

**A definition of exponential functions**

For each  $b > 0$  the function  $\exp_b x := b^x$  is defined for all  $x \in \mathbb{R}$ . We call  $b$  the *base* and  $x$  the *exponent*, or *power*. We showed before that if  $b > 1$  and  $r < s$  are rational, then  $b^r < b^s$ . The same is true for arbitrary real exponents  $x$  and  $y$ . This is a consequence of our definition of  $b^x$  as the limit of a sequence of rational powers of  $b$  and the inequality just mentioned.

**Some properties of exponential functions**

We have just seen that

if  $b > 1$  then  $x \rightarrow b^x$  is a strictly increasing function.

In case  $0 < b < 1$  then by the rules for exponents  $b^x = b^{(-1)(-x)} = (b^{-1})^{-x} = 1/(b^{-1})^x$ , so that

if  $0 < b < 1$  then  $x \rightarrow b^x$  is a strictly decreasing function.

It is to be expected that exponential functions are continuous and that they each take on all positive values, so their monotonicity will give them continuous inverses, unless  $b = 1$ , in which case we have  $1^x \equiv 1$ .

It is to be expected that exponential functions are differentiable.

For continuity we will use an extension, to the case of real exponents, of a theorem we proved before.

**Theorem (Comparison of Powers):** If  $t$  and  $u$  are real numbers,  $k$  is an integer at least as large as  $\max(|t|, |u|, 1)$ ,  $a > 1$  and  $a^{-1} \leq x \leq a$ , then

$$|x^t - x^u| \leq a^{5k}|t - u|.$$

*Proof:* Left to you;  $k$  no longer needs to be an integer.

**Corollary:** For each  $b > 1$ , the exponential function  $\exp_b x$  is locally Lipschitz continuous.

*Proof:* Let us be given  $c < d \in \mathbb{R}$ . We will show that there is a constant  $L$  such that  $|\exp_b x - \exp_b y| \leq L|x - y|$  whenever  $x$  and  $y$  belong to  $[c, d]$ . In the context of the Comparison of Powers Theorem, we set  $a := b + 1$ , and change notation there, replacing  $x$  by  $b$  and setting  $k := \lceil |d| \rceil + 1$  (this choice of  $k$  may be unnecessarily large). This easily gives

$$|b^x - b^y| \leq (b + 1)^{5k}|x - y|, \text{ if } x, y \in [c, d].$$

**With one exception, the exponential functions map  $\mathbb{R}$  onto  $\mathbb{R}^+$** 

The exception is  $\exp_1 x \equiv 1$ . Because of the identity  $b^x = 1/(b^{-1})^x$ , we may assume  $b > 1$ . The proof of the Archimedean Property can be modified to show that the set  $\{b^n : n \in \mathbb{Z}\}$  has supremum  $+\infty$  and infimum 0. Since  $x \rightarrow b^x$  is continuous, it follows using the Intermediate Value Theorem that  $\exp_b x$  maps  $\mathbb{R}$  onto  $\mathbb{R}^+$ .

We can summarize these properties of exponential function in a Theorem.

**Theorem:** If  $b$  is positive and not equal to 1 then  $\exp_b x$  is a strictly monotone continuous function that maps  $\mathbb{R}$  onto  $\mathbb{R}^+$ . In particular,

if  $1 < b < \infty$  then  $x \rightarrow b^x = \exp_b x$  is a strictly increasing function

$$\text{and: } \lim_{x \rightarrow +\infty} b^x = +\infty, \quad \lim_{x \rightarrow -\infty} b^x = 0;$$

if  $0 < b < 1$  then  $x \rightarrow b^x = \exp_b x$  is a strictly decreasing function

$$\text{and: } \lim_{x \rightarrow +\infty} b^x = 0, \quad \lim_{x \rightarrow -\infty} b^x = +\infty.$$

**On the differentiability of  $\exp_b x$** 

We will gradually approach the proof of differentiability by taking advantage of the rules for exponents. We will use the Mean Value Theorem repeatedly.

