1. EXAMPLES OF METRICS

Notation: R and N, respectively, are the sets of real and natural numbers. The Cartesian product $A \times B$ of sets A and B is the set of all ordered pairs (a,b) with a from A and b from B.

Definitions: A metric $d$ on a set S is a function $d: S \times S \to \text{Reals}$ satisfying, for all $x, y, z \in S$,

(i) $d(x, y) \geq 0$,
(ii) $d(x, y) = 0$ if and only if $x = y$,
(iii) $d(x, y) = d(y, x)$ and (importantly);
(iv) The Triangle Inequality: $d(x, y) \leq d(x, z) + d(z, y)$.

The open ball of radius $r$ and center $a$ is $B(a, r) = \{ x \in S : d(x, a) < r \}$ and the closed ball with the same center and radius is $CB(a, r) = \{ x \in S : d(x, a) \leq r \}$.

Example (Usual Metric on Reals): The absolute value function $|x| = \max\{x, -x\}$ give rise to a natural metric $d(x, y) = |x - y|$. To check, say, the triangle inequality, note that for each sign $s = \pm 1$,

$s(x - y) = s(x - z) + s(z - y) \leq |x - z| + |z - y|$. In the natural metric $B(a, r)$ is the open interval $(a - r, a + r)$ and $CB(a, r)$ is the closed interval $[a - r, a + r]$.

Example (Discrete Metric): For S non-empty define $d(x, y) = 0$ iff $x = y$, and $d(x, y) = 1$ otherwise. Properties (i) - (iii) in the metric definition are immediate. For the triangle inequality let $x, y, z \in S$ and consider two cases. First, if $x = y$ the $d(x, y) = 0 \leq d(x, z) + d(z, y)$. Second, if $x \neq y$ then either $x \neq z$ or $z \neq y$. Thus in this case either $d(x, z) = 1$ or $d(z, y) = 1$, implying $d(x, y) = 1 \leq d(x, z) + d(z, y)$. In the discrete metric $B(a, r) = \{ a \}$ when $0 < r \leq 1$, and $B(a, r) = S$ when $1 < r$.

Example (Railroad Metric): On the reals define $d(x, y) = 0$ iff $x = y$, and $d(x, y) = |x| + |y|$ otherwise. Properties (i) - (iii) in the metric definition are immediate. For the triangle inequality first note that $d(u, v)$ for all $u$ and $v$, $|u - v| \leq d(u, v)$. In fact, when $u = v$, $d(u, v) = 0 \leq |u - v|$, and when $u \neq v$

$|u - v| \leq |u| + |v| = d(u, v)$. Turning to the triangle inequality let $x, y$ and $z$ be any reals. If $x = y$ then $d(x, y) = 0 \leq d(x, z) + d(z, y)$. On the other hand if $x \neq y$ then either $x \neq z$ or $z \neq y$. For convenience suppose $z \neq y$. In that case $d(x, y) = |(x - z) + (z - y)| \leq |x - z| + |z - y| \leq d(x, z) + d(z, y)$, the last inequality from (*).

Example: For fixed $a > 0$ the railroad metric has these open balls.

(1) If $0 < r \leq a$ then $B(a, r) = \{ a \}$.
(2) If $a < r$ then $B(a, r) = (a - r, r + a) \cup \{ a \}$.

Here are two special instances of (2);

(3) If $a < r < 2a$ then $B(a, r) = (a - r, r + a) \cup \{ a \} = B(0, r - a) \cup B(a, a)$ is the union of two disjoint open balls.
(4) If $r = 2a$ then $B(a, r) = (-a, a]$. 
Notation: \( A^B \) is the set of all functions \( f : B \to A \). In particular \( A^N \) is the set of all functions \( f : N \to A \), that is, the set of all sequences \( f = (f(1), f(2), \ldots, f(n), \ldots) \) of elements of \( A \). In case \( B = \{1, 2, \ldots, n\} \), \( A^B \) is just written \( A^n \) and is the set of all \( n \)-tuples \( x = (x(1), x(2), \ldots, x(n)) \) of elements from \( A \).

Three Famous Metric Spaces: For \( f, g \in \mathbb{R}^N \) real sequences write
\[
d(f, g) = \sum_{n \geq 1} 2^{-n} \min \{ |f(n) - g(n)|, 1 \}.
\]
A result of the first problem set is that this is a metric on \( \mathbb{R}^N \). \( \mathbb{R}^N \) with this metric will be called Frechet's Space. Notice that if \( f \) and \( g \) have all their values in the closed interval \([0, 1]\) then
\[
d(f, g) = \sum_{n \geq 1} 2^{-n} |f(n) - g(n)|.
\]
\([0, 1]^N\) with this metric is the Hilbert Cube. \( \{0, 1\}^N\) with this metric is the standard Cantor Set.

Example (Hamming Metric): The Hamming metric is defined on \( \{0, 1\}^N \) by \( d(x, y) = \text{card}(N(x, y)) \), where \( N(x, y) = \{ k : x(k) \neq y(k) \} \). To check the triangle inequality for \( x, y \) and \( z \), consider the case \( x \neq y \).
Notice \( N(x, y) \subseteq N(x, z) \cup N(z, y) \) because if \( x(k) \neq y(k) \) then either \( x(k) \neq z(k) \) or \( z(k) \neq y(k) \).
Estimating cardinalities, \( \text{card}(N(x, y)) \leq \text{card}(N(x, z) \cup N(z, y)) \leq \text{card}(N(x, z)) + \text{card}(N(z, y)) \).

2. NORMS AND INNER PRODUCTS

Definitions: Let \( V \) be a real vector space. A norm on \( V \) is a function from \( V \) into the reals so that, for all \( x \) and \( z \) in \( V \) and scalars \( c \), (i) \( \|x\| \geq 0 \); (ii) \( \|x\| = 0 \) iff \( x = 0 \); (iii) \( \|cx\| = |c| \|x\| \) and; (iv, The Triangle Inequality) \( \|x + z\| \leq \|x\| + \|z\| \). The simplest and most basic example is absolute value on the reals. A semi-norm is a function satisfying only (i), (iii) and (iv).

Theorem: For \( \|\cdot\| \) a norm, \( d(x, z) = \|x - z\| \) is a metric and \( \|-x\| = \|x\| \) for each vector.
Pf: (0) From the identity \( -x = (-1)x \) and norm property (iii), \( \|-x\| = \|x\| \).
(1) \( 0 < \|x - z\| = d(x, z) \) by norm property (i).
(2) \( 0 = d(x, z) = \|x - z\| \Rightarrow x - z = 0 \Rightarrow x = z \), the first implication by norm property (ii).
(3) Using (0), \( d(x, z) = \|x - z\| = \|-z - x\| = \|z - x\| = d(z, x) \).
(4) By the norm's triangle inequality
\[
d(x, y) = \|x - y\| = \|(x - z) + (z - y)\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y) \).

One way to generate norms and metrics is via inner products.
Definition: Let \( V \) be a real vector space. An inner product on \( V \) is a function from \( V \times V \) into the reals so that, for all \( x, y \) and \( z \) in \( V \) and scalars \( c \),

(i) \( 0 \leq \langle x, x \rangle \);  (ii) \( 0 = \langle x, x \rangle \) iff \( x = 0 \);  (iii) \( \langle x, y \rangle = \langle y, x \rangle \);
(iv) \( \langle cx, y \rangle = c \langle x, y \rangle \) and;  (v) \( \langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle \).

Properties (iii), (iv) and (v) define a symmetric bilinear form. Bilinearity refers to linearity in the second variable, which holds because

\[
\langle y, cx \rangle = c \langle x, y \rangle \quad \text{and} \quad \langle y, x + z \rangle = \langle y, x \rangle + \langle y, z \rangle.
\]

A symmetric bilinear form satisfying (i) is called positive. If in addition (ii) holds, the form is called positive definite.

Example (Standard Inner Product): \( \mathbb{R}^n \) is a vector space under the usual coordinate operations. The standard inner product is \( \langle x, z \rangle = \sum_{k=1}^{n} x(k)z(k) \).

Theorem (Cauchy-Schwarz Inequality) Let \( \langle \cdot, \cdot \rangle \) be a positive symmetric bilinear form on \( V \), and write \( \| x \| = \sqrt{\langle x, x \rangle} \). For any \( x \) and \( y \) in \( V \), \( \langle x, y \rangle \leq \| x \| \| y \| \).

\textbf{Pf:} A consequence of the quadratic formula is that if \( 0 \leq at^2 + bt + c \) for all real \( t \), then \( b^2 - 4ac \leq 0 \). Now for each \( t \)

\[
\langle tx + y, tx + y \rangle = t \langle x, tx + y \rangle + \langle y, tx + y \rangle = t^2 \langle x, x \rangle + t \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = t^2 \| x \|^2 + 2t \langle x, y \rangle + \| y \|^2.
\]

Since the form is positive the last expression is non-negative, leading to \( 4 \langle x, y \rangle^2 - 4 \| x \|^2 \| y \|^2 \leq 0 \).

