Chapter 4: Continuous functions (121 2004–05)

While we have introduced continuous functions in the last chapter (functions that are continuous at each point of their domain), we have basically dealt with properties of continuity at single points. We now go on to deal with properties that are harder to prove and are more significant. They are 'global' properties, meaning that they depend on knowing that the function is continuous at each point of its domain, but they also depend in an important way on properties of the domain (usually an interval or a finite closed interval in these theorems).

The theorems are true because of a combination of properties of continuity and the interval.

Definition 4.1. *A subset $S \subset \mathbb{R}$ is called (sequentially)* compact *if each sequence $(x_n)_{n=1}^\infty$ in S has a subsequence $(x_{n_j})_{j=1}^\infty$ which converges to a limit $\lim_{j\to\infty} x_{n_j} \in S$ (in the set S).*

Note that this is a property that considers **all possible** sequences of terms in S. The reason to mention the *sequentially* is that there is a more abstract way of defining compactness which applies in more abstract settings. However, this definition is equivalent in the setting we are in of subsets of $\mathbb{R}$.

Theorem 4.2 (Heine–Borel). *Finite closed intervals $[a, b] \subset \mathbb{R}$ are compact sets.*

Proof. Consider any sequence $(x_n)_{n=1}^\infty$ in $[a, b]$.

Being a bounded sequence, by the Bolzano-Weierstrass Theorem (2.19) there is a subsequence $(x_{n_j})_{j=1}^\infty$ which converges to a limit $\ell \in \mathbb{R}$. We claim that $\ell \in [a, b]$.

If not $\ell < a$ or $\ell > b$.

If $\ell < a$ we can arrive at a contradiction as follows. Take $\varepsilon = (a - \ell)/2$ in the ε-N definition of $\lim_{j\to\infty} x_{n_j} = \ell$. We can find J so that $|x_{n_j} - \ell| < \varepsilon$ for $j \geq J$. But $a \leq x_n$ for all n, so $|x_{n_j} - \ell| = x_{n_j} - \ell \geq a - \ell = 2\varepsilon > \varepsilon$ so that $j = J$ cannot satisfy $|x_{n_j} - \ell| < \varepsilon$. This is a contradiction.

We can similarly rule out $\ell > b$ by taking $\varepsilon = (\ell - b)/2$ (if $\ell > b$) and showing that no j satisfies $|x_{n_j} - \ell| < \varepsilon$. $\square$

(Because of the way this theorem follows from the Bolzano-Weierstrass theorem without any really hard work, people sometimes consider them as more or less the same theorem. There is this additional aspect of the limit being in the set. Indeed the set does not come into the Bolzano-Weierstrass, as that is about a bounded sequence.)

Definition 4.3. *If $f: S \to \mathbb{R}$ is a real-valued function (on any set S), then f is called* bounded above *if its range $f(S) \subset \mathbb{R}$ is bounded above. In other words, f is bounded above if there exists a number $M \in \mathbb{R}$ so that $f(x) \le M$ for all $x \in S$.*

Similarly f is called bounded below *if there exists $L \in \mathbb{R}$ so that $L \le f(x)$ holds for all $x \in S$ (or, equivalently, if $f(S)$ is bounded below).*

A function $f: S \to \mathbb{R}$ is called bounded *if it is both bounded above and bounded below.*

Theorem 4.4. *If $f: [a,b] \to \mathbb{R}$ is a continuous function on a finite closed intervals $[a,b] \subset \mathbb{R}$ then f is bounded.*

Proof. If f is not bounded, then it is not true for any $n \in \mathbb{N}$ that $-n \le f(x) \le n$ for all $x \in [a,b]$. (If $-n \le f(x) \le n$ for all $x \in [a,b]$, then f is bounded above by n and below by $-n$.) In other words it is not the case that $|f(x)| \le n$ holds for all $x \in [a,b]$ and so there must be at least one $x_n \in [a,b]$ with $|f(x_n)| > n$. Chose one such x_n for each $n \in \mathbb{N}$ and we now have a sequence $(x_n)_{n=1}^\infty$ in $[a,b]$.

By Theorem 4.2, there is a subsequence $(x_{n_j})_{j=1}^\infty$ which converges to a limit $\ell \in [a,b]$.

Now f is continuous at $\ell \in [a,b]$ and $(x_{n_j})_{j=1}^\infty$ is a sequence in $[a,b]$ converging to ℓ. So $\lim_{j\to\infty} f(x_{n_j}) = f(\ell)$ and, as a convergent sequence, $\big(f(x_{n_j})\big)_{j=1}^\infty$ must be a bounded sequence (see the proof of Theorem 2.9 (iii) for a proof of this). But we have $|f(x_{n_j})| > n_j \ge j$ for all j and this is incompatible with the sequence being bounded. (If $L, M \in \mathbb{R}$ satisfy $L \le f(x_{n_j}) \le M$ for all j, then $|f(x_{n_j})| \le \max(|L|, |M|)$ for all j and this would imply $j < \max(|L|, |M|)$ for all $j \in \mathbb{N}$ — impossible as $\mathbb{N}$ is not bounded above.) $\square$

Example 4.5. *There are simple examples that show that the result is false if the interval is not closed or if there is even one point in the interval where the function fails to be continuous.*

For example $f: (0,1] \to \mathbb{R}$ given by $f(x) = 1/x$ is not bounded above, though it is continuous. [Not bounded above because $f(1/n) = n$, so that $\mathbb{N} \subset f((0,1])$, and $\mathbb{N}$ not bounded above.]

We can make a discontinuous example by considering $g: [0,1] \to \mathbb{R}$ with

$$g(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1/x & \text{if } x \ne 0 \end{cases}$$

Another example is $h: [0,\infty) \to \mathbb{R}$, $h(x) = -x$ which is continuous but not bounded (below). The domain $[0,\infty)$ is considered a closed interval because it includes its only finite endpoint 0.

Theorem 4.6. *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function on a finite closed interval $[a,b]$ where $a \le b$. Then f has a largest value and a smallest value.*

That is there exist $x_L, x_M \in [a,b]$ so that $f(x_L) \le f(x) \le f(x_M)$ holds for all $x \in [a,b]$.

Proof. We use Theorem 4.4. To start with, we prove that there is some $x_M \in [a,b]$ with $f(x) \le f(x_M)$ for all $x \in [a,b]$.

