**Section 2.6 (cont.) Properties of Real Functions**

Here we first study properties of functions from $\mathbf{R}$ to $\mathbf{R}$, making use of the additional structure we have in $\mathbf{R}$ as opposed to general metric spaces.

Let $f : X \to \mathbf{R}$ where $X \subseteq \mathbf{R}$. We say $f$ is *bounded above* if

$$f(X) = \{r \in \mathbf{R} : f(x) = r \text{ for some } x \in X\}$$

is bounded above. Similarly, we say $f$ is *bounded below* if $f(X)$ is bounded below. Finally, $f$ is *bounded* if $f$ is both bounded above and bounded below, that is, if $f(X)$ is a bounded set.

**Theorem 1 (Thm. 6.23, Extreme Value Theorem)** *Let $a, b \in \mathbf{R}$ with $a \le b$ and let $f : [a, b] \to \mathbf{R}$ be a continuous function. Then $f$ assumes its minimum and maximum on $[a, b]$. That is, if*

$$M = \sup_{t \in [a,b]} f(t) \qquad m = \inf_{t \in [a,b]} f(t)$$

*then $\exists t_M, t_m \in [a, b]$ such that $f(t_M) = M$ and $f(t_m) = m$.*

*In particular, $f$ is bounded above and below.*

**Proof:** Let

$$M = \sup\{f(t) : t \in [a, b]\}$$

If $M$ is finite, then for each $n$, we may choose $t_n \in [a, b]$ such that $M \ge f(t_n) \ge M - \frac{1}{n}$ (if we couldn't make such a choice, then $M - \frac{1}{n}$ would be an upper bound and $M$ would not be

the supremum). If $M$ is infinite, choose $t_n$ such that $f(t_n) \geq n$. By the Bolzano-Weierstrass

Theorem, $\{t_n\}$ contains a convergent subsequence $\{t_{n_k}\}$, with

$$\lim_{k \to \infty} t_{n_k} = t_0 \in [a, b]$$

Since $f$ is continuous,

$$
\begin{aligned}
f(t_0) &= \lim_{t \to t_0} f(t) \\
&= \lim_{k \to \infty} f(t_{n_k}) \\
&= M
\end{aligned}
$$

so $M$ is finite and

$$f(t_0) = M = \sup\{f(t) : t \in [a, b]\}$$

so $f$ attains its maximum and is bounded above.

The argument for the minimum is similar. ■

**Theorem 2 (Thm. 6.24, Intermediate Value Theorem)** *Suppose $f : [a, b] \to \mathbf{R}$ is continuous, and $f(a) < d < f(b)$. Then there exists $c \in (a, b)$ such that $f(c) = d$.*

**Proof:** We did a hands-on proof already. Now, we can simplify it a bit. Let

$$B = \{t \in [a, b] : f(t) < d\}$$

$a \in B$, so $B \neq \emptyset$. By the Supremum Property, $\sup B$ exists and is real so let $c = \sup B$. Since $a \in B$, $c \geq a$. $B \subseteq [a, b]$, so $c \leq b$. Therefore, $c \in [a, b]$. We claim that $f(c) = d$.

Let

$$t_n = \min\left\{c + \frac{1}{n}, b\right\} \geq c$$

Either $t_n > c$, in which case $t_n \notin B$, or $t_n = c$, in which case $t_n = b$ so $f(t_n) > d$, so again $t_n \notin B$; in either case, $f(t_n) \geq d$. Since $f$ is continuous at $c$, $f(c) = \lim_{n \to \infty} f(t_n) \geq d$ (Theorem 3.5 in de la Fuente).

Since $c = \sup B$, we may find $s_n \in B$ such that

$$c \geq s_n \geq c - \frac{1}{n}$$

Since $s_n \in B$, $f(s_n) < d$. Since $f$ is continuous at $c$, $f(c) = \lim_{n \to \infty} f(s_n) \leq d$ (Theorem 3.5 in de la Fuente).

