1 Continuous Functions on Metric Spaces

1.1 The Definition of Continuity

Up to this point we have studied functions $f : D \to E$ where D and E are subsets of or ^{n}. We now extend our discussion to functions $f : M \to N$ where M and N are metric spaces. Many concepts and proofs that apply to Euclidean spaces can be easily transferred to general metric spaces.

Definition 1 *A function $f : M \to N$ is* continuous *if $f^{-1}(A)$ is open for all open sets $A \subset N$.*

This definition is compatible with our previous definition of continuous functions on Euclidean spaces, see Theorem **??**.

Theorem 1 *The following are equivalent.*

1. *$f : M \to N$ is continuous.*
2. *$\forall\, x_0 \in M$, and $\forall\, 1/m$, $\exists\, 1/n$ such that*
$$d_M(x, x_0) < \frac{1}{n} \Rightarrow d_N(f(x), f(x_0)) < \frac{1}{m}$$
3. *$\forall\, \{x_j\} \subset M$, if $x_j \to x_0$ in M, then $f(x_j) \to f(x_0)$ in N.*

$1 \Rightarrow 2$: Let $x_0 \in M$ and let $1/m$ be given. By assumption, the inverse image of an open ball is open, so

$$\begin{aligned} f^{-1}(B_n(f(x_0), \frac{1}{m})) &= \{x \in M \mid f(x) \in B_N(f(x_0), \frac{1}{m})\} \\ &= \{x \in M \mid d_N(f(x), f(x_0)) < \frac{1}{m}\} \end{aligned}$$

is *open* and contains x_0. Thus, $\exists\, 1/n$ such that

$$B_M(x_0, 1/n) \subset f^{-1}(B_N(f(x_0), 1/m))$$

i.e., if $d_M(x, x_0) < 1/n$ then $d_N(f(x), f(x_0)) < 1/m$.

$2 \Rightarrow 3$: Let $x_j \to x_0$ in M, and let $1/m$ be given. By assumption, $\exists\, 1/n$ such that $D_n(f(x_j), f(x_0)) < 1/m$ whenever $d_M(x_j, x_0) < 1/n$. However, since $x_j \to x_0$, $\exists\, J$ such that $d_M(x_j, x_0) < 1/n$ whenever $j \geq J$. So, if $j \geq J$, then $d_N(f(x_j), f(x_0)) < 1/m$ showing that $f(x_j) \to f(x_0)$.

$3 \Rightarrow 1$: Let $A \subset N$ be open and suppose $f^{-1}(A)$ is *not* open. Then $\exists\, x_0 \in f^{-1}(A)$ such that no open ball about x_0 is contained in $f^{-1}(A)$. Therefore, $\forall\, 1/j$, $\exists\, x_j \in B_M(x_0, 1/j)$ such that $x_j \notin f^{-1}(A)$. Clearly, $x_j \to x_0$, so by assumption, $f(x_j) \to f(x_0)$. Since A is open, $\exists\, 1/m$ such that $B(f(x_0), 1/m) \subset A$. But since $f(x_j) \to f(x_0)$, $\exists\, J$ such that $f(x_j) \in B(f(x_0), 1/m) \subset A$, $\forall\, j \geq J$, and this implies $x_j \in f^{-1}(A)$, a contradiction.

1.1.1 Examples

1) Let M be $C[a,b]$ with the sup-norm. Then $I : M \to$ defined by $I(f) = \int_a^b f(x)dx$ is continuous. To prove this, let $f_n \to f$ with respect to the sup-norm. As we have seen this means that $f_n \to f$ uniformly. Therefore, f is integrable and $I(f_n) \to I(f)$ by Theorem **??**, showing that I is continuous.

2) Let $M = C^1[a,b]$ with norm $\|f\|_* = \|f\|_\infty + \|f'\|_\infty$ and let $N = C[a,b]$ with the sup-norm. Then $D : M \to N$ defined by $D(f) = f'$ is continuous. To see this, note that if f_n converges to f with respect to the $*$-norm, then $f_n \to f$ and $f'_n \to f'$ uniformly, see Theorems **??** and **??**. Therefore, $D(f_n) \to D(f)$ with respect to the sup-norm, proving that D is continuous.