The range $f([a,b])$ is a nonempty set ($a \in [a,b] \Rightarrow f(a) \in f([a,b])$ for example) and by Theorem 4.4 it is bounded above. By the least upper bound principle, there is a least upper bound $U \in \mathbb{R}$ for $f([a,b])$. So $f(x) \le U$ for all $x \in [a,b]$ (and no number in $\mathbb{R}$ strictly smaller that U has this property). We want to find x_M so that $f(x_M) = U$. If no such x_M exist then $f(x) < U$ holds for all $x \in [a,b]$ and we can define a continuous function $g\colon [a,b] \to \mathbb{R}$ by

$$g(x) = \frac{1}{U - f(x)}$$

(because the denominator is never 0, this makes sense for all $x \in [a,b]$ and defines a continuous function). Applying Theorem 4.4 to g we find there is some $U_0 \in \mathbb{R}$ with $g(x) \le U_0$ for all $x \in [a,b]$. Then, as $g(x) > 0$ (reason: $f(x) < U$) we have $1/g(x) = U - f(x) \ge 1/U_0$ and so

$$U - \frac{1}{U_0} \ge f(x)$$

for all $x \in [a,b]$. As $U_0 \ge g(a) > 0$ we have $U - 1/U_0$ an upper bound for f strictly smaller than the least upper bound U — a contradiction.

Hence $x_M \in [a,b]$ with $f(x_m) = U$ must exist.

For the existence of x_L we could repeat a similar argument using the greatest lower bound, or we can apply what we have just proved to the continuous function $h\colon [a,b] \to \mathbb{R}$ given by $h(x) = -f(x)$. If $x_L \in [a,b]$ is such that $h(x) \le h(x_L)$ for all $x \in [a,b]$, then $f(x) \ge f(x_L)$ for all x. □

Theorem 4.7 (Intermediate Value Theorem). *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function on a finite closed interval with $f(a) < 0 < f(b)$. Then there is some $c \in (a,b)$ with $f(c) = 0$.*

(This theorem is the one that makes the word ‘continuous’ especially convincing. In graphical terms it says that the graph of a continuous function on an interval cannot jump from being negative to positive without being actually 0 at some place. If you imagine drawing a graph of a function starting at $x = a$ and a negative value of y, ending at $x = b$ with a positive value of y, and if you don't allow yourself to lift your pen in between, you will be fairly convinced that the graph has to cross the axis $y = 0$.

However, that is not a proof. To convince yourself that the result is not obvious, imagine that we only had rational numbers instead of all the real numbers. Then the function $f(x) = x^2 - 2$ on $[0, 2] \cap \mathbb{Q}$ starts out at $y = f(0) = -2 < 0$ and ends up at $y = f(2) = 2 > 0$ but (as there is no $\sqrt{2}$ in $\mathbb{Q}$) is never 0. Thus the Intermediate Value Theorem is actually a result of continuity plus the fact that there are no holes in the real axis (no numbers left out that 'should' be there, or no points on the line that do not correspond to a real number).

Proof. Let $S = \{t \in [a, b] : f(x) < 0 \text{ for all } x \in [a, t]\}$. Notice that $a \in S$ and so $S \neq \emptyset$. Being bounded above by b, S must have a least upper bound c. We claim that $a < c < b$ and that $f(c) = 0$.

To help with the proof we divide part of it out as a lemma, and then we will return to the proof. □

Lemma 4.8. *Let $f: S \to \mathbb{R}$ be a function on a set $S \subset \mathbb{R}$ that is continuous at a point $x_0 \in S$ and that satisfies $f(x_0) > 0$. Then there is a $\delta > 0$ so that*

$$x \in S, |x - x_0| < \delta \Rightarrow f(x) > 0.$$

Proof. If $f(x_0) > 0$ then $\varepsilon = f(x_0)/2 > 0$ and we can use the definition of continuity of f at x_0 to find $\delta > 0$ so that

$$x \in S, |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

But if $|f(x) - f(x_0)| < \varepsilon$, then it follows that $f(x) - f(x_0) > -\varepsilon = -f(x_0)/2 \Rightarrow f(x) > f(x_0)/2 > 0$. Hence, for this $\delta > 0$ we have

$$x \in S, |x - x_0| < \delta \Rightarrow f(x) > 0.$$

□

Lemma 4.9. *Let $f: S \to \mathbb{R}$ be a function on a set $S \subset \mathbb{R}$ that is continuous at a point $x_0 \in S$ and that satisfies $f(x_0) < 0$. Then there is a $\delta > 0$ so that*

$$x \in S, |x - x_0| < \delta \Rightarrow f(x) < 0.$$

Proof. Apply Lemma 4.8 to $-f(x)$. □

Proof. (of Theorem 4.7 continued)

First note that $f(a) < 0$ and so by Lemma 4.9 there is $\delta > 0$ so that $x \in [a,b], |x-a| < \delta \Rightarrow f(x) < 0$. As $f(b) > 0$, we must have $\delta \leq b-a$. Also $a \leq x < a+\delta \Rightarrow x \in [a,b], |x-a| < \delta \Rightarrow f(x) < 0$. Hence $a \leq x < a+\delta \Rightarrow f(t) < 0$ for all $t \in [a,x]$. Thus $[a, a+\delta) \subset S$.

Hence $c = \text{lub}(S) \geq a+\delta > a$.

Also $f(b) > 0$ and so by Lemma 4.8 there is $\delta_0 > 0$ so that $x \in [a,b], |x-b| < \delta_0 \Rightarrow f(x) > 0$. Since $f(a) < 0$ we must have $\delta_0 < b-a$. Also, $b-\delta_0 < x \leq b \Rightarrow x \in [a,b], |x-b| < \delta_0 \Rightarrow f(x) > 0 \Rightarrow x \notin S$. Thus $S \subset [a, b-\delta_0]$ and $c \leq b-\delta_0 < b$. Hence we have $c \in (a,b)$.

Finally, we claim that $f(c) = 0$. If not either $f(c) > 0$ or $f(c) < 0$.

If $f(c) > 0$ we can apply Lemma 4.8 to find $\delta_1 > 0$ so that $x \in [a,b], |x-c| < \delta_1 \Rightarrow f(x) > 0$. But then $\delta_1 > c-a$ since $f(a) < 0$ and we have $c-\delta_1 < x \leq c \Rightarrow x \in [a,b], |x-c| < \delta_1 \Rightarrow f(x) > 0 \Rightarrow x \notin S$. As $S \subset [a,c]$ (because $c = \text{lub}(S)$ and $S \subset [a.b]$) we must have $S \subset [a, c-\delta_1]$ and $c-\delta_1$ is an upper bound for S strictly less than the least upper bound. This contradiction rules out $f(c) > 0$.