\textbf{Theorem:} If \( \langle \cdot, \cdot \rangle \) is a positive symmetric bilinear form on \( V \), then \( \| x \| = \sqrt{\langle x, x \rangle} \) is a semi-norm. If in addition the form is positive definite then \( \| x \| = \sqrt{\langle x, x \rangle} \) is a norm on \( V \).

\textbf{Pf:} To check (iii), \( \| c x \|^2 = \langle c x, c x \rangle = c^2 \langle x, x \rangle = c^2 \| x \|^2 \).

To check (iv), taking \( t = 1 \) in equality (**),

\[
\| x + y \|^2 = \langle x + y, x + y \rangle = \| x \|^2 + 2 \langle x, y \rangle + \| y \|^2 \leq \| x \|^2 + 2 \| x \| \| y \| + \| y \|^2 = (\| x \| + \| y \|)^2.
\]

Finally, if the form is positive definite, \( 0 = \| x \|^2 = \langle x, x \rangle \Rightarrow x = 0 \).

\textbf{Corollary:} \( \| x \|_2 = \left( \sum_{k=1}^{n} x(k)^2 \right)^{1/2} \) is a norm on \( \mathbb{R}^n \), called the Euclidean norm.
Corollary: For a norm determined by an inner product, \( \| x \| = \max \{ \langle x, z \rangle : \| z \| = 1 \} \). For \( x \) non-zero the maximum is attained at \( z = \| x \|^{-1} x \).

Pf: The equality is trivial when \( x = 0 \). For \( \| z \| = 1 \), Cauchy-Schwartz gives \( \langle x, z \rangle \leq \| x \| \| z \| = \| x \| \) and so \( \sup \{ \langle x, z \rangle : \| z \| = 1 \} \leq \| x \| \). For \( x \neq 0 \), \( \| x \|^{-1} x = \| x \|^{-1} \| x \| = 1 \) and
\[ \langle x, x \rangle = \| x \|^{-1} x \geq \| x \|^{-1} \langle x, x \rangle = \| x \|^{-1} \| x \| = 1 \], meaning that the supremum is attained when \( z = \| x \|^{-1} x \).

A geometric re-statement of the last corollary involves half-spaces. For \( z \in \mathbb{R}^n \) the set \( H_z = \{ x : \langle x, z \rangle \leq 1 \} \) is a closed half-space determined by \( z \). For instance if \( n = 2 \), \( z = (a, b) \) and a typical vector is \( x = (s, t) \), then \( H_z \) is the half-plane in the \( (s, t) \)-plane that contains the origin and is bounded by the line \( as + bt = 1 \).

Corollary: For a norm determined by an inner product, \( B(0,1) = \cap \{ H_z : \| z \| = 1 \} \).

We’ll close this section with two examples of infinite dimensional inner product spaces. Temporarily denote by \( R[a,b] \) the vector space of all bounded Riemann integrable functions on \([a,b]\). An advanced calculus result tells us that for \( f, g \in R[a,b] \), the pointwise product \( fg \) is also in \( R[a,b] \) and consequently the integral \( \langle f, g \rangle = \int_a^b f(s)g(s)ds \) is meaningful. Using elementary properties of the integral establishes that it defines a positive bilinear form with associated semi-norm \( \| f \|_2 = \left( \int_a^b (f(s))^2 ds \right)^{1/2} \). This form is not positive definite on \( R[a,b] \) but something can be said about functions with \( \langle f, f \rangle = 0 \).

Theorem: If \( \langle f, f \rangle = \int_a^b (f(s))^2 ds = 0 \) then \( f(u) = 0 \) at each point of continuity of \( f \) in \([a,b]\).

Pf: For each point \( u \) in \([a,b]\), \( 0 \leq \int_a^u (f(s))^2 ds \leq \int_a^b (f(s))^2 ds = 0 \), so \( F(t) = \int_a^t (f(s))^2 ds \) is zero everywhere on \([a,b]\). By the Fundamental Theorem \( F'(u) = f(u) \) at each point of continuity of \( f \).

Corollary: Let \( C[a,b] \) be the set of all continuous real functions on \([a,b]\). On \( C[a,b] \)
\[ \langle f, g \rangle = \int_a^b (f(s)g(s)ds \) is an inner product and \( \| f \|_2 = \left( \int_a^b (f(s))^2 ds \right)^{1/2} \) is a norm.

Example (A Separable Hilbert Space): Let \( l_2 \) be the set of all real sequences \( x = (x(1), x(2), ..., x(n), ...) \) for which \( \sum_{k \geq 1} x(k)^2 < \infty \). Under the coordinate operations \( l_2 \) is a vector space with a natural inner product and associated norm, viz,
\[ \langle x, z \rangle = \sum_{k \geq 1} x(k)z(k) \quad \text{and} \quad \| x \|_2 = \left( \sum_{k \geq 1} x(k)^2 \right)^{1/2} \]
Pf: Verifying the last example involves several steps. We'll check two and leave the rest as exercises.

(1) If \(x\) and \(y\) are in \(l_2\) then so in their pointwise sum \(x+y\). Why? For each index \(k\)
\[
0 \leq (x(k)-y(k))^2 \leq (x(k)-y(k))^2 + (x(k)+y(k))^2 = 2x(k)^2 + 2y(k)^2
\]
so that \(\sum_{k \geq 1} [x(k) + y(k)]^2\) converges by comparison with \(\sum_{k \geq 1} [x(k)^2 + y(k)^2]\).

(2) If \(x\) and \(z\) are in \(l_2\) then \(\sum_{k \geq 1} x(k)z(k)\) is absolutely convergent. Why? Fix \(n\).

Applying the Cauchy-Schwartz Inequality in \(\mathbb{R}^n\) with the \(n\)-tuples
\[
x_n = (|x(1)|, |x(2)|, ..., |x(n)|) \text{ and } z_n = (|z(1)|, |z(2)|, ..., |z(n)|)
\]
shows
\[
\sum_{1 \leq k \leq n} |x(k)z(k)| \leq \left( \sum_{1 \leq k \leq n} x(k)^2 \right)^{1/2} \left( \sum_{1 \leq k \leq n} z(k)^2 \right)^{1/2} \leq \|x\|_2 \|z\|_2.
\]
Since the partial sums \(\sum_{1 \leq k \leq n} |x(k)z(k)|\) are bounded above, \(\sum_{k \geq 1} x(k)z(k)\) converges absolutely.

3. CONVERGENCE AND COMPLETENESS

Definitions: Let \((S,d)\) be a metric space. A sequence \((f_n)\) in \(S\) converges to a point \(f\) in \(S\) iff \(\forall \varepsilon > 0 \exists m \in \mathbb{N}\) so that \(n \geq m \Rightarrow d(f_n,f) < \varepsilon\). Convergence to \(f\) is denoted by writing \(f_n \to f\) or \(\lim f_n = f\).

Convergence to \(f\) is equivalent to each of these statements.

(i) For any open ball \(B\) centered at \(f\), \(\exists m \in \mathbb{N}\) so that \(n \geq m \Rightarrow f_n \in B\).

(ii) The real sequence \(d(f_n,f)\) converges to zero, i.e., \(d(f_n,f) \to 0\).

Example: In Frechet space \(\mathbb{R}^n\) if \(f_n = (1,2,3,...,n,1,1,...\text{etc...})\) for each \(n\) and \(f = (1,2,3,...,n,n+1,n+2,...\text{etc...})\) then \(f_n \to f\).

Pf. For this example \(d(f_n,f)\) can be calculated exactly. \(f(k) - f_n(k) = 0\) for \(1 \leq k \leq n\), and \(f(k) - f_n(k) = k-1\) for \(k > n\). Thus \(\min \{|f(k) - f_n(k)|, 1\} = 0\) for \(1 \leq k \leq n\) and \(\min \{|f(k) - f_n(k)|, 1\} = 1\) for \(k > n\), giving
\[
d(f,f_n) = \sum_{k \geq 1} 2^{-k} \min \{|f(k) - f_n(k)|, 1\} = \sum_{k > n} 2^{-k} = 2^{-n}.
\]

Example: In \(C[0,1]\) with the 2-norm the sequence \(f_n(x) = e^{-nx}\) converges to the constantly zero function. However note \(f_n(0) = 1\) for all \(n\).
Pf. \[ \| f_n - 0 \|_2^2 = \int_0^1 e^{-2ns} \, ds = \frac{1}{2n} \left[ 1 - e^{-2n} \right] \to 0 \]

Definitions: Let \((S,d)\) be a metric space and \((f_n)\) be a sequence in \(S\).

(a) \((f_n)\) is Cauchy iff \(\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \text{ so that } n, k \geq m \Rightarrow d(f_n, f_k) < \varepsilon\).

(b) \(S\) is complete iff each Cauchy sequence in \(S\) converges to some point in \(S\).

(c) A set \(E\) in \(S\) is bounded iff \(E\) is contained in some open ball.

(d) \((f_n)\) is bounded iff its set of values \(\{ f_1, f_2, \ldots, f_n, \ldots \}\) is bounded in \(S\).

