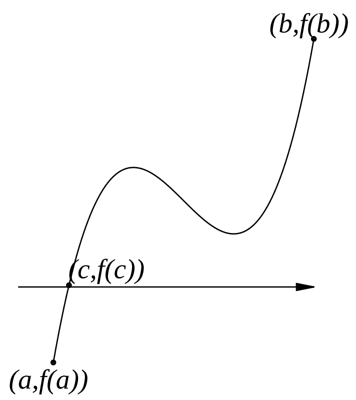
Chapter 14

The Inverse Function Theorem

14.1 The Intermediate Value Property

14.1 Assumption (Intermediate value property 1.) *Let a, b be real numbers with $a < b$, and let f be a continuous function from $[a, b]$ to $\mathbf{R}$ such that $f(a) < 0$ and $f(b) > 0$. Then there is some number $c \in (a, b)$ such that $f(c) = 0$.*



The intermediate value *theorem* was first proved in 1817 by Bernard Bolzano (1781–1848). However Bolzano published his proof in a rather obscure Bohemian journal, and his work did not become well known until much later. Before the nineteenth century the theorem was often assumed implicitly, i.e. it was used without stating that it was an assumption.

14.2 Definition (c is between a and b.) Let a, b and c be real numbers with $a \neq b$. We say that c is *between* a and b if either $a < c < b$ or $b < c < a$.

14.3 Corollary (Intermediate value property 2.) *Let f be a continuous function from some interval $[a, b]$ to $\mathbf{R}$, such that $f(a)$ and $f(b)$ have opposite signs. Then there is some number c between a and b such that $f(c) = 0$.*

Proof: If $f(a) < 0 < f(b)$ the result follows from assumption 14.1. Suppose that $f(b) < 0 < f(a)$. Let $g(x) = -f(x)$ for all $x \in [a, b]$. then g is a continuous function on $[a, b]$ and $g(a) < 0 < g(b)$. It follows that there is a number $c \in (a, b)$ such that $g(c) = 0$, and then $f(c) = -g(c) = 0$. ||

14.4 Corollary (Intermediate value property 3.) *Let a, b be real numbers with $a < b$, and let $f : [a, b] \to \mathbf{R}$ be a continuous function such that $f(a) \neq f(b)$. Let p be any number between $f(a)$ and $f(b)$. Then there is a number $c \in (a, b)$ such that $f(c) = p$.*

14.5 Exercise. Prove Corollary 14.4. You may assume that $f(a) < f(b)$.

14.2 Applications

14.6 Example. We know that ln is continuous on $\mathbf{R}^+$, and that $\ln(2) \leq 1 \leq \ln(4)$.(Cf equation (5.78).) It follows that there is a number e in $[2, 4]$ such that $\ln(e) = 1$.

14.7 Example. Two points P, Q on a sphere are called *antipodal points* if P and Q are opposite ends of the same diameter of the sphere. We will consider the surface of the earth to be a sphere of radius R. At any fixed time, let $T(p)$ denote the temperature of the earth at the point p on the surface of the earth. (More precisely, let $T(p)$ be the number such that the temperature at p is $T(p)^\circ C$). We will show that there are two antipodal points P, Q on the surface of the earth such that $T(P) = T(Q)$. In fact, we will show that there are two antipodal points on the equator with the same temperature. We first introduce a coordinate system so that the center of the earth is at the origin, and the plane of the equator is the x-y plane, and the point on the equator passing through the Greenwich meridian is the point $(R, 0)$. Then the points on the equator are the points

$$(R\cos(\theta), R\sin(\theta)) \text{ where } \theta \in \mathbf{R}.$$

Define a function $f : [0, \pi] \to \mathbf{R}$ by

$$f(\theta) = T(R\cos(\theta), R\sin(\theta)) - T(-R\cos(\theta), -R\sin(\theta)).$$

Thus

$$f(0) = T(R, 0) - T(-R, 0).$$

We suppose that f is a continuous function on $[0, \pi]$. If $f(0) = 0$ then $T(R, 0) = T(-R, 0)$, so $(R, 0)$ and $(-R, 0)$ are a pair of antipodal points with the same temperature. Now

$$f(\pi) = T(-R, 0) - T(R, 0) = -f(0),$$

so if $f(0) \neq 0$ then $f(0)$ and $f(\pi)$ have opposite signs. Hence by the intermediate value property, there is a number $c \in (0, \pi)$ such that $f(c) = 0$, i.e.

$$T(R\cos(c), R\sin(c)) = T(-R\cos(c), -R\sin(c)).$$

Then $(R\cos(c), R(\sin(c))$ and $(-R\cos(c), -R(\sin(c))$ are a pair of antipodal points with the same temperature. $\|$

14.8 Example. Let

$$P = a_0 + a_1X + a_2X^2 + a_3X^3$$

where a_0, a_1, a_2, and a_3 are real numbers, and $a_3 \neq 0$. Then there exists some number $r \in \mathbf{R}$ such that $P(r) = 0$.