**Lemma:** The exponential function  $\exp_b x$  is differentiable at  $x_o \in \mathbb{R}$  if and only if  $\exp_b x$  is differentiable at 0, and then

$$\frac{d}{dx} \exp_b x \Big|_{x=x_o} = \exp_b x_o \cdot \frac{d}{dx} \exp_b x \Big|_{x=0}.$$

*Proof:* The difference quotient we want is

$$\frac{b^x - b^{x_o}}{x - x_o} = b^{x_o} \cdot \frac{b^{x-x_o} - 1}{x - x_o},$$

by the rules for exponents. We can now take limits and revert to the more formal notation in the statement of the Theorem.

**Lemma:** The exponential function  $\exp_b x$  is differentiable at 0.

*Proof:* We may assume without loss of generality that  $b > 1$ , since  $b^x = (b^{-1})^{-x}$ , or  $\exp_b x = \exp_{1/b}(-x)$ , so we can use the Chain Rule to handle the case  $0 < b < 1$ . By the way, the case  $b = 1$  is immediate: the derivative of  $1^x$  is identically zero.

We need to show that

$$\lambda(b) := \lim_{h \rightarrow 0} \frac{b^h - 1}{h}$$

exists. We have another reduction to make, in case  $h$  is negative:

$$\frac{b^h - 1}{h} = b^h \frac{1 - b^{-h}}{h} = -b^h \frac{1 - b^{-h}}{-h} = b^h \frac{b^{-h} - 1}{-h},$$

so the difference quotient at a negative  $h$  is proportional to that for the positive number  $-h$ . By continuity,  $b^h \rightarrow 1$  as  $h \rightarrow 0$ , so all we need to do is show that

$$\lim_{h \downarrow 0} \frac{b^h - 1}{h}$$

exists. To do so, let us use the Mean Value Theorem to show that  $\frac{b^h - 1}{h}$  is an increasing function of  $h > 0$ . This would do, since  $\frac{b^h - 1}{h} > 0$ , and now the ratio would decrease and be bounded below as  $h \downarrow 0$ .

We momentarily change our point of view back to the “power” case, in which we regard  $h$  as a constant, and regard  $b$  as the variable (“frozen” at  $b$ ). We will also choose another exponent,  $k$ ,  $0 < k < h$ . To pass from  $k$  to  $h$  we use one of the rules for exponents:  $b^h = (b^k)^{h/k}$ . To apply the Mean Value Theorem we will work with the function  $x \rightarrow x^{h/k}$ . Our interval will be  $[1, b^k]$ , so there exists  $1 < \beta < b^k$  such that

$$\frac{b^h - 1}{h} = \frac{(b^k)^{h/k} - 1}{h} = \frac{(h/k) [(b^k)^{h/k} - 1]}{h} = (h/k) \frac{\beta^{h/k-1} [b^k - 1]}{h} = \beta^{h/k-1} \frac{b^k - 1}{k} > \frac{b^k - 1}{k}$$

since (because  $h/k > 1$ )  $\beta^{h/k-1} > 1$ . Thus the desired limit exists, and the proof of the Lemma is complete.

We have no *immediate* information about the value of this limit, except that it is non-negative. However, this Lemma (including its reduction to the case  $b > 1$ ), the previous Lemma, and the knowledge that  $b^x$  is strictly increasing when  $b > 1$ , and strictly decreasing when  $b < 1$ , together imply that the limit is positive when  $b > 1$ , negative when  $0 < b < 1$ , and (as we already knew) zero when  $b = 1$ . We can put all this together into yet another Lemma.

**Definition and Lemma:** For  $b > 0$  we set  $\lambda(b) := \frac{d}{dx} \exp_b x \Big|_{x=0}$ . If  $a$  and  $b$  are positive and  $x$  is real,

$$\lambda(1) = 0, \lambda(ab) = \lambda(a) + \lambda(b), \lambda(b) > 0 \iff b > 0, \lambda(b^x) = x\lambda(b).$$

The quantity  $\lambda(b)$  is called the natural logarithm of  $b$ .

