_n x^{r_n-1})$ converges to rx^{r-1}. We can do all this without ever setting foot in Advanced Calculus; the fact that "the limit of a product is the product of the limits" (for convergent real-valued sequences) can easily be proven even in the most elementary treatment of sequences, such as in Calculus 2.

But justifying the interchange-of-limits equation (1.10) is another matter entirely. The entire notion of "interchange of limits" is far beyond the level of Calculus 1. To my knowledge, justifying (1.10) is impossible at the level of Calculus 1, and would be difficult even in Advanced Calculus without relying on the fact that the elementary definition of a^r for *rational* r is consistent with Definition 1.7.

Epilog

The (once standard) approach to defining exponentiation presented in the first few pages of these notes is a triumph of calculus, a true gem.[5] It unifies the definitions of a^r for positive integer, negative integer, non-integer rational, and irrational r; Definition 1.7 is the same for all exponents. It leads *easily* to the derivative formula (1.5) for *all* real exponents. By showing that Definition 1.7 agrees with the "elementary" definition for rational exponents, we see why our elementary derivations of $\frac{d}{dx}x^r$ for rational r (the optional reading in Remark 1.16) *had* to keep giving the same formula for all exponents. With Definition 1.7, continuity of the function exp *guarantees* that for any real $a > 0$, any real number r, and any rational (or even real) sequence (r_n) converging to r, the sequence

[5]This may be difficult for modern students to appreciate, especially if they've been taught out of an "early transcendentals" calculus textbook. Calculus textbooks that define exponential functions early, through a "filling in the holes in the graph" idea—the same idea as the limit-procedures discussed in Remark 1.12—rather than waiting until the groundwork has been laid for Definition 1.7, often say shortly after deriving equation (1.5) for non-negative integer exponents that "We will show later" that (1.5) holds for all real r. These textbooks slip under the rug the fact that they do *not* derive (1.5) from their first definition of x^r. Rather, they wait until they have *re*defined x^r using Definition 1.7, after which they derive (1.5) exactly as we did in the proof of Proposition 1.14. The "derivation" of (1.5) in at least one popular early-transcendentals textbook starts by writing "$x^r = (e^{\ln x})^r = e^{r \ln x}$", and then uses the Chain Rule to compute the derivative. But the reasoning is completely reversed from the correct logic! Starting with "$x^r = (e^{\ln x})^r = e^{r \ln x}$" suggests that $x^r = e^{r \ln x}$ *because* $(x^b)^c = x^{bc}$ for all real b, c. The truth is exactly the opposite: "$x^r = e^{r \ln x}$" is the *definition* of x^r, a definition from which we *derive* the fact that $(x^b)^c = x^{bc}$ (a fact that, in "early transcendentals" textbooks, is often relegated to an appendix).

There is nothing wrong in a student's using the familiar formula $(x^b)^c = x^{bc}$ (familiar for *rational exponents only*) as a *mnemonic device* to help remember that $x^r = e^{r \ln x}$, but the above so-called derivation encourages the student to think, wrongly, that this formula for x^r is a *consequence* of "$(x^b)^c = x^{bc}$".

(a^{r_n}) converges to a limit that is independent of which sequence (r_n) we choose. (The continuity of exp follows from our proven differentiability, but can also be proven directly by using LN Proposition 25.4(b).) And at the core of all this were three major theorems from the theory of integration: the integrability of continuous functions, the Fundamental Theorem of Calculus, and the change-of-variables theorem.

References

[1] G. B. Thomas, *Calculus and Analytic Geometry*, alternate edition, Addison-Wesley, 1972. 9