Chapter 1
The Banach Contraction Principle

**Definition 1.1** Let $X$ be a metric space equipped with a distance $d$. A map

$$f : X \rightarrow X$$

is said to be **Lipschitz continuous** if there is $\lambda \geq 0$ such that

$$d(f(x), f(y)) \leq \lambda d(x, y), \quad \forall x, y \in X.$$ 

In this case, it is readily seen that there exists the smallest value $\lambda$ for which the inequality holds, called the **Lipschitz constant** of $f$.

- If $\lambda = 1$, then $f$ is said to be **non-expansive**.
- If $\lambda < 1$, then $f$ is said to be a **contraction**.

Perhaps the most important result in the theory of fixed points is the celebrated Banach contraction principle (BCP), stated and proved by Banach [4] in 1922 and subsequently -and independently- rediscovered by Caccioppoli [13].

**Theorem 1.1** (BCP) Let $f$ be a contraction on a complete metric space $X$. Then $f$ has a unique fixed point $\bar{x} \in X$.

*Proof* Note first that if $x, y \in X$ are fixed points of $f$, then

$$d(x, y) = d(f(x), f(y)) \leq \lambda d(x, y),$$

which implies $x = y$. Choose now any $x_0 \in X$, and define the iterated sequence

$$x_{n+1} = f(x_n).$$
By induction on $n$,
\[ d(x_{n+1}, x_n) \leq \lambda^n d(f(x_0), x_0), \quad \forall n \geq 1. \]

Thus, if $n \in \mathbb{N}$ and $m \geq 1$, recalling that $\lambda < 1$ we have
\[
\begin{align*}
  d(x_{n+m}, x_n) &\leq d(x_{n+m}, x_{n+m-1}) + \cdots + d(x_{n+1}, x_n) \\
  &\leq (\lambda^{n+m-1} + \cdots + \lambda^n) d(f(x_0), x_0) \\
  &\leq \lambda^n \left( \sum_{j=0}^{\infty} \lambda^j \right) d(f(x_0), x_0) \\
  &= \frac{\lambda^n}{1-\lambda} d(f(x_0), x_0).
\end{align*}
\]

Hence $x_n$ is a Cauchy sequence, and admits a limit $\bar{x} \in X$, for $X$ is complete. Since $f$ is continuous, we have
\[
  f(\bar{x}) = \lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} x_{n+1} = \bar{x},
\]
as claimed. $\square$

The construction of the approximating sequence $x_n$ above is known as Picard iteration method [62].

**Corollary 1.1** The approximating sequence $x_n$ fulfills the estimate
\[
  d(\bar{x}, x_n) \leq \frac{\lambda^n}{1-\lambda} d(f(x_0), x_0).
\]

**Proof** From the previous proof,
\[
  d(x_{n+m}, x_n) \leq \frac{\lambda^n}{1-\lambda} d(f(x_0), x_0).
\]

Knowing now that $x_n \to \bar{x}$, the claim follows by letting $m \to \infty$. $\square$

**Remark** As noted in [56], the inequality established in Corollary 1.1 is important also for the following reason: assume we want to find the fixed point up to an “error” $\varepsilon > 0$, namely, we want to find a point $\hat{x}$ such that
\[
  d(\bar{x}, \hat{x}) < \varepsilon,
\]
where $\bar{x}$ is the actual fixed point. Then, the inequality above allows to find an explicit value of $n \in \mathbb{N}$ for which $\hat{x} = x_n$ will do. Indeed, such an $n$ has to comply with the relation $d(x_n, \bar{x}) < \varepsilon$. Accordingly, we have to take $n$ large enough that
The Banach Contraction Principle 5

\[
\frac{\lambda^n}{1 - \lambda} d(f(x_0), x_0) < \varepsilon.
\]

The quantity \( \varrho = d(f(x_0), x_0) \) is something known right after the first iteration. Hence, recalling that \( \log \lambda < 0 \), the desired \( n \) fulfills

\[
n > \frac{\log \varepsilon + \log(1 - \lambda) - \log \varrho}{\log \lambda}.
\]

The completeness of \( X \) plays a crucial role in the BCP. Indeed, contractions on incomplete metric spaces may fail to have fixed points.