The notions of "Cauchy sequence" and "bounded" depend heavily on the particular metric.

Example: (1) \(d(x, y) = |1/x - 1/y|\) define a new metric on \(S = (0,1)\). Most of the defining properties for this \(d\)-metric are apparent. The \(d\)-triangle inequality is the usual trick:
\[ |1/x - 1/y| = |(1/x - 1/z) + (1/z - 1/y)| \leq |1/x - 1/z| + |1/z - 1/y| \]

(2) \((S, d)\) has the same convergent sequences as \((0,1)\) with the usual metric. Why? When each \(a_n \) and \(a\) are in \((0,1)\), \(a_n \to a\) implies \(1/a_n \to 1/a\), which in turn implies \(a_n = 1/(1/a_n) \to 1/(1/a) = a\).

(3) \(w_n = 1/n\) is Cauchy in the usual metric but isn't \(d\)-Cauchy, because \(d(w_{n+1}, w_n) = 1\) each \(n\).

(4) \(S\) is bounded in the usual metric but \(S\) isn't \(d\)-bounded. Why? Consider any \(d\)-open ball \(B_d(a, r)\) with \(a\) in \((0,1)\) and \(r > 0\). For large \(n\) \(d(a, a/n) = (n - 1)/n > r\), so \(a/n \notin S\) but \(a/n \notin B_d(a, r)\).

Theorem: In a metric space a convergent sequence is Cauchy, and a Cauchy sequence is bounded.

Pf: Suppose that \(f_n \to f\), let \(\varepsilon > 0\) and find \(m\) so that \(n \geq m \Rightarrow d(f_n, f) < \varepsilon/2\). Clearly \(n, k \geq m \Rightarrow d(f_n, f_k) \leq d(f_n, f) + d(f, f_k) < \varepsilon\). For the second assertion suppose that \((f_n)\) is Cauchy.

Find \(m\) so that \(n, k \geq m \Rightarrow d(f_n, f_k) < 1\). If \(r = 2 + \sum_{n=1}^{m} d(f_n, f_m)\) then \(d(f_n, f_m) < r\) for all \(n\).

Definition: The sign or signum function is defined by \(\text{sgn}(x) = 1\) if \(x \geq 0\), and \(\text{sgn}(x) = -1\) otherwise.

Example: Consider \(C[-1,1]\) with the 2-norm \(\| f \|_2 = \left( \int_1 f(x)^2 \, dx \right)^{1/2}\).

Define \(f_n \in C[-1,1]\) by \(f_n(x) = nx\) for \(|x| \leq 1/n\) and \(f_n(x) = \text{sgn}(x)\) otherwise. Then

1. \(\| f_n - \text{sgn} \|_2^2 = 2 \int_0^{1/n} (nx - 1)^2 \, dx = 2/(3n) \to 0\).

2. \((f_n)\) is Cauchy in the 2-norm, for by the triangle inequality
\[ \| f_n - f_k \|_2 \leq \| f_n - \text{sgn} \|_2 + \| f_k - \text{sgn} \|_2 \to 0 \text{ as } n, k \to \infty.\]
(3) \((f_n)\) doesn't converge in 2-norm to any continuous function. **Why?** To get a contradiction suppose \(f\) is continuous on \([-1,1]\) and \(\|f_n - f\|_2 \to 0\). Then \(\|f - \text{sgn}\|_2 = 0\) because

\[
\|f - \text{sgn}\|_2 \leq \|f - f_n\|_2 + \|f_n - \text{sgn}\|_2 \to 0.
\]

By previous result \(f(u) - \text{sgn}(u) = 0\) at each point of continuity in \([-1,1]\), that is, for all \(u \neq 0\). But then \(f\) can't be continuous at zero. Finally

(4) \(C[-1,1]\) isn't complete under the 2 norm.

4. MORE ON COMPLETENESS

**Definition:** A metric space \((S,d)\) is complete iff each Cauchy sequence of points in \(S\) converges to a point in \(S\). More generally a set \(K\) in a metric space is complete iff each Cauchy sequence of points in \(K\) converges to a point in \(K\).

**Fact of Life (from Advanced Calculus):** With the usual metric the reals are complete and every closed bounded interval in \(R\) is complete.

Recall that \(R^N\) is the set of all sequences of reals with the metric

\[
d(f,g) = \sum_{k=1}^{\infty} 2^{-k} \min\{1, |f(k) - g(k)|\}.
\]

Our first goal is to prove the completeness of \(R^N\).

**Theorem:** For a sequence \((f_n)\) of points in \(R^N\) statements (1)-(4) are equivalent.

1. \((f_n)\) is convergent in \(R^N\).
2. \((f_n)\) is Cauchy in \(R^N\).
3. For each coordinate \(k\), \((f_n(k))_{n=1}^{\infty}\) is a Cauchy sequence of reals.
4. For each coordinate \(k\), \((f_n(k))_{n=1}^{\infty}\) is a convergent sequence of reals.

**Pf** (1) \(\Rightarrow\) (2) is a general fact.

**Pf** (2) \(\Rightarrow\) (3): For each \(r > 0\) choose an index \(m(r) \in N\) so that \(s \& t \geq m(r) \Rightarrow d(f_s, f_t) < r\). Now fix \(k \geq 1\) and let \(\varepsilon > 0\). Put \(r = 2^{-k} \min\{1, \varepsilon\}\). \(s \& t \geq m(r) \Rightarrow\)

\[
2^{-k} \min\{1, |f_s(k) - f_t(k)|\} \leq d(f_s, f_t) < r = 2^{-k} \min\{1, \varepsilon\}
\]

That is possible only if the minimum on the left-hand side is \(|f_s(k) - f_t(k)|\), forcing

\[
|f_s(k) - f_t(k)| < \min\{1, \varepsilon\} \leq \varepsilon.
\]

**Pf** (3) \(\Rightarrow\) (4) holds since the reals are complete.
Pf (4) ⇒ (1): For each \( k \) define \( f(k) = \lim_n f_n(k) \). Let \( \varepsilon > 0 \) and choose \( T \) so that \( \sum_{k=T+1}^{\infty} 2^{-k} < \varepsilon/2 \).

For each \( k \leq T \), \( f_n(k) \rightarrow f(k) \) implies \( \min\{1, |f_n(k) - f(k)|\} \rightarrow 0 \) and consequently, by the usual limit arithmetic, \( \lim_n \sum_{k=1}^{T} 2^{-k} \min\{1, |f_n(k) - f(k)|\} = 0 \). The last implies there is an \( m \) so that

\[
\sum_{k=m}^{\infty} 2^{-k} \min\{1, |f_n(k) - f(k)|\} < \frac{\varepsilon}{2}.
\]

But also

\[
\sum_{k=T+1}^{\infty} 2^{-k} \min\{1, |f_n(k) - f(k)|\} < \sum_{k=T+1}^{\infty} 2^{-k} < \frac{\varepsilon}{2}.
\]

Combining the last two displayed inequalities establishes \( d(f_n,f) < \varepsilon \) whenever \( n \geq m \).

**Corollary:** \( \mathbb{R}^N \) is a complete metric space.

The proof that (4) ⇒ (1) gives this characterization of convergence in \( \mathbb{R}^N \).

**Corollary:** Let \( \{f_n\} \) be a sequence of points in \( \mathbb{R}^N \) and let \( f \) be a point in \( \mathbb{R}^N \). \( f_n \rightarrow f \) in \( \mathbb{R}^N \) iff \( f_n(k) \rightarrow f(k) \) for each coordinate \( k \). More succinctly, convergence in \( \mathbb{R}^N \) means coordinate-wise convergence.

It turns out that \( \{0,1\}^N \) and \([0,1]^N \) are also complete. Checking that uses the notion of closed sets and a little topology. \( \mathbb{R}^n \) and \( l_2 \) are also complete metric spaces. (See problem set).

**Definitions:** A set \( U \) in a metric space \((S,d)\) is open iff \( \forall x \in U \ \exists r > 0 \) so that \( B(x,r) \subset U \). A set \( M \) is closed iff its complement \( M^c = \{ x \in S : x \not\in M \} \) is open.

**Examples:** In any metric space \((S,d)\) (1) each open ball is an open set; (2) each closed ball is a closed set, and; (3) finite sets are closed.

**Pf (1):** Given \( x \in B(a,p) \) let \( r = p - d(a,x) \). \( B(x,r) \subset B(a,p) \) because if \( d(x,z) < r = p - d(x,a) \) then \( d(a,z) \leq d(a,x) + d(x,z) < d(a,x) + p - d(a,x) = p \).

**Pf (2):** Given \( x \in CB(a,p)^c \) let \( r = d(a,x) - p \). \( B(x,r) \subset CB(a,p)^c \) because if \( d(x,z) < r = d(x,a) - p \) then \( d(a,x) \leq d(a,z) + d(z,x) < d(a,z) + [d(x,a) - p] \) yielding \( p < d(a,z) \).

**Pf (3):** For \( F \) finite and \( x \in F^c \), let \( r = \min\{d(x,f) : f \in F\} \). Clearly \( r > 0 \) and \( d(x,z) < r \) implies \( z \not\in F \).