Since $d \leq f(c) \leq d$, $f(c) = d$. Since $f(a) < d$ and $f(b) > d$, $a \neq c \neq b$, so $c \in (a, b)$. ∎

**Monotonic Functions:**

**Definition 3** A function $f : \mathbf{R} \to \mathbf{R}$ is *monotonically increasing* if

$$y > x \Rightarrow f(y) \geq f(x)$$

**Theorem 4 (Thm. 6.27)** *Let $a, b \in \mathbf{R}$ with $a < b$, and let $f : (a, b) \to \mathbf{R}$ be monotonically increasing. Then the one-sided limits*

$$f(t^+) = \lim_{u \to t^+} f(u)$$
$$f(t^-) = \lim_{u \to t^-} f(u)$$

*exist and are real numbers for all $t \in (a, b)$.*

**Proof:** This is analogous to the proof that a bounded monotone sequence converges. ∎
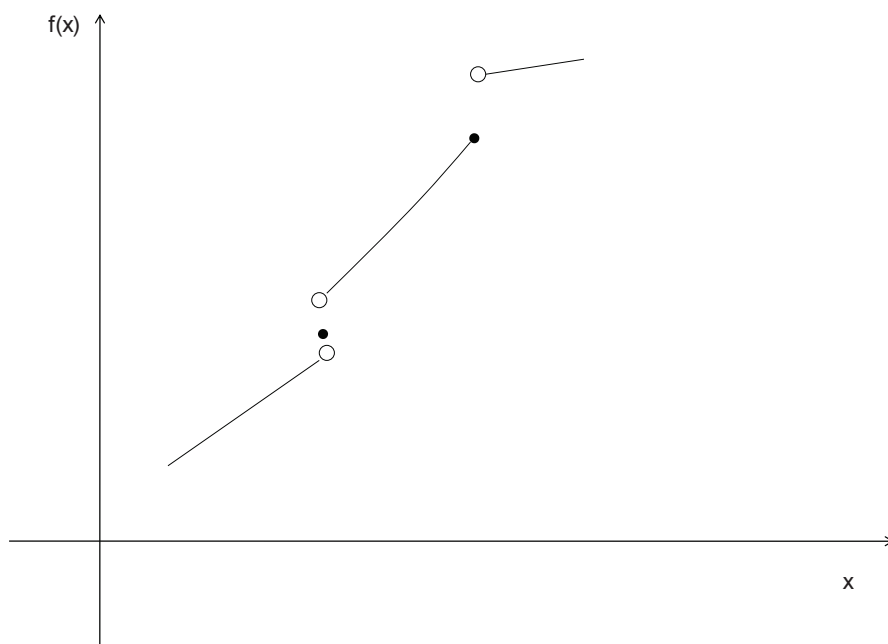
Figure 1: A monotonic function has only simple jump discontinuities.

We say that $f$ has a *simple jump discontinuity* at $t$ if the one-sided limits $f(t^-)$ and $f(t^+)$ both exist but $f$ is not continuous at $t$.[1] Note that there are two ways $f$ can have a simple jump discontinuity at $t$: either $f(t^+) \neq f(t^-)$, or $f(t^+) = f(t^-) \neq f(t)$.

The previous theorem says that monotonic functions have only simple jump discontinuities; note that monotonicity also implies that $f(t^-) \leq f(t) \leq f(t^+)$ for every $t$. See Figure 1.

---

[1]This is also called a discontinuity of the first kind; otherwise, if $f$ is not continuous at $t$, it is called a discontinuity of the second kind. An example of a discontinuity of the second kind is given by the indicator function of the rational numbers, that is $f(t) = 1$ if $t \in \mathbf{Q}$ and $f(t) = 0$ if $t \notin \mathbf{Q}$. $f$ is discontinuous at every $t$ and each discontinuity is of the second kind: $f(t^+), f(t^-)$ do not exist at every $t$.

Monotonic functions are particularly well-behaved.

