Real Analysis

July 10, 2006

1 Introduction

These notes are intended for use in the warm-up camp for incoming Berkeley Statistics graduate students. Welcome to Cal! The real analysis review presented here is intended to prepare you for Stat 204 and occasional topics in other statistics courses. We will not cover measure theory topics and some other material that you should be very familiar with if you intend to take Stat 205. If you have never taken a real analysis course, you are strongly encouraged to do so by taking math 104 (or the honors version of it). Math 105 (usually offered in the spring) will provide you with necessary measure-theoretical background essential for Stat 205. The presentation follows closely and borrows heavily from 'Real Mathematical Analysis' by C.C. Pugh, the standard textbook for honors version of math 104. The emphasis is on metric space concepts and the pertinent results on the reals are presented as specific cases of more general results, and a lot of them are presented together as exersises in section

3.8. We do not expect you to be familiar with the metric space concepts but we do expect you to be familiar with specific results on real line, as that is usually the approach taken in most real analysis courses. We hope that you find these notes helpful! Go Bears!

2 Some Definitions

We will denote by $\mathbb{R}, \mathbb{Q}, \mathbb{Z}$, and $\mathbb{N}$ the sets of all real numbers, rational numbers, integers and positive integers, respectively. We will take for granted the familiarity with notions of finite, countably infinite and uncountably infinite sets. Given a set S$\subset \mathbb{R}$, M$\in \mathbb{R}$ is an **upper bound** for S if $\forall s \in S$ it is true that $s \leq M$. S is said to be bounded above by M. M^* is said to be the **least upper bound** (or l.u.b) for S if for all upper bounds M it is true that $M^* \leq M$ (note that it implies that l.u.b. is unique). The lower bounds and the greatest lower bound (g.l.b) are defined similarly. For example, if S=[0,1], then 1 and 4 are upper bounds, and 0, -4 are lower bounds with 0 and 1 being g.l.b and l.u.b, respectively. A set is said to be **bounded** if it is bounded from above and below. A set S is said to be **unbounded from above** if $\forall N\ \exists s \in S$ s.t. $s > N$. The definition of a set unbounded below is similar. A set is unbounded if it's either unbounded from above or below. The **supremum** (sup) of a set $S \subset \mathbb{R}$ is defined to be l.u.b for S if S is bounded from above, and to be $+\infty$ otherwise. The **infinum** (inf) is defined to be g.l.b for a set bounded form below and to be $-\infty$ otherwise.The following results will be assumed about $\mathbb{R}$ (proofs can be looked up in any analysis textbook):

1. If $S \subset \mathbb{R}$ is bounded from above/below then the l.u.b/g.l.b for S exists and is unique

2. **Triangle Inequality**: $|x + y| \leq |x| + |y|$

3. ϵ**-principle**: If x,y $\in \mathbb{R}$, and $\forall\epsilon > 0$, $x \leq y + \epsilon$ then $x \leq y$. Also if $\forall\epsilon > 0$, $|x - y| \leq \epsilon$, then x=y.

4. Every interval (a,b) contains countably infinitely many rationals and uncountably infinitely many irrationals.

LIMINF AND LIMSUP???

3 Metric Spaces

3.1 Definition

A **metric space** M is a set of elements together with a function d:$M \times M \rightarrow R$ (known as **metric**) that satisfies the following 3 properties. For all x, y, z $\in$ M:

1. d(x,y)$\geq$ 0 and d(x,y)=0 iff x=y

2. d(x,y) = d(y,x)

3. d(x,y) $\leq$ d(x,z) + d(z,y)

When metric d is understood, we refer to M as the metric space. When we want to specify that the metric is in M , we might use notation $d_M(x, y)$. It also helps sometimes to think of d as the distance function, since it makes the 3 properties more intuitive (we usually think of distance as being nonnegative, and the distance from A to B should be the same as the distance from B to A). Here are some examples:

1. $\mathbb{R}$ with usual distance function: d(x,y) = $|x - y|$
2. $\mathbb{Q}$ with the same metric
3. $\mathbb{R}^n$ with Euclidean distance d(x,y) = $\|x - y\|$
4. Any metric space (for ex. $\mathbb{R}$ or $\mathbb{N}$) with the **discrete** metric: d(x,y)=1 if $x \neq y$, d(x,y)=0 otherwise. This metric makes the distance from a point to itself be 0 and the distance between any two distinct points be 1.