If, on the other hand, $f(c) < 0$, then we can apply Lemma 4.9 to find $\delta_2 > 0$ so that $x \in [a,b], |x-c| < \delta_2 \Rightarrow f(x) < 0$. Since $f(b) > 0$ we must have $c+\delta_2 \leq b$.

As $c-\delta_2 < c = \text{lub}(S)$, $c-\delta_2$ cannot be an upper bound for S and so there is $t \in S$ with $t > c-\delta_2$. Certainly $t \leq c$ and so $t \in (c-\delta_2, c]$. Consider $c+\delta_2/2$. We know $f(x) < 0$ for all $x \in [a,t]$ (since $t \in S$) and also for all $x \in [t, c+\delta_2/2]$ because $x \in [t, c+\delta_2/2] \Rightarrow x \in [a,b]$ and $|x-c| < \delta_2$. Putting these together we have $f(x) < 0$ for all $x \in [a,t] \cup [t, c+\delta_2/2] = [a, c+\delta_2/2]$ and so $c+\delta_2/2 \in S$. But $c+\delta_2/2 > c = \text{lub}(S)$ is then a contradiction.

These leaves $f(c) = 0$ as the only possibility. □

Example 4.10. *There is a positive number* $x \in \mathbb{R}$ *with* $x^2 = 3$.

Consider $f\colon [1,2] \to \mathbb{R}$ given by $f(x) = x^2 - 3$. We have $f(1) = -2 < 0 < 1 = f(2)$ and also that f is continuous on $[0,1]$. (In fact $x \mapsto x^2-3$ is continuous on $\mathbb{R}$ since it is a polynomial and then f is the restriction of the polynomial to $[1,2]$, hence continuous.)

By the Intermediate Value Theorem, there is $x \in (1,2)$ with $f(x) = 0$, that is $x^2 - 3 = 0$.

(You might like to compare this with the proof in 1.19 that there is a number $\sqrt{2} \in \mathbb{R}$. Of course we have a lot more theory involved now, but the proof is almost effortless now.)

Corollary 4.11 (Intermediate Value Theorem, marginally improved). *Let* $f\colon [a,b] \to \mathbb{R}$ *be a continuous function on a finite closed interval, and let* $y_0 \in \mathbb{R}$ *with* $f(a) < y_0 < f(b)$*. Then there is some* $c \in (a,b)$ *with* $f(c) = y_0$*.*

Proof. Apply Theorem 4.7 to the function $g\colon [a,b] \to \mathbb{R}$ where $g(x) = f(x) - y_0$. We have $g(a) < 0 < g(b)$. If $g(c) = 0$ then $f(c) = y_0$. □

Corollary 4.12 (Intermediate Value Theorem, slight variation). *Let* $f\colon [a,b] \to \mathbb{R}$ *be a continuous function on a finite closed interval, and let* $y_0 \in \mathbb{R}$ *with* $f(a) > y_0 > f(b)$*. Then there is some* $c \in (a,b)$ *with* $f(c) = y_0$*.*

Proof. Apply Corollary 4.11 to the function $g\colon [a,b] \to \mathbb{R}$ where $g(x) = -f(x)$. We have $g(a) < -y_0 < g(b)$. If $g(c) = -y_0$ then $f(c) = y_0$. □

Proposition 4.13. *A subset* $S \subset \mathbb{R}$ *is an interval if and only if it has the following "between-ness" property:*

If $\alpha, \beta \in S$ *with* $\alpha < \beta$*, then* $[\alpha, \beta] \subset S$

Proof. $\Rightarrow$: It is easy to see that every interval (see 3.2) has this property.

$\Leftarrow$: Assume $S \subset \mathbb{R}$ is a subset with the "between-ness" property above. If $S = \emptyset$, the $S = (a,a)$ for any $a \in \mathbb{R}$. So S is an interval. If $S = \{a\}$ has just one point, then $S = [a,a]$ is also an interval.

From now on we assume that S has more than one point. The proof is based on considering several cases about whether S is bounded above or not, bounded below or not.

First assume S is both bounded above and bounded below. Let $a = \text{glb}(S)$, $b = \text{lub}(S)$. They exist since S is not empty. We must have $a < b$ since if $a = b$ then $S \subset [a,a]$ could only have one point. We claim that $(a,b) \subset S$. To show this, let $x \in (a,b)$. Then $x < b = \text{lub}(S) \Rightarrow x$ not an upper bound for $S \Rightarrow$ there is $\beta \in S$ with $x < \beta$. Also, $x > a = \text{glb}(S) \Rightarrow x$ not a lower bound for $S \Rightarrow$ there is $\alpha \in S$ with $x > \alpha$. Now $\alpha < x < \beta$ and so by the "between-ness" property of S we have $x \in [\alpha, \beta] \subset S$. Since this is true about each $x \in (a,b)$ we have $(a,b) \subset S$. Also, from the definitions of a and b, $S \subset [a,b]$. This leaves 4 possibilities for S, $[a,b]$, $[a,b)$, $(a,b]$ and (a,b) depending on whether $a \in S$ and/or $b \in S$.

Next consider the case where S is bounded above but not below. Put $b = \text{lub}(S)$. We claim that $(-\infty, b) \subset S \subset (-\infty, b]$. The proof is essentially identical the the argument we have just given. We then have two possibilities $S = (-\infty, b]$ and $S = (-\infty, b)$ depending on whether $b \in S$ or not.

If S is bounded below but not above, put $a = \text{glb}(S)$ and then a similar argument shows $S = [a, \infty)$ or $S = (a, \infty)$.

Finally if S is not bounded above and not bounded below, we can show that $\mathbb{R} \subset S$ (because $x \in S \Rightarrow x$ is neither an upper bound not a lower bound for $S \Rightarrow$ there are $\beta, \alpha \in S$ with $x < \beta$ and $x > \alpha$ — hence $x \in [\alpha, \beta] \subset S$ by the "between-ness" property of S.

In all cases, we end up with S being an interval. ☐

Theorem 4.14. *If $I \subset \mathbb{R}$ is an interval and $f \colon I \to \mathbb{R}$ is a continuous function, then the range $f(I)$ is an interval.*

Proof. According to Proposition 4.13, it is sufficient to show that $f(I)$ has the "between-ness" property.