Proof: I will suppose that $P(t) \neq 0$ for all $t \in \mathbf{R}$ and derive a contradiction. Let

$$Q(x) = \frac{P(x)}{P(-x)} \text{ for all } x \in \mathbf{R}.$$

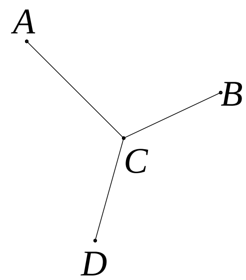
Since $P(x) \neq 0$ for all $x \in \mathbf{R}$, Q is continuous on $\mathbf{R}$. We know that

$$\begin{aligned} \lim\{Q(n)\} &= \lim\{\frac{a_0 + a_1n + a_2n^2 + a_3n^3}{a_0 - a_1n + a_2n^2 - a_3n^3}\} \\ &= \lim\left\{\frac{\frac{a_0}{n^3} + \frac{a_1}{n^2} + \frac{a_2}{n} + a_3}{\frac{a_0}{n^3} - \frac{a_1}{n^2} + \frac{a_2}{n} - a_3}\right\} = -1. \end{aligned}$$

Hence $Q(N) < 0$ for some $N \in \mathbf{Z}^+$. Then $P(N)$ and $P(-N)$ have opposite signs, so by the intermediate value property there is a number $r \in [-N, N]$ such that $P(r) = 0$. This contradicts our assumption that $P(t) \neq 0$ for all $t \in \mathbf{R}$. $\|$

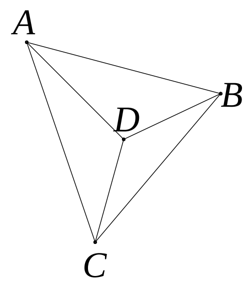
14.9 Exercise. Let $p(x) = x^3 - 3x + 1$. Show that there are at least three different numbers a, b, c such that $p(a) = p(b) = p(c) = 0$.

14.10 Exercise. Three wires AC, BC, DC are joined at a common point C.



Let S be the Y-shaped figure formed by the three wires. Prove that at any time there are two points in S with the same temperature.

14.11 Exercise. Six wires are joined to form the figure F shown in the diagram.



Show that at any time there are three points in F that have the same temperature. To simplify the problem, you may assume that the temperatures at A,B, C, and D are all distinct.

14.3 Inverse Functions

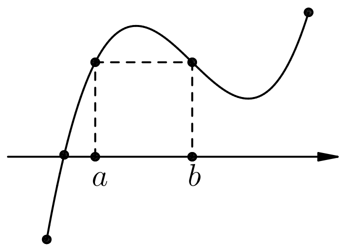
14.12 Definition (Injective.) Let A and B be sets. A function $f : A \to B$ is called *injective* or *one-to-one* if and only if for all points a, b in A

$$(a \neq b) \implies (f(a) \neq f(b)),$$

or equivalently if and only if

$$(f(a) = f(b)) \implies (a = b).$$

If f is a function whose domain and codomain are subsets of $\mathbf{R}$ then f is injective if and only if each horizontal line intersects the graph of f at most once.



$a \neq b$, $f(a) = f(b)$, $\qquad f$ is not injective

14.13 Examples. Let $f : [0, \infty) \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ be defined by

$$f(x) = x^2 \text{ for all } x \in [0, \infty)$$

$$g(x) = x^2 \text{ for all } x \in (-\infty, \infty).$$

Then f is injective, since for all $x, y \in [0, \infty)$ we have $x + y > 0$, and hence

$$\left(x^2 = y^2\right) \implies \left(x^2 - y^2 = 0\right) \implies \left((x-y)(x+y) = 0\right) \implies \left(x = y\right).$$

However g is not injective, since $g(-1) = g(1)$.

14.14 Remark (Strictly monotonic functions are injective.) If h is strictly increasing on an interval J, then h is injective on J, since for all $x, y \in J$

$$\begin{aligned} x \neq y &\implies ((x < y) \text{ or } (y < x)) \\ &\implies ((h(x) < h(y)) \text{ or } ((h(y) < h(x)) \\ &\implies h(x) \neq h(y). \end{aligned}$$

Similarly, any strictly decreasing function on J is injective.

14.15 Definition (Surjective.) Let A, B be sets and let $f : A \to B$. We say that f is *surjective* if and only if $B = \text{image}(f)$, i.e. if and only if for every $b \in B$ there is at least one element a of A such that $f(a) = b$.