**Example:** For each of \( a = 0 \) and \( a = 1 \), \( U(a) = \{ f \in \{0,1\}^N : f(1) = a \} \) is an open and closed subset of the Cantor Set.
**Pf:** \( U(0) \) and \( U(1) \) are complements so it's enough to check that \( U(a) \) is open. For \( f \in U(a) \), \( B(f,1/4) \subseteq U(a) \) because \( d(f,g) < 1/4 \) implies \( 2^{-1} | f(1) - g(1) | \leq d(f,g) < 1/4 \) and \( g(1) = f(1) = a \).

**Theorem:** In a metric space \((S,d)\) a set \( M \) is closed iff the following sequential condition holds; (s) for every sequence \( (x_n) \) of points in \( M \), if \( x_n \to x \) then \( x \in M \).

**Pf that \( M \) closed implies (s):** To get a contradiction suppose not and let \( (x_n) \) be a sequence in \( M \) with \( x_n \to x \) and \( x \not\in M \). Since \( x \in M^c \) and \( M^c \) is open, \( \exists r > 0 \) so that \( B(x,r) \subseteq M^c \). Using the limit definition, \( \exists m \in \mathbb{N} \) so that \( n \geq m \Rightarrow x_n \in B(x,r) \subseteq M^c \), contradicting \( x_m \in M^c \).

**Pf that \( M \) not closed implies not (s):** If \( M \) isn't closed, then \( M^c \) isn't open and \( \exists x \in M^c \quad \forall r > 0 \exists z \in B(x,r) \) with \( z \not\in M^c \). Taking \( r = 1/n \) produces a point \( z_n \in B(x,1/n) \) with \( z_n \not\in M^c \). A sequence \( (z_n) \) of such points is in \( M \) and converges to \( x \), but \( x \not\in M \).

**Example:** \([0,1]^\mathbb{N}\) and \( \{0,1\}^\mathbb{N}\) are both closed in \( \mathbb{R}^\mathbb{N}\).

**Pf:** For the first let \((f_n)\) be a sequence in \([0,1]^\mathbb{N}\) with \( f_n \to f \) in \( \mathbb{R}^\mathbb{N}\). By the characterization of convergence in \( \mathbb{R}^\mathbb{N}\), \( f_n(k) \to f(k) \) for each \( k \in \mathbb{N} \). For each \( n \) and \( k \), \( 0 \leq f_n(k) \leq 1 \) and thus \( 0 \leq f(k) = \lim_n f_n(k) \leq 1 \), showing \( f \in [0,1]^\mathbb{N}\). The proof that \( \{0,1\}^\mathbb{N}\) is closed in \( \mathbb{R}^\mathbb{N}\) is similar.

---

**5. TOTALLY BOUNDED AND COMPACT SETS**

**Definitions:** Let \( E \) be a subset of a metric space \((S,d)\).

1. For a given \( \varepsilon > 0 \), a set \( F \subseteq S \) in an \( \varepsilon \)-net for \( E \) iff \( \forall x \in E \ \exists z \in F \) so that \( d(x,z) < \varepsilon \).
2. Equivalently, \( F \) in an \( \varepsilon \)-net for \( E \) iff \( E \subseteq \bigcup_{z \in F} B(z, \varepsilon) \).

3. \( E \) is **totally bounded** iff \( E \) has a finite \( \varepsilon \)-net for each \( \varepsilon > 0 \).

**Note:** In case \( E \) is totally bounded the \( \varepsilon \)-nets can be taken from \( E \). In fact, given \( \varepsilon > 0 \), let \( \{z_1, z_2, \ldots, z_n\} \) be an \( \varepsilon/2 \)-net for \( E \). From each non-empty intersection \( E \cap B(z_k, \varepsilon/2) \) select a point \( w_k \). Points \( w_k \) so chosen are an \( \varepsilon \)-net for \( E \). Why? If \( x \in E \) then \( d(x, z_s) < \varepsilon/2 \) for some index \( s \) with \( E \cap B(z_s, \varepsilon/2) \neq \emptyset \). Any \( w_s \) satisfies \( d(w_s, z_s) < \varepsilon/2 \), and by the triangle inequality \( d(x, w_s) < \varepsilon \).

**Example:** In the reals with the usual metric each closed bounded interval \([a,b]\) is totally bounded. Why? Given \( \varepsilon > 0 \), choose \( n \) so satisfy \( (b-a)/n < \varepsilon \) and for \( 0 \leq k \leq n \) let \( z_k = a + k(b-a)/n \).
Example: The Cantor Set \( C = \{0,1\}^N \) is totally bounded. To display an \( \varepsilon \)-net, choose \( n \) so that \( 2^{-n} < \varepsilon \) and let \( F = \{ f \in C : f(k) = 1 \text{ for all } k > n \} \). \( F \) has \( 2^n \) elements. Given \( g \in C \),
\[
f = (g(1), g(2), ..., g(n), 1, 1, ...) \in F \text{ and } d(g, f) = \sum_{k > n} 2^{-k} |g(k) - f(k)| < \sum_{k > n} 2^{-k} = 2^{-n}.
\]

**Theorem:** A totally bounded set is bounded.

**Pf:** Let \( F = \{ z_1, z_2, ..., z_n \} \) be a \( 1 \)-net for \( E \) and \( b = 2 + \max \{ d(z_k, z_{k+1}) : 1 \leq k \leq n \} \). Then \( E \subset B(z_1, b) \) because if \( x \in E \) and \( d(z_m, x) < 1 \) then \( d(x, z_1) \leq d(x, z_m) + d(z_m, z_1) < b \).

The converse is false. There are many bounded sets failing to be totally bounded.

**Definition:** A set \( V \) in a metric space is \( \varepsilon \)-separated iff \( d(x, y) \geq \varepsilon \) for every pair of distinct elements of \( V \). Note that if \( E \) contains an infinite \( \varepsilon \)-separated set for some \( \varepsilon > 0 \), then \( E \) can't have a finite \( \varepsilon/2 \)-net and \( E \) can't be totally bounded.

**Example:** \( B(0, 1) \) is not totally bounded in \( \mathbb{R}^N \).

**Pf:** By Problem 3.4 Frechet's space contains a sequence \( (f_n) \) with \( d(f_n, f_m) = 1 - 2^{-m} \) whenever \( 1 \leq n \leq m \). Let \( g_m = f_m - f_1 \). \( g_m \in B(0, 1) \) because \( d(g_m, 0) = d(f_m, f_1) = 1 - 2^{-m} < 1 \), and the sequence \( (g_m) \) is \( 1/2 \)-separated since \( d(g_n, g_m) = d(f_n, f_m) = 1 - 2^{-m} \geq 1/2 \) for \( 1 \leq n \leq m \).

**Example:** \( CB(0, 1) \) is not totally bounded in \( l_2 \).

**Pf:** In \( l_2 \) define \( e_n \) by \( e_n(n) = 1 \) and \( e_n(k) = 0 \) for \( k \neq n \). \( e_n \) is called the \( n \)th unit vector. For \( 1 \leq n < m \), \( e_n - e_m = (0, 0, ..., 0, 1, 0, ..., 0, -1, 0, ...) \) and \( \|e_n - e_m\|_2 = \sqrt{2} \).

**Theorem:** Let \( (S, d) \) be a metric space and \( E \subset S \). \( E \) is totally bounded iff every sequence of points in \( E \) has a Cauchy subsequence.

**Pf:** We'll first show that if there is an \( \varepsilon > 0 \) for which \( E \) has no finite \( \varepsilon \)-net, then \( E \) contains a non-Cauchy, \( \varepsilon \)-separated sequence \( (e_n) \). To find the sequence recursively let \( e_1 \in E \) be any element. If \( e_1, e_2, ..., e_n \in E \) have been chosen then \( E \) can't be contained in \( \bigcup_{k=1}^n B(e_k, \varepsilon) \). There must be an \( e_{n+1} \in E \) with \( e_{n+1} \in B(e_k, \varepsilon) \) for each \( k = 1, 2, ..., n \). Continuing in this manner generates an \( \varepsilon \)-separated sequence.

For the other direction, assume that \( E \) is totally bounded and let \( (e_n) \) be any sequence in \( E \). Finding a Cauchy subsequence uses a diagonalization argument which we give in some detail.
Claim (a) \( \forall \varepsilon > 0 \ \forall A \subset \mathbb{N} \) infinite \( \exists B \subset A \) infinite so that \( n \& k \in B \Rightarrow d(e_n, e_k) < \varepsilon \).

Pf (a): Let \( f_1, f_2, \ldots, f_n \in E \) be an \( \varepsilon/2 \)-net and write \( A_i = \{ k : d(e_k, f_i) < \varepsilon/2 \} \). \( A = A_1 \cup A_2 \cup \ldots \cup A_n \) is infinite so some \( A_s \) is also infinite. That's the desired set \( B \) since \( n \& k \in A_s = B \Rightarrow d(e_n, e_k) \leq d(e_n, f_s) + d(f_s, e_k) < \varepsilon \).