Let $\alpha, \beta \in f(I)$ with $\alpha < \beta$. Then there must be $a, b \in S$ with $f(a) = \alpha$ and $f(b) = \beta$. To show $(\alpha, \beta) \subset f(I)$, take $y_0 \in (\alpha, \beta)$. Then there are two situations.

$a < b$: Then we have $f(a) = \alpha < y_0 < \beta = f(b)$ and we can apply Corollary 4.11 to get $c \in (a, b) \subset I$ with $f(c) = y_0$. Hence $y_0 \in f(I)$.

$a > b$: Then we have $f(b) = \beta > y_0 > \alpha = f(a)$ and we can apply Corollary 4.12 on the interval $[b, a]$ to get $c \in (b, a) \subset I$ with $f(c) = y_0$. Hence $y_0 \in f(I)$.

Thus, we have shown $(\alpha, \beta) \subset f(I)$. Since $\alpha, \beta \in f(I)$, we have $[\alpha, \beta] \subset f(I)$.

This establishes that $f(I)$ has the "between-ness" property. ☐

(In an abstract sense, this theorem is another incarnation of the Intermediate Value Theorem.)

We now apply some of the results to the situation of *polynomials*. A polynomial is a a finite sum $p(x) = a + 0 + a_1 x + \cdots + a_n x^n = \sum_{j=1}^n a_j x^j$ for some constants $a_j \in \mathbb{R}$ ($0 \leq j \leq n$). If $a_n \neq 0$ we say that n is the *degree* of the polynomial, while if $a_n = 0$, we can rewrite the sum without any x^n term.

Sometimes the term *monic polynomial* is used for polynomials $p(x) = x^n + \sum_{j=0}^{n-1} a_j x^j$ where the coefficient of x^n is 1. Most questions about polynomial equations $p(x) = 0$ can be reduced to questions where $p(x)$ is monic (by dividing across by a_n if $a_n \neq 1$, with n the degree)

Proposition 4.15. *If $p(x) = \sum_{j=0}^n a_j x^j$ is a polynomial of degree n, then there is $R > 0$ so that*

$$|p(x) - a_n x^n| < \frac{1}{2} |a_n x^n|$$

holds for all x with $|x| > R$.

(In other words for $|x|$ large, the highest term $a_n x^n$ dominates all the others.)

Proof. It is enough to prove the result for $p(x)$ monic (as can be seen by dividing $p(x)$ by a_n.

Assuming $a_n = 1$, take $R = \max\left(1, 2\sum_{j=0}^{n-1}|a_j|\right)$. Then we have

$$p(x) - x^n = \sum_{j=0}^{n-1} a_j x^j$$

and for $|x| > R$. Thus

$$|p(x) - x^n| \leq \sum_{j=0}^{n-1} |a_j||x|^j \leq \sum_{j=0}^{n-1} |a_j||x|^{n-1}$$

using the triangle inequality first and then the observation $|x| > 1 \Rightarrow |x|^j \leq |x|^{j+1}$. Thus we have

$$|p(x) - x^n| \leq \left(\sum_{j=0}^{n-1} |a_j|\right)|x|^{n-1} \leq (R/2)|x|^{n-1} < (|x|/2)|x|^{n-1} = |x|^n/2.$$

$\square$

Proposition 4.16. *Any odd degree polynomial equation $\sum_{j=0}^{n} a_j x^j = 0$, or*

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0,$$

(with n odd and $a_n \neq 0$) has a solution $x \in \mathbb{R}$.

Proof. Fix an odd degree polynomial equation, $p(x) = \sum_{j=0}^{n} a_j x^j = 0$. Dividing by a_n, it is enough to how that in the case $a_n = 1$ (monic polynomial case) the equation has a solution. So take $p(x) = \sum_{j=0}^{n-1} a_j x^j + x^n$.

By Proposition 4.15, we can find R so that

$$|x| > R \Rightarrow |p(x) - x^n| \leq \frac{1}{2}|x|^n.$$

This means that, for $x = R + 1$ we have

$$\begin{aligned} |p(R+1) - (R+1)^n| &< \frac{1}{2}(R+1)^n \\ \Rightarrow (R+1)^n - \frac{1}{2}(R+1)^n < p(R+1) &< (R+1)^n + \frac{1}{2}(R+1)^n \end{aligned}$$

or

$$\frac{1}{2}(R+1)^n < p(R+1) < \frac{3}{2}(R+1)^n.$$

In particular $p(R+1) > 0$.

Taking now $x = -(R+1)$ we get

$$|p(-(R+1)) - (-(R+1))^n| < \frac{1}{2}| - (R+1)^n|$$

and since n is odd this means

$$\begin{aligned} |p(-(R+1)) + (R+1)^n| &< \frac{1}{2}(R+1)^n \\ \Rightarrow -(R+1)^n - \frac{1}{2}(R+1)^n < p(-(R+1)) &< -(R+1)^n + \frac{1}{2}(R+1)^n \end{aligned}$$

and from this we conclude that $p(-(R+1)) < -\frac{1}{2}(R+1)^n < 0$.

Now we can apply the Intermediate Value Theorem (4.7) to the continuous function $p(x)$ on the interval $[-(R+1), R+1]$ to conclude that there is some $x \in (-(R+1), R+1)$ with $p(x) = 0$. $\square$

Note: A lot of mathematics has arisen from the desire to understand how to solve equations. Here we have quite a general results about equations. It implies that every odd degree polynomial $p(x)$ with real numbers as coefficients can be factored, if you use the remainder theorem. The remainder theorem tells you that if a polynomial $p(x)$ has $p(c) = 0$ the the polynomial $x - c$ of degree 1 divides $p(x)$.

Corollary 4.17. *If $p(x)$ is an odd degree polynomial, then the function $p\colon \mathbb{R} \to \mathbb{R}$ given by the polynomial is surjective.*

Proof. If $y_0 \in \mathbb{R}$, then the equation $p(x) = y_0$ can be rewritten $p(x) - y_0 = 0$ and so it is an odd degree polynomial to which the previous result applies. So there is $x \in \mathbb{R}$ with $p(x) - y_0 = 0$ or $p(x) = y_0$. This shows that the function is surjective. $\square$

Proposition 4.18. *For any even degree polynomial $p(x) = \sum_{j=0}^n a_j x^j = 0$ with leading coefficient $a_n > 0$ (n even), there is $x_L \in \mathbb{R}$ so that*

$$p(x_L) = \sum_{j=0}^n a_j x_L^j \leq \sum_{j=0}^n a_j x^j = p(x)$$

holds for all $x \in \mathbb{R}$.