14.16 Examples. Let $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to [0, \infty)$ be defined by

$$f(x) = x^2 \text{ for all } x \in (-\infty, \infty)$$

$$g(x) = x^2 \text{ for all } x \in [0, \infty).$$

Then g is surjective, since if $x \in [0, \infty)$, then $x = g(\sqrt{x})$, but f is not surjective, since -1 is not in the image of f.

14.17 Exercise. Give examples of functions with the following properties, or else show that no such functions exist.
$f : \mathbf{R} \to \mathbf{R}$, f is injective and surjective.
$g : \mathbf{R} \to \mathbf{R}$, g is injective but not surjective.
$h : \mathbf{R} \to \mathbf{R}$, h is surjective but not injective.
$k : \mathbf{R} \to \mathbf{R}$, k is neither injective nor surjective.

14.18 Definition (Bijective.) Let A, B be sets. A function $f : A \to B$ is called *bijective* if and only if f is both injective and surjective.

14.19 Examples. If $f : [0, \infty) \to [0, \infty)$ is defined by

$$f(x) = x^2 \text{ for all } x \in [0, \infty),$$

then f is bijective.

The function ln is a bijective function from $\mathbf{R}^+$ to $\mathbf{R}$. We know that ln is strictly increasing, and hence is injective. If y is any real number we know that ln takes on values greater than y, and values less that y, so by the intermediate value property (here we use the fact that ln is continuous) it also takes on the value y, i.e. ln is surjective.

14.20 Remark. Let A and B be sets, and let $f : A \to B$ be a bijective function. Let b be a generic element of B. Since f is surjective, there is an element a in A such that $f(a) = b$. Since f is injective this element a is unique, i.e. if a and c are elements of A then

$$\Big(f(a) = b \text{ and } f(c) = b\Big) \Longrightarrow \Big(f(a) = f(c)\Big) \Longrightarrow \Big(a = c\Big).$$

Hence we can define a function $g : B \to A$ by the rule

$$g(b) = \text{ the unique element } a \in A \text{ such that } f(a) = b.$$

Then by definition

$$f(g(b)) = b \text{ for all } b \in B.$$

Now let $a \in A$, so that $f(a) \in B$. It is clear that the unique element s in A such that $f(s) = f(a)$ is $s = a$, and hence

$$g(f(a)) = a \text{ for all } a \in A.$$

14.21 Definition (Inverse function.) Let A, B be sets, and let $f : A \to B$. An *inverse function* for f is a function $g : B \to A$ such that

$$\Big(f(g(b)) = b \text{ for all } b \in B\Big) \text{ and } \Big(g(f(a)) = a \text{ for all } a \in A\Big).$$

14.22 Remark (Bijective functions have inverses.) Notice that in the definition of inverse functions, both the domain and the codomain of f enter in a crucial way. It is clear that if g is an inverse function for f, then f is an inverse function for g. Remark 14.20 shows that every bijective function $f : A \to B$ has an inverse.

14.23 Example. Let $f : [0, \infty)$ be defined by

$$f(x) = x^2 \text{ for all } x \in [0, \infty).$$

We saw above that f is bijective, and hence has an inverse. If

$$g(x) = \sqrt{x} \text{ for all } x \in [0, \infty)$$

Then it is clear that g is an inverse function for f.

We also saw that $\ln : \mathbf{R}^+ \to \mathbf{R}$ is bijective, and so it has an inverse. This inverse is not expressible in terms of any functions we have discussed. We will give it a name.

14.24 Definition ($E(x)$.) Let E denote the inverse of the logarithm function. Thus E is a function from $\mathbf{R}$ to $\mathbf{R}^+$, and it satisfies the conditions

$$\ln(E(x)) = x \text{ for all } x \in \mathbf{R},$$

$$E(\ln(x)) = x \text{ for all } x \in \mathbf{R}^+.$$

We will investigate the properties of E after we have proved a few general properties of inverse functions.

In order to speak of *the inverse* of a function, as we did in the last definition, we should note that inverses are unique.

14.25 Theorem (Uniqueness of inverses.) *Let A, B be sets and let $f : A \to B$. If g and h are inverse functions for f, then $g = h$.*

Proof: If g and h are inverse functions for f then

$$\text{dom}(g) = \text{dom}(h) = \text{ codomain}(f) = B,$$

and

$$\text{codomain}(g) = \text{codomain}(h) = \text{dom}(f) = A.$$

Also for all $x \in B$

$$h(x) = g(f(h(x))) = g(x).$$

(I have used the facts that $y = g(f(y))$ for all $y \in A$, and $f(h(x)) = x$ for all $x \in B$).