You should check for yourself that the metrics above satisfy the 3 conditions.

3.2 Sequences

We will use the notation (x_n) for the sequence of points $x_1, x_2, \ldots, x_n, \ldots$ in metric space M. The members of a sequence are not assumed to be distinct, thus 1,1,1,1,... is a legitimate sequence of points in $\mathbb{Q}$. A sequence (y_k) is a **subsequence** of (x_n) if there exists sequence $1 \leq n_1 < n_2 < n_3 < \ldots$ s.t. $y_k = x_{n_k}$. Some subsequences of the sequence 1,1,2,1,2,3,1,2,3,4,1,2,3,4,5,... are:

1. twos: 2,2,2,2,2,...

2. odds: 1,3,5,7,9,...

3. primes: 2,3,5,7,11,...

4. original sequence with duplicates removed: 1,2,3,4,5,6,7,...

5. the previous subsequence with first 3 elements removed: 4,5,6,7,...

A sequence (x_n) of points in M is said to **converge to the limit** x in M if $\forall \epsilon > 0$ $\exists$ N s.t. $n \geq N \implies d(x_n, x) < \epsilon$. We then say that $x_n \to x$. Notice that if our metric space is $\mathbb{R}$, then replacing $d(x_n, x)$ by the usual metric $|x_n - x|$ gives the familiar definition of a limit.

Theorem: The limit of a sequence, if it exists, is unique

Proof: Let (x_n) be a sequence in M that converges and suppose its limit is not unique. Let x,y denote two (of possibly even more) limits. Let $\epsilon > 0$ be given. Then $\exists N_1$ s.t. $n \geq N_1$ $\implies d(x_n, x) < \epsilon/2$. Similarly $\exists N_2$ s.t. $n \geq N_2 \implies d(x_n, y) < \epsilon/2$. Let N = $max(N_1, N_2)$, and let $n \geq N$. Then by the 3rd property of metric function:

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < \epsilon/2 + \epsilon/2 = \epsilon.$$

Since this is true for every ϵ we have x=y by the ϵ-principle

Theorem: Every subsequence of a convergent sequence converges, and it converges to the same limit as the original sequence.

Proof: Easy.

A sequence (x_n) in M is said to be **Cauchy** if $\forall \epsilon > 0\ \exists N$ s.t. $n, m \geq N \Longrightarrow d(x_n, x_m) < \epsilon$. In other words a sequence is Cauchy if eventually all the terms are all very close to each other.

Theorem: Every convergent sequence is Cauchy

Proof: Suppose $x_n \to x$ in M. Let $\epsilon > 0$ be given. Then $\exists N$ s.t. $n \geq N \Longrightarrow d(x_n, x)$ $< \epsilon/2$. Let $n, m \geq N$. Then

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \epsilon/2 + \epsilon/2 < \epsilon \Longrightarrow (x_n) \text{ is Cauchy.}$$

Does every Cauchy sequence converge to a limit? Consider the sequence 3, 3.14, 3.141, 3.1415, ...This sequence is clearly Cauchy. When considered as a sequence in $\mathbb{R}$, it does converge to π. However, we can also think of it as a sequence in $\mathbb{Q}$, in which case it doesn't converge, since $\pi \notin \mathbb{Q}$. The following definition formalizes the difference between $\mathbb{Q}$ and $\mathbb{R}$: the metric space M is said to be **complete** if all Cauchy sequences in M converge to a limit in M.

Theorem: $\mathbb{R}$ is complete.

Proof: Let (a_n) be a Cauchy sequence in $\mathbb{R}$ and let A be the set of elements of the sequence, i.e.

$$\text{A}=\{x \in \mathbb{R} : \exists n \in \mathbb{N} \text{ and } a_n = x\}$$

Let ϵ=1. Then since (a_n) is Cauchy, $\exists N_1$ s.t. $\forall n, m \geq N_1$, $|a_n - a_m| < 1$. Therefore, $\forall n \geq N_1$ we have $|a_n - a_{N_1}| < 1$ and thus $n \geq N_1 \implies a_n \in [a_{N_1} - 1, a_{N_1} + 1]$.