**Theorem 5 (Thm. 6.28)** *Let $a, b \in \mathbf{R}$ with $a < b$, and let $f : (a, b) \to \mathbf{R}$ be monotonically increasing. Then*

$$D = \{t \in (a, b) : f \text{ is discontinuous at } t\}$$

*is finite (possibly empty) or countable.*

As this result shows, a monotonic function is continuous "almost everywhere".[2]

**Proof:** If $t \in D$, then $f(t^-) < f(t^+)$ (if the left- and right-hand limits agreed, then by monotonicity they would have to equal $f(t)$, so $f$ would be continuous at $t$). $\mathbf{Q}$ is dense in $\mathbf{R}$, that is, if $x, y \in \mathbf{R}$ and $x < y$ then $\exists r \in \mathbf{Q}$ such that $x < r < y$.[3] So for every $t \in D$ we may choose $r(t) \in \mathbf{Q}$ such that

$$f(t^-) < r(t) < f(t^+)$$

This defines a function $r : D \to \mathbf{Q}$.[4] Notice that

$$s > t \Rightarrow f(s^-) \geq f(t^+)$$

so

$$s > t, s, t \in D \Rightarrow r(s) > f(s^-) \geq f(t^+) > r(t)$$

---

[2]This statement is also formally correct, as a finite or countable set has Lebesgue measure zero. We return to formalize this in lectures 12 and 13.

[3]This can be shown as a consequence of the Archimedean property. More generally, denseness can be defined in an arbitrary metric space. For a metric space $(X, d)$, $E \subseteq X$ is *dense* if $\bar{E} = X$.

[4]Here we have used the Axiom of Choice, which says that if we can choose such a rational $r$ for each $t \in D$, then we can can choose a *function* $r : D \to \mathbf{Q}$.

so $r(s) \neq r(t)$. Therefore, $r$ is one-to-one, so it is a bijection from $D$ to a subset of $\mathbf{Q}$. Thus $D$ is finite or countable. ∎

**Section 2.7. Complete Metric Spaces, Contraction Mapping Theorem**

Roughly, a metric space is complete if "every sequence that ought to converge to a limit has a limit to converge to."

To begin to formalize this, recall that $x_n \to x$ means

$$\forall \varepsilon > 0 \; \exists N(\varepsilon/2) \text{ s.t. } n > N(\varepsilon/2) \Rightarrow d(x_n, x) < \frac{\varepsilon}{2}$$

Observe that if $n, m > N(\varepsilon/2)$, then

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

This motivates the following definition:

**Definition 6** A sequence $\{x_n\}$ in a metric space $(X, d)$ is *Cauchy* if

$$\forall \varepsilon > 0 \; \exists N(\varepsilon) \text{ s.t. } n, m > N(\varepsilon) \Rightarrow d(x_n, x_m) < \varepsilon$$

A Cauchy sequence is trying really hard to converge, but there may not be anything for it to converge to. Any sequence that does converge must be Cauchy, however, by the argument above.

**Theorem 7 (Thm. 7.2)** *Every convergent sequence in a metric space is Cauchy.*

**Proof:** We just did it.∎

**Example:** Let $X = (0, 1]$ and $d$ be the Euclidean metric. Let $x_n = \frac{1}{n}$. Then $x_n \to 0$ in $\mathbf{E}^1$, so $\{x_n\}$ is Cauchy in $\mathbf{E}^1$. But the Cauchy property depends only on the sequence and the metric $d$, not on the ambient metric space. So $\{x_n\}$ is Cauchy in $(X, d)$, but $\{x_n\}$ does not *converge* in $(X, d)$ because the point it is trying to converge to $(0)$ is not an element of $X$.

**Definition 8** A metric space $(X, d)$ is *complete* if every Cauchy sequence $\{x_n\} \subseteq X$ converges to a limit $x \in X$.

**Definition 9** A *Banach space* is a normed space that is complete in the metric generated by its norm.

**Example:** Consider the earlier example of $X = (0, 1]$ with $d$ the usual Euclidean metric. Since $x_n = \frac{1}{n}$ is Cauchy but does not converge, $((0, 1], d)$ is not complete.