14.26 Theorem (Reflection theorem.) *Let $f : A \to B$ be a function which has an inverse function $g : B \to A$. Then for all $(a, b) \in A \times B$*

$$(a, b) \in \text{graph}(f) \iff (b, a) \in \text{graph}(g).$$

Proof: Let $f : A \to B$ be a function that has an inverse function $g : B \to A$. Then for all $(a, b) \in A \times B$

$$\Big(b = f(a)\Big) \implies \Big(g(b) = g(f(a)) = a\Big) \implies \Big(g(b) = a\Big)$$

and

$$\Big(g(b) = a\Big) \implies \Big(b = (f(g(b)) = f(a)\Big) \implies \Big(b = f(a)\Big).$$

Thus

$$\Big(b = f(a)\Big) \iff \Big(a = g(b)\Big).$$

Now

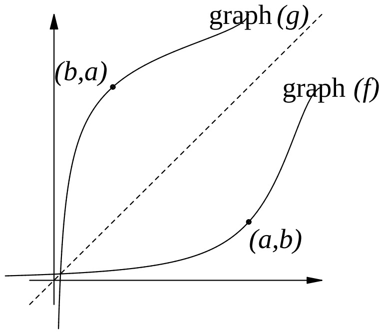
$$\Big(b = f(a)\Big) \iff \Big((a, b) \in \text{ graph}(f)\Big),$$

and

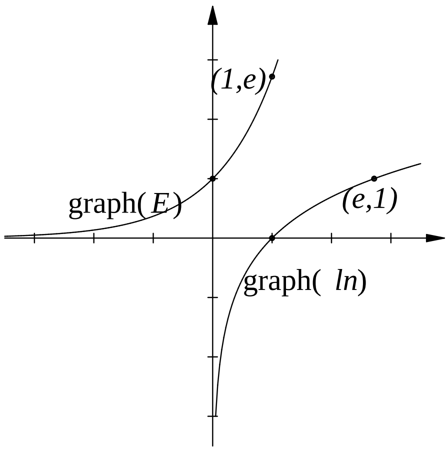
$$\Big(a = g(b)\Big) \iff \Big((b, a) \in \text{ graph}(g)\Big),$$

and the theorem now follows. |||

Remark: If f is a bijective function with $\text{dom}(f) \subset \mathbf{R}$ and $\text{codomain}(f) \subset \mathbf{R}$ Then the reflection theorem says that if g is the inverse function for f, then $\text{graph}(g) = D_+(\text{graph}(f))$ where D_+ is the reflection about the line $y = x$.



Since we know what the graph of ln looks like, we can make a reasonable sketch of graph(E).



It is a standard notation to denote the inverse of a function f by f^{-1}. However since this is also a standard notation for the function $\frac{1}{f}$ which is an entirely different object, I will not use this notation.

We have shown that if $f : A \to B$ is bijective, then f has an inverse function. The converse is also true.

14.27 Theorem. *Let A, B be sets and let $f : A \to B$. If f has an inverse function, then f is both injective and surjective.*

Proof: Suppose f has an inverse function $g : B \to A$. Then for all s, t in A we have

$$\Big(f(s) = f(t)\Big) \implies \Big(g(f(s)) = g(f(t))\Big) \implies \Big(s = t\Big) \tag{14.28}$$

and hence f is injective. Also, for each $b \in B$

$$b = f(g(b)),$$

so $b \in \text{image}(f)$, and f is surjective. ‖

14.4 The Exponential Function

14.29 Example. We will now derive some properties of the inverse function E of the logarithm.

We have

$$\begin{aligned} \ln(1) = 0 &\implies E(0) = 1, \\ \ln(e) = 1 &\implies E(1) = e. \end{aligned}$$

For all a and b in $\mathbf{R}$,

$$a + b = \ln(E(a)) + \ln(E(b)) = \ln(E(a)E(b)).$$

If we apply E to both sides of this equality we get

$$E(a + b) = E(a)E(b) \text{ for all } a, b \in \mathbf{R}.$$

For all $a \in \mathbf{R}$ we have

$$1 = E(0) = E(a + (-a)) = E(a)E(-a),$$

from which it follows that

$$E(-a) = (E(a))^{-1} \text{ for all } a \in \mathbf{R}.$$

If $a \in \mathbf{R}$ and $q \in \mathbf{Q}$ we have

$$\ln((E(a))^q) = q \ln(E(a)) = qa.$$

If we apply E to both sides of this identity we get

$$(E(a))^q = E(qa) \text{ for all } a \in \mathbf{R}^+, q \in \mathbf{Q}.$$

In particular,

$$e^q = (E(1))^q = E(q) \text{ for all } q \in \mathbf{Q}. \tag{14.30}$$

Now we have defined $E(x)$ for all $x \in \mathbf{R}$, but we have only defined x^q when $x \in \mathbf{R}^+$ and $q \in \mathbf{Q}$. (We know what $2^{\frac{1}{2}}$ is, but we have not defined $2^{\sqrt{2}}$.) Because of relation (14.30) we often write e^x in place of $E(x)$. E is called the *exponential function*, and is written

$$E(x) = e^x = \exp(x) \text{ for all } x \in \mathbf{R}.$$

We can summarize the results of this example in the following theorem:

14.31 Theorem (Properties of the exponential function.) *The exponential function is a function from* $\mathbf{R}$ *onto* $\mathbf{R}^+$. *We have*

$$\begin{aligned}
e^{a+b} &= e^a e^b \textit{ for all } a, b \in \mathbf{R}. \\
e^{a-b} &= \frac{e^a}{e^b} \textit{ for all } a, b \in \mathbf{R}. && (14.32) \\
(e^a)^q &= e^{aq} \textit{ for all } a \in \mathbf{R}, \textit{ and for all } q \in \mathbf{Q}. \\
(e^a)^{-1} &= e^{-a} \textit{ for all } a \in \mathbf{R}. \\
e^{\ln(x)} &= x \textit{ for all } x \in \mathbf{R}^+. \\
\ln(e^a) &= a \textit{ for all } a \in \mathbf{R}. \\
e^0 &= 1. \\
e^1 &= e. && (14.33)
\end{aligned}$$

Proof: We have proved all of these properties except for relation (14.32). The proof of (14.32) is the next exercise.

14.34 Exercise. Show that $e^{a-b} = \dfrac{e^a}{e^b}$ for all $a, b \in \mathbf{R}$.

14.35 Exercise. Show that if $a \in \mathbf{R}^+$ and $q \in \mathbf{Q}$, then

$$a^q = e^{q \ln(a)}.$$

14.36 Definition (a^x.) The result of the last exercise motivates us to make the definition

$$a^x = e^{x \ln(a)} \text{ for all } x \in \mathbf{R} \text{ and for all } a \in \mathbf{R}^+.$$

14.37 Exercise. Prove the following results:

$$\begin{aligned}
a^x a^y &= a^{x+y} \text{ for all } a \in \mathbf{R}^+ \text{ and for all } x, y \in \mathbf{R}. \\
(a^x)^y &= a^{xy} \text{ for all } a \in \mathbf{R}^+ \text{ and for all } x, y \in \mathbf{R}. \\
(ab)^x &= a^x b^x \text{ for all } a, b \in \mathbf{R}^+ \text{ and for all } x \in \mathbf{R}. \\
\ln(a^x) &= x \ln(a) \text{ for all } a \in \mathbf{R}^+ \text{ and for all } x \in \mathbf{R}.
\end{aligned}$$

14.5 Inverse Function Theorems

14.38 Lemma. *Let f be a strictly increasing continuous function whose domain is an interval $[a, b]$. Then the image of f is the interval $[f(a), f(b)]$, and the function $f : [a, b] \to [f(a), f(b)]$ has an inverse.*

Proof: It is clear that $f(a)$ and $f(b)$ are in $\text{image}(f)$. Since f is continuous we can apply the intermediate value property to conclude that for every number z between $f(a)$ and $f(b)$ there is a number $c \in [a, b]$ such that $z = f(c)$, i.e. $[f(a), f(b)] \subset \text{image}(f)$. Since f is increasing on $[a, b]$ we have $f(a) \leq f(t) \leq f(b)$ whenever $a \leq t \leq b$, and thus $\text{image}(f) \subset [f(a), f(b)]$. It follows that $f : [a, b] \to [f(a), f(b)]$ is surjective, and since strictly increasing functions are injective, f is bijective. By remark (14.22) f has an inverse.

14.39 Exercise. State and prove the analogue of lemma 14.38 for strictly decreasing functions.

14.40 Exercise. Let f be a function whose domain is an interval $[a, b]$, and whose image is an interval. Does it follow that f is continuous?

14.41 Exercise. Let f be a continuous function on a closed bounded interval $[a, b]$. Show that the image of f is a closed bounded interval $[A, B]$.

14.42 Exercise. Let J and I be non-empty intervals and let $f : J \to I$ be a continuous function such that $I = \text{image}(f)$.

a) Show that if f is strictly increasing, then the inverse function for f is also strictly increasing.

b) Show that if f is strictly decreasing, then the inverse function for f is also strictly decreasing.

