Second Supplement to Chapter 12 of Sutherland,

Introduction to Metric and Topological Spaces (Second Edition)

At various points in the course we have stated that if f is a continuous and strictly increasing function from some interval $J \subset \mathbb{R}$ to $\mathbb{R}$, then f has a continuous and strictly increasing inverse function. The proof of this fact depends upon the concept of connectedness, and since the latter was developed in Chapter 12 we are now in a position to give a rigorous proof of this fact. Here is the formal statement of the result when J is an open interval.

THEOREM. *Let a and b be real numbers or $\pm\infty$ such that $a < b$, and let $f : (a,b) \to \mathbb{R}$ be a continuous function which is strictly increasing. Then the image of f is an open interval (c,d), where c and d are real numbers or $\pm\infty$, and there is an inverse function $g : (c,d) \to \mathbb{R}$ such that $g(f(x)) = x$ for all $x \in (a,b)$ and $f(g(y)) = y$ for all $y \in (c,d)$; in other words, $x = g(y)$ if and only if $y = f(x)$.*

There is a similar result for continuous functions which are strictly decreasing, and it is an immediate consequence of the theorem by the following argument:

> If we are given a continuous function $f : (a,b) \to \mathbb{R}$ which is strictly decreasing, then $F = -f$ is increasing, so by the theorem it has an inverse G, and we obtain the inverse g to f by setting g equal to $-G$.

Proof of the theorem. Since (a,b) is a connected subset of $\mathbb{R}$, its image must also be a connected subset of $\mathbb{R}$, which means that the image is some interval. In fact, the image must be an open interval, for (a,b) has no minimal or maximal element, and since f is strictly increasing the same must be true for its image. Therefore the image must be an interval of the form (c,d), where c and d are real numbers or $\pm\infty$.

Since a strictly increasing function is 1–1 (because $x < x'$ implies $f(x) < f(x')$), we know that there is some set-theoretic inverse function $g : (c,d) \to (a,b)$. This inverse is also strictly increasing, for if $y < y'$ and we have $y = f(x)$ and $y' = f(x')$, then we must have $g(y) = x < x' = g(y')$; if the latter did not hold, then we would have $xz \geq x'$, which would imply $y \geq y'$ because f is strictly increasing. To complete the proof of the theorem, we need to verify that g is continuous.

The continuity of g will follow if we can find a base $\mathcal{B}$ for the topology of (a,b) such that for each basic open set $V \in \mathcal{B}$ the set $g^{-1}[V]$ is open in (c,d). We shall take the base consisting of all open intervals $(u,v) \subset (a,b)$, where $a < u < v < b$, and we claim that $g^{-1}[\,(u,v)\,] = (\,f(u), f(v)\,)$. Since $g(y) \in (u,v)$ implies that $y = f \circ g(y) \in (\,f(u), f(v)\,)$ because f is strictly increasing, it follows immediately that $g^{-1}[\,(u,v)\,] \subset (\,f(u), f(v)\,)$, so at this point we only need to prove that the reverse inclusion also holds. But if $f(u) < y < f(v)$ then we can use the strictly increasing behavior of the inverse function g to conclude that $u = g \circ f(u) < g(y) < g \circ f(v) = v$ which yields $g(y) \in (u,v)$ and hence implies that $g^{-1}[\,(u,v)\,] \supset (\,f(u), f(v)\,)$. By the previous remarks, this concludes the proof that g is continuous.■

Generalization to other types of intervals

Similar results hold if the domain of the continuous and strictly increasing function f is either a closed interval or a half-open interval which contains an end point either on the left or on the right. One quick way of proving this is to extend the original function f to a larger interval so that the new function F is still continuous and strictly increasing.

Specifically, we can construct the extension F as follows: Let J be an interval of the form $[a,b)$, $(a,b]$ or $[a,b]$, where we allow $a=-\infty$ or $b=+\infty$ if a or b (respectively) is not contained in J.

(1) If $a \in J$, extend f to $(a-1,b)$ by the linear formula $F(x) = x - a + f(a)$ on $(a-1,a]$. The combination of these maps is continuous because $F(a) = f(a)$.

(2) If $b \in J$, extend f to $(a,b+1)$ by the linear formula $F(x) = x - b + f(b)$ on $[b,b+1)$. The combination of these maps is continuous because $F(b) = f(b)$.

(3) If $a,b \in J$, extend f to $(a-1,b+1)$ as in (1) and (2). For the same reasons as above, the combination of the maps defined on the subintervals $(a-1,a]$, $[a,b]$ and $[b,b+1)$ will be continuous.

In all cases the extended function F has a continuous, strictly increasing inverse by the theorem, and we shall denote this continuous, strictly increasing inverse by G. We can then retrieve the inverse g to f by taking the restriction of G to $(f(a),f(b)]$, $[f(a),f(b))$ and $[f(a),f(b)]$ in the respective three cases, and in each instance the continuity of g follows from the continuity of G.■

As before, there are also similar results for continuous and strictly **decreasing** functions defined on closed or half-open intervals.■