**Example:** $\mathbf{Q}$ is not complete in the Euclidean metric. To see this, let

$$x_n = \frac{\lfloor 10^n \sqrt{2} \rfloor}{10^n}$$

where as before, $\lfloor y \rfloor$ is the greatest integer less than or equal to $y$; $x_n$ is just equal to the decimal expansion of $\sqrt{2}$ to $n$ digits past the decimal point. Clearly, $x_n$ is rational. $|x_n - \sqrt{2}| \leq 10^{-n}$, so $x_n \to \sqrt{2}$ in $\mathbf{E}^1$, so $\{x_n\}$ is Cauchy in $\mathbf{E}^1$, hence Cauchy in $\mathbf{Q}$; since $\sqrt{2} \notin \mathbf{Q}$, $\{x_n\}$ is not convergent in $\mathbf{Q}$, so $\mathbf{Q}$ is not complete.

**Theorem 10 (Thm. 7.10)** $\mathbf{R}$ *is complete with the usual metric (so $\mathbf{E}^1$ is a Banach space).*

**Proof:**[5] Suppose $\{x_n\}$ is a Cauchy sequence in $\mathbf{R}$. Fix $\varepsilon > 0$. Find $N(\varepsilon/2)$ such that

$$n, m > N(\varepsilon/2) \Rightarrow |x_n - x_m| < \frac{\varepsilon}{2}$$

Let

$$\alpha_n = \sup\{x_k : k \geq n\}$$

$$\beta_n = \inf\{x_k : k \geq n\}$$

Fix $m > N(\varepsilon/2)$. Then

$$k \geq m \Rightarrow k > N(\varepsilon/2) \Rightarrow x_k < x_m + \frac{\varepsilon}{2}$$

$$\Rightarrow \alpha_m = \sup\{x_k : k \geq m\} \leq x_m + \frac{\varepsilon}{2}$$

Since $\alpha_m < \infty$,

$$\limsup x_n = \lim_{n \to \infty} \alpha_n \leq \alpha_m \leq x_m + \frac{\varepsilon}{2}$$

since the sequence $\{\alpha_n\}$ is decreasing. Similarly,

$$\liminf x_n \geq x_m - \frac{\varepsilon}{2}$$

Therefore,

$$0 \leq \limsup_{n \to \infty} x_n - \liminf_{n \to \infty} x_n \leq \varepsilon$$

Since $\varepsilon$ is arbitrary,

$$\limsup_{n \to \infty} x_n = \liminf_{n \to \infty} x_n \in \mathbf{R}$$

so $\lim_{n \to \infty} x_n$ exists and is real. Thus $\{x_n\}$ is convergent. $\blacksquare$

**Theorem 11 (Thm. 7.11)** $\mathbf{E}^n$ *is complete for every* $n \in \mathbf{N}$.

---

[5]This proof is different from the one in de la Fuente.

**Proof:** See de la Fuente. ∎

**Theorem 12 (Thm. 7.9)** *Suppose $(X, d)$ is a complete metric space and $Y \subseteq X$. Then $(Y, d) = (Y, d|_Y)$ is complete if and only if $Y$ is a closed subset of $X$.*

**Proof:** Suppose $(Y, d)$ is complete. We need to show that $Y$ is closed. Consider a sequence $\{y_n\} \subseteq Y$ such that $y_n \to_{(X,d)} x \in X$. Then $\{y_n\}$ is Cauchy in $X$, hence Cauchy in $Y$; since $Y$ is complete, $y_n \to_{(Y,d)} y$ for some $y \in Y$. Therefore, $y_n \to_{(X,d)} y$. By uniqueness of limits, $y = x$, so $x \in Y$. Thus $Y$ is closed.