Claim (b) There is a sequence \( (B_n) \) of infinite subsets of \( N \) so that
(i) \( B_1 \supset B_2 \supset \ldots \supset B_n \supset B_{n+1} \supset \ldots \), and
(ii) \( n \& k \in B_n \Rightarrow d(e_n, e_k) < 1/n \)

Pf (b): Generate the sets \( B_n \) recursively. To select \( B_1 \) use Claim (a) with \( A = N \) and \( \varepsilon = 1 \). When \( B_1, B_2, \ldots, B_n \) have been chosen, use Claim (a) again with \( A = B_n \) and \( \varepsilon = 1/(n+1) \). Take \( B_{n+1} \) to be any set \( B \) of the sort guaranteed by Claim (a).

Abbreviate "least element of set \( B \)" by "least \( B \)." To select the Cauchy subsequence recursively let \( n(1) = \text{least } B_1 \), \( n(2) = \text{least } \{ b \in B_2 : b > n(1) \} \), and in general \( n(k) = \text{least } \{ b \in B_k : b > n(k-1) \} \). This process is never forced to stop because each of the displayed sets is infinite. Note \( n(k) \in B_k \) for each \( k \), and \( n(k) < n(k+1) \) for each \( k \).

Claim (c) \( m \leq s < t \Rightarrow d(e_{n(t)}, e_{n(s)}) < 1/m \). Why? Since the \( B \)'s are chained
\( m < t \Rightarrow n(t) \in B_t \subset B_{t-1} \subset \ldots \subset B_m \), and \( m \leq s \Rightarrow n(s) \in B_s \subset B_{s-1} \subset \ldots \subset B_m \).

By Claim (b) \( n(t), n(s) \in B_m \Rightarrow d(e_{n(s)}, e_{n(t)}) < 1/m \).

Definitions: Let \( K \) be a subset of a metric space \( (S,d) \). An open cover of \( K \) is a collection \( C \) of open sets whose union contains \( K \). A subcover of \( K \) is a subcollection of \( C \) which also covers \( K \). \( K \) is compact iff every open cover has a finite subcover.

Theorem: For \( S \) a complete metric space and \( K \subset S \), the following are equivalent.
(1) \( K \) is compact.
(2) \( K \) is closed and totally bounded.
(3) Every sequence in \( K \) has a subsequence that converges to a point in \( K \).

Corollary (to proof): Let \( K \) be a compact subset of a complete metric space. If \( C \) is an open cover of \( K \) then \( \exists \varepsilon > 0 \ \forall x \in K \ \exists U \in C \) so that \( B(x, \varepsilon) \subset U \). Such an epsilon is called a \( \text{Lebesgue number} \) for \( C \).

Pf (1) implies (2): For fixed \( \varepsilon > 0 \) the collection of all open balls \( B(x, \varepsilon) \), \( x \in K \), forms an open cover of \( K \). Since \( K \) is compact it can be covered by a finite number of these balls. As an exercise prove that a compact set in a metric space must be closed.
Pf (2) implies (3): Let \( (f_n) \) be a sequence in \( K \). Since \( K \) is totally bounded the last theorem says there is a Cauchy subsequence \( (f_{n(k)}) \). \( S \) is complete so \( (f_{n(k)}) \) converges to some point \( f \) in \( S \), and \( f \) is in \( K \) because \( K \) is closed.

**Intermediate Step:** (3) implies each open cover \( C \) of \( K \) has a Lebesgue number, i.e., \( \exists \varepsilon > 0 \ \forall x \in K \exists U \in C \) so that \( B(x, \varepsilon) \subset U \). To get a contradiction suppose not. Then \( \forall n \in \mathbb{N} \ \exists x_n \in K \forall U \in C, B(x_n, 1/n) \not\subset U \). Without loss of generality assume \( x_n \to x \) for some \( x \) in \( K \).

\[ C \text{ covers } K \text{ so } \exists V \in C, x \in V. \]

\[ \text{V is open so } \exists \varepsilon > 0, B(x, \varepsilon) \subset V. \]

\[ x_n \to x \text{ so } \exists n > 2/\varepsilon, d(x_n, x) < \varepsilon/2 \]

The contradiction is \( B(x_n, 1/n) \subset B(x, \varepsilon) \subset V \), which holds since \( d(x_n, y) < 1/n \Rightarrow d(y, x) < d(y, x_n) + d(x_n, x) < 1/n + \varepsilon/2 < \varepsilon \)

Pf (3) implies (1): Let \( C \) be an open cover of \( K \) and \( \varepsilon > 0 \) be a Lebesgue number for \( C \). \( K \) is totally bounded because every sequence in \( K \) has a convergent (hence Cauchy) subsequence. Let \( \{ z_1, z_2, \ldots, z_n \} \) be an \( \varepsilon \)-net and find, for each \( k \), a set \( V_k \subset C \) with \( B(z_k, \varepsilon) \subset V_k \). \( \{ V_1, V_2, \ldots, V_n \} \) is the required finite subcover because \( K \subset \bigcup_{k=1}^n B(z_k, \varepsilon) \subset \bigcup_{k=1}^n V_k \).

**Theorem:** A set \( M \) in \( \mathbb{R}^N \) is totally bounded iff, for each \( k \geq 1 \), \( \{ f_k : f \in M \} \) is bounded.

**Pf:** Suppose that \( \{ f_k : f \in M \} \) is bounded for each \( k \) and let \( \varepsilon > 0 \). There is an index \( T \) so that \( \sum_{k \geq 1} 2^{-k} < \varepsilon/2 \). Write \( b = \sup \{ |f_k| : f \in M \text{ and } 1 \leq k \leq T \} \) and let \( F \) be a finite \( \min \{ 1, \varepsilon/2 \} \)-net for \([-b, b] \). Define \( J = \{ g \in \mathbb{R}^N : g(k) = 0 \text{ for } k > T \text{ and } g(k) \in F \text{ for } 1 \leq k \leq T \} \). \( J \) has \( \text{card}(F)^T \) elements. To see that \( J \) is an \( \varepsilon \)-net for \( M \), let \( f \in M \). For each \( k \leq T \), \( f(k) \in [-b, b] \) and there is an \( x_k \in F \) with \( |f(k) - x_k| < \min \{ 1, \varepsilon/2 \} \). \( g = (x_1, x_2, \ldots, x_T, 0, 0, \ldots) \in J \) and

\[ d(f, g) = \sum_{k \leq T} 2^{-k} \min \{ 1, |f(k) - x_k| \} + \sum_{k > T} 2^{-k} \min \{ 1, |f(k) - 0| \} \]

\[ < \sum_{k \leq T} 2^{-k}(\varepsilon/2) + \sum_{k > T} 2^{-k}(1) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon \]

**Pf:** Conversely suppose that \( M \) is totally bounded. Fix \( k \). Let \( L \subset M \) be a finite \( 2^{-k-1} \)-net for \( M \) and write \( b = l + \max \{ |g(k)| : g \in L \} \). For \( f \in M \) there is a \( g \in L \) so that \( d(f, g) < 2^{-k-1} \) so that \( 2^{-k} \min \{ 1, |f(k) - g(k)| \} \leq d(f, g) < 2^{-k-1} \) and \( \min \{ 1, |f(k) - g(k)| \} < 2^{-1} \). The last can be true only if \( |f(k) - g(k)| < 2^{-1} < 1 \) and thus \( |f(k)| \leq 1 + |g(k)| \leq b \).

**Corollary:** \([0, 1]^N \) and \( \{0, 1\}^N \) are totally bounded in \( \mathbb{R}^N \).
Which spaces are complete? So far here's what we know from class notes and problem sets.
(a) The reals $\mathbb{R}$ and every closed bounded interval $[a,b]$ are complete.
(b) $\mathbb{R}^n$, $\mathbb{R}^n$, $l_2$ and $l_1$ are complete.
(c) $[0,1]^n$ and $\{0,1\}^n$ are closed in $\mathbb{R}^n$ and hence complete.
(d) $C[a,b]$ is not complete under the 2-norm.

Which spaces are compact? So far here's what we know from class notes and problem sets.
(a) $[0,1]^n$ is totally bounded and closed in $\mathbb{R}^n$, hence compact.
(b) $\{0,1\}^n$ is totally bounded and closed in $\mathbb{R}^n$, hence compact.
(c) (Heine Borel Theorem) A closed bounded set in $\mathbb{R}^n$ is compact.
(d) $CB(0,1)$ is not totally bounded in $\mathbb{R}^n$ or in $l_2$. It turns out that $CB(0,1)$ is non-compact in every infinite dimensional normed space.

6. CONTINUOUS FUNCTIONS

Definition: For $(S,d)$ and $(T,p)$ metric spaces and $x$ a given point in $S$, a function $f : S \to T$ is continuous at $x$ iff $\forall \varepsilon > 0 \ \exists \delta > 0$ so that $d(x,z) < \delta \Rightarrow p(f(x), f(z)) < \varepsilon$. The last implication is equivalent to $z \in B(x,\delta) \Rightarrow f(z) \in B(f(x),\varepsilon)$. $f$ is continuous iff $f$ is continuous at each point in $S$.