Recall that the uniform convergence of functions does not necessarily imply the convergence of their derivatives. Hence, D would not be continuous with respect to the sup-norm on $C^1[a,b]$. Continuity depends on the distance function under consideration.

3) Let M be a metric space and fix $x_0 \in M$. Then $f : M \to$ defined by $f(x) = d(x, x_0)$ is continuous: If $x_n \to x$ in M, then, by definition, $\forall\, 1/m$, $\exists\, N$ such that $d(x_n, x) < 1/m$, $\forall\, n \geq N$. Since $d(x_n, x_0) \leq d(x_n, x) + d(x, x_0)$ we find that

$$|f(x_n) - f(x)| = |d((x_n, x_0) - d(x, x_0)| \leq d(x_n, x) < \frac{1}{m}, \quad \forall n \geq N$$

so $f(x_n) \to f(x_0)$. Therefore, f is continuous.

Remark: The usual ways of combining continuous functions lead to new continuous functions. For example, if $f : M \to N$ and $g : N \to P$ are continuous, then it is easy to prove that the composition, $f \circ g : M \to P$, $f \circ g(x) = f(g(x))$, is continuous. If $f, g : M \to^n$ are continuous, then $f + g : M \to^N$, $(f+g)(x) = f(x) + g(x)$, is continuous. If $f : M \to^n$ and $g : M \to$ are continuous, then $f \cdot g : M \to^n$, $(f \cdot g)(x) = f(x)g(x)$, and $f/g : M \to^n$, $(f/g)(x) = f(x)/g(x)$ (for x such that $g(x) \neq 0$), are continuous. If $f_1, \ldots, f_n : M \to$ are continuous, then $f : M \to^n$, $f(x) = (f_1(x), \ldots, f_n(x))$ is continuous and conversely.

1.2 Continuous Functions on Compact Domains

Let M and N be metric spaces with M compact. In this section we establish theorems for continuous functions $f : M \to N$ that are analogous to familiar theorems about continuous real-valued functions on compact intervals.

Definition 2 *$f : M \to N$ is* uniformly continuous *if $\forall\, 1/m$, $\exists\, 1/n$ such that $\forall\, x, y \in M$*

$$d_M(x,y) < \frac{1}{n} \Rightarrow d_N(f(x), f(y)) < \frac{1}{m}$$

Theorem 2 *If M is compact and $f : M \to N$ is continuous, then f is uniformly continuous.*

Given $1/m$ and $x \in M$, $\exists\, 1/n_x$ (depending on $1/m$ and x) such that

$$d_M(x,y) < \frac{2}{n_x} \Rightarrow d_N(f(x), f(y)) < \frac{1}{2m} \tag{1}$$

by the continuity of f. The open balls $B_M(x, 1/n_x)$, for $x \in M$, cover M, and since M is compact, a finite number will cover M:

$$M = B_M(x_1, 1/n_1) \cup \cdots B_M(x_t, 1/n_t)$$

Let $n = \max\{n_1, \ldots, n_t\}$.

For any $x, y \in M$, x must lie in one of the open balls, $x \in B_M(x_i, 1/n_i)$ for some i. If $d_M(x,y) < 1/n$ then

$$d_M(x_i, y) \le d_M(x_i, x) + d_M(x, y) < \frac{1}{n_i} + \frac{1}{n} \le \frac{2}{n_i}$$

so

$$d_N(f(x), f(y)) \le d_N(f(x), f(x_i)) + d_N(f(x_i), f(y)) < \frac{1}{2m} + \frac{1}{2m} = \frac{1}{m}$$

by (**??**).

Theorem 3 *If M is compact and $f : M \to$ is continuous, then f has a maximum and minimum in M, that is, there exists points $u, v \in M$ such that*

$$f(u) = \sup_{x \in M} f(x) \quad \textit{and} \quad f(v) = \inf_{x \in M} f(x)$$

There is a sequence $\{x_n\} \subset M$ such that $\lim f(x_n) = \sup f(x)$ (or to ∞ if $\sup f(x) = \infty$). Since M is compact, there is a convergent subsequence, $x_{n'} \to b$. Then $\sup f(x) = \lim f(x_{n'}) = f(b)$ by the continuity of f, showing that the supremum is finite and achieved at a point $b \in M$. The proof for the infimum is similar.