**Example** Let \( X = (0, 1] \) with the usual distance. Define \( f : X \to X \) as \( f(x) = x/2 \).

**Corollary 1.2** Let \( X \) be a complete metric space and \( W \) be a topological space. Let \( f : X \times W \to X \) be a continuous function. Assume that \( f \) is a contraction on \( X \) uniformly in \( W \), that is,

\[
d(f(x, w), f(y, w)) \leq \lambda d(x, y), \quad \forall x, y \in X, \forall w \in W,
\]

for some \( \lambda < 1 \). Then, for every fixed \( w \in W \), the map \( x \mapsto f(x, w) \) has a unique fixed point \( \varphi(w) \). Moreover, the function \( w \mapsto \varphi(w) \) is continuous from \( W \) to \( X \).

Note that if \( f : X \times W \to X \) is continuous on \( W \) for every fixed \( x \in X \) and is a contraction on \( X \) uniformly in \( W \), then \( f \) is automatically continuous on \( X \times W \).

**Proof** In light of Theorem 1.1, we only have to prove the continuity of \( \varphi \). For \( w, w_0 \in W \), we have

\[
d(\varphi(w), \varphi(w_0)) = d(f(\varphi(w), w), f(\varphi(w_0), w_0))
\leq d(f(\varphi(w), w), f(\varphi(w_0), w)) + d(f(\varphi(w_0), w), f(\varphi(w_0), w_0))
\leq \lambda d(\varphi(w), \varphi(w_0)) + d(f(\varphi(w_0), w), f(\varphi(w_0), w_0)),
\]

which implies

\[
d(\varphi(w), \varphi(w_0)) \leq \frac{1}{1 - \lambda} d(f(\varphi(w_0), w), f(\varphi(w_0), w_0)).
\]

Since the above right-hand side goes to zero as \( w \to w_0 \), we have the desired continuity.

**Remark** If in addition \( W \) is a metric space and \( f \) is Lipschitz continuous in \( W \), uniformly with respect to \( X \), with Lipschitz constant \( L \geq 0 \), then the function \( w \mapsto \varphi(w) \) is Lipschitz continuous with Lipschitz constant less than or equal to \( L/(1 - \lambda) \).
Theorem 1.1 establishes a sufficient condition in order for $f$ to have a unique fixed point.

**Example** Consider the map

$$f(x) = \begin{cases} 
\frac{1}{2} + 2x & x \in [0, 1/4], \\
\frac{1}{2} & x \in (1/4, 1] 
\end{cases}$$

mapping $[0, 1]$ onto itself. Then $f$ is not even continuous, but it has the unique fixed point $\bar{x} = 1/2$.

The next corollary takes into account this situation, providing existence and uniqueness of a fixed point under more general conditions. We first need a definition.

**Definition 1.2** For $f: X \to X$ and $n \in \mathbb{N}$, we denote by $f^n$ the $n$th-iteration of $f$, namely,

$$f^n = f \circ \cdots \circ f \bigg|_{n \text{ times}}$$

where $f^0$ is understood to be the identity map.

**Corollary 1.3** Let $X$ be a complete metric space, and let $f: X \to X$. If $f^m$ is a contraction for some $m \geq 1$, then $f$ has a unique fixed point $\bar{x} \in X$. Moreover, for every $x_0 \in X$, the sequence $f^n(x_0)$ converges to $\bar{x}$.

Note that in the previous example $f^2(x) \equiv 1/2$.

**Proof** Let $\bar{x}$ be the unique fixed point of $f^m$, given by Theorem 1.1. Then

$$f^m(f(\bar{x})) = f(f^m(\bar{x})) = f(\bar{x}),$$

which implies $f(\bar{x}) = \bar{x}$. Since a fixed point of $f$ is clearly a fixed point of $f^m$, we have uniqueness as well. The proof of the convergence $f^n(x_0) \to \bar{x}$ is left as an exercise. □

We conclude the chapter discussing the converse to the BCP. Assume we are given a set $X$ and a map $f: X \to X$. We are interested to find a metric $d$ on $X$ such that $(X, d)$ is a complete metric space and $f$ is a contraction on $X$. Clearly, in light of Theorem 1.1, a necessary condition is that each iteration $f^n$ has a unique fixed point. Surprisingly enough, the condition turns out to be sufficient as well.

