THE CONTRACTION MAPPING THEOREM

KEITH CONRAD

1. Introduction

A mapping $f\colon X \to X$ from a set X to itself has a *fixed point* if there is an $x \in X$ such that $f(x) = x$. The simplest fixed point theorem is that a continuous function $f\colon [a,b] \to [a,b]$ has at least one fixed point. This is a consequence of the intermediate value theorem from calculus, as follows. Since $f(a) \geq a$ and $f(b) \leq b$, we have $f(b) - b \leq 0 \leq f(a) - a$. The difference $f(x) - x$ is continuous, so by the intermediate value theorem 0 is a value of $f(x) - x$ for some $x \in [a,b]$, and that x is a fixed point of f. (Of course, there could be more than one fixed point.) There are higher-dimensional generalizations of this result, such as the Brouwer fixed point theorem and the Lefschetz fixed point theorem. These generalizations and their proofs belong to algebraic topology.

We are interested here in the most basic fixed point theorem in analysis. It is due to Banach and appeared in his Ph.D. thesis (1920, published in 1922).

Theorem 1.1. *Let (X,d) be a complete metric space and $f\colon X \to X$ be a map such that*

$$d(f(x), f(x')) \leq c d(x,x')$$

for some $0 \leq c < 1$ and all x and x' in X. Then f has a unique fixed point in X. Moreover, for any $x_0 \in X$ the sequence of iterates $x_0, f(x_0), f(f(x_0)), \ldots$ converges to the fixed point of f.

When $d(f(x), f(x')) \leq c d(x,x')$ for some $0 \leq c < 1$ and all x and x' in X, f is called a *contraction*. A contraction shrinks distances by a uniform factor c less than 1 for all pairs of points. Theorem 1.1 is called the contraction mapping theorem or Banach's fixed-point theorem.

A contraction mapping is uniformly continuous. This is clear when $c = 0$ since then f is a constant function. If $0 < c < 1$ and we are given $\varepsilon > 0$, setting $\delta = \varepsilon/c$ implies that if $d(x,x') < \delta$ then $d(f(x), f(x')) \leq c d(x,x') < c\delta = \varepsilon$.

An elementary account of the contraction mapping theorem and some applications, including its role in solving nonlinear ordinary differential equations, is in [3]. The contraction mapping theorem can be used to prove the inverse function theorem [6], [9, pp. 221-223] and to construct fractals [11, Chap. 5]. Its role in Google's page-rank algorithm is in [8], [13]. A survey of fixed point theorems in metric spaces is in [7].

2. Proof of the Contraction Mapping Theorem

Recalling the notation, $f\colon X \to X$ is a contraction with contraction constant c. We want to show f has a unique fixed point, which can be obtained as a limit through iteration of f from any initial value. To show f has at most one fixed point in X, let a and a' be fixed points of f. Then

$$d(a,a') = d(f(a), f(a')) \leq c d(a,a').$$

If $a \neq a'$ then $d(a, a') > 0$ so we can divide by $d(a, a')$ to get $1 \leq c$, which is false. Thus $a = a'$.

Next we want to show, for any $x_0 \in X$, that the recursively defined iterates $x_n = f(x_{n-1})$ for $n \geq 1$ converge to a fixed point of f. The key idea is that iterating the function several times contracts distances by an increasing power of the contraction constant. This brings points together through iteration at a geometric rate, and that will be enough to force convergence of the iterates because X is complete.

How close is x_n to x_{n+1}? For any $n \geq 1$, $d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq cd(x_{n-1}, x_n)$. Therefore

$$d(x_n, x_{n+1}) \leq cd(x_{n-1}, x_n) \leq c^2 d(x_{n-2}, x_{n-1}) \leq \cdots \leq c^n d(x_0, x_1).$$

Using the expression on the far right as an upper bound on $d(x_n, x_{n+1})$ shows the x_n's are getting consecutively close at a geometric rate. This implies the x_n's are Cauchy: for any $m > n$, using the triangle inequality several times shows

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m) \\ &\leq c^n d(x_0, x_1) + c^{n+1} d(x_0, x_1) + \cdots + c^{m-1} d(x_0, x_1) \\ &= (c^n + c^{n+1} + \cdots + c^{m-1}) d(x_0, x_1) \\ &\leq (c^n + c^{n+1} + c^{n+2} + \cdots) d(x_0, x_1) \\ &= \frac{c^n}{1-c} d(x_0, x_1). \qquad (2.1) \end{aligned}$$