Conversely, suppose $Y$ is closed. We need to show that $Y$ is complete. Let $\{y_n\}$ be a Cauchy sequence in $Y$. Then $\{y_n\}$ is Cauchy in $X$, hence convergent, so $y_n \to_{(X,d)} x$ for some $x \in X$. Since $Y$ is closed, $x \in Y$, so $y_n \to_{(Y,d)} x \in Y$. Thus $Y$ is complete. ∎

**Theorem 13 (Thm. 7.12)** *Given $X \subseteq \mathbf{R}^n$, let $C(X)$ be the set of bounded continuous functions from $X$ to $\mathbf{R}$ with*

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}$$

*Then $C(X)$ is a Banach space.*

**Contractions:**

**Definition 14** Let $(X, d)$ be a nonempty complete metric space. An *operator* is a function $T : X \to X$. An operator $T$ is a *contraction of modulus $\beta$* if $\beta < 1$ and

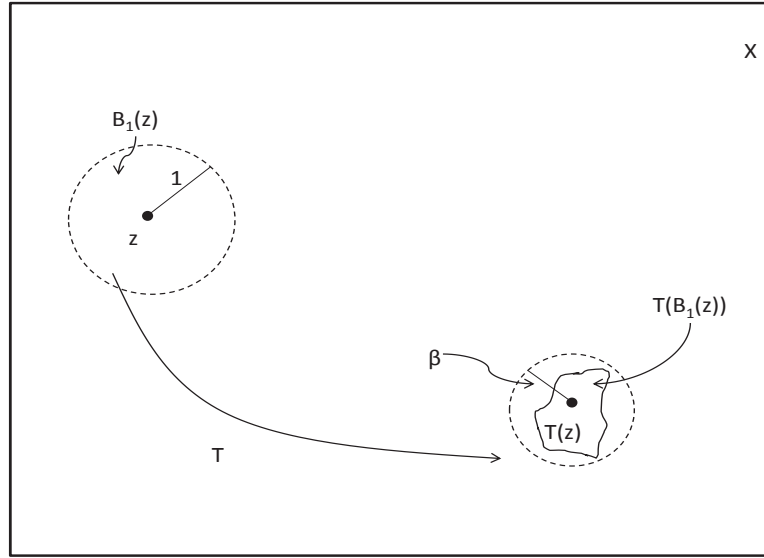$$d(T(x), T(y)) \leq \beta d(x, y) \quad \forall x, y \in X$$

Figure 2: A contraction mapping

A contraction shrinks distances by a *uniform* factor $\beta < 1$. See Figure 2.

**Theorem 15** *Every contraction is uniformly continuous.*

**Proof:** Let $\delta = \frac{\varepsilon}{\beta}$. ∎

In fact, note that a contraction is Lipschitz continuous with Lipschitz constant $\beta < 1$.
This also shows that contractions are uniformly continuous.

**Definition 16** A *fixed point* of an operator $T$ is element $x^* \in X$ such that $T(x^*) = x^*$.

**Theorem 17 (Thm. 7.16, Contraction Mapping Theorem)** *Let $(X, d)$ be a nonempty
complete metric space and $T : X \to X$ a contraction with modulus $\beta < 1$. Then*

1. *$T$ has a unique fixed point $x^*$.*

2. *For every $x_0 \in X$, the sequence defined by*

$$x_1 = T(x_0)$$

$$x_2 = T(x_1) = T(T(x_0)) = T^2(x_0)$$

$$\vdots$$

$$x_{n+1} = T(x_n) = T^n(x_0)$$

*converges to $x^*$.*

Note that the theorem asserts both the existence and uniqueness of the fixed point, as well as giving an algorithm to find the fixed point of a contraction. Later in the course we will discuss more general fixed point theorems which, in contrast, only guarantee existence. See Figure 3.

**Proof:** Define the sequence $\{x_n\}$ as above by first fixing $x_0 \in X$ and then letting $x_n = T(x_{n-1}) = T^n(x_0)$ for $n = 1, 2, \ldots$, where $T^n = T \circ T \circ \ldots \circ T$ is the $n$-fold iteration of $T$.