**Theorem 1.2** Let $X$ be an arbitrary set, and let $f: X \to X$ be a map such that $f^n$ has a unique fixed point $\bar{x} \in X$ for every $n \geq 1$. Then for every $\varepsilon \in (0, 1)$, there is a metric $d = d_\varepsilon$ on $X$ that makes $X$ a complete metric space, and $f$ is a contraction on $X$ with Lipschitz constant less than or equal to $\varepsilon$. 


Proof Choose $\varepsilon \in (0, 1)$. Let $Z$ be the subset of $X$ consisting of all elements $x$ such that $f^k(x) = \bar{x}$ for some $k \in \mathbb{N}$. We define the following equivalence relation on the (possibly empty) set $X \setminus Z$: we say that $x \sim y$ if and only if $f^n(x) = f^m(y)$ for some $n, m \in \mathbb{N}$. Note that if 

\[ f^n(x) = f^m(y) \quad \text{and} \quad f'^n(x) = f'^m(y), \]

then 

\[ f^{n+m'}(x) = f^{m+n'}(x). \]

But since $x \not\in Z$, this yields $n + m' = m + n'$, that is, 

\[ n - m = n' - m'. \]

At this point, by means of the axiom of choice, we select an element from each equivalence class. We now define the distance of $\bar{x}$ from a generic $x \in X$ by setting 

\[
\begin{align*}
0 & \quad \text{if } x = \bar{x}, \\
\varepsilon^{-k} & \quad \text{if } \bar{x} \neq x \in Z, \\
\varepsilon^{n-m} & \quad \text{if } x \not\in Z,
\end{align*}
\]

where $k = \min \{ p \geq 1 : f^p(x) = \bar{x} \}$, while $n, m \in \mathbb{N}$ are such that $f^n(\hat{x}) = f^m(x)$, where $\hat{x}$ is the selected representative of the equivalence class $[x]$. The definition is unambiguous, due to the discussion above. Finally, for any $x, y \in X$, we set 

\[
d(x, y) = \begin{cases} 
0 & \quad \text{if } x = y, \\
d(x, \bar{x}) + d(y, \bar{x}) & \quad \text{if } x \neq y.
\end{cases}
\]

It is straightforward to verify that $d$ is a metric. To see that $d$ is complete, observe that the only Cauchy sequences which do not converge to $\bar{x}$ are ultimately constant. We are left to show that $f$ is a contraction with Lipschitz constant equal to $\varepsilon$. Let then $x \in X$, with $x \neq \bar{x}$. We shall distinguish three cases:

- If $x \in Z$ and $f(x) = \bar{x}$, then 

\[ 0 = d(f(x), \bar{x}) < \varepsilon d(x, \bar{x}). \]

- If $x \in Z$ and $f(x) \neq \bar{x}$, then there is the smallest $k \geq 2$ such that 

\[ \bar{x} = f^k(x) = f^{k-1}(f(x)). \]

Hence, 

\[ d(f(x), \bar{x}) = \varepsilon^{-k+1} = \varepsilon d(x, \bar{x}). \]
• If \( x \notin Z \), then there are \( n, m \in \mathbb{N} \), and we can take \( m \geq 1 \), such that

\[
\hat{x}^n = \hat{x}^m = \hat{x}^{m-1}(f(x)),
\]

where \( \hat{x} \) is the representative of \([x]\). Hence,

\[
d(f(x), \bar{x}) = \epsilon^{n-m+1} = \epsilon d(x, \bar{x}).
\]

In summary, for every \( x \in X \),

\[
d(f(x), \bar{x}) \leq \epsilon d(x, \bar{x}).
\]

From the definition of the distance, given any \( x \neq y \in X \) we conclude that

\[
d(f(x), f(y)) \leq d(f(x), \bar{x}) + d(f(y), \bar{x})
\leq \epsilon [d(x, \bar{x}) + d(y, \bar{x})]
= \epsilon d(x, y),
\]

as desired. \( \square \)

Theorem 1.2 is due to Bessaga [6]. An alternative proof can be found in the book [18] (pp. 191–192). The argument presented here, indeed much simpler, is due to Peirone [61]. There is an interesting related result for compact metric spaces due to Janoš [37], where an equivalent metric that makes \( f \) a contraction is constructed.