Note: If $f$ is continuous at $x$ and $x_n \to x$ in $S$, then $f(x_n) \to f(x)$ in $T$. Why? For $\varepsilon > 0$ choose $\delta > 0$ as in the definition of continuity at $x$. Since $x_n \to x$, $\exists m \in \mathbb{N}$ so that $n \geq m$ implies $d(x_n, x) < \delta$, which gives $d(f(x_n), f(x)) < \varepsilon$.

Definitions: Let $(S,d)$ and $(T,p)$ be metric spaces and $f : S \to T$.
(a) $f$ is a Lipschitz function iff there is a constant $L > 0$ so that $p(f(x), f(z)) \leq L d(x, z)$ for all $x$ and $z$ in $S$. Note that a Lipschitz function is continuous (take $\delta = \varepsilon/(2L)$).
(b) The Lipschitz constant of $f$ is $\text{Lip}(f) = \sup \{ p(f(x), f(z)) / d(x, z) : x \neq z \}$.

Examples: (1) For $x, z$ and $w$ in $(S,d)$, $|d(x, w) - d(z, w)| \leq d(x, z)$. Why? If, say, $d(z, w) \leq d(x, w)$, then $d(x, w) \leq d(x, z) + d(z, w)$ and $|d(x, w) - d(z, w)| = d(x, w) - d(z, w) \leq d(x, z)$.
(2) $f(x) = d(x, w)$ is Lipschitz with $\text{Lip}(f) = 1$. Why? The preceding shows $\text{Lip}(f) \leq 1$. If $|d(x, w) - d(z, w)| \leq L d(x, z)$ for all $x$ and $z$, taking $z = w$ shows $1 \leq L$.
(3) In a normed space $f(x) = \| x \| = d(x, 0)$ is Lipschitz with $\text{Lip}(f) = 1$.

Theorem: (a) If $f : S \to T$ is continuous at $x$ and $g : T \to U$ is continuous at $f(x)$, then the composition $g \circ f : S \to U$ is continuous at $x$. 
(b) If \( f : S \to T \) is continuous and \( g : T \to U \) is continuous, then the composition \( g \circ f : S \to U \) is continuous.

**Pf (a):** Let \( \varepsilon > 0 \). First find \( \alpha > 0 \) so that \( d_T(f(x), y) < \alpha \Rightarrow d_U(g(f(x)), g(y)) < \varepsilon \), and then find \( \delta > 0 \) so that \( d_S(x, z) < \delta \Rightarrow d_T(f(x), f(z)) < \alpha \). Chaining the two implication checks the definition.

**Theorem:** (a) If \( f : S \to T \) and \( g : U \to V \) are continuous and \([f, g] : S \times U \to T \times V\) is defined by \([f, g](s, u) = (f(s), g(u))\), then \([f, g]\) is also continuous. The notation \([f, g]\) is non-standard.

(b) If \( f : S \to T \) and \( g : S \to V \) are continuous and \( f \times g : S \to T \times V \) is defined by \((f \times g)(s) = (f(s), g(s))\), then \((f \times g)\) is also continuous.

**Pf (a):** Use the box metrics on the product spaces. Fix \((s_0, u_0) \in S \times U\) and let \( \varepsilon > 0 \). Find \( \delta_T > 0 \) and \( \delta_\varepsilon > 0 \) so that \( d_S(s_0, s) < \delta_T \Rightarrow d_T(f(s_0), f(s)) < \varepsilon \) and \( d_U(u_0, u) < \delta_\varepsilon \Rightarrow d_V(g(u_0), g(u)) < \varepsilon \). Then
\[
\max \{ d_S(s_0, s), d_U(u_0, u) \} = d_{S \times U}( (s_0, u_0), (s, u) ) < \min \{ \delta_T, \delta_\varepsilon \} \Rightarrow \\
d_S(s_0, s) < \delta_T \Rightarrow \text{and} \ d_U(u_0, u) < \delta_\varepsilon \Rightarrow \ d_T(f(s_0), f(s)) < \varepsilon \text{and} \ d_V(g(u_0), g(u)) < \varepsilon \Rightarrow \\
d_{S \times U} \left( [f, g](s_0, u_0), [f, g](s, u) \right) = d_{T \times V} \left( (f(s_0), g(u_0)), (f(s), g(u)) \right) = \max \{ d_T(f(s_0), f(s)), d_V(g(u_0), g(u)) \} < \varepsilon.
\]

**Pf (b):** A direct proof is similar to that of part (a). A fancy proof is to observe that \( f \times g = [f, g] \circ D \), where \( D : S \to S \times S \) is the (obviously) continuous function \( D(s) = (s, s) \).

**Theorem:** Let \( V \) be a normed space.

(a) Addition \( A(x, y) = x + y \) is a continuous function \( A : V \times V \to V \).

(b) Scalar multiplication \( M(t, x) = tx \) is a continuous function \( M : R \times V \to V \).

**Pf for addition:** Let \( \varepsilon > 0 \).
\[
|| (x + y) - (u + v) || = || (x - u) + (y - v) || \leq || x - u || + || y - v || < \varepsilon
\]
provided
\[
\max \{ || x - u ||, || y - v || \} = d_{V \times V}( (x, y), (u, v) ) < \varepsilon/2.
\]

**Pf for multiplication:** Fix a vector \( x \), a scalar \( t \) and let \( \varepsilon > 0 \). If
\[
\max \{ |t - s|, || x - u || \} = d_{R \times V}( (t, x), (s, u) ) < \delta \text{ for sufficiently small } \delta > 0 \text{ then}
\]
\[
|| tx - su || = || t(x - u) + (t - s)(x - u) || \\
\leq |t| || x - u || + |t - s| || x - u || \\
\leq |t| \delta + || x || \delta + \delta^2 < \varepsilon
\]

**Corollary:** Let \( S \) be a metric space, \( x \in S \) and \( V \) be a normed space.

(a) If \( f : S \to V \) and \( g : S \to V \) are continuous at \( x \) (resp, are continuous) then so is the sum \( f + g \).
(b) If \( f : S \rightarrow V \) and \( m : S \rightarrow \text{Reals} \) are continuous at \( x \) (resp, are continuous) then so is the scalar product \( mf \).

**Pf:** Writing \( A \) and \( M \) for addition and scalar multiplication on \( V \), \( f + g = A \circ (f \times g) \) and \( mf = M \circ (m \times f) \).

**Corollary:** Let \( S \) be a metric space. If \( f \) and \( g \) are continuous real valued functions on \( S \), then so are \( f + g \), \( fg \) and \( cf \) for any scalar \( c \).

**Pf:** For the product, \( R \) is itself a real vector space and \( fg = M \circ (f \times g) \).

7. UNIFORM CONVERGENCE; THE SPACES B(S) AND C(S)

**Definitions:** Let \( S \) be a set, \( f : S \rightarrow \mathbb{R} \) be a function, and \( (f_n) \) be a sequence of functions from \( S \) into \( \mathbb{R} \).

(a) \( (f_n) \) converges pointwise on \( S \) to \( f \) iff, for each \( x \) in \( S \), \( f_n(x) \rightarrow f(x) \). Equivalently,

\[
\forall x \in S \forall \varepsilon > 0 \exists m(x) \in \mathbb{N}, n \geq m(x) \Rightarrow \left| f_n(x) - f(x) \right| < \varepsilon.
\]

(b) \( (f_n) \) converges uniformly on \( S \) to \( f \) iff \( \forall \varepsilon > 0 \exists m \in \mathbb{N} \forall x \in S, n \geq m \Rightarrow \left| f_n(x) - f(x) \right| < \varepsilon \).

**Notes:** (1) If \( (f_n) \) converges uniformly on \( S \) to \( f \) then \( (f_n) \) converges pointwise on \( S \) to \( f \).

(2) Pointwise convergence need not imply uniform convergence. On \([0,1]\) the functions \( f_n(x) = x^n \) have as their pointwise limit the function with \( f(1) = 1 \) and \( f(x) = 0 \) otherwise. Convergence isn’t uniform because, say, \( f_n(2^{-1/n}) - f(2^{-1/n}) = 2^{-1} \) for all \( n \).

**Definitions:** Let \( S \) be a set.

(a) A function \( f : S \rightarrow \mathbb{R} \) is **bounded** iff there is a number \( b \) so that \( |f(x)| \leq b \) for all \( x \) in \( S \).

(b) The **uniform norm** of a bounded function \( f : S \rightarrow \mathbb{R} \) is

\[
\left\| f \right\|_u = \sup \{ |f(x)| : x \in S \} = \min \{ b : \forall x \in S, |f(x)| \leq b \}.
\]

(c) \( B(S) \) denotes the set of all bounded functions \( f : S \rightarrow \mathbb{R} \).

**Theorem:** (a) \( B(S) \) is a vector space under the pointwise operations

\[ (f + g)(x) = f(x) + g(x) \quad \text{and} \quad (cf)(x) = cf(x). \]

(b) The uniform norm is a norm on \( B(S) \).