We first show that $\{x_n\}$ is Cauchy, and hence converges to a limit $x$. Then

$$d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1}))$$

$$\leq \beta d(x_n, x_{n-1})$$

$$\leq \beta^2 d(x_{n-1}, x_{n-2})$$

$$\vdots$$

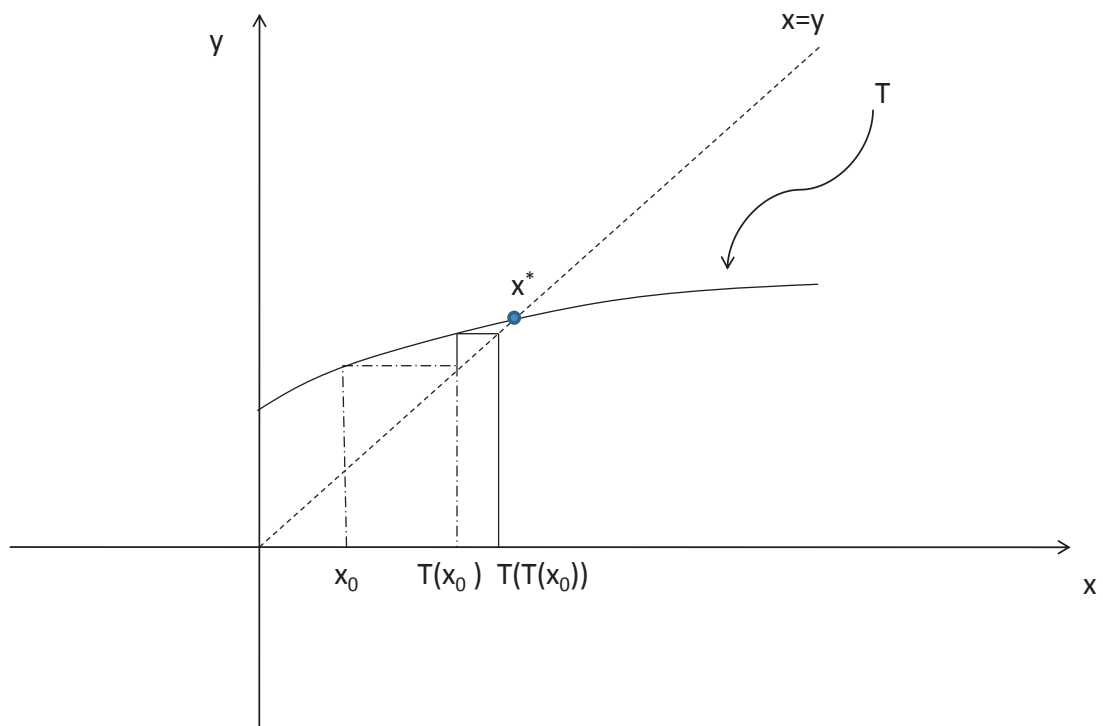$$\leq \beta^n d(x_1, x_0)$$

11

Figure 3: The Contraction Mapping Theorem

Then for any $n > m$,

$$
\begin{aligned}
d(x_n, x_m) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \cdots + d(x_{m+1}, x_m) \\
&\leq (\beta^{n-1} + \beta^{n-2} + \cdots + \beta^m) d(x_1, x_0) \\
&= d(x_1, x_0) \sum_{\ell=m}^{n-1} \beta^\ell \\
&< d(x_1, x_0) \sum_{\ell=m}^{\infty} \beta^\ell \\[2em]
&= \frac{\beta^m}{1-\beta} d(x_1, x_0) \quad \text{(sum of a geometric series)}
\end{aligned}
$$

Fix $\varepsilon > 0$. Since $\frac{\beta^m}{1-\beta} d(x_1, x_0) \to 0$ as $m \to \infty$, choose $N(\varepsilon)$ such that for any $m > N(\varepsilon)$,

$\frac{\beta^m}{1-\beta} d(x_1, x_0) < \varepsilon$. Then for $n, m > N(\varepsilon)$,

$$
d(x_n, x_m) \leq \frac{\beta^m}{1-\beta} d(x_1, x_0) < \varepsilon
$$

Therefore, $\{x_n\}$ is Cauchy. Since $(X, d)$ is complete, $x_n \to x^*$ for some $x^* \in X$.

Next, we show that $x^*$ is a fixed point of $T$.

$$
\begin{aligned}
T(x^*) &= T\left(\lim_{n \to \infty} x_n\right) \\
&= \lim_{n \to \infty} T(x_n) \text{ since } T \text{ is continuous} \\
&= \lim_{n \to \infty} x_{n+1} \\
&= x^*
\end{aligned}
$$

so $x^*$ is a fixed point of $T$.