14.43 Theorem (Inverse function theorem.) *Let f be a continuous strictly increasing function on an interval $J = [a, b]$ of positive length, such that $f'(x) > 0$ for all $x \in \text{interior}(J)$. Let I be the image of J and let*

$$g : I \to J$$

be the inverse function for f. Then g is differentiable on the interior of I and

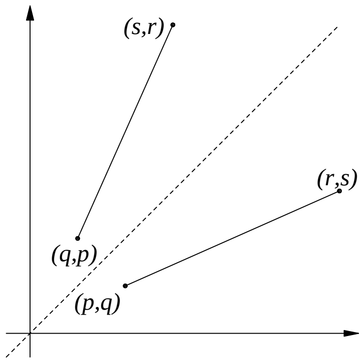
$$g'(s) = \frac{1}{f'(g(s))} \text{ for all } s \in \text{interior}(I) \tag{14.44}$$

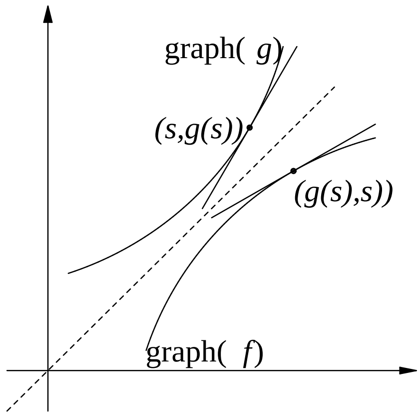
Remark: If l is a nonvertical line joining two points (p, q) and (r, s) then the slope of l is

$$m = \frac{s-q}{r-p}.$$

The reflection of l about the line whose equation is $y = x$ passes through the points (q, p) and (s, r), so the slope of the reflected line is

$$\frac{r-p}{s-q} = \frac{1}{m}.$$





$$g'(s) = \frac{1}{f'(g(s))}$$

Thus theorem 14.43 says that the tangent to graph(g) at the point $(s, g(s))$ is obtained by reflecting the tangent to graph(f) at $(g(s), s)$ about the line whose equation is $y = x$. This is what you should expect from the geometry of the situation.

Proof of theorem 14.43: The first thing that should be done, is to prove that g is continuous. I am going to omit that proof and just assume the continuity of g, and then show that g is differentiable, and that g' is given by formula (14.44).

Let s be a point in the interior of dom(g). then

$$\begin{aligned} \lim_{t\to s}\frac{g(t)-g(s)}{t-s} &= \lim_{t\to s}\frac{g(t)-g(s)}{f(g(t))-f(g(s))} \\ &= \lim_{t\to s}\frac{1}{\frac{f(g(t))-f(g(s))}{g(t)-g(s)}}. \end{aligned} \tag{14.45}$$

(Observe that we have not divided by zero). Let $\{t_n\}$ be a sequence in $\mathrm{dom}(g) \setminus \{s\}$ such that $\{t_n\} \to s$. Then $\{g(t_n)\} \to g(s)$ (since g is assumed to be continuous), and $g(t_n) \neq g(s)$ for all $n \in \mathbf{Z}^+$ (since g is injective). Since f is differentiable at $g(s)$, it follows that

$$\left\{\frac{f(g(t_n)) - f(g(s))}{g(t_n) - g(s)}\right\} \to f'(g(s)).$$

Since $f'(g(s)) \neq 0$ it follows that

$$\left\{\frac{1}{\frac{f(g(t_n))-f(g(s))}{g(t_n)-g(s)}}\right\} \to \frac{1}{f'(g(s)}.$$

It follows that

$$\lim_{t\to s} \frac{g(t) - g(s)}{t - s} = \frac{1}{f'(g(s))}$$

and the theorem is proved. $\|$

Remark: The inverse function theorem also applies to continuous functions f on J such that $f'(s) < 0$ for all $s \in$ interior (a, b). Formula (14.44) is valid in this case also.

Remark: Although we have stated the inverse function theorem for functions on intervals of the form $[a, b]$, it holds for functions defined on any interval. Let J be an interval, and let f be a continuous strictly increasing function from J to $\mathbf{R}$ such that $f'(x) > 0$ for all x in the interior of J. Let p be a point in the interior of image(J). Then we can find points r and s in image(J) such that $r < p < s$. Now f maps the interval $[g(r), g(s)]$ bijectively onto $[r, s]$, and since $p \in (r, s)$ we can apply the inverse function theorem on the interval $[g(r), g(s)]$ to conclude that $g'(p) = \dfrac{1}{f'(g(p))}$. It is not necessary to remember the formula for $g'(p)$. Once we know that g is differentiable, we can calculate g' by using the chain rule, as illustrated by the examples in the next section.