(c) \( (f_n) \) converges to \( f \) in the metric space \( B(S) \) iff \( (f_n) \) converges uniformly on \( S \) to \( f \).

(d) \( B(S) \) is a complete under the uniform norm. A complete normed space is called a **Banach space**.
\textbf{Pf (a) and (b):} Once we have B(S) closed under the pointwise operations the vector space axioms hold automatically. For simplicity the subscript \( u \) on the norm is omitted. Clearly \( \| f \| \geq 0 \) for all \( f \), and \( \| f \| = 0 \) iff \( f(x) = 0 \) for all \( x \in S \) iff \( f = \) zero function.

\textbf{Claim:} B(S) is closed under scalar multiplication and \( \| cf \| = |c| \| f \| \). Let \( f \in B(S) \) and \( c \) be a scalar.

The case \( c = 0 \) is trivial. For all \( x \) \( \| (c f)(x) \| = |c| \| f(x) \| \leq |c| \| f \| \), which shows that \( cf \in B(S) \) and \( \| cf \| \leq |c| \| f \| \). For the reverse inequality, \( \| f \| = \| (1/c)(cf) \| \leq 1/c \| cf \| \), hence \( |c| \| f \| \leq \| cf \| \).

\textbf{Claim:} B(S) is closed under addition and the triangle inequality holds. Let \( f, g \in B(S) \). For all \( x \)
\[ |(f + g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \| f \| + \| g \|. \]
Thus \( f + g \) is bounded and \( \| f + g \| \leq \| f \| + \| g \| \).

\textbf{Pf (c):} \( \| f_n - f \| \to 0 \) iff \( \forall \varepsilon > 0 \ \exists m \in \mathbb{N} \) so that \( n \geq m \Rightarrow \| f_n - f \| \leq \varepsilon \) iff \( \forall \varepsilon > 0 \ \exists m \in \mathbb{N} \) so that \( \forall x \in S \ \forall n \geq m \Rightarrow \| f_n(x) - f(x) \| \leq \varepsilon \).

\textbf{Pf (d):} Let \( (f_n) \) be Cauchy in norm. For each \( \varepsilon > 0 \) choose once and for all an index \( m(\varepsilon) \) so that \( \forall x \in S, n \geq k \geq m(\varepsilon) \Rightarrow |f_n(x) - f_k(x)| \leq \varepsilon/2 \). Consequently \( \forall s \in S, (f_n(s)) \) is Cauchy in the reals.
Define \( f(s) = \lim_n f_n(s) \). By the choice of \( m(\varepsilon) \)
\[ \forall x \in S, n \geq m(\varepsilon) \Rightarrow |f_n(x) - f(x)| = \lim_k |f_n(x) - f_k(x)| \leq \varepsilon/2 \].
In particular the last shows that \( \forall n \geq m(\varepsilon), f_n - f \in B(S) \) and that \( \| f_n - f \| \leq \varepsilon/2 \). But since B(S) is closed under addition, \( f = (f - f_{m(1)}) + f_{m(1)} \in B(S) \) and \( f_n \to f \) in norm.

\textbf{Theorem:} If \( f : S \to T \) is continuous and \( K \subset S \) is compact, then \( f(K) = \{ f(x) : x \in K \} \) is a compact subset of \( T \).

\textbf{Pf:} It's enough to prove that every sequence \( (y_n) \) in \( f(K) \) has a subsequence converging to a point in \( f(K) \). Write \( y_n = f(x_n) \) for \( x_n \in K \). \( K \) is compact so some subsequence \( (x_{n(k)}) \) converges to a point \( x \in K \). \( f \) is continuous and \( x_{n(k)} \to x \) so also \( y_{n(k)} = f(x_{n(k)}) \to f(x) \) and \( f(x) \in f(K) \).

\textbf{Corollary:} If \( f : S \to \mathbb{R} \) is continuous and \( S \) is compact then there are points \( p \) and \( q \) in \( S \) so that \( f(q) \leq f(x) \leq f(p) \) for all \( x \) in \( S \). In particular \( f \) is bounded on \( S \).

\textbf{Pf:} By the preceding \( f(S) \) is compact and thus bounded. Let \( c = \sup \{ f(x) : x \in S \} \). For each \( n \) there is an \( x_n \in S \) with \( f(x_n) > c - 1/n \). The compactness of \( S \) produces a subsequence \( (x_{n(k)}) \) converging to a point \( p \in K \). Again \( f \) continuous at \( p \) and \( x_{n(k)} \to p \) implies \( f(p) = \lim_k f(x_{n(k)}) \geq c \). The reverse inequality \( f(p) \leq c \) holds by the definition of \( c \), so \( f(p) = c \). A similar procedure produces \( q \).
Definition and Theorem: For $S$ a compact metric space let $C(S)$ denote the set of all continuous functions $f : S \to \mathbb{R}$.

(a) $C(S)$ is a vector space under the pointwise operations.
(b) If $(f_n)$ is a sequence in $C(S)$ and $(f_n)$ converges uniformly on $S$ to $f$, then $f$ is continuous.
(c) $C(S)$ is a closed subset of $B(S)$ and thus is complete under the uniform norm.

Proof (a): $C(S)$ is a vector subspace of $B(S)$ since, by a previous corollary, it is closed under pointwise addition and scalar multiplication.

Proof (b): Let $z \in S$ and let $\varepsilon > 0$. By uniform convergence $\exists m \in \mathbb{N}$ so that

$\forall x \in S, \ n \geq m \Rightarrow |f_n(x) - f(x)| < \varepsilon/3$. $f_m$ is continuous at $z$ so

$\exists \delta > 0 \ \forall x \in S, \ d(x,z) < \delta \Rightarrow |f_m(z) - f_m(x)| < \varepsilon/3$. Combining inequalities, for $d(x,z) < \delta$

$|f(z) - f(x)| \leq |f(z) - f_m(z)| + |f_m(z) - f_m(x)| + |f_m(x) - f(x)| < \varepsilon$.

Proof (c): Recall that a closed subset of a complete space is complete.

Corollary: If $S$ is any metric space and $(f_n)$ is a sequence of continuous functions on $S$ that converges uniformly to $f$, then $f$ is continuous.

Proof: The compactness of $S$ isn’t used in the proof of (b).

Theorem: Let $(S,d)$ and $(T,p)$ be metric spaces. If $f : S \to T$ is continuous and $K \subset S$ is compact, then $f$ is uniformly continuous on $K$, that is, $\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall z \in K \ \forall x \in S, \ d(x,z) < \delta \Rightarrow p(f(z),f(x)) < \varepsilon$.

Proof: To get a contradiction suppose not. Then

$\exists \varepsilon > 0 \ \forall n \in \mathbb{N} \ \exists z_n \in K \ \exists x_n \in S, \ d(x_n,z_n) < 1/n$ and $p(f(z_n),f(x_n)) \geq \varepsilon$.

By the compactness of $K$ we may assume that $(z_n)$ converges to a some point $z$ in $K$. Then $(x_n)$ also converges to $z$ because $d(x_n,z) \leq d(x_n,z_n) + d(z_n,z) < 1/n + d(z_n,z) \to 0$. The continuity of $f$ at $z$ produced the contradiction $0 = p(f(z),f(z)) = \lim_n p(f(z_n),f(x_n)) \geq \varepsilon$.

7.5. PROPERTIES OF UNIFORM CONVERGENCE

Lemma: Suppose $f$, $g$, and each $f_n$ are continuous on $[a,b]$.

1. $\int_a^b g(x) \, dx \leq \sqrt{b-a} \|g\|_2$ and $\|g\|_2 \leq \sqrt{b-a} \|g\|_u$.

2. If $f_n \to f$ uniformly on $[a,b]$ then $f_n \to f$ pointwise and in 2-norm.
Pf: By the Cauchy-Schwartz Inequality \[ \left| \int_a^b g(x) \, dx \right| = |<1, g>\|_2 \leq \|1\|_2 \|g\|_2 = \sqrt{b-a} \|g\|_2. \]

Since \( |g(x)| \leq \|g\|_u \) for all \( x \) in \([a, b]\), \( \|g\|_2 = \left( \int_a^b |g(x)|^2 \, dx \right)^{1/2} \leq \left( \int_a^b \|g\|_u^2 \, dx \right)^{1/2} = \sqrt{b-a} \|g\|_u \).

For the last statement using \( g = f_n - f \) shows \( \|f_n - f\|_2 \leq \sqrt{b-a} \|f_n - f\|_u \).

Example: Pointwise convergence doesn't imply uniform or 2-norm convergence. On \([0, 1]\) let \( f_n(x) = n \sin(\pi nx) \) for \( 0 \leq x \leq 1/\), and \( f_n(x) = 0 \) otherwise. \( f_n \to 0 \) pointwise and \( \|f_n\|_2^2 = n/2 \).