The finite set $a_1, a_2, \ldots, a_{N_1}, a_{N_1} - 1, a_{N_1} + 1$ is bounded (as is every finite subset of $\mathbb{R}$), and therefore all of its elements belong to some interval $[-L, L]$. Since both $a_{N_1} - 1, a_{N_1} + 1$ $\in [-L, L]$, we conclude that $[a_{N_1} - 1, a_{N_1} + 1] \subset [-L, L]$ and therefore $a_n \in [-L, L]$ $\forall$n, and A is bounded. Now, consider set

$$S = \{s \in [-L, L] : \exists \text{ infinitely many } n \in \mathbb{N}, \text{ for which } a_n \geq s\}$$

Obviously, $-L \in S$ and S is bounded from above by L. We then know that the l.u.b for S exists (call it b). We will show that $a_n \to b$.

Let $\epsilon > 0$ be given. Then since (a_n) is Cauchy, $\exists N_2$ s.t. $m, n \geq N_2 \implies |a_n - a_m| < \epsilon/2$. Since $\forall s \in S$, $s \leq b$, we have that $b + \epsilon/2 \notin S$. That means that a_n exceeds $b + \epsilon/2$ only finitely often and $\exists N_3 \geq N_2$ s.t. $n \geq N_3 \Longrightarrow a_n \leq b + \epsilon/2$. Since b is the least upper bound for S we have that $b - \epsilon/2$ is **not** an upper bound for S and therefore $\exists s \in S$ s.t. $s > b - \epsilon/2$ and $a_n \geq s > b - \epsilon/2$ infinitely often. In particular that gurantees that there exists some $N_4 \geq N_3$ s.t. $a_{N_4} > b - \epsilon/2$. Moreover, since $N_4 \geq N_3$ we have $a_{N_4} \in (b - \epsilon/2, b + \epsilon/2]$. And since $N_4 > N_2$, we get $n \geq N_4 \Longrightarrow$

$$|a_n - b| \leq |a_n - a_{N_4}| + |a_{N_4} - b| < \epsilon/2 + \epsilon/2 < \epsilon$$

and therefore $a_n \to b$.

It is an easy exersise to show that every discrete metric space (for example, $\mathbb{N}$ with discrete metric) is complete.

3.3 Open and Closed Sets

Let M be a metric space and let S be a subset of M. We say that $x \in M$ is a **limit**(point) of S (note that S is not a sequence here, but a subset of M) if there exists a sequence (x_n) in S s.t. $x_n \to x$. For example, let M = $\mathbb{R}$ and S=$\mathbb{Q}$. Then 2 is an example of a limit point of S ($(x_n) = 2, 2, 2, 2, 2, \ldots$ or $(x_n) = 1, 1 + 1/2, 1 + 2/3, 1 + 3/4, \ldots$) that belongs to S, and π is an example of a limit point of S ($(x_n) = 3, 3.1, 3.14, 3.141, 3.1415, \ldots$) that does not belong to S (of course, it still belongs to M). A set is said to be **closed** (in underlying metric space

M) if it contains all of its limits. For example, a singleton set x is closed in M (assuming M is non-empty), since the only possible sequence is $(x_n) = x, x, x, \ldots \to x$. Clearly, M is a closed subset of itself. A set S is said to be **open** (in underlying mertric space M) if $\forall x \in S\ \exists r > 0$ s.t. $d(x, y) < r \implies y \in S$. The set of points $\{y \in M : d(x, y) < r\}$ is called (open) r-neighborhood of x and is denoted $B_r(x)$ (for example, $B_2(5)$ in $\mathbb{R}$ is simply the open interval $(3, 7)$). Thus a set S is open (in M) if for every point in S there exists some small neighborhood of that point in M contained entirely in S. It's easy to see that interval (a,b) is an open set in $\mathbb{R}$, and clearly every metric space M is an open subset of itself (since every r-neighborhood of x is still in M). The following theorem provides a connection between open and closed sets:

Theorem: The compliment of an open set is closed and the compliment of a closed set is open.