(In other words even degree polynomials with positive leading term have a smallest value. It follows that those with even degree polynomials with a negative leading term have a largest value.)

Proof. By Proposition 4.15 (and using $a_n > 0$), we can find R so that

$$|x| > R \Rightarrow |p(x) - a_n x^n| \leq \frac{1}{2} a_n |x|^n.$$

As n is even, we can simplify this a little because $|x|^n = x^n$.

It implies that, for $|x| = R + 1$ we have

$$\begin{aligned} |p(R+1) - a_n(R+1)^n| &\leq \frac{1}{2} a_n (R+1)^n \\ \Rightarrow \quad & \frac{1}{2} a_n (R+1)^n < p(R+1) < \frac{3}{2} a_n (R+1)^n. \end{aligned}$$

Also for $|x| > 2(R+1)$

$$\begin{aligned} |p(x) - x^n| &< \frac{1}{2} x^n \\ \Rightarrow \quad & \frac{1}{2} x^n < p(x) \\ \Rightarrow \quad & \frac{1}{2}(2(R+1))^n = 2(R+1)^n < p(x) \\ \Rightarrow \quad & p(R+1) < p(x). \end{aligned}$$

It follows that if we are looking for a smallest value, then no x with $|x| > 2(R+1)$ could provide it. We should concentrate on looking then for $x \in [-2(R+1), 2(R+1)]$. But now we have a continuous function $p(x)$ on a finite closed interval $[a, b] = [-2(R+1), 2(R+1)]$ and Theorem 4.6 assures us that there is $x_L \in [-2(R+1), 2(R+1)]$ with

$$p(x_L) \leq p(x) \text{ for all } x \in [-2(R+1), 2(R+1)].$$

In particular $p(x_L) \leq p(R+1)$ and so $p(x_L) \leq p(x)$ holds for all x with $|x| > 2(R+1)$ (recall $p(R+1) < p(x)$ for those x).

So now we have
$$p(x_L) \leq p(x) \text{ for all } x \in \mathbb{R}.$$

□

An example of the Proposition would be $p(x) = 2(x-1)^2 - 5 = 2x^2 - 4x - 3$ which has $p(1) = -5 \leq 2(x-1)^2 - 5 = p(x)$ for all x. In terms of equations, it means we cannot solve $2x^2 - 4x - 3 = y_0$ if $y_0 < -5$.

The Proposition does not imply it, but we can solve $2x^2 - 4x - 3 = y_0$ for all $y_0 \geq -5$. To prove that we could use that the largest term ($2x^2$ in this case) dominates for $|x|$ large, and so $2x^2 - 4x - 3 > \frac{1}{2}(2x^2) = x^2$ for $|x|$ large enough. This means that there is some x with $2x^2 - 4x - 3 > y_0$ (any fixed y_0 we choose). From this we can establish that the interval $p(\mathbb{R})$ which is the range of $p(x) = 2x^2 - 4x - 3$ must be $[-5, \infty)$ in this case.

In fact for a general even degree polynomial $p(x)$ with positive leading term the range $p(\mathbb{R}) = [p(x_L), \infty)$ as long as the degree is not 0 (constant polynomial).

A Appendix — April

Theorem A.1. *If $I \subset \mathbb{R}$ is an interval and $f\colon I \to \mathbb{R}$ is a continuous injective function, then f must be strictly monotone.*

The proof of this result is more technical than difficult, as the main point is that the intermediate value theorem holds.

As an example to indicate the importance of the domain being an interval, consider the function $f\colon S \to \mathbb{R}$ where $S = [0,1] \cup [2,3]$ and f is given by the rule

$$f(x) = \begin{cases} x & \text{if } x \in [0,1] \\ 5-x & \text{if } x \in [2,3] \end{cases}$$

It is fairly easy to see that f is injective with $f(S) = [0,1] \cup [2,3]$ but f is monotone increasing on $[0,1]$, monotone decreasing on $[2,3]$ and not monotone on S.

Our proof of the theorem will be broken down into small steps, via the following lemmas.

Lemma A.2. *Suppose $I \subset \mathbb{R}$ is an interval, $f\colon I \to \mathbb{R}$ is a continuous injective function, $a, b \in I$, $a < b$ and $f(a) < f(b)$.*

Then $f([a,b]) = [f(a), f(b)]$.

Proof. We know that $f([a,b])$ is an interval and so $f([a,b]) \supseteq [f(a), f(b)]$.

To show it is equal, suppose there is $y_0 \in f([a,b]) \setminus [f(a), f(b)]$. Then $y_0 = f(x_0)$ for some $x_0 \in (a,b)$ and either $y_0 > f(b)$ or $y_0 < f(a)$.

If $y_0 > f(b)$, then $f([a,x_0]) \subset [f(a), f(x_0)]$ and contains $f(b)$. So there is $x_1 \in (a, x_0)$ with $f(x_1) = f(b)$ — contradicting injectivity of f (as $x_1 \neq b$, $x_1, b \in I$).

If $y_0 < f(a)$, then $f([x_0,b]) \subset [f(x_0), f(b)]$ and contains $f(a)$. So there is $x_1 \in (x_0, b)$ with $f(x_1) = f(a)$ — contradicting injectivity of f (as $x_1 \neq a$, $x_1, a \in I$).

As both are ruled out, we must have $f([a,b]) = [f(a), f(b)]$. □

Lemma A.3. *Suppose $I \subset \mathbb{R}$ is an interval, $f\colon I \to \mathbb{R}$ is a continuous injective function, $a, b \in I$, $a < b$ and $f(a) < f(b)$.*

Then

(i) $\alpha \in I, \alpha < a \Rightarrow f(\alpha) < f(a)$, *and*

(ii) $\beta \in I, \beta > b \Rightarrow f(\beta) > f(b)$.

Proof. To prove (i), note first from Lemma A.2 that $f([a, b]) = [f(a), f(b)]$ and so, since f is injective we have $f(\alpha) \notin [f(a), f(b)]$. Thus either $f(\alpha) < f(a)$ as desired or $f(\alpha) > f(b)$. But if $f(\alpha) > f(b)$, then $f(\alpha) > f(b) > f(a)$ and so, by the Intermediate Value Theorem, there is $x_0 \in (\alpha, a)$ with $f(x_0) = f(b)$ — contradicting injectivity of f (as $x_0 \neq b$). So we must have $f(\alpha) < f(a)$.