This is true when $m = n$ too. To prove from this bound that the x_n's are Cauchy, for any $\varepsilon > 0$ pick $N \geq 1$ such that $(c^N/(1-c))d(x_0, x_1) < \varepsilon$. Then for any $m > n \geq N$,

$$d(x_n, x_m) \leq \frac{c^n}{1-c} d(x_0, x_1) \leq \frac{c^N}{1-c} d(x_0, x_1) < \varepsilon.$$

This proves $\{x_n\}$ is a Cauchy sequence. Since X is *complete*, the x_n's converge in X. Set $a = \lim_{n \to \infty} x_n$ in X.

We noted before that contractions are uniformly continuous, so the continuity of f tells us that from $x_n \to a$ we get $f(x_n) \to f(a)$. Since $f(x_n) = x_{n+1}$, $f(x_n) \to a$ as $n \to \infty$. Then $f(a)$ and a are both limits of $\{x_n\}_{n \geq 0}$. From the uniqueness of limits, $a = f(a)$. This concludes the proof of the contraction mapping theorem.

Example 2.1. Consider the cosine function on $[0, 1]$. Graphs of $y = \cos x$ and $y = x$ intersect once over $[0, 1]$, which means the cosine function has a fixed point in $[0, 1]$. We will show this point can be obtained through iteration.

Since cosine is decreasing on $[0, 1]$ and $\cos 1 \approx .54$, $\cos([0, 1]) \subset [0, 1]$. To show $\cos x$ is a contraction mapping on $[0, 1]$, we will use the mean-value theorem: for any differentiable function f, $f(x) - f(y) = f'(t)(x - y)$ for some t between x and y, so bounding the derivative of f will give us a contraction constant. Taking $f(x) = \cos x$,

$$\cos x - \cos y = \cos'(t) \cdot (x - y) = (-\sin t)(x - y)$$

for some t between x and y. Thus

$$|\cos x - \cos y| = |\sin t||x - y|.$$

Since sine is increasing on $[0, 1]$, $|\sin t| = \sin t \leq \sin 1 \approx .84147$. Therefore

$$|\cos x - \cos y| \leq .8415|x - y|,$$

so cosine is a contraction mapping on $[0, 1]$, which is complete. Hence there is a unique $a \in [0, 1]$ with $\cos a = a$, and we can estimate a by iteration: repeatedly pressing the `cos` button on a calculator starting from any seed value in $[0, 1]$, we obtain the approximation $a \approx .739$.

Remark 2.2. When $f: X \to X$ is a contraction with constant $c < 1$, any iterate f^n is a contraction with constant $c^n < 1$, so each f^n has a unique fixed point. It has to be the unique fixed point of f, since the fixed point of f is fixed by f^n, so they are equal by uniqueness.

More broadly, if a function has a fixed point then every iterate of it has the same fixed point, and if some iterate of a function has a unique fixed point then the function has that same fixed point[1], but if a function has a unique fixed point then an iterate of it need not have a unique fixed point. For example, the function $f: [0, 1] \to [0, 1]$ given by $f(x) = 1 - x$ has the unique fixed point $\frac{1}{2}$, but $f^2(x) = f(f(x)) = x$ for all x in $[0, 1]$.

The proof of the contraction mapping theorem yields useful information about the rate of convergence towards the fixed point, as follows.

Corollary 2.3. *Let f be a contraction mapping on a complete metric space X, with contraction constant c and fixed point a. For any $x_0 \in X$, with f-iterates $\{x_n\}$, we have the estimates*

$$d(x_n, a) \le \frac{c^n}{1-c} d(x_0, f(x_0)), \tag{2.2}$$

$$d(x_n, a) \le c\,d(x_{n-1}, a), \tag{2.3}$$

and

$$d(x_n, a) \le \frac{c}{1-c} d(x_{n-1}, x_n). \tag{2.4}$$

Proof. From (2.1), for $m \ge n$ we have

$$d(x_n, x_m) \le \frac{c^n}{1-c} d(x_0, x_1) = \frac{c^n}{1-c} d(x_0, f(x_0)).$$

The right side is independent of m. Let $m \to \infty$ to get $d(x_n, a) \le (c^n/(1-c))d(x_0, f(x_0))$.