Theorem 4 *Suppose $f : M \to N$ is continuous. If $A \subset M$ is compact, then $f(A) \subset N$ is compact.*

The argument is word-for-word the same as given in Theorem **??**: Let $f(A) \subset \bigcup U_\alpha$ be an open cover of $f(A)$. Then $\bigcup f^{-1}(U_\alpha)$ is an open cover of A. Since A is compact, there exists a finite subcover, $A \subset f^{-1}(U_{\alpha_1}) \cup \ldots \cup f^{-1}(U_{\alpha_n})$. But then $f(A) \subset U_{\alpha_1} \cup \ldots \cup U_{\alpha_n}$ is a finite subcover of $f(A)$ proving that $f(A)$ is compact.

1.3 Connectedness

Recall the Intermediate Value Theorem:

If $f : [a, b] \to$ is continuous, then $f([a, b]) = [c, d]$ where $c = \min f(x)$ and $d = \max f(x)$.

The point of this theorem is that if there are no gaps in the domain of a continuous function, then there are no gaps in the values of the function. To generalize this theorem to metric spaces we need a way to specify that there are "no gaps." There are two ways to define such a concept: being "connected," and being "arc-wise connected." The first is the more general and is defined in terms of open sets. The second concept is more restrictive, but is often more practical to apply.

Definition 3 *Let M be a metric space.*

1. *M is* connected *if there does not exist a pair of disjoint, non-empty, open sets A and B with $M = A \cup B$.*
2. *M is* arc-wise connected *is there is a curve connecting any two points in M. More precisely, given any two points $x, y \in M$, there exists a continuous function $f : [a, b] \to M$ such that $f(a) = x$ and $f(b) = y$.*

Any interval of real numbers is connected.

Suppose I is an interval and $I = A \cup B$ with A and B disjoint, non-empty, open subsets of I. Fix $a \in A$ and $b \in B$. We may assume $a < b$. To produce a contradiction, we look for a dividing point between A and B. Let

$$d = \sup\{x \in A \mid x \leq b\}$$

Let $\{x_n\} \subset A$ such that $x_n \to d$. Since A is the complement of the open set B in I, A is closed in I. Therefore, $d \in A$ unless $d \notin I$. The latter is only possible if d is an endpoint of the interval I, which is not the case here since $a \leq d \leq b$. Therefore, $d \in A$. But A is also open, so there exists an open interval around d contained in A, which contradicts d being the supremum of points in A that are $\leq b$.

Theorem 5 *Let M be a metric space.*

1. *M is arc-wise connected $\Rightarrow$ M is connected.*
2. *M is connected $\iff$ the only subsets of M that are both open and closed are $\emptyset$ and M.*

1) Suppose M is arc-wise connected but not connected. Then $M = A \cup B$ where A and B are disjoint, non-empty and open. Let $x \in A$ and $y \in B$ and let $f : [a.b] \to M$ be a continuous function such that $f(a) = x$ and $f(b) = y$. Then

$$[a, b] = f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$$

decomposes $[a, b]$ into disjoint, non-empty, open sets, contradicting Proposition **??**. Therefore, M must be connected.

2) ($\Rightarrow$) Suppose M is connected and $A \subset M$ is both open and closed. Then $B = A^c$ is open and closed by Theorem **??**. Both A and B cannot be non-empty, since otherwise $M = A \cup B$ with A and B disjoint, non-empty, open subsets. Therefore, either $A = \emptyset$ and $B = M$, or $B = \emptyset$ and $A = M$.

($\Leftarrow$) Conversely, suppose the only subsets of M that are both open and closed are $\emptyset$ and M. If $M = A \cup B$ with A and B disjoint and open in M, then $A = B^c$ and $B = A^c$ are also be closed by Theorem **??**. By assumption, either $A = \emptyset$ and $B = M$, or $B = \emptyset$ and $A = M$, proving that M cannot be decomposed into a disjoint union of *non-empty* open sets. Therefore, M is connected.