To prove (ii), we similarly deduce from Lemma A.2 that $f(\beta) \notin [f(a), f(b)]$. To rule out $f(\beta) < f(a)$ we show that this situation would result in $f(a) = f(x)$ for some $x \in (b, \beta)$. □

Lemma A.4. *Suppose $I \subset \mathbb{R}$ is an interval, $f: I \to \mathbb{R}$ is a continuous injective function, $a, b \in I$, $a < b$ and $f(a) > f(b)$.*

Then

(i) $\alpha \in I, \alpha < a \Rightarrow f(\alpha) > f(a)$*, and*

(ii) $\beta \in I, \beta > b \Rightarrow f(\beta) < f(b)$*.*

Proof. Apply Lemma A.3 to $g: I \to \mathbb{R}$ given by $g(x) = -f(x)$. Then g is injective and $g(a) < g(b)$. □

Lemma A.5. *Suppose $I \subset \mathbb{R}$ is an interval, $f: I \to \mathbb{R}$ is a continuous injective function, $a, b \in I$, $a < b$ and $f(a) < f(b)$.*

Then f is strictly monotone increasing (on I).

Proof. The complete proof is rather detailed, but more tedious than difficult. The hardest parts are done already.

We consider $\alpha, \beta \in I$ with $\alpha < \beta$ and we claim that $f(\alpha) < f(\beta)$ must hold. (When we establish the claim we will have shown that f is strictly monotone increasing, since $\alpha, \beta \in I$ are quite arbitrary with $\alpha < \beta$.)

The proof is divided into several cases, depending on where α, β lie with respect to a, b.

Case 1: $\alpha < a$.

subcase 1A: $\alpha < \beta \leq a$.

In this situation, since $f(\alpha) \neq f(\beta)$ (by injectivity) we either have $f(\alpha) < f(\beta)$ (as we claim) or $f(\alpha) > f(\beta)$.

If $f(\alpha) > f(\beta)$ then Lemma A.4 (ii) (using α and β where 'a' and 'b' occur in the Lemma), implies $f(\beta) > f(b)$ and so $f(\beta) > f(b) > f(a)$. Lemma A.3 (i) tells us $f(\beta) < f(a)$, a contradiction.

So $f(\alpha) < f(\beta)$ holds in this situation.

subcase 1B: $\alpha < a < \beta < b$.

By Lemma A.3 (i), in this case $f(\alpha) < f(a)$ while by Lemma A.2, $f(\beta) \in f([a, b]) = [f(a), f(b)]$ and so $f(\alpha) < f(a) \le f(\beta) \Rightarrow f(\alpha) < f(\beta)$.

subcase 1C: $\alpha < a < b \le \beta$.

By Lemma A.3 (i) and (ii), in this case $f(\alpha) < f(a)$ and $f(b) \le f(\beta)$. Since $f(a) < f(b)$ we deduce $f(\alpha) < f(\beta)$.

Case 2: $a \le \alpha \le b$.

subcase 2A: $a \le \alpha < \beta \le b$.

As in case 1A, what we have to exclude is $f(\alpha) > f(\beta)$ (since $f(\alpha) \neq f(\beta)$ by injectivity of f). But if $f(\alpha) > f(\beta)$ then by Lemma A.4 (i) and (ii) (using α and β where 'a' and 'b' occur in the Lemma) , $f(a) \ge f(\alpha)$ and $f(\beta) \ge f(b)$. So

$$f(a) \ge f(\alpha) > f(\beta) \ge f(b) \Rightarrow f(a) > f(b),$$

a contradiction. We are forced therefore to have $f(\alpha) < f(\beta)$, as claimed.

subcase 2B: $a \le \alpha < b \le \beta$.

By Lemma A.2 and Lemma A.3 (ii), in this situation we have $f(\alpha) \in f([a, b]) = [f(a), f(b)] \Rightarrow f(\alpha) \le f(b)$ and $f(b) \le f(\beta)$. Since $f(\alpha) \neq f(b)$ by injectivity, $f(\alpha) < f(b) \le f(\beta) \Rightarrow f(\alpha) < f(\beta)$, as claimed.

Case 3: $a < b < \alpha < \beta$.

Again we must rule out $f(\alpha) > f(\beta)$, but if that happens Lemma A.4 (i) (using α and β where 'a' and 'b' occur in the Lemma) implies that $f(a) > f(\alpha)$ while Lemma A.3 (ii) implies that $f(b) < f(\alpha)$. So $f(\alpha) < f(a) < f(b) < f(\alpha)$, a contradiction.

□

Proof. (of Theorem A.1)

We first dispose of the case where I might be empty or a singleton. In these cases f is vacuously strictly monotone increasing (and decreasing also) because there are not two values of f to be compared.

Now, if I has at last two points, pick $a, b \in I$ with $a \neq b$ and assume we have labeled them so that $a < b$. Since f is injective, we know $f(a) \neq f(b)$ and so either $f(a) < f(b)$ or else $f(a) > f(b)$.

If $f(a) < f(b)$, then Lemma A.5 tells us that f is strictly monotone increasing.

On the other hand if $f(a) > f(b)$ we can consider $g: I \to \mathbb{R}$ with $g(x) = -f(x)$. This g will also be injective and continuous on I and $g(a) < g(b)$. By the Lemma A.5 g must be strictly monotone increasing, and this implies that f is strictly monotone decreasing ($\alpha, \beta \in I, \alpha < \beta \Rightarrow g(\alpha) < g(\beta) \Rightarrow -f(\alpha) < -f(\beta) \Rightarrow f(\alpha) > f(\beta)$). □

Theorem A.6. *Let $I \subset \mathbb{R}$ be an interval and $f: I \to \mathbb{R}$ an injective continuous function. For $J = f(I) =$ the range of f, we take f^{-1} to mean the inverse of the bijection $f: I \to J$ (same values as $f: I \to \mathbb{R}$ but a different co-domain, so technically a different function). Then $f^{-1}: J \to I$ is continuous.*

Proof. We know by Theorem A.1 that f is strictly monotone and we give the proof in the case where f is strictly monotone increasing. The other case (decreasing) is similar or can be deduced by considering $g(x) = -f(x)$ (and $g^{-1}(y) = f^{-1}(-y)$).

We know that $J = f(I)$ is an interval (corollary to the Intermediate Value Theorem).

Fix $y_0 \in J$ and $\varepsilon > 0$. We aim to verify that there is $\delta > 0$ satisfying the implication

$$y \in J, |y - y_0| < \delta \Rightarrow |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon.$$

This will establish continuity of f^{-1} at y_0 (and since $y_0 \in J$ is arbitrary, continuity of f^{-1} on J).