*Proof:* The first and third items in the list have already been proved. In the second we “subtract and add:”

$$\lambda(ab) \leftarrow \frac{(ab)^h - 1}{h} = \frac{a^h b^h - 1}{h} = \frac{a^h b^h - a^h + a^h - 1}{h} = a^h \frac{b^h - 1}{h} + \frac{a^h - 1}{h} \rightarrow \lambda(a) + \lambda(b)$$

as  $h \rightarrow 0$ , since  $a^h \rightarrow 1$  as  $h \rightarrow 0$ .

The last item falls into two cases: when  $x = 0$ , both sides are zero. When  $x \neq 0$ ,

$$\lambda(b^x) \leftarrow \frac{(b^x)^h - 1}{h} = x \frac{b^{xh} - 1}{xh} \rightarrow x\lambda(b).$$

We can summarize, in two Theorems, what we have done in connection with the differentiability of exponential functions. The second Theorem will include further properties of the natural logarithm, as a function in its own right.

**Theorem:** For each  $b > 0$  the exponential function  $\exp_b x$  is differentiable at every  $x_o \in \mathbb{R}$ , and

$$\frac{d}{dx} \exp_b x \Big|_{x=x_o} = \log b \cdot \exp_b x_o,$$

where

$$\log b := \lim_{h \rightarrow 0} \frac{b^h - 1}{h} = \frac{d}{dx} \exp_b x \Big|_{x=0}.$$

*Proof:* Already done.

**Theorem:** The function  $\log : \mathbb{R}^+ \rightarrow \mathbb{R}$  is continuous, strictly increasing, onto and differentiable. The function  $y \rightarrow \log y$  has these properties:

$$\begin{aligned} \log 1 &= 0; \\ \log ab &= \log a + \log b, \text{ for } a > 0 \text{ and } b > 0; \\ \log b^x &= x \log b, \text{ for } b > 0 \text{ and } x \in \mathbb{R}; \\ \frac{d}{dy} \log y &= \frac{1}{y}, \text{ for } y > 0. \end{aligned}$$

*Proof:* The first three properties were proved when we still used  $\lambda(y)$  to denote  $\log y$ . For differentiability it will be handy to use a consequence of the first two properties:  $\log(1/y) = -\log y$ .

We let  $y_o > 0$ ,  $\delta > 0$ , and let  $y$  satisfy

$$\frac{y_o}{1+\delta} < y < y_o(1+\delta), \text{ so that } y = y_o\theta \text{ with } \frac{1}{1+\delta} < \theta < 1+\delta.$$

This is a clever choice of bounds, because we also have  $\frac{1}{1+\delta} < 1/\theta < 1+\delta$ . We now examine the difference quotient, and once more use the Mean Value Theorem:

$$\begin{aligned} \frac{\log y - \log y_o}{y - y_o} &= \frac{\log(y/y_o)}{y - y_o} \\ &= \frac{1}{y - y_o} \lim_{h \rightarrow 0} \frac{(y/y_o)^h - 1}{h} \\ &= \frac{1}{y - y_o} \lim_{h \rightarrow 0} \frac{\theta^h - 1}{h} = \frac{1}{y - y_o} \lim_{h \rightarrow 0} \frac{h\eta^{h-1}(\theta - 1)}{h} \\ &= \frac{1}{y - y_o} \lim_{h \rightarrow 0} \eta^{h-1}((y/y_o) - 1) = \frac{1}{y_o} \lim_{h \rightarrow 0} \eta^{h-1}. \end{aligned}$$

Here,  $\eta$  is a number strictly between  $\theta$  and 1, that also depends on  $h$ . We know the limit exists, but we have no direct way to evaluate it (why does the limit exist?). Let  $\rho$  denote  $1 + \delta$ , which exceeds the maximum of  $\theta$  and  $1/\theta$ . Then  $1/\rho < \eta < \rho$ , and  $h - 1 < 0$  for  $h$  small in absolute value, so

$$\rho^{h-1} < \eta^{h-1} < (1/\rho)^{h-1}.$$

We know the limits on  $\rho$  and  $1/\rho$ , so we conclude (how do we keep the  $<$  below?) that

$$1/\rho < \lim_{h \rightarrow 0} \eta^{h-1} < \rho.$$

But  $1 < \rho = 1 + \delta$  so  $\delta < \frac{-\delta}{1 + \delta} < \lim_{h \rightarrow 0} \eta^{h-1} - 1 < \delta$ . That is,  $|\lim_{h \rightarrow 0} \eta^{h-1} - 1| < \delta$ . Thus for  $\frac{y_0}{1 + \delta} < y < y_0(1 + \delta)$ ,

$$\left| \frac{\log y - \log y_0}{y - y_0} - \frac{1}{y_0} \right| = \frac{1}{y_0} \left| \lim_{h \rightarrow 0} \eta^{h-1} - 1 \right| < \frac{\delta}{y_0}.$$