To derive (2.3), from a being a fixed point we get

$$d(x_n, a) = d(f(x_{n-1}), f(a)) \le c\,d(x_{n-1}, a). \tag{2.5}$$

Applying the triangle inequality to $d(x_{n-1}, a)$ on the right side of (2.5) using the three points x_{n-1}, x_n, and a,

$$d(x_n, a) \le c(d(x_{n-1}, x_n) + d(x_n, a)),$$

and isolating $d(x_n, a)$ yields

$$d(x_n, a) \le \frac{c}{1-c} d(x_{n-1}, x_n).$$

That's (2.4), so we're done. □

[1]Suppose f^n has the unique fixed point x_0. Applying f to the equation $f^n(x_0) = x_0$ makes it $f^{n+1}(x_0) = f(x_0)$, so $f^n(f(x_0)) = f(x_0)$. Therefore $f(x_0)$ is a fixed point of f^n, so $f(x_0) = x_0$ by uniqueness.

The three inequalities in Corollary 2.3 serve different purposes. The inequality (2.2) tells us, in terms of the distance between x_0 and $f(x_0) = x_1$, how many times we need to iterate f starting from x_0 to be certain that we are within a specified distance from the fixed point. This is an upper bound on how long we need to compute. It is called an *a priori* estimate. Inequality (2.3) shows that once we find a term by iteration within some desired distance of the fixed point, all further iterates will be within that distance. However, (2.3) is not so useful as an error estimate since both sides of (2.3) involve the unknown fixed point. The inequality (2.4) tells us, after each computation, how much closer we are to the fixed point in terms of the previous two iterations. This kind of estimate, called an *a posteriori* estimate, is very important because if two successive iterations are nearly equal, (2.4) guarantees that we are very close to the fixed point. For example, if $c = 2/3$ and $d(x_{n-1}, x_n) < 1/10^{10}$ then (2.4) tells us $d(x_n, a) < 2/10^{10} < 1/10^9$.

Example 2.4. Returning to Example 2.1, how many iterates of cosine are needed to be sure we have found the solution to $\cos a = a$ in $[0, 1]$ accurately to 3 decimal places? On $[0, 1]$, cosine is a contraction with constant $c = .8415$. Here are some iterations of $\cos x$ starting from $x_0 = 0$:

n	x_n
0	0
1	1
2	.5403023
$\vdots$	$\vdots$
19	.7393038
20	.7389377
21	.7391843
22	.7390182

The estimates (2.2) and (2.4) with $x_0 = 0$ become

$$|x_n - a| \leq \frac{.8415^n}{1 - .8415}, \quad |x_n - a| \leq \frac{.8415}{1 - .8415}|x_{n-1} - x_n|. \tag{2.6}$$

We can use these bounds to know with certainty a value of n for which $|x_n - a| < .001$. The first bound falls below .001 for the first time when $n = 51$, while the second one falls below .001 much earlier, at $n = 22$.

Taking $n = 22$, the second upper bound on $|x_{22} - a|$ in (2.6) is a little less than .0009, so a lies between

$$x_{22} - .0009 \approx .7381 \text{ and } x_{22} + .0009 \approx .7399.$$

The iterative process achieves the same accuracy in fewer steps if we begin the iteration closer to the fixed point. Let's start at $x_0 = .7$ instead of at $x_0 = 0$. Here are some iterations:

n	x_n
0	.7
1	.7648421
2	.7214916
$\vdots$	$\vdots$
14	.7389324
15	.7391879
16	.7390158

Now (2.2) and (2.4) with $x_0 = .7$ say

$$|x_n - a| \le \frac{.8415^n}{1 - .8415}|.7 - \cos(.7)|, \quad |x_n - a| \le \frac{.8415}{1 - .8415}|x_{n-1} - x_n|.$$