Let $x_0 = f^{-1}(y_0)$. There are a few different cases to consider, depending on whether or not $x_0 + \varepsilon \in I$ and $x_0 - \varepsilon \in I$.

If both $x_0 + \varepsilon \in I$ and $x_0 - \varepsilon \in I$, put $y_L = f(x_0 - \varepsilon)$ and $y_R = f(x_0 + \varepsilon)$. Take $\delta = \min(y_R - y_0, y_0 - y_L)$. Since f is strictly monotone increasing $y_0 = f(x_0) < f(x_0 + \varepsilon) = y_R$ and $y_L = f(x_0 - \varepsilon) < f(x_0) = y_0$ and so $\delta > 0$. Also

$$\begin{aligned} |y - y_0| < \delta &\Rightarrow y_0 - \delta < y < y_0 + \delta \\ &\Rightarrow y_L < y < y_R \\ &\Rightarrow y \in J \text{ and } f^{-1}(y_L) < f^{-1}(y) < f^{-1}(y_R) \end{aligned}$$

(The last step is justified because f^{-1} must actually be strictly monotone increasing as well as f. For example if $f^{-1}(y_L) > f^{-1}(y)$ then $f(f^{-1}(y_L)) >$

$f(f^{-1}(y)) \Rightarrow y_L > y$, which is false.) Now we have

$$\begin{aligned} |y - y_0| < \delta &\Rightarrow x_0 - \varepsilon = f^{-1}(y_L) < f^{-1}(y) < f^{-1}(y_R) = x_0 + \varepsilon \\ &\Rightarrow |f^{-1}(y) - x_0| < \varepsilon \end{aligned}$$

and this is what we require because $x_0 = f^{-1}(y_0)$.

Another extreme is the case when both $x_0 + \varepsilon \notin I$ and $x_0 - \varepsilon \notin I$. Then $I \subset (x_0 - \varepsilon, x_0 + \varepsilon)$ and every $y \in J$ satisfies $|f^{-1}(y) - x_0| = |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon$. Thus for any $\delta > 0$ positive (say $\delta = 1$ for the sake of being specific), we have

$$y \in J, |y - y_0| < \delta \Rightarrow |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon.$$

There are two other cases. If $x_0 + \varepsilon \notin I$ but $x_0 - \varepsilon \in I$, then $I \subset (-\infty, x_0 + \varepsilon)$. Take $y_L = f(x_0 - \varepsilon)$ and $\delta = y_0 - y_L$. Then

$$\begin{aligned} y \in J, |y - y_0| < \delta &\Rightarrow y \in J, y > y_L \\ &\Rightarrow f^{-1}(y) > f^{-1}(y_L) = x_0 - \varepsilon \\ &\Rightarrow x_0 - \varepsilon < f^{-1}(y) < x_0 + \varepsilon \\ &\qquad \text{(since } f^{-1}(y) \in I\text{)} \\ &\Rightarrow |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon. \end{aligned}$$

In the final case, $x_0 + \varepsilon \in I$ but $x_0 - \varepsilon \notin I$. In this case $I \subset (x_0 - \varepsilon, \infty)$. Take $y_R = f(x_0 + \varepsilon)$ and $\delta = y_R - y_0$. Then

$$\begin{aligned} y \in J, |y - y_0| < \delta &\Rightarrow y \in J, y < y_R \\ &\Rightarrow f^{-1}(y) < f^{-1}(y_R) = x_0 + \varepsilon \\ &\Rightarrow x_0 - \varepsilon < f^{-1}(y) < x_0 + \varepsilon \\ &\qquad \text{(since } f^{-1}(y) \in I\text{)} \\ &\Rightarrow |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon. \end{aligned}$$

□

Example A.7. If $f\colon \mathbb{R} \to \mathbb{R}$ is $f(x) = x^3$, then f is a bijection and $f^{-1}\colon \mathbb{R} \to \mathbb{R}$ is given by $f(x) = x^{1/3}$ (cube root of x). By the theorem f^{-1} is continuous.

Notice that since $f(\mathbb{R})$ is an interval and contains $n^3 > n$ and $-n^3 < -n$ for each $n \in \mathbb{N}$ we can see that $f(\mathbb{R}) = \mathbb{R}$.

Theorem A.8. *Let $I \subset \mathbb{R}$ be an open interval (finite or infinite) and let $f\colon I \to \mathbb{R}$ be a differentiable injective function. Take $J = f(I)$. Then J is also an open interval and the inverse function $f^{-1}\colon J \to I$ is differentiable at a point $y_0 \in J$ if and only if $f'(f^{-1}(y_0)) \neq 0$.*

Moreover

$$f'(y_0) = \frac{1}{f'(f^{-1}(y_0))}$$

when $f'(f^{-1}(y_0)) \neq 0$.

Notice that a way to remember the result is to use Leibniz notation, $y = f(x)$ for $x \in I, y \in J$, $x = f^{-1}(y)$ and $\frac{dy}{dx} = f'(x)$. Then the theorem says

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$$

with the added explanation that dx/dy is normally evaluated at a point y while the dy/dx on the other side is evaluated a point x related to y by $y = f(x)$ or $x = f^{-1}(y)$.

Proof. As differentiable functions are continuous, f most be strictly monotone by Theorem A.1. We know that $J = f(I)$ is an interval and J has to be an open interval since I is. The reason is that if J contained any endpoint y_0 of itself, then $x_0 = f^{-1}(y_0) \in I$ and cannot be an endpoint of I because I is open. So we can find $x_1, x_2 \in I$ with $x_1 < x_0 < x_2$. If f is monotone increasing then $f(x_1) < f(x_0) = y_0 < f(x_2)$ while if f is monotone decreasing $f(x_1) > f(x_0) = y_0 > f(x_2)$. Either way we have points $f(x_1), f(x_2) \in J = f(I)$ on either side of y_0 and y_0 can't be an endpoint of J.