Aside: this can be cleaned up to be a useful inequality for  $\log(1 + x)$ , when  $x$  is close to 0.

The proof of differentiability is completed by letting  $\delta \rightarrow 0$ . Continuity follows from differentiability, strict increase follows from the positivity of the derivative and it only remains to show that for every  $x \in \mathbb{R}$ ,  $x = \log y$  for some  $y > 0$ .

We will use Rudin's 3.30 and 3.31, Definition and Theorem, respectively:

$$e := \sum_{n=0}^{\infty} \frac{1}{n!} \quad \text{and} \quad e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

We may now use continuity, and part of the Theorem we just proved, in 3.31:

$$\log(e) = \log \left[ \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \right] = \lim_{n \rightarrow \infty} \log \left[ \left(1 + \frac{1}{n}\right)^n \right] = \lim_{n \rightarrow \infty} n \log \left(1 + \frac{1}{n}\right).$$

In the last expression, inside the limit, we recognize that

$$n \log \left(1 + \frac{1}{n}\right) = \frac{\log \left(1 + \frac{1}{n}\right) - \log 1}{\frac{1}{n}},$$

the difference quotient for  $\log y$  at  $y = 1$ , evaluated at  $h = 1/n$ . Taking the limit then leads to the limit being 1, and we conclude that

$$\log e = 1.$$

### An adjustment in notation

We will, from now on, write  $\exp x = \exp(x) := \exp_e(x) = e^x$ . The notation will be useful when the exponent is a complicated expression. When someone says "the exponential" or "the logarithm,"  $e^x$  and  $\log y$  are usually meant. The notation  $\ln y$  is often used for  $\log y$ .

**A brief digression** We could have done this differently, by showing *directly* that there is a number whose natural logarithm is 1. For example, we can verify by calculations that the last inequality in

$$0 < \log 2 = \lim_{h \rightarrow 0} \frac{2^h - 1}{h} < \frac{2^{1/2} - 1}{1/2} < 1$$

is true. The next-to-last inequality is part of the proof that the limit defining  $\log b$  exists. Then, by the Archimedean Property, there exists an integer  $N$  such that  $\log 2^N = N \log 2 > 1$ . The Intermediate Value Theorem then assures us that 1 is a value of  $\log y$ . **end of digression**

We can easily now prove that  $y \rightarrow \log y$  is onto  $\mathbb{R}$ . We select  $b = e$  in the third property stated in the Theorem on the natural logarithm, and set  $y := e^x$ . Then

$$\log y = \log(e^x) = x \log e = x.$$

This completes the proof of the Theorem.

**Corollary:** The inverse function of  $e^x$  (as a function defined on  $\mathbb{R}$ , taking on all positive values) is  $\log y$  (as a function defined on  $\mathbb{R}^+$ , taking on all real values).

*Proof:* We have shown that  $\log e^x \equiv x$ , for  $x$  real, and that each function (exponential and logarithm) is one-to-one and onto. Thus  $e^{\log y} \equiv y$  (why?! ) for  $y$  positive.

**Remark:** The inverse pairs corresponding to the other exponential functions are  $\exp_b x$  and

**Definition:**  $\log_b y := \log y / \log b$ . The formulas are:

$$\exp_b(\log_b y) = \exp_b\left(\frac{\log y}{\log b}\right) = y \text{ for all } y > 0, \text{ and } \log_b(\exp_b x) = \frac{\log(\exp_b x)}{\log b} = x \text{ for all } x \in \mathbb{R}.$$

### Further properties of exponential and logarithmic functions

#### 1. Formulas for the exponential and logarithmic functions in terms of $e^x$ and $\log y$

$$\exp_b x = b^x = e^{x \log b}; \quad \log_b y = \log y / \log b$$

Next we list some useful limits. Several of them say that one function tends to its limit faster than another. Also, some can be expressed as “little-oh” equations. For example, the second one means  $\log y = o(y)$  as  $y \rightarrow \infty$ .