Proof:Suppose $S \subset M$ is open. Suppose $x_n \to x$ in M, and moreover $x_n \in S^c\ \forall n$. We need to show that $x \in S^c$. Suppose not, then $x \in S$. Since S is open, $\exists r > 0$ s.t. $d(x, y) < r \implies y \in S$. Now, since $x_n \to x\ \exists N$ s.t. $n \geq N \implies d(x_n, x) < r \implies x_n \in S$. That is clearly impossible since no point could be in both S and S^c, and we have reached a condtradiction. Therefore S^c is closed.

Now suppose that $S \subset M$ is closed. Suppose S^c is not open, then $\exists x \in S^c$ s.t. $\forall r_n > 0\ \exists x_n$

s.t. $d(x, x_n) < r_n$ but $x_n \in S$. Now let $r_n = 1/n$ and pick x_n as above. Then $x_n \in S\ \forall n$ and $x_n \to x \in S^c \implies S$ is not closed (since it fails to contain all of its limits) $\implies$ contradiction. Therefore S^c is open.

Notice that some sets (like the space M itself) are both closed and open, they are referred to as **clopen** sets. By theorem above, $M^c = \varnothing$ is clopen. Now consider interval $[0, 1) \subset \mathbb{R}$. It is neither open (every r-neighborhood of 0 includes points in $[0, 1)^c \subset \mathbb{R}$) nor closed (it fails to include 1, which is the limit of the sequence $x_n = 1 - 1/n$ in $[0, 1)$). Thus subsets of a metric space can be open, closed, both, or neither.

Theorem: Arbitrary union of open sets is open.

Proof: Suppose $\{U_\alpha\}$is a collection of open sets in M, and let $U = \cup U_\alpha$. Then $x \in U$ $\implies x \in U_\alpha$ for some α and $\exists r > 0$ s.t. $d(x, y) < r \implies y \in U_\alpha \implies y \in U$. Therefore U is open.

Theorem: Intersection of finitely many open sets is open.

Proof: Suppose $U_1, U_2, \ldots, U_n$ are open sets in M. Let $U = \cap U_k$. If $U = \varnothing$ then U is open. Now suppose $x \in U$, then $x \in U_k$ for k = 1, 2, 3, . . . , n. Since each U_k is open, $\exists r_k > 0$

s.t. $d(x,y) < r_k \implies y \in U_k$. Let $r = min(r_1, r_2, \ldots, r_n)$, then $d(x,y) < r \implies y \in U_k \ \forall k$ $\implies y \in U$. Therefore U is open.

Notice that the infinite intersection of open sets is not necessarily open. For example, it's easy to see that $U_k = (-1/k, 1/k)$ is an open subset of $\mathbb{R}$, but $\cap U_k = \{0\}$ is clearly not open in $\mathbb{R}$.

Theorem:Arbitrary intersection of closed sets is closed. Also, finite union of closed sets is closed.

Proof: We'll use DeMorgan's Laws: $(\cup U_k)^c = \cap({U_k}^c)$. Let $\{K_\alpha\}$ be a collection of closed sets in M. Then $\cap K_\alpha = (\cup K_\alpha^c)^c$ and since each ${K_\alpha}^c$ is open, their union is open and it's complement is closed. The proof of the second part of the theorem is similar.

Notice that the infinite union of closed sets is not guaranteed to be closed. For example, even though each $K_k = [0, 1 - 1/k]$ is closed in $\mathbb{R}$, the intersection $\cup K_k = [0, 1)$ is not.

Theorem:Let M be a complete metric space, $N \subset M$ be closed. Then N is complete as a metric space in its own right.

Proof: Let (x_n) be a Cauchy sequence in N. Since (x_n) is also a Cauchy sequence in M, and M is complete, we have $x_n \to x \in M$. But N is closed in M, therefore $x \in N$ and we conclude that N is complete.