The first upper bound on $|x_n - a|$ falls below .001 for $n = 35$, and the second upper bound falls below .001 when $n = 16$. Using the second bound, a lies between

$$x_{16} - .00092 \approx .7381 \text{ and } x_{16} + .00092 \approx .7399.$$

Surprisingly, the contraction mapping theorem admits a converse, due to Bessaga [2], [10, pp. 523–526]. If X is any set (not yet a metric space), $c \in (0, 1)$, and $f \colon X \to X$ is a function such that each iterate $f^n \colon X \to X$ has a unique fixed point then there is a metric on X making it a complete metric space such that, for this metric, f is a contraction with contraction constant c. For instance, the function $f \colon \mathbf{R} \to \mathbf{R}$ given by $f(x) = 2x$ has 0 as its unique fixed point, and the same applies to its iterates $f^n(x) = 2^n x$. Therefore there is a metric on $\mathbf{R}$ with respect to which it is *complete* and the function $x \mapsto 2x$ is a contraction, so in particular $2^n \to 0$ in this metric: how strange! More practically, this converse theorem motivated the discovery of a new proof of the Cauchy–Kovalevskaya theorem in partial differential equations [12].

It might seem strange to have the hypothesis in Bessaga's theorem be that each iterate of f has a unique fixed point, rather than just f itself having a unique fixed point, but this detail is important. First of all, it is a necessary consequence of the contraction mapping theorem that any iterate of a contraction on a complete metric space has a unique fixed point since iterates of a contraction are contractions also. Second of all, there is an example of a function with a unique fixed point whose second iterate has more than one fixed point: the function $f(x) = 1 - x$ on $[0, 1]$ (from Remark 2.2) has for its second iterate the identity map, which has every point of $[0, 1]$ as a fixed point, so f is not a contraction with respect to any metric on $[0, 1]$ even though it has one fixed point in $[0, 1]$.

Pete Clark pointed out the following nice connection between "potential" fixed points of iterates of a function and whether the function is a contraction for some metric (see also [4, Theorem 17.5]).

Theorem 2.5. *Let X be a set and $f \colon X \to X$ be a function. The following two conditions are equivalent.*

(1) *Every iterate f^n has at most one fixed point in X.*
(2) *There is a metric on X with respect to which f is a contraction.*

We are not assuming in the second condition that the metric on X is complete, which corresponds to the possibility in the first condition that iterates of f may have no fixed points in X. This will become clearer in the proof.

Proof. (1) $\Rightarrow$ (2): We consider two cases: f has a fixed point and f has no fixed point.

If f has a fixed point in X then this is also a fixed point of every f^n, so the iterates of f have a unique fixed point by the hypothesis that the number of fixed points for f^n is 0 or 1. Bessaga's theorem now equips X with a (complete) metric with respect to which f is a contraction.

If f does not have a fixed point in X then no iterate of f can have a fixed point: that iterate would then have a unique fixed point by hypothesis, so f has a fixed point by the footnote in Remark 2.2, which is a contradiction. Let $X' = X \cup \{\infty\}$ and extend f to a map $X' \to X'$ by $f(\infty) = \infty$. Then every iterate of f on X' has unique fixed point ∞, so

Bessaga's theorem applied to X' tells us that X' admits a complete metric with respect to which f is a contraction. On the subset X, f is still a contraction, although the metric on X is not complete (the point ∞ in X' is a limit of many sequences in X by the contraction mapping theorem on X').

(2) $\Rightarrow$ (1): If there is a metric on X with respect to which f is a contraction, then every iterate of f is a contraction and we saw at the start of Section 2 that a contraction has at most one fixed point. □

3. Weakening the Hypotheses

In some situations a function is not a contraction but an iterate of it is. This turns out to suffice to get the conclusion of the contraction mapping theorem for the original function.

Theorem 3.1. *If X is a complete metric space and $f\colon X \to X$ is a mapping such that some iterate $f^N\colon X \to X$ is a contraction, then f has a unique fixed point. Moreover, the fixed point of f can be obtained by iteration of f starting from any $x_0 \in X$.*

Proof. By the contraction mapping theorem, f^N has a unique fixed point. Call it a, so $f^N(a) = a$. To show a is the only possible fixed point of f, observe that a fixed point of f is a fixed point of f^N, and thus must be a. To show a really is a fixed point of f, we note that $f(a) = f(f^N(a)) = f^N(f(a))$, so $f(a)$ is a fixed point of f^N. Therefore $f(a)$ and a are both fixed points of f^N. Since f^N has a unique fixed point, $f(a) = a$.