Example: It is possible for a set to be connected but not arc-wise connected. The graph of the function

$$f(x) = \begin{cases} \sin(1/x) & x > 0 \\ 0 & x = 0 \end{cases}$$

is a connected subset of ^{2}: Any decomposition of the graph, $G = \{(x, f(x) \mid x \geq 0\}$, into disjoint, non-empty, open subsets would give a similar decomposition of the part of the graph, G_0, for $x > 0$ which is impossible since G_0 is clearly arc-wise connected. However, there is no curve from $(0, 0)$ to any other point on the graph. Suppose $f : [0, 1] \to G$ is continuous with $f(0) = (0, 0)$ and $f(1) = (a, \sin(1/a))$ for some $a > 0$. Then $f(t) = (x(t), \sin(x(t))) \to (0, 0)$ as $t \to 0$. This implies $x(t) \to 0$ and $\sin(1/x(t)) \to 0$, which is impossible since $\lim_{u \to 0} \sin(1/u)$ does not exist.

Theorem 6 *Let M and N be metric spaces and suppose $f : M \to N$ is continuous. If M is connected then $f(M)$ is connected. If M is arc-wise connected, then $f(M)$ is arc-wise connected.*

Suppose M is connected. If $f(M) = A \cup B$ with A and B disjoint, non-empty, open subsets, then

$$M = f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$$

decomposes M into disjoint, non-empty, open subsets, a contradiction. Therefore, no such decomposition of $f(M)$ exists and $f(M)$ is connected.

Next assume M is arc-wise connected. Let $x, y \in f(M)$. Then there are $u, v \in M$ such that $x = f(u)$ and $y = f(v)$. Let $g : [a, b] \to M$ be a continuous function such that $g(a) = u$ and $g(b) = v$. Then $f \circ g : [a, b] \to f(M)$ is continuous with $f \circ g(a) = x$ and $f \circ g(b) = y$, proving that $f(M)$ is arc-wise connected.

1.4 The Contractive Mapping Principle

In this section we consider a continuous function of a complete metric space M to itself, $f : M \to M$, that "shrinks" distance. We shall prove that any such function must have a unique *fixed point*, that is, a point $x_0 \in M$ such that $f(x_0) = x_0$. This theorem has many applications in analysis. In fact, we shall use it to prove two fundamentally important theorems:

- The existence of solutions to ordinary differential equations.
- The Implicit Function Theorem.

Definition 4 *Let M and N be metric spaces.*

1. *A* mapping *of M to N is a continuous function $f : M \to N$.*
2. *A mapping $f : M \to M$ is* contractive *if $\exists\, r$, $0 \le r < 1$, such that*

$$d(f(x), f(y)) \le r\, d(x, y), \quad \forall\, x, y \in M$$

We introduce the notation f^n for the iterated mapping $f \circ f \circ \cdots \circ f$ (n times).

Theorem 7 (Contractive Mapping Principle) *Let M be a complete metric space and assume $f : M \to M$ is a contractive mapping. Then there exists a unique fixed point $x_0 \in M$, $f(x_0) = x_0$. Moreover, $\forall\, x \in M$, $\exists\, c > 0$ such that*

$$d(f^n(x), x_0) \le cr^n$$

In particular, $\lim_{n\to\infty} f^n(x) = x_0$.

Let $x \in M$. We first show the sequence $\{f^n(x)\}$ is Cauchy. The definition of a contractive mapping implies

$$d(f^{n+1}(x), f^n(x)) \le r\, d(f^n(x), f^{n-1}(x)) \le \cdots \le r^n d(f(x), x)$$

For any $j > n$, the above inequality, along with the triangle inequality, implies

$$\begin{aligned} d(f^j(x), f^n(x)) &\le d(f^j(x), f^{j-1}(x)) + \cdots + d(f^{n+1}(x), f^n(x)) \\ &\le (r^{j-1} + \cdots + r^n) d(f(x), x) \\ &\le \frac{r^n}{1-r} d(f(x), x) \end{aligned}$$

Since $r < 1$, $r^n d(f(x), x)/(1-r) \to 0$ as $n \to \infty$, proving that the sequence $\{f^n(x)\}$ is Cauchy.