Let **lim** S denote the set of all limit points of S in M. It is pretty clear that $x \in$ lim S $\iff \forall r > 0$, $B_r(x) \cap S \neq \varnothing$ (you can easily prove it as an exersise). We can also show that lim S is a closed set. Let $\epsilon > 0$ be given and let (y_n) be a sequence of points in lim S s.t. $y_n \to y$ in M. Then $\exists N$ s.t. $n \geq N \implies d(y_n, y) < \epsilon/2$. Since each $y_n \in$ lim S, $\exists x_n \in S$ s.t. $d(y_n, x_n) < \epsilon/2$ $\forall n$. Then we have $n \geq N \implies$

$$d(x_n, y) \leq d(x_n, y_n) + d(y_n, y) < \epsilon/2 + \epsilon/2 = \epsilon$$

Therefore $x_n \to y$ and $y \in$ lim S. We conclude that lim S is closed.

We can also show that $B_r(x)$ is an open subset of M as follows: let $y \in B_r(x)$. Let s $=$ $r - d(x, y) > 0$. Then $d(y, z) < s \implies$

$$d(z, x) \leq d(x, y) + d(y, z) < d(x, y) + (r - d(x, y)) = r$$

$\implies z \in B_r(x)$. Thus if $y \in B_r(x)$ $\exists s$ s.t. $B_s(y) \subset B_r(x)$, and therefore $B_r(x)$ is open.

Theorem: Every open set $U \subset \mathbb{R}$ can be expressed as a countable disjoint union of open intervals of the form (a, b), where a is allowed to take on the value $-\infty$ and b is allowed to take on the vaule $+\infty$.

Proof:If $U = \varnothing = (0,0)$, then the statement is vacuously true. If U is not empty, $\forall x \in U$ define $a_x = inf\{a : (a,x) \subset U\}$, $b_x = sup\{b : (x,b) \subset U\}$. Then $I_x = (a_x, b_x)$ is a (possibly unbounded) open interval, containing x. It is maximal in the following sense: suppose $b_x \in U$, then by construction above $\exists J \subset U$, an open interval s.t. $b_x \in J$. You can prove for yourself that the union of two open intervals with non-empty intersection is in fact, an open interval (you just have to take care of a number of base cases for endpoints). We then have that $I_x \cup J$ is an open interval containing x, and moreover, since J is open, and $b_x \in J$, $\exists b^* \in J$ s.t $b^* > b_x$. But then $b^* \in \{b : (x,b) \subset U\}$, which contradicts b_x being the supremum of such a set. Thereofre $b_x \notin U$, and similarly $a_x \notin U$. Now let $x, y \in U$, and suppose $I_x \cap I_y \neq \varnothing$. Then once again, $I_x \cup I_y$ is an open interval containing both x and y and by maximality we have $I_x = I_x \cup I_y = I_y$. Thus we conclude that $\forall x, y \in U$ either $I_x = I_y$ or the two intervals are disjoint. So, U is a disjoint union of open intervals. To show that the union is countable, pick a rational number in each interval. Since the intervals are disjoint, the numbers are distinct, and their collection is countable (since rationals are countable).

A few more definitions are in order. As before, S is some subset of a metric space M. The **closure** of S is $\bar{S} = \cap K_\alpha$ where $\{K_\alpha\}$ is the collection of all closed sets that contain S. The **interior** of S is $\text{int}(S) = \cup U_\alpha$, where $\{U_\alpha\}$ is the collection of all open sets contained in S. Finally, the **boundary** of S is $\partial S = \bar{S} \cap \bar{S^c}$. For example, if M=$\mathbb{R}$ and S=$[a, b]$, then $\bar{S} = [a, b]$, $\text{int}(S) = (a, b)$, and $\partial S = a \cup b$. If S=$\mathbb{Q}$, then $\bar{S} = \mathbb{R}$, $\text{int}(S) = \varnothing$, and $\partial S = \mathbb{R}$. Notice that the closure of a closed set S is S istelf, and so is the interior of an open set S. Also notice that both $\bar{S}$ and ∂S are closed (as intersections of closed sets), and $\text{int}(S)$ is open as a union of open sets.

Theorem: $\bar{S}$=lim S.