We now show that for any $x_0 \in X$ the points $f^n(x_0)$ converge to a as $n \to \infty$. Consider the iterates $f^n(x_0)$ as n runs through a congruence class modulo N. That is, pick $0 \le r \le N-1$ and look at the points $f^{Nk+r}(x_0)$ as $k \to \infty$. Since

$$f^{Nk+r}(x_0) = f^{Nk}(f^r(x_0)) = (f^N)^k(f^r(x_0)),$$

these points can be viewed (for each r) as iterates of f^N starting at the point $y_0 = f^r(x_0)$. Since f^N is a contraction, these iterates of f^N (from any initial point, such as y_0) must tend to a by the contraction mapping theorem. This limit is independent of the value of r in the range $\{1, \ldots, N-1\}$, so all N sequences $\{f^{Nk+r}(x_0)\}_{k\ge1}$ tend to a as $k \to \infty$. This shows

$$f^n(x_0) \to a \tag{3.1}$$

as $n \to \infty$. □

Example 3.2. The graphs of $y = e^{-x}$ and $y = x$ intersect once, so $e^{-a} = a$ for a unique real number a. However, the function $f(x) = e^{-x}$ is not a contraction on $\mathbf{R}$ (for instance, $|f(-2)-f(0)| \approx 6.38 > |-2-0|$), so the contraction mapping theorem itself does not justify finding the fixed point of f on $\mathbf{R}$ by iteration. But the second iterate $g(x) = f^2(x) = e^{-e^{-x}}$ is a contraction: by the mean-value theorem

$$g(x) - g(y) = g'(t)(x - y)$$

for some t between x and y, where $|g'(t)| = |e^{-e^{-t}}e^{-t}| = e^{-(t+e^{-t})} \le e^{-1}$ (since $t + e^{-t} \ge 1$ for all real t). Hence f^2 has contraction constant $1/e \approx .367 < 1$. By Theorem 3.1 the solution to $e^{-a} = a$ can be approximated by iteration of f starting with any real number. Iterating f enough times with $x_0 = 0$ suggests $a \approx .567$. To prove this approximation is correct, one can generalize the error estimates in Corollary 2.3 to apply to a function having an iterate as a contraction. Alternatively, check $f([0, 1]) \subset [0, 1]$ and f (not just f^2) is a contraction on this interval, even though it isn't on all of $\mathbf{R}$, so Corollary 2.3 can be applied directly to f by using $x_0 \in [0,1]$.

Remark 3.3. When f^N is a contraction, f^N is continuous, but that does not imply f is continuous. For example, let $f\colon [0,1] \to [0,1]$ by $f(x) = 0$ for $0 \leq x \leq 1/2$ and $f(x) = 1/2$ for $1/2 < x \leq 1$. Then $f(f(x)) = 0$ for all x, so f^2 is a contraction but f is discontinuous. The unique fixed point of f is $x = 0$ and trivially $f^2(x_0)$ is this fixed point for any $x_0 \in [0,1]$.

It was important in the proof of the contraction mapping theorem that the contraction constant c be strictly less than 1. That gave us control over the rate of convergence of $f^n(x_0)$ to the fixed point since $c^n \to 0$ as $n \to \infty$. If instead of f being a contraction we suppose $d(f(x), f(x')) < d(x, x')$ whenever $x \neq x'$ in X then we lose that control and indeed a fixed point need not exist.

Example 3.4. Let I be a closed interval in $\mathbf{R}$ and $f\colon I \to I$ be differentiable with $|f'(t)| < 1$ for all t. Then the mean-value theorem implies $|f(x) - f(x')| < |x - x'|$ for $x \neq x'$ in I. The following three functions all fit this condition, where $I = [1, \infty)$ in the first case and $I = \mathbf{R}$ in the second and third cases:

$$f(x) = x + \frac{1}{x}, \quad f(x) = \sqrt{x^2 + 1}, \quad f(x) = \log(1 + e^x).$$

In each case, $f(x) > x$, so none of these functions has a fixed point.

Despite such examples, there is a fixed-point theorem when $d(f(x), f(x')) < d(x, x')$ for all $x \neq x'$ provided the space is *compact*, which is not the case in the previous example.