#### 2. Some important logarithmic limits

$$\lim_{y \rightarrow \infty} \log y = +\infty; \quad \lim_{y \rightarrow \infty} \frac{\log y}{y} = 0; \quad \lim_{y \rightarrow 0} y \log y = 0; \quad \lim_{y \rightarrow 0} \log y = -\infty;$$

$$\lim_{h \rightarrow 0} \frac{\log(1+h)}{h} = \lim_{y \rightarrow 1} \frac{\log y}{y-1} = 1; \quad \lim_{y \rightarrow 0} y^y = 1.$$

In this list of limits, we already know the first, fourth and fifth (in the fifth, insert “ $-\log 1$ ” into the numerators). The second can be made an application of l’Hospital’s Rule. The third is really the second with the literal change of variable  $y \rightarrow 1/y$ . The sixth is an application of the third:  $y^y = e^{\log y \cdot y} = e^{y \log y}$ .

The second limit,  $\lim_{y \rightarrow \infty} \frac{\log y}{y} = 0$ , can be extended:

$$\text{For all } p > 0, \quad \lim_{y \rightarrow \infty} \frac{\log y}{y^p} = 0.$$

The verification, an application of one of the properties of the logarithm, is left to you. One more useful result concerns zero limits of positive functions.

**Theorem:** A function that is positive in a deleted neighborhood of a point  $x_0$  has limit zero at  $x_0$  if and only if

$$\lim_{x \rightarrow x_0} \log f(x) = -\infty.$$

*Proof:* Given the limit,

$$\lim_{x \rightarrow x_0} f(x) = \exp(\log f(x)) = \lim_{t \rightarrow -\infty} e^t = 0.$$

The other part of the proof is similar.

#### 3. Some important exponential limits

$$\text{For all } p > 0, \quad \lim_{x \rightarrow \infty} x^p e^{-x} = 0 \text{ and } \lim_{x \rightarrow \infty} x^{-p} e^x = +\infty; \quad e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

To prove the first we take the logarithm:

$$\log(x^p e^{-x}) = x \left( p \frac{\log x}{x} - 1 \right).$$

As  $x \rightarrow \infty$ , the expression in parentheses tends to  $-1$ , so  $\log(x^p e^{-x}) \rightarrow -\infty$  as  $x \rightarrow \infty$ . We can apply the Theorem above. The second limit is a consequence of the first.

To prove the third limit we take the natural logarithm again: we are to show that

$$x = \lim_{n \rightarrow \infty} \log \left( 1 + \frac{x}{n} \right)^n.$$

But as we have done before,

$$\log \left( 1 + \frac{x}{n} \right)^n = n \log \left( 1 + \frac{x}{n} \right) = x \frac{\log \left( 1 + \frac{x}{n} \right) - \log 1}{\frac{x}{n}} \rightarrow x$$

as  $n \rightarrow \infty$ . We can take the exponential now to finish the proof.

#### 4. Some characterizations of exponential and logarithmic functions

Functions that satisfy certain *functional equations*, and suitable side conditions, *must* be power, exponential or logarithmic functions. We will derive these functional equations here. They all reduce to the one for powers.

**Theorem:** If  $f(x)$  is real-valued, defined for positive  $x$ , continuous and not identically zero, and

$$f(xy) = f(x)f(y) \text{ for all } x > 0 \text{ and } y > 0,$$

then  $f(x) = x^t$  for some real  $t$ .

*Proof:* Suppose  $f(x_0) \neq 0$ . Then  $f(x_0) = f(x_0 \cdot 1) = f(x_0)f(1)$ . Thus  $f(1) \neq 0$ . Then  $f(1) = f(1 \cdot 1) = f(1)f(1)$ , so  $f(1) = 1$ . But then  $f(x) \neq 0$  for every  $x > 0$  since  $f(1) = f(x)f(1/x)$ . More is true. Since  $f(x) = f(\sqrt{x} \cdot \sqrt{x}) = f(\sqrt{x})^2 > 0$ , the function is everywhere positive.