14.6 Some Derivative Calculations

14.46 Example (Derivative of exp.**)** We know that

$$\ln(E(t)) = t \text{ for all } t \in \mathbf{R}.$$

If we differentiate both sides of this equation, we get

$$\frac{1}{E(t)}E'(t) = 1,$$

i.e.

$$E'(t) = E(t) \text{ for all } t \in \mathbf{R}.$$

14.47 Example (Derivative of x^r.) Let r be any real number and let $f(x) = x^r$ for all $x \in \mathbf{R}^+$. Then

$$f(x) = x^r = E(r\ln(x)),$$

so by the chain rule

$$f'(x) = E'(r\ln(x)) \cdot \frac{r}{x} = E(r\ln(x)) \cdot rx^{-1} = x^r rx^{-1} = rx^{r-1}.$$

(Here I have used the result of exercise 14.37.) Thus the formula

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

which we have known for quite a while for rational exponents, is actually valid for all real exponents.

14.48 Exercise (Derivative of a^x.) Let $a \in \mathbf{R}^+$. Show that

$$\frac{d}{dx}(a^x) = a^x \ln(a)$$

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

14.49 Example (Derivative of x^x.)

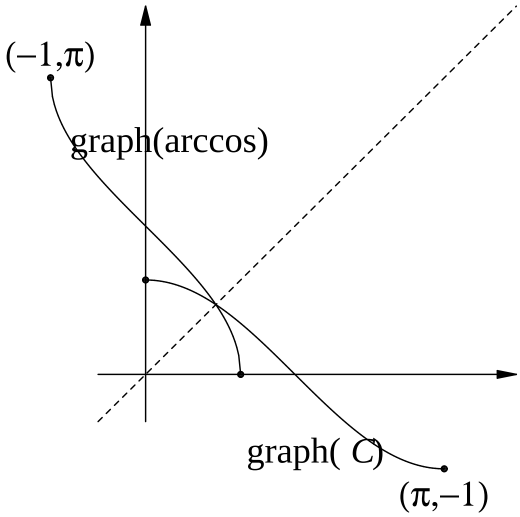
$$\begin{aligned}\frac{d}{dx}x^x &= \frac{d}{dx}e^{x\ln(x)} = e^{x\ln(x)}\frac{d}{dx}(x\ln(x)) \\ &= x^x(x \cdot \frac{1}{x} + \ln(x)) = x^x(1+\ln(x)).\end{aligned}$$

Hence

$$\frac{d}{dx}x^x = x^x(1+\ln(x)) \text{ for all } x \in \mathbf{R}^+.$$

14.50 Example (Derivative of arccos.**)** Let $C : [0, \pi] \to [-1, 1]$ be defined by

$$C(x) = \cos(x) \text{ for all } x \in [0, \pi].$$



We have

$$C'(x) = -\sin(x) < 0 \text{ for all } x \in (0, \pi),$$

so C has an inverse function which is denoted by arccos. By the inverse function theorem arccos is differentiable on $(-1, 1)$. and we have

$$\cos(\arccos(t)) = C(\arccos(t)) = t \text{ for all } t \in [-1, 1].$$

By the chain rule

$$-\sin(\arccos(t)) \arccos'(t) = 1 \text{ for all } t \in (-1, 1).$$

Now since the sine function is positive on $(0, \pi)$ we get

$$\sin(s) = \sqrt{1 - \cos^2(s)}$$

for all $s \in (0, \pi)$, so

$$\sin(\arccos(t)) = \sqrt{1 - (\cos(\arccos(t)))^2} = \sqrt{1 - t^2} \text{ for all } t \in (-1, 1).$$

Thus

$$\arccos'(t) = \frac{-1}{\sin(\arccos(t))} = \frac{-1}{\sqrt{1 - t^2}} \text{ for all } t \in (-1, 1).$$

14.51 Exercise (Derivative of arcsin.**)** Let

$$S(t) = \sin(t) \text{ for all } t \in [-\frac{\pi}{2}, \frac{\pi}{2}].$$

Show that S has an inverse function that is differentiable on the interior of its domain. This inverse functions is called arcsin. Describe the domain of arcsin, sketch the graphs of S and of arcsin, and show that