Theorem 3.5. *Let X be a compact metric space. If $f\colon X \to X$ satisfies $d(f(x), f(x')) < d(x, x')$ when $x \neq x'$ in X, then f has a unique fixed point in X and the fixed point can be found as the limit of $f^n(x_0)$ as $n \to \infty$ for any $x_0 \in X$.*

Proof. To show f has at most one fixed point in X, suppose f has two fixed points $a \neq a'$. Then $d(a, a') = d(f(a), f(a')) < d(a, a')$. This is impossible, so $a = a'$.

To prove f actually has a fixed point, we will look at the function $F\colon X \to [0, \infty)$ given by $F(x) = d(x, f(x))$. This measures the distance between each point and its f-value. A fixed point of f is where F takes the value 0.

Since F is continuous and X is compact, the function F takes on a minimum value: there is an $a \in X$ such that $F(a) \leq F(x)$ for all $x \in X$. We'll show by contradiction that a is a fixed point for f. If $f(a) \neq a$ then the hypothesis about f in the theorem (taking $x = a$ and $x' = f(a)$) says

$$F(f(a)) = d(f(a), f(f(a))) < d(a, f(a)) = F(a),$$

which contradicts the minimality of $F(a)$ among all values of F. So $f(a) = a$ (and $F(a) = 0$).

Finally, we show for any $x_0 \in X$ that the sequence $x_n = f^n(x_0)$ converges to a as $n \to \infty$. This can't be done as in the proof of the contraction mapping theorem since we don't have the uniform control coming from the contraction constant. Instead we will exploit compactness.

If for some $k \geq 0$ we have $x_k = a$ then $x_{k+1} = f(x_k) = f(a) = a$, and more generally $x_n = a$ for all $n \geq k$, so $x_n \to a$ since the terms of the sequence equal a for all large n. Now we may assume instead that $x_n \neq a$ for all n. Then

$$0 < d(x_{n+1}, a) = d(f(x_n), f(a)) < d(x_n, a),$$

so the sequence of numbers $d(x_n, a)$ is decreasing and positive. Thus it has a limit $\ell = \lim_{n\to\infty} d(x_n, a) \geq 0$. We will show $\ell = 0$ (so $d(x_n, a) \to 0$, which means $x_n \to a$ in X). By compactness of X, the sequence $\{x_n\}$ has a convergent subsequence x_{n_i}, say $x_{n_i} \to y \in X$. Then, by continuity of f, $f(x_{n_i}) \to f(y)$, which says $x_{n_i+1} \to f(y)$ as $i \to \infty$. Since $d(x_n, a) \to \ell$ as $n \to \infty$, $d(x_{n_i}, a) \to \ell$ and $d(x_{n_i+1}, a) \to \ell$ as $i \to \infty$. By continuity of the metric, $d(x_{n_i}, a) \to d(y, a)$ and $d(x_{n_i+1}, a) = d(f(x_{n_i}), a) \to d(f(y), a)$. Having already shown these limits are ℓ,

$$d(y, a) = \ell = d(f(y), a) = d(f(y), f(a)). \tag{3.2}$$

If $y \neq a$ then $d(f(y), f(a)) < d(y, a)$, but this contradicts (3.2). So $y = a$, which means $\ell = d(y, a) = 0$. That shows $d(x_n, a) \to 0$ as $n \to \infty$. $\square$

Theorem 3.5 is due to Edelstein [5]. The proof of Theorem 3.5 does not yield any error estimate on the rate of convergence to the fixed point, since we proved convergence to the fixed point without having to make estimates along the way.

It is natural to wonder if the compactness of X might force f in Theorem 3.5 to be a contraction after all, so the usual contraction mapping theorem would apply. For instance, the ratios $d(f(x), f(x'))/d(x, x')$ for $x \neq x'$ are always less than 1, so they should be less than or equal to some definite constant $c < 1$ from compactness (a continuous real-valued function on a compact set achieves its supremum as a value). But this reasoning is bogus, because $d(f(x), f(x'))/d(x, x')$ is defined not on the compact set $X \times X$ but rather on $X \times X - \{(x, x) : x \in X\}$ where the diagonal is removed, and this is not compact (take a look at the special case $X = [0, 1]$, for instance). There is no way to show f in Theorem 3.5 has to be a contraction since there are examples where it isn't.

