

MATHEMATICAL OLYMPIADS LECTURE NOTES

Review of logarithms

Greg Gamble

Definition

The *logarithm to base a* of b , written $\log_a(b)$ is the power m that a must be raised to get b . For this definition to make sense we need $a > 0$ but $a \neq 1$ and $b > 0$. (Usually a is also an integer but this is not necessary.) More precisely then, for $1 \neq a > 0$ and $b > 0$,

$$\log_a(b) = m \iff a^m = b.$$



The \log_a function should be seen as the inverse function of the \exp_a function defined by:

$$\exp_a(x) := a^x, \quad x \in \mathbb{R},$$

where $a > 0$.



There are two questions here:

1. Why can we allow x to be in \mathbb{R} ? Shouldn't x be in \mathbb{Q} ?
2. Why do we need $a > 0$?

Suppose first that $x = p/q \in \mathbb{Q}$, where $p, q \in \mathbb{Z}$ and $q \neq 0$, and $a > 0$. Then we define

$$\exp_a(x) := (a^p)^{1/q} = \sqrt[q]{a^p},$$

where for $b = a^p > 0$, $b^{1/q} = \sqrt[q]{b}$ is the *positive real q^{th} root* of b and


$$a^p = \underbrace{a \cdot a \cdot \dots \cdot a}_p$$

(the usual definition). We can *extend* this definition to have $x \in \mathbb{R}$ by *continuity* – essentially this means for any sequence of rational numbers $p_1/q_1, p_2/q_2, \dots$ that approach a given $x \in \mathbb{R}$, the sequence $a^{p_1/q_1}, a^{p_2/q_2}, \dots$ approaches some real number y – a^x is then defined to be y .

Now, why do we need $a > 0$? ... Well, if $a = 0$ then $\exp_a(x)$ is undefined for nonpositive x ; and we avoid negative values of a since otherwise we encounter conflicts like the following:

- $a^{1/3}$ exists if we interpret this as the real cube root of a (then the value of $a^{1/3}$ is *negative* for negative a);
- $a^{2/6}$ ought to be interpreted as $(a^2)^{1/6}$ which is the positive 6th root of the positive number a^2 (i.e. $a^{2/6}$ would necessarily be *positive* for negative a); and
- $\frac{1}{3} = \frac{2}{6}$.

So, insisting that $a > 0$ ensures that \exp_a is *well-defined*.

 **What is a *function*?** Firstly, a more general concept than a *function* is a *map* (or *mapping*). We define f to be a *map* (or *mapping*) if it is a rule that takes elements of one set, called the *domain*, to elements of another set, called the *codomain*. We say that f is a *map* from its *domain* to its *codomain*; or that f *maps* elements of its *domain* to its *codomain*. Now a *function* is a *map* with further properties. We say a map f is a *function* with domain D and codomain Y if

- f is defined for all $x \in D$; and
- for each $x \in D$, $f(x)$ is just *one* element of Y .

A particularly nice notation that emphasises this way of thinking has the following form

$$\begin{aligned} f &: \text{domain} \rightarrow \text{codomain} \\ &: x \mapsto f(x). \end{aligned}$$

The value $f(x)$ is called the *image* of x under f . For example, the *real square* function may be represented by

$$\begin{aligned} g &: \mathbb{R} \rightarrow \mathbb{R} \\ &: x \mapsto x^2. \end{aligned}$$


This says

*g [is the function that]
maps x in \mathbb{R} to x^2 in \mathbb{R} .*

The *image* of 3 under g is 9, and more generally the *image* of x under g is x^2 . You may be more familiar with defining the function g by

$$g(x) = x^2, \quad x \in \mathbb{R}.$$

This notation, however, fails to tell us what we are considering to be the *codomain*. Notice that we have not insisted that f maps at least one point of the *domain* to each point of the *codomain*. For the *real square* function g , notice that there are no points of the *domain* that are mapped to negative elements of the *codomain*. We reserve the term *range* for the subset of the *codomain* containing only the *images* of points of the *domain*. For g the *range* is $\mathbb{R}_{\geq 0}$, the set of all *nonnegative* elements of \mathbb{R} . (Note that *nonnegative* means *positive or zero*).

 **What is an *inverse function*?** Basically the *inverse function* of a function f (if one exists) should take elements of the *codomain* of f back to elements of the *domain* of f ; and it should itself be a *function*. We usually write f^{-1} for the *inverse function* of the function f (when it exists). Suppose f is a function, where

$$\begin{aligned} f : D &\rightarrow Y \\ &: x \mapsto f(x), \end{aligned}$$

and suppose R is the *range* of f . Then f^{-1} is defined by

$$\begin{aligned} f^{-1} : Y &\rightarrow D \\ &: f(x) \mapsto x, \end{aligned}$$

so long as this definition defines a function. That is, the domain of f^{-1} is the *codomain* of f and the codomain of f^{-1} is the *domain* of f ; and f^{-1} is a function if

- f^{-1} is defined for all y in its domain Y , i.e. each $y \in Y$ must be an image under f of a point in X – this can only happen if $Y = R$, in which case f is said to be *onto* (that is, if $Y = R$ we say f maps D *onto* Y); and
- for each $y \in Y = R$, $f^{-1}(y)$ is just *one* element of its codomain D , i.e. only *one* element of the domain D of f maps to any *one* element of the range R of f – if f has this property then f is said to be *one-to-one*, (this property can often be achieved by reducing the domain of f).