Proof: We have shown before that lim S is a closed set, also since each point of S is also a limit of point of S (it's the limit of a sequence $s, s, s, \ldots$), we have $S \subset$lim S, and we conclude that $\bar{S} \subset$lim S. Also, since $S \subset \bar{S}$, and $\bar{S}$ is closed, it must contain all the limit points of S, thus lim S $\subset \bar{S}$. We conclude that $\bar{S} =$ lim S.

You may check for yourself at this point that every subset of a discrete metric space M (for example, $\mathbb{N}$ with discrete metric) is clopen (why would it suffice to show that a singleton $\{x\}$ is open?) and that therefore $\forall S \subset M$, $\text{int}(S) = S = \bar{S}$ and $\partial S = \varnothing$.

A subset S of a metric space M is said to **cluster** at point $x \in M$ if $\forall r > 0$ $B_r(x)$ contains infinitely many points of S. S is said to **condense** at x, if each $B_r(x)$ contains uncountably many points of S. For example if M=$\mathbb{R}$, then every point of S $= [a, b]$ is a condensation point (also a cluster point), and no other point is a cluster point. If S=$\{1/n : n \in \mathbb{N}\}$then the only cluster point is $0 \notin S$, and no point is a condensation point (no point could have uncountably many members of S in its r-neighborhood, since S itself is countable). If S=$\mathbb{Q}$, then x is a cluster point of S $\forall x \in \mathbb{R}$. An open interval (a, b) is an example of a subset of $\mathbb{R}$ that contains some but not all of its cluster/condensation points. Finally, $\mathbb{N}$ with either discrete metric or with the metric it inherits as subset of $\mathbb{R}$ has no cluster points. We emphasize that if x is a cluster point of S, each r-neighborhood of x must contain infinitely many **distinct** points of S. It is easy to see that x is a cluster point of S iff $\exists (x_n)$, a sequence of distinct points in S, s.t. $x_n \to x$. We denote the set of all cluster points of S by S'.

Theorem: $S \cup S' = \bar{S}$. S is closed iff $S' \subset S$.

Proof: We already know that $S \subset \bar{S}$. Moreover, by above we have that a cluster point is a limit point and therefore $S' \subset$lim S $= \bar{S}$. Thus $S \cup S' \subset \bar{S}$. Also, if $x \in \bar{S} =$ lim S, then either $x \in S$ or $\exists (x_n)$, a sequence of distinct points in S, s.t. $x_n \to x$, but that would make x a cluster point. So we have $\bar{S} \subset S \cup S'$ and we conclude that $S \cup S' = \bar{S}$. Now, we know that S is closed iff $S = \bar{S} = S \cup S'$ iff $S' \subset S$.

We notice here that if $S \subset \mathbb{R}$ is a bounded from above/below, then it's l.u.b/g.l.b are its cluster points (you can check it for yourself), and therefore by above theorem belong to the closure of S. Thus a closed and bounded subset of $\mathbb{R}$ contains both its l.u.b and g.l.b.

Up to this point, instead of just saying ‘S is open’ or ‘S is closed’, we often said ‘S is open in M’ or ‘S is closed in M’. We can drop the mention of M, when it is understood what the underlying metric space is, but we point out that it is essential to the openness/closedness of S. For example, both $\mathbb{Q}$ and the half-open interval $[a, b)$ are clopen when considered as metric spaces in their own right. Neither one, of course, is either open or closed when treated as a subset of $\mathbb{R}$. Another example is a set $S = \mathbb{Q} \cap (-\pi, \pi)$, a set of all rational numbers in the interval $(-\pi, \pi)$. As a subset of metric space $\mathbb{Q}$ S is both closed (if (x_n) is a sequence in S, and $x_n \to x \in \mathbb{Q}$ then $x \in S$) and open (check for yourself). As a subset of $\mathbb{R}$, however, it is neither open (if $x \in S$ then every neighborhood of x contains some $y \notin \mathbb{Q}$) nor closed (there are sequences in S converging to $\pi \in \mathbb{R}$). The following few theorems establish the relationship between being open/closed in metric space M and some metric subspace N of M that inherits its metric from M (i.e. $d_N(x, y) = d_M(x, y)$. We will denote the clousre of set S in M by $cl^M(S)$, and the closure of S in N by $cl_N(S)$. Note that $cl_N(S) = cl_M(S) \cap N$.