To derive the formula, we will use an identity. This is an idea similar to “write zero or one in a useful way, so it includes the quantity you are interested in.” We will write  $e$  in such a way:

$$\text{if } 0 < x \neq 1 \text{ then, from the equation } x = e^{\log x} \text{ we have } e = x^{1/\log x}.$$

We have (using induction)  $f(x) = f((x^{1/n})^n) = f(x^{1/n})^n$  and (simultaneously?)  $f(x^m) = f(x)^m$ , so that  $f(x^{m/n}) = f(x)^{m/n}$ , or  $f(x^r) = f(x)^r$  for all rational  $r$ . By continuity,  $f(x^u) = f(x)^u$  for all real  $u$ . If  $x \neq 1$ , let us set  $u = 1/\log x \neq 0$ . Then

$$f(x)^u = f(x)^{1/\log x} = f(x^{1/\log x}) = f(e) = e^{\log f(e)} =: e^t, \text{ where } t = \log f(e).$$

Hence

$$f(x) = (e^t)^{\log x} = (e^t)^{\log x} = (e^{\log x})^t = x^t, \text{ where } t = \log f(e).$$

**Theorem:** If  $f(y)$  is real-valued, defined for positive  $y$ , continuous, and

$$f(xy) = f(x) + f(y) \text{ for all } x > 0 \text{ and } y > 0,$$

then  $f(y) = A \log y$  for some real  $A$ .

*Proof:* Set  $F(y) := e^{f(y)}$ . Then  $F(xy) = e^{f(xy)} = e^{f(x)+f(y)} = F(x)F(y)$ , and  $F$  satisfies the side conditions in the theorem on powers. Thus  $F(y) = y^A$  for some real  $A$ . On taking the (natural!) logarithm we have  $f(y) = A \log y$ . In this case,  $A = f(e)$ .

**Theorem:** If  $f(x)$  is positive, defined for all real  $x$ , continuous and

$$f(x+y) = f(x)f(y) \text{ for all real } x \text{ and } y,$$

then  $f(x) = b^x$  for some positive  $b$ .

*Proof:* For  $y > 0$  let us define  $F(y) := f(\log y)$ . Then  $F(xy) = f(\log x)f(\log y) = F(x)F(y)$  for all positive  $x$  and  $y$ . Hence by the powers theorem  $F(y) = y^A$  for some real  $A$ . That is,  $f(x) = F(e^x) = (e^x)^A = (e^A)^x =: b^x$ , where  $b = f(1)$ .

### Another functional equation: additive

There is another functional equation of interest that does not really belong here. I include it because it is important at the “foundations” of mathematics. A few words about “why” will follow the theorem and its proof.

**Theorem (on continuous additive functions):** *If  $f(x)$  is defined for all real  $x$ , continuous, and*

$$f(x + y) = f(x) + f(y) \text{ for all real } x \text{ and } y,$$

*then  $f(x) = Ax$  for some real  $A$ . Comment:  $A = f(1)$ .*

*Proof:* To begin, we show that  $f(0) = 0$ , for  $f(0) = f(0 + 0) = f(0) + f(0)$ . Thus  $0 = f(x - x) = f(x) + f(-x)$ , so  $f(-x) = -f(x)$  for all  $x$ . By induction,  $f(m/n) = f((1/n) + \dots + (1/n)) = mf(1/n)$ , whenever  $m$  and  $n$  are positive integers. By taking  $m = n$  and dividing by  $n$  we have  $f(1/n) = f(1)/n$ . Thus for every positive rational number  $r = m/n$ , where  $m$  and  $n$  are positive integers,  $f(r) = f(m/n) = f(1)(m/n) = f(1)r$ . By continuity,  $f(x) = f(1)x$  for every  $x \geq 0$ . For negative  $x$ ,  $f(x) = -f(-x) = -f(1)(-x) = f(1)x$ . Thus  $f(x) = Ax$  for all  $x$ , where  $A = f(1)$ .