For the *square function* g , the codomain is \mathbb{R} rather than $\mathbb{R}_{\geq 0}$ (so g is not *onto*), and if $x \neq 0$ then $-x$ and x are two points that map to the same image x^2 under g (so g is not *one-to-one*). We can make g *onto* by redefining its *codomain* to be $\mathbb{R}_{\geq 0}$ and we can make g *one-to-one* by reducing its domain to say $\mathbb{R}_{\geq 0}$. Really, this is a *new function* – so let's call it g^* , i.e.

$$\begin{aligned} g^* : \mathbb{R}_{\geq 0} &\rightarrow \mathbb{R}_{\geq 0} \\ &: x \mapsto x^2. \end{aligned}$$

Now g^* has *inverse function* g^{*-1} defined by

$$\begin{aligned} g^{*-1} : \mathbb{R}_{\geq 0} &\rightarrow \mathbb{R}_{\geq 0} \\ &: y \mapsto \sqrt{y}. \end{aligned}$$

Now let's consider the inverse function of \exp_a . We will need its codomain to be equal to its range, which is $\mathbb{R}_{>0}$, the set of all *positive real numbers*. So let's define, for $a > 0$

$$\begin{aligned} \exp_a : \mathbb{R} &\rightarrow \mathbb{R}_{>0} \\ &: x \mapsto a^x. \end{aligned}$$

This function is *one-to-one*, for all $1 \neq a > 0$. (Note that, $\exp_1(x) = 1^x = 1$ for all $x \in \mathbb{R}$, so that \exp_1 is not *one-to-one*.) Now finally, we can define \log_a to be the inverse function of \exp_a , i.e. for $1 \neq a > 0$,

$$\begin{aligned} \log_a : \mathbb{R}_{>0} &\rightarrow \mathbb{R} \\ &: a^x \mapsto x. \end{aligned}$$

Properties

Suppose $1 \neq a > 0$, $1 \neq b > 0$, $x, y > 0$ and $m, n \in \mathbb{R}$. Then the function \log_a has the following properties.

1. $\log_a(1) = 0$

since ... $a^0 = 1$ and $a \neq 0$ (in fact, we assumed $1 \neq a > 0$).

2. $\log_a(a) = 1$

since ... $a^1 = a$.

3. $\log_a(a^n) = n$

4. $\log_a(xy) = \log_a(x) + \log_a(y)$

since ...

The statement is essentially

$$a^m \cdot a^n = a^{m+n}$$

in disguise. Let $x = a^m$, $y = a^n$. Then

$$\begin{aligned} \log_a(xy) &= \log_a(a^m \cdot a^n) = \log_a(a^{m+n}) \\ &= m + n \\ &= \log_a(x) + \log_a(y) \end{aligned}$$

5. $\log_a(x^n) = n \log_a(x)$

since ...

The statement is essentially

$$(a^m)^n = a^{mn}$$

in disguise. Let $x = a^m$. Then

$$\begin{aligned} \log_a(x^n) &= \log_a((a^m)^n) = \log_a(a^{mn}) \\ &= mn \\ &= nm \\ &= n \log_a(x) \end{aligned}$$

6. $\log_b(x) = \frac{\log_a(x)}{\log_a(b)}$

since ...

$$\begin{aligned} b^{\log_b(x)} = x &\iff (a^{\log_a(b)})^{\log_b(x)} = x \\ &\iff a^{\log_a(b) \cdot \log_b(x)} = x \\ &\iff \log_a(b) \cdot \log_b(x) = \log_a(x) \\ &\iff \log_b(x) = \frac{\log_a(x)}{\log_a(b)}. \end{aligned}$$

Additional properties

Suppose $\boxed{1 \neq a > 0}$, $\boxed{1 \neq b > 0}$ and $\boxed{x, y > 0}$. Then the function \log_a has the following additional properties that are corollaries of the previous properties.

$$7. \quad \boxed{\log_a \left(\frac{x}{y} \right) = \log_a(x) - \log_a(y)}$$

since ...

$$\begin{aligned} \log_a \left(\frac{x}{y} \right) &= \log_a(x \cdot y^{-1}) \\ &= \log_a(x) + \log_a(y^{-1}), \quad \text{by Property 4.} \\ &= \log_a(x) - \log_a(y), \quad \text{by Property 5.} \end{aligned}$$

$$8. \quad \boxed{\log_a \left(\frac{1}{y} \right) = -\log_a(y)}$$

since ... $\log_a \left(\frac{1}{y} \right) = \log_a(1) - \log_a(y) = -\log_a(y)$, by Properties 7 and 1.

$$9. \quad \boxed{\log_b(a) = \frac{1}{\log_a(b)}}$$

since ... $\log_b(a) = \frac{\log_a(a)}{\log_a(b)} = \frac{1}{\log_a(b)}$, by Properties 6 and 2.

Special properties

Below we will be assuming $\boxed{a > 1}$, and $\boxed{x, y > 0}$.

10. If $a > 1$ then \log_a is an increasing function,

since ...

Let $a > 1$. Then

$$a^m > a^n \iff m > n.$$

Now let $x = a^m$, $y = a^n$. Then

$$x > y \iff \log_a(x) > \log_a(y).$$

In particular,

$$\log_a(x) > \log_a(y) \implies x > y,$$

which is equivalent to saying \log_a is an *increasing function*.

11. If $a > 1$ and $x > 1$ then $\log_a(x) > 0$.

since ...

$a > 1$ implies \log_a is an increasing function. So, since $x > 1$,

$$\log_a(x) > \log_a(1) = 0.$$