Because M is complete, the sequence $\{f^n(x)\}$ must converge to some point $x_0 \in M$. This point is a fixed point of f:

$$f(x_0) = f(\lim_{n\to\infty} f^n(x)) = \lim_{n\to\infty} f^{n+1}(x) = x_0$$

The previous inequality gives us the rate of convergence:

$$d(x_0, f^n(x)) = \lim_{j\to\infty} d(f^j(x), f^n(x)) \le cr^n$$

where $c = d(f(x), x)/(1-r)$.

The fixed point is unique since if $f(x_1) = x_1$ then

$$d(x_0, x_1) = d(f(x_0), f(x_1)) \le rd(x_0, x_1)$$

But $r < 1$, so $d(x_0, x_1) = 0$ and hence $x_1 = x_0$.

It can be difficult to apply the Contractive Mapping Principle. In order for a mapping $f : M \to M$ to have the strong property of shrinking distances, one must often first restrict the domain to a smaller subset $M_0 \subset M$. But then one must verify that $f(M_0) \subset M_0$ in order to apply the theorem. For example, if $f(x) = x^2 + 1$, then $|f(x) - f(y)| = |x^2 - y^2| = |x + y| \cdot |x - y|$. If we restrict the domain to $[-1/3, 1/3]$, then $|f(x) - f(y)| \le (2/3)|x - y|$ which looks like the condition for a contractive mapping. However, $f([-1/3, 1/3]) = [1, 10/9]$, so f does not map $[-1/3, 1/3]$ to itself.

Example: Suppose $f : [a, b] \to [a, b]$ satisfies $|f'(x)| < 1$ for all $x \in [a, b]$. Then f is contractive because $\forall\, x, y \in [a, b]$, the Mean Value Theorem implies there is a z between x and y such that

$$|f(x) - f(y)| = |f'(z)| \cdot |x - y| \le r \cdot |x - y|, \quad r = \max_{z \in [a,b]} |f'(z)| < 1$$

The Contractive Mapping Principle tells us that f has a fixed point in $[a, b]$ that can be found by repeatedly applying f to any starting number in $[a, b]$. To get a numerical approximation to the fixed point, one could program f into a calculator, evaluate f on a, then repeatedly evaluate f on the previous output until the numbers on the display do not change.

For example, consider $f(x) = e^{-x}$ on the interval $[.3, .8]$. Since f is decreasing, $f([.3, .8]) = [f(.8), f(.3)] = [e^{-.8}, e^{-.3}] \subset [.3, .8]$. Moreover, $|f'(x)| \le e^{-.3} < 1$ for $x \ge .3$, so f is contractive. The Contractive Mapping Principle implies that there exists a unique number $x_0 \in [.3, .8]$ such that $f(x_0) = x_0$, i.e., such that $e^{-x_0} = x_0$. If we start with 0.5 and repeatedly press the "button" for the function e^{-x} on the calculator, we obtain the sequence: 0.500000, 0.606531, 0.545239, 0.579703, 0.560065, 0.571172, 0.564863, 0.568438, 0.566409, 0.567560, 0.566907, 0.567277, 0.567067, 0.567186, 0.567119, 0.567157, 0.567135, 0.567148, 0.567141, 0.567145, 0.567142, 0.567144, 0.567143, ... The digits stabilize at 0.567143, and we find that $e^{-0.567143} = 0.567143$.

Actually, we do not need a contractive mapping $f : [a, b] \to [a, b]$ in order to find a fixed point.

Any mapping $f : [a, b] \to [a, b]$ has a fixed point.

To prove this, consider $g(x) = f(x) - x$. Since $g(a) = f(a) - a \ge 0$ and $g(b) = f(b) - b \le 0$, the Intermediate Value Theorem implies that there exists $c \in [a, b]$ such that $g(c) = 0$ and hence $f(c) = c$. This is a special case of a more general fact known as the *Brouwer Fixed Point Theorem*:

Any mapping of a closed ball in ^{n} into itself has a fixed point.

It is necessary to have a *closed* ball for this theorem, even in the case of intervals. The function $f(x) = x/2$ maps $(0, 1)$ to itself, but does not have a fixed point in $(0, 1)$. The disadvantage of Brouwer's theorem over the Contractive Mapping Principle is that it is not constructive. We know the fixed point must exist, but the theorem does not tell us how to find it.