Example 3.6. Let $f\colon [0, 1] \to [0, 1]$ by $f(x) = \frac{1}{1+x}$, so $|f(x)-f(y)| = |x-y|/|(1+x)(1+y)|$. When $x \neq y$ we have $|f(x) - f(y)|/|x - y| = 1/(1 + x)(1 + y) < 1$, so $|f(x) - f(y)| < |x - y|$ and $|f(x) - f(y)|/|x - y|$ gets arbitrarily close to 1 when x and y are sufficiently close to 0. Therefore f is not a contraction on $[0, 1]$ with respect to the usual metric.

Theorem 3.5 says f has a unique fixed point in $[0, 1]$ and $f^n(x_0)$ tends to this point as $n \to \infty$ for any choice of x_0. Of course, it is easy to find the fixed point: $x = 1/(1 + x)$ in $[0, 1]$ at $x = (-1 + \sqrt{5})/2 \approx .61803$.

This example does not actually require Theorem 3.5. One way around it is to check that if $1/2 \leq x \leq 1$ then $1/2 \leq f(x) \leq 2/3$, so $f\colon [1/2, 1] \to [1/2, 1]$ with $\max_{x\in[1/2,1]} |f'(x)| = 4/9$. We could apply the contraction mapping theorem to f on $[1/2, 1]$ to find the fixed point. A second method is to check that $f^2(x) = (1 + x)/(2 + x)$ is a contraction on $[0, 1]$ (the derivative of $(1+x)/(2+x)$ on $[0, 1]$ has absolute value at most $1/4$), so we could apply the contraction mapping theorem to f^2.

Example 3.7. Following [1], let $f\colon [0, 1] \to [0, 1]$ by $f(x) = x/(1 + x)$. By the same calculation as in the previous example, $|f(x) - f(y)|/|x - y| = 1/(1 + x)(1 + y)$ for $x \neq y$, so the ratio $|f(x) - f(y)|/|x - y|$ for $x \neq y$ can be made arbitrarily close to 1 by taking x and y sufficiently close to 0. What makes this example different from the previous one is that, since 0 is now the fixed point, f does not restrict to a contraction on any neighborhood of its fixed point. Moreover, since $f^n(x) = x/(1+nx)$, for $x \neq y$ the ratio $|f^n(x)-f^n(y)|/|x-y| = 1/(1 + nx)(1 + ny)$ is arbitrarily close to 1 when x and y are sufficiently close to 0, so no iterate f^n is a contraction on any neighborhood of the fixed point in $[0, 1]$.

A similar example on $[0, 1]$ with 1 as the fixed point uses $g(x) = 1 - f(1 - x) = 1/(2 - x)$.

References

[1] D. G. Bennett and B. Fisher, *On a Fixed Point Theorem for Compact Metric Spaces*, Math. Magazine **47** (1974), 40–41.

[2] C. Bessaga, *On the Converse of the Banach Fixed-point Principle*, Colloq. Math. **7** (1959), 41–43.

[3] V. Bryant, "Metric Spaces: Iteration and Application," Cambridge Univ. Press, Cambridge, 1985.

[4] K. Deimling, "Nonlinear functional analysis," Springer–Verlag, Berlin, 1985.

[5] M. Edelstein, *On Fixed and Periodic Points Under Contractive Mappings*, J. London Math. Soc. **37** (1962), 74–79.

[6] M. Freeman, *The Inverse as a Fixed Point in Function Space*, Amer. Math. Monthly **83** (1976), 344–348.

[7] K. Goebel and W. A. Kirk, "Topics in metric fixed point theory," Cambridge Univ. Press, Cambridge, 1990.

[8] D. J. Higham and A. Taylor, The Sleekest Link Algorithm, `http://www.maths.strath.ac.uk/~aas96106/rep20_2003.pdf`.

[9] W. Rudin, "Principles of Mathematical Analysis," 3rd ed., McGraw–Hill, New York, 1976.

[10] E. Schechter, "Handbook of Analysis and its Applications," Academic Press, San Diego, 1997.

[11] E. Scheinerman, "Invitation to Dynamical Systems," Prentice-Hall, Upper Saddle River, NJ, 1995.

[12] W. Walter, *An Elementary Proof of the Cauchy–Kowalevsky Theorem*, Amer. Math. Monthly **92** (1985), 115–126.

[13] B. White, How Google Ranks Web Pages, `http://www.cfm.brown.edu/people.jansh/Courses/PageRank.pdf`.