To prove the result about $(f^{-1})'(y_0)$ for $y_0 \in J$ fixed, consider the definition of the derivative in the form

$$(f^{-1})'(y_0) = \lim_{k \to 0} \frac{f^{-1}(y_0 + k) - f^{-1}(y_0)}{k}$$

(where we use k instead of the usual h because we want to keep h for later). To prove that this limit exists and is what it is claimed to be when $f'(f^{-1}(y_0)) \neq 0$, we use sequences. Consider a sequence (k_n) with $k_n \neq 0$ (for all n), all k_n small enough that $y_0 + k_n \in J$ and so that $\lim_{n\to\infty} k_n = 0$. Write

$$h_n = f^{-1}(y_0 + k_n) - f^{-1}(y_0).$$

Taking $x_0 = f^{-1}(y_0)$, we have

$$x_0 + h_n = f^{-1}(y_0 + k_n) \Rightarrow f(x_0 + h_n) = y_0 + k_n = f(x_0) + k_n$$

so that

$$f(x_0 + h_n) - f(x_0) = k_n.$$

By continuity of f^{-1} we can say that

$$\lim_{n\to\infty} h_n = \lim_{n\to\infty} f^{-1}(y_0 + k_n) - f^{-1}(y_0) = 0.$$

Because $f(x_0 + h_n) - f(x_0) = k_n \neq 0$ we must also have $h_n \neq 0$ (all n).

So

$$f'(x_0) = \lim_{n\to\infty} \frac{f(x_0 + h_n) - f(x_0)}{h_n} = \lim_{n\to\infty} \frac{k_n}{f^{-1}(y_0 + k_n) - f^{-1}(y_0)}.$$

If $f'(x_0) \neq 0$ we are able to take reciprocals and deduce

$$\lim_{n\to\infty} \frac{f^{-1}(y_0 + k_n) - f^{-1}(y_0)}{k_n} = \frac{1}{f'(x_0)}.$$

As this is true for all sequences $(k_n)_{n=1}^\infty$ with $k_n \neq 0$ and $\lim_{n\to\infty} k_n = 0$, we have

$$\lim_{k\to 0} \frac{f^{-1}(y_0 + k) - f^{-1}(y_0)}{k} = \frac{1}{f'(x_0)}.$$

This proves the theorem in the case $f'(x_0) = f'(f^{-1}(y_0)) \neq 0$.

To show finally that the derivative $(f^{-1})'(y_0)$ does not exist if $f'(f^{-1}(y_0)) = 0$ consider the chain rule applied to

$$f(f^{-1}(y)) = y$$

If $(f^{-1})'(y_0)$ existed, since we know f is differentiable, the chain rule would say

$$f'(f^{-1}(y_0))(f^{-1})'(y_0) = 1$$

and that leads to the contradiction $0 = 1$. □

Example A.9. If we return to the previous example $f(x) = x^3$, $f^{-1}(y) = y^{1/3}$ then we have

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} = \frac{1}{3(f^{-1}(y))^2} = \frac{1}{3y^{2/3}}$$

as long as $y \neq 0$. In other words

$$\frac{d}{dy} y^{1/3} = \frac{1}{3y^{2/3}} \qquad (y \neq 0)$$

but the derivative does not exist at $y = 0$.

Considering the cube root function in its own right (rather than as the inverse of the cube function) we would normally use x for the independent variable and write say $g(x) = x^{1/3}$. Our formula is

$$g'(x) = \frac{d}{dx} x^{1/3} = \frac{1}{3x^{2/3}} \qquad (x \neq 0)$$

(and g is not differentiable at 0).

We don't have to stop with cube roots. The same argument (more or less) applies to any odd root (inverse of $x \mapsto x^5$ is $x \mapsto x^{1/5}$, for example) and we get

$$\frac{d}{dx} x^{1/n} = \frac{1}{nx^{(n-1)/n}} \qquad (x \neq 0)$$

for n odd.

For the case of n even we have $(-x)^n = x^n$ and so $x \mapsto x^n$ is not injective on $\mathbb{R}$. We can however restrict to $x > 0$ and get a bijection $f \colon (0, \infty) \to (0, \infty)$ with $f(x) = x^n$ and inverse $f^{-1}(x) = x^{1/n}$. Since $f'(x) = nx^{n-1} \neq 0$ for $x > 0$ we can see from the theorem that

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} = \frac{1}{n(f^{-1}(y))^{n-1}} = \frac{1}{ny^{(n-1)/n}} \qquad (y > 0)$$

We have (without being as careful as usual) made a definition in the middle of all this. We have defined $x^{1/n}$ for all $x \in \mathbb{R}$ if $n \in \mathbb{N}$ is odd, and for $x > 0$ if $n \in \mathbb{N}$ is odd. Here is how we defined them. And then we can go on to define any rational power (at least of positive numbers).

Definition A.10. (i) *For $n \in \mathbb{N}$ odd, we define the function $x \mapsto x^{1/n} \colon \mathbb{R} \to \mathbb{R}$ (called the nth root function) to be the inverse of the function $x \mapsto x^n \colon \mathbb{R} \to \mathbb{R}$.*

(ii) *For $n \in \mathbb{N}$ even, we define the function $x \mapsto x^{1/n} \colon [0, \infty) \to [0, \infty)$ to be the inverse of the function $x \mapsto x^n \colon [0, \infty) \to [0, \infty)$.*

(iii) For $m \in \mathbb{Z} \setminus \{0\}$ *relatively prime to* $n \in \mathbb{N}$ *we define*

$$x^{m/n} = (x^{1/n})^m$$

(for all $x \in \mathbb{R}$ *if* n *odd and* $m > 0$*, and for* $x > 0$ *otherwise.)*

In case it was not clear $x^{-m} = 1/x^m$ and $x = x^1$, $x^2 = xx$ are defined by induction ($x^{m+1} = xx^m$). We can define $x^0 = 1$ as long as $x \neq 0$, but 0^0 is best left undefined.

One can check that with this definition the law of exponents work. So if $p, q \in \mathbb{Q}$ we can say $x^{p+q} = x^p x^q$ and $(x^p)^q = x^{pq}$ as long as we avoid dividing by 0 and even roots of negative numbers.

Proposition A.11. *For* $x > 0$ *and* $p \in \mathbb{Q} \setminus \{0\}$*,*

$$\frac{d}{dx} x^p = p x^{p-1}.$$

The same is true for $x < 0$ *if* $p = m/n$ *with* n *odd.*

Proof. We can use the chain rule on $h\colon (0, \infty) \to (0, \infty)$ given by $h(x) = x^{m/n} = (x^{1/n})^m$) to get

$$h'(x) = m(x^{1/n})^{m-1} \frac{d}{dx}(x^{1/n}) = n x^{(m-1)/n} \frac{1}{n x^{(n-1)/n}} = \frac{n}{m} x^{n/m-1}$$

$\square$