Finally, we show that there is at most one fixed point. Suppose $x^*$ and $y^*$ are both fixed points of $T$, so $T(x^*) = x^*$ and $T(y^*) = y^*$. Then

$$
d(x^*, y^*) = d(T(x^*), T(y^*))
$$

13

$$\leq \quad \beta d(x^*, y^*)$$

$$\Rightarrow (1 - \beta)d(x^*, y^*) \quad \leq \quad 0$$

$$\Rightarrow d(x^*, y^*) \quad \leq \quad 0$$

So $d(x^*, y^*) = 0$, which implies $x^* = y^*$. ∎

**Theorem 18 (Thm. 7.18', Continuous Dependence on Parameters)** *Let $(X, d)$ and $(\Omega, \rho)$ be two metric spaces and $T : X \times \Omega \to X$. For each $\omega \in \Omega$ let $T_\omega : X \to X$ be defined by*

$$T_\omega(x) = T(x, \omega)$$

*Suppose $(X, d)$ is complete, $T$ is continuous in $\omega$, that is $T(x, \cdot) : \Omega \to X$ is continuous for each $x \in X$, and $\exists \beta < 1$ such that $T_\omega$ is a contraction of modulus $\beta$ $\forall \omega \in \Omega$. Then the fixed point function $x^* : \Omega \to X$ defined by*

$$x^*(\omega) = T_\omega(x^*(\omega))$$

*is continuous.*

**Remark:** See the comments in the *Corrections* handout. Theorem 7.18 in de la Fuente only requires that each map $T_\omega$ be a contraction of modulus $\beta_\omega < 1$. However, his proof assumes that there is a single $\beta < 1$ such that each $T_\omega$ is a contraction of modulus $\beta$. I do not know whether de la Fuente's Theorem 7.18 is correct as stated.

An important result due to Blackwell gives a set of sufficient conditions for an operator to be a contraction that is particularly useful in dynamic programming problems.

Let $X$ be a set, and let $B(X)$ be the set of all bounded functions from $X$ to $\mathbf{R}$. Then $(B(X), \|\cdot\|_\infty)$ is a normed vector space.

Notice that below we use shorthand notation that identifies a constant function with its constant value in $\mathbf{R}$, that is, we write interchangeably $a \in \mathbf{R}$ and $a : X \to \mathbf{R}$ to denote the function such that $a(x) = a \ \forall x \in X$.

**Theorem 19 (Blackwell's Sufficient Conditions)** *Consider $B(X)$ with the sup norm $\|\cdot\|_\infty$. Let $T : B(X) \to B(X)$ be an operator satisfying*

1. *(monotonicity) $f(x) \le g(x) \ \forall x \in X \Rightarrow (Tf)(x) \le (Tg)(x) \ \forall x \in X$*

2. *(discounting) $\exists \beta \in (0,1)$ such that for every $a \ge 0$ and $x \in X$,*

$$(T(f+a))(x) \le (Tf)(x) + \beta a$$

*Then $T$ is a contraction with modulus $\beta$.*

**Proof:** Fix $f, g \in B(X)$. By the definition of the sup norm,

$$f(x) \le g(x) + \|f - g\|_\infty \ \forall x \in X$$

Then

$$(Tf)(x) \le (T(g + \|f - g\|_\infty))(x) \le (Tg)(x) + \beta\|f - g\|_\infty \ \ \forall x \in X$$

where the first inequality above follows from monotonicity, and the second from discounting. Thus

$$(Tf)(x) - (Tg)(x) \le \beta\|f - g\|_\infty \ \ \forall x \in X$$

Reversing the roles of $f$ and $g$ above gives

$$(Tg)(x) - (Tf)(x) \leq \beta \|f - g\|_\infty \quad \forall x \in X$$

Thus

$$\|T(f) - T(g)\|_\infty \leq \beta \|f - g\|_\infty$$

Thus $T$ is a contraction with modulus $\beta$ ∎