The point of all this is that “there exist” functions  $f(x)$  that are *NOT* continuous that satisfy the functional equation

$$f(x + y) = f(x) + f(y) \text{ for all real } x \text{ and } y.$$

The “existence” depends on the use of something often called the Axiom of Choice. It is a statement about families of sets, so it is part of Set Theory, which we use all the time in Analysis. It says, loosely, that if we are given a family of sets, no two different ones of which have any points in common, then *there exists a set  $C$*  that contains exactly one element from each of the sets in the given family of sets. According to J. L. Kelley in *General Topology*, Chapter 0, item 25, this actually is the Zermelo Postulate. Item 25 is a slightly confusing theorem that asserts that 8 statements are true, each following from a ninth statement called the Hausdorff Maximal Principle. Kelley then points out that each of the eight statements implies the Hausdorff Maximal Principle and hence all the others.

Here is Zermelo’s Postulate, seventh in the list:

**ZERMELO POSTULATE** *If  $\mathcal{A}$  is a disjoint family of non-void sets, there is a set  $C$  such that  $A \cap C$  consists of a single point for every  $A$  in  $\mathcal{A}$ .*

Here is what Kelley calls the Axiom of choice, the sixth statement in the list:

**AXIOM OF CHOICE** *If  $X_a$  is a non-void set for each member  $a$  of an index set  $A$ , then there is a function  $c$  on  $A$  such that  $c(a) \in X_a$  for each  $a$  in  $A$ .*

The way this is used: we regard  $\mathbb{R}$  as a vector space with scalars  $\mathbb{Q}$ . This is an infinite-dimensional vector space. In fact, the dimension is uncountable. The “Axiom of Choice” is used to show that every vector space has a basis, namely a subset with the property that every vector in the vector space can be expressed as a linear combination of basis elements. Specifically this means that, given a vector  $v$ , there exist *finitely many* basis elements,  $v_1, \dots, v_n$  and corresponding scalars  $c_1, \dots, c_n$  such that  $v = \sum_{i=1}^n c_i v_i$ , and the finite set of basis elements and scalars are uniquely determined by  $v$  and the basis.

A basis for  $\mathbb{R}$  as a vector space with scalars  $\mathbb{Q}$  being obtained, we call it a Hamel basis and construct a “linear transformation”  $f : \mathbb{R} \rightarrow \mathbb{R}$  as a linear mapping on this infinite-dimensional vector space. A linear transformation is completely determined by what it does on basis elements. We choose 2 different basis elements,  $b$  and  $c$  (both must be non-zero) and define  $f(x)$  this way:  $f(b) = 1 = f(c)$  and  $f(v) = 0$  for all other basis elements  $v$ . To verify that  $f$  is linear with respect to rational scalars, let  $x = \sum_{i=1}^n x_i v_i$ ,  $y = \sum_{i=1}^n y_i v_i$ . We can do this by combining

the linear combinations for  $x$  and  $y$ , and using zero coefficients as needed. Then

$$\begin{aligned} f(x + y) &= f\left(\sum_{i=1}^n x_i v_i + \sum_{i=1}^n y_i v_i\right) = f\left(\sum_{i=1}^n (x_i + y_i) v_i\right) \\ &= \sum_{i=1}^n (x_i + y_i) f(v_i) = \sum_{i=1}^n x_i f(v_i) + \sum_{i=1}^n y_i f(v_i) \\ &= \sum_{i=1}^n (x_i + y_i) f(v_i) = f(x) + f(y). \end{aligned}$$

The proof of the Theorem on continuous additive functions did not use continuity to show that  $f(rx) = rf(x)$  for rational numbers  $r$ . Hence linearity with rational scalars is the same as “additive.”

But this  $f$ , even though it satisfies the equation  $f(x + y) = f(x) + f(y)$ , cannot be continuous because if it were then  $f(x) = Ax$  for some real  $A$ . However,  $f(b) = Ab = 1 = Ac = f(c)$  is impossible. Hence using a powerful *and controversial* axiom from one of many versions of set theory, there is an additive discontinuous function on  $\mathbb{R}$ .

We can then construct discontinuous functions that satisfy the other functional equations we have examined. This is left to you.