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

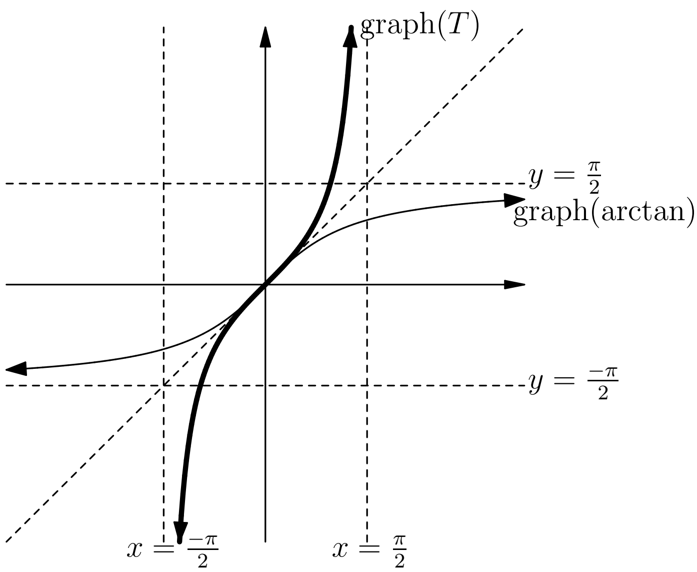
14.52 Example (Derivative of arctan.**)** Let

$$T(x) = \tan(x) \text{ for all } x \in (-\frac{\pi}{2}, \frac{\pi}{2}).$$

Then T is continuous, and the image of T is unbounded both above and below, so image$(T) = \mathbf{R}$. Also

$$T'(x) = \sec^2(x) > 0 \text{ for all } x \in (-\frac{\pi}{2}, \frac{\pi}{2})$$

so T has an inverse function, which we denote by arctan.



For all $x \in \mathbf{R}$
$$\tan(\arctan(x)) = T(\arctan(x)) = x,$$
so by the chain rule
$$\sec^2(\arctan(x))\arctan'(x) = 1 \text{ for all } x \in \mathbf{R}.$$
Now
$$\sec^2(t) = 1 + \tan^2(t) \text{ for all } t \in \text{dom}(\sec),$$
so
$$\sec^2(\arctan(x)) = 1 + \tan^2(\arctan(x)) = 1 + x^2 \text{ for all } x \in \mathbf{R}.$$
Thus
$$\arctan'(x) = \frac{1}{\sec^2(\arctan(x))} = \frac{1}{1+x^2} \text{ for all } x \in \mathbf{R}.$$

14.53 Exercise (Derivative of arccot.**)** Let
$$V(x) = \cot(x) \text{ for all } x \in (0, \pi).$$
Show that V has an inverse function arccot, and that
$$\frac{d}{dx}\text{arccot}(x) = -\frac{1}{1+x^2}.$$
What is dom(arccot)? Sketch the graphs of V and of arccot.

Remark The first person to give a name to the inverse trigonometric functions was Daniel Bernoulli (1700-1792) who used AS for arcsin in 1729. Other early notations included arc(cos. $= x$) and ang(cos. $= x$)[15, page 175]. Many calculators and some calculus books use $\cos^{-1}$ to denote arccos. (If you use your calculator to find inverse trigonometric functions, make sure that you set the degree-radian-grad mode to radians.)

14.54 Exercise. Calculate the derivatives of the following functions, and simplify your answers (Here a is a constant.)

a) $f(x) = x\sqrt{a^2 - x^2} + a^2 \arcsin(\frac{x}{a})$.

b) $g(x) = \arcsin(x) + \dfrac{x}{1-x^2}$.

c) $h(x) = x\arccos(ax) - \frac{1}{a}\sqrt{1-a^2x^2}$.

d) $k(x) = \arctan(e^x + e^{-x})$.

e) $m(x) = x\sqrt{1-x^2} + \arcsin(x)(2x^2 - 1)$.

f) $n(x) = e^{ax}(a\sin(bx) - b\cos(bx))$. Here a and b are constants.

g) $p(x) = e^{ax}(a^2x^2 - 2ax + 2)$. Here a is a constant..

14.55 Exercise. Let

$$l(x) = \arctan(\tan(x)).$$

Calculate the derivative of l. What is the domain of this function? Sketch the graph of l.

14.56 Exercise (Hyperbolic functions.) We define functions sinh and cosh on $\mathbf{R}$ by

$$\begin{aligned} \cosh(x) &= \frac{e^x + e^{-x}}{2} \text{ for all } x \in \mathbf{R}. \\ \sinh(x) &= \frac{e^x - e^{-x}}{2} \text{ for all } x \in \mathbf{R}. \end{aligned}$$

These functions are called the *hyperbolic sine* and the *hyperbolic cosine* respectively. Show that

$$\frac{d}{dx}\cosh(x) = \sinh(x),$$

and

$$\frac{d}{dx}\sinh(x) = \cosh(x).$$

Calculate

$$\frac{d}{dx}\Big(\cosh^2(x) - \sinh^2(x)\Big),$$

and simplify your answer as much as you can. What conclusion can you draw from your answer? Sketch the graphs of cosh and sinh on one set of coordinate axes.