Example: 2-norm convergence doesn't imply uniform or pointwise convergence. On \([0, 1]\) \( f_n(x) = x^n \) has \( \|f_n\|_2^2 = 1/(2n+1) \to 0 \) but \( f_n(1) = 1 \) for all \( n \).

Question One: If \( (f_n) \) is a sequence of continuous functions on \([a, b]\) and \( f_n \to f \), is the limit function \( f \) continuous?

Answers: The answers are yes for uniform convergence, no for pointwise convergence and no for 2-norm convergence. \( f_n(x) = x^n \) on \([0, 1]\) is a counterexample for pointwise convergence, and there is a sequence of continuous functions on \([-1, 1]\) which converges in 2-norm to \( \text{sgn}(x) \), but doesn't converge in 2-norm to any continuous function.

There are several interesting cases in which 2-norm convergence does imply uniform convergence.

Lemma: If \( g \) is continuously differentiable on \([a, b]\) then \( \|g\|_u^2 \leq 2 \|g\|_2 \|g'\|_2 + (b-a)^{-1} \|g\|_2^2. \)

Pf: Let \( x \) and \( z \) be in \([a, b]\). By the Cauchy-Schwartz Inequality

\[
\left| g(x)^2 - g(z)^2 \right| = \left| \int_x^z [g(t)^2]' \, dt \right| = \left| \int_x^z g(t) g'(t) \, dt \right| \\
\leq 2 \int_{\min(x, z)}^{\max(x, z)} \left| g(t) \right| \|g'(t)\|_2 \, dt \leq 2 \|g\|_2 \|g'\|_2.
\]

This shows \( g(x)^2 \leq g(z)^2 + 2 \|g\|_2 \|g'\|_2 \). Now fixing \( x \) and integrating \( dz \) over \([a, b]\),

\( g(x)^2 (b-a) \leq \|g\|_2^2 + 2 \|g\|_2 \|g'\|_2 (b-a) \). Since \( x \) was arbitrary the result follows.

Theorem: If \( (f_n) \) is a sequence of continuously differentiable functions on \([a, b]\), \( (f_n) \) is Cauchy in 2-norm and the sequence of derivatives \( (f_n') \) is bounded in 2-norm, then \( (f_n) \) converges uniformly to a continuous function.
Pf: Let \( c = \sup_{n \geq 1} \| f_n' \| \). Using the lemma with \( g = f_n - f_m \),
\[
\| f_n - f_m \|_u^2 \leq 2 \| f_n - f_m \|_2 \| f_n' - f_m' \|_2 + (b - a)^{-1} \| f_n - f_m \|_2^2 \\
\leq 2 \| f_n - f_m \|_2 (2c) + (b - a)^{-1} \| f_n - f_m \|_2^2
\]
The last inequality says \( f_n' \) is Cauchy in the Banach space \( C(S) \).

**Corollary:** If \( (f_n) \) is a sequence of continuously differentiable functions on \([a,b] \), \( f_n \rightarrow f \) in 2-norm and \( (f_n') \) is bounded in 2-norm, then there is a continuous function \( h \) on \([a,b] \) with \( \| f - h \|_2 = 0 \).

**Pf:** For some continuous \( h \), \( f_n \rightarrow h \) uniformly and hence in 2-norm, and
\[
\| f - h \|_2 \leq \| f - f_n \|_2 + \| f_n - h \|_2 \rightarrow 0.
\]

**Question Two:** If \( (f_n) \) is a sequence of continuously differentiable functions on \([a,b] \) and \( f_n \rightarrow f \), is the limit function \( f \) differentiable? The **answer** is no for uniform, pointwise and 2-norm convergence.

**Example:** For \( f_n(x) = \sqrt{x^2 + 1/n} \) and any \( x \), \( \| f_n(x) - |x| \|_2 = x^2 + 1/n - 2|x|\sqrt{x^2 + 1/n} \leq 1/n \). On \([-1,1] \), \( \| f_n - \text{abs} \|_u \leq 1/\sqrt{n} \) and \( f_n(x) \rightarrow |x| \) uniformly.

**Question Three:** If \( (f_n) \) is a sequence of continuously differentiable functions on \([a,b] \), \( f_n \rightarrow f \), and the limit function \( f \) differentiable, does \( f_n' \rightarrow f' \) is some sense? Again the **answer** is no for uniform, pointwise and 2-norm convergence.

**Example:** On \([0,1] \) let \( f_n(x) = n^{-1} \sin(n^2 \pi x) \). \( \| f_n \|_u = 1/n \rightarrow 0 \), but \( f_n'(0) = n \pi \cos(0) \rightarrow \infty \) and
\[
\| f_n' \|_2^2 = (n \pi)^2 / 2 \rightarrow \infty.
\]

Although "\( f_n \rightarrow f \Rightarrow f_n' \rightarrow f' \)" is false a version of the converse "\( f_n' \rightarrow f' \Rightarrow f_n \rightarrow f \)" is true.

**Theorem:** Let \( (f_n) \) be a sequence of continuously differentiable functions on \([a,b] \). If \( \lim_n f_n(a) \) exists and \( f_n' \rightarrow g \) in 2-norm for some continuous \( g \), then there is a differentiable \( f \) on \([a,b] \) so that \( f' = g \) and \( f_n \rightarrow f \) uniformly on \([a,b] \).

**Pf:** Write \( c = \lim_n f_n(a) \) and define \( f(x) = c + \int_a^x g(t) \, dt \). By the Fundamental Theorem \( f' = g \). For \( x \) in \([a,b] \), \( f_n(x) - f(x) = f_n(a) - c + \int_a^x (f_n'(t) - g(t)) \, dt \). Using the Cauchy-Schwartz Inequality
\[ |f_n(x) - f(x)| \leq |f_n(a) - c| + \left( \int_a^x [f_n'(t) - g(t)] \, dt \right) \leq |f_n(a) - c| + \sqrt{b - a} \left( \int_a^x [f_n'(t) - g(t)]^2 \, dt \right)^{1/2} \leq |f_n(a) - c| + \sqrt{b - a} \left\| f_n' - g \right\|_2. \]

Since \( x \) was arbitrary, \( \left\| f_n - f \right\|_u \leq |f_n(a) - c| + \sqrt{b - a} \left\| f_n' - g \right\|_2. \)

**Corollary:** With the uniform norm on \( C[a,b], \ G = \{ (f,f') : f \text{ is continuously differentiable on } [a,b] \} \) is a closed set and a vector subspace of \( C[a,b] \times C[a,b] \). \( G \) is the graph of a differentiation operation defined on a vector subspace of \( C[a,b] \).

**Pf:** The product has the box metric. If \((f_n,f'_n) \to (h,g) \in C[a,b] \times C[a,b], \) then \( f_n \to h \) uniformly, \( f'_n \to g \) uniformly, and \( f'_n \to g \) in 2-norm. By the Theorem there is a continuously differentiable \( f \) with \( f' = g \) and \( f_n \to f \) uniformly on \([a,b] \). Since \( f_n \to h \) and \( f_n \to f \) in \( C[a,b] \) and limits are unique, \( h = f \) and \((h,g) = (f,f') \in G \).

**Question Four:** If \((f_n)\) is a sequence of continuous functions on \([a,b]\) and \( f_n \to f \), is the limit function \( f \) (Riemann) integrable? The answer is yes for uniform convergence (the proof from Advanced Calculus won't be repeated) and no for pointwise convergence.

**Example:** For \( 0 < x \leq 1, \) define \( f_n(x) = \min \{ n, 1/x \} \) and \( f(x) = 1/x \), with \( f_n(0) = 0 \) and \( f(0) = 0 \). \( f_n \to f \) pointwise on \([0,1]\) but \( f(x) = 1/x \) is not integrable on \([0,1]\).

**Non-continuous Example:** On \([0,1]\) define \( f_n(x) = 1 \) if \( x \) is an integer, and \( f_n(0) = 0 \) otherwise. \( f \) is only finitely non-zero, hence integrable, but the pointwise limit is \( f(x) = 1 \) for \( x \) rational, and \( f(x) = 0 \) for irrational \( x \). The limit function is not Riemann integrable.

**Question Five:** If \((f_n)\) is a sequence of continuous functions on \([a,b]\), \( f_n \to f \), and the limit function \( f \) is integrable, does \( \int_a^b f_n(x) \, dx \to \int_a^b f(x) \, dx \)? The answer is yes for uniform and 2-norm convergence.

Using \( g = f_n - f_m \) in the first lemma of this section:

\[ \left| \int_a^b f_n(x) \, dx - \int_a^b f(x) \, dx \right| = \left| \int_a^b f_n(x) - f(x) \, dx \right| \leq \sqrt{b - a} \left\| f_n - f \right\|_2 \leq (b - a) \left\| f_n - f \right\|_u. \]

**Example:** Pointwise convergence doesn't imply convergence of the integrals. On \([0,1]\) let 
\( f_n(x) = n \sin(n \pi x) \) for \( 0 \leq x \leq 1/n \), and \( f_n(x) = 0 \) otherwise. \( f_n \to 0 \) pointwise and \( \int_0^1 f_n(x) \, dx = 2\pi \).