Theorem: If $S \subset N \subset M$, then S is closed in N iff $\exists L \subset M$ s.t. L is closed in M and $S = L \cap N$.

Proof: Suppose S is closed in N, then let $L = cl_M(S)$. Clearly L is closed in M, and

$L \cap N = cl_N(S) = S$ (since S is closed in N). Now suppose that L as in the statement of theorem exists. Since L is closed, it contains all of its limit points and $S = L \cap N$ contains all of its limit points in N, therefore S is closed in N.

Theorem If $S \subset N \subset M$, then S is open in N iff $\exists L \subset M$ s.t. L is open in M and $S = L \cap N$.

Proof:Notice that the complement of $L \cap N$ in N is $L^c \cap N$, where L^c si the complement of L in M. Now take complements and apply previous theorem.

A popular way to summarize the preceding two theorems is to say that metric subspace N **inherits** its opens and closeds from M.

We also introduce here the notion of boundedness. $S \subset M$ is **bounded** if $\exists x \in M$, $\exists r > 0$ s.t. $S \subset B_r(x)$, i.e S is bounded if there's some point $x \in M$ s.t. S is contained in some neighborhood of x. For example, $[-1, 1]$ is bounded in $\mathbb{R}$ since it's contained in $B_5(2)$ or $B_2(0)$. On the other hand, the graph of function $f(x) = sin(x)$ is an unbounded subset of $\mathbb{R}^2$, although the range of f is a bounded subset of $\mathbb{R}$ (range $= [-1, 1]$). In general, we say that f is a **bounded function** if its range is a bounded subset of the target space.

Theorem: Let (x_n) be a Cauchy sequence in M. Then $S = \{x \in M : x = x_n \text{ for some } n\}$ is bounded. In other words, Cauchy sequences are bounded.

Proof: Let $\epsilon = 1$. Then $\exists N$ s.t. $n, m \geq N \implies d(x_n, x_m < 1$, in particular $n \geq N \implies d(x_n, x_N) < 1$. Now, let $r = 1 + max\{d(x_1, x_2), d(x_1, x_3), \ldots, d(x_1, x_N)\}$. Clearly $\forall 1 \leq k \leq N$ $d(x_1, x_k) < r - 1 < r$. For $k \geq N$, $d(x_1, x_k) \leq d(x_1, x_N) + d(x_N, x_k) < r - 1 + 1 = r$. Thus $S \subset B_r(x_1)$ and the sequence is bounded.

As a consequence of preceding theorem, all convergent sequences are bounded (since convergence implies Cauchy).

3.4 Continuous Functions

Let M, N be two metric spaces. A function $f : M \to N$ is **continous at** $x \in M$ if $\forall \epsilon > 0$ $\exists \delta > 0$ s.t. $y \in M$ and $d_M(x, y) < \delta \implies d_N(f(x), f(y)) < \epsilon$. We say that f is **continuous on M** if it's continuous at every $x \in M$. Notice that a specific δ depends on **both** x and ϵ. If the choice of δ does **not** depend on x, then we have the following definition: a function $f : M \to N$ is **uniformly continuous** if $\forall \epsilon > 0$, $\exists \delta > 0$s.t. $y \in M$ and $d_M(x, y) < \delta \implies d_N(f(x), f(y)) < \epsilon$.

For example, function 1/x on the interval $(0, 1)$ is continuous (note that 0 is outside of the domain), but is not uniformly continuous (given ϵ, no matter how small we choose δ to be, there are always points x, y in the interval $(0, \delta)$ s.t. $|f(x) - f(y)| > \epsilon$).

The following theorem provides some alternative characterizations of continous functions:

Theorem: The following definitions of continuity are equivalent:

1. ϵ, δ definition

2. $f : M \to N$ is continuous if for each convergent sequence $x_n \to x$ in M, we have $f(x_n) \to f(x)$ in N. Thus continous functions send convergent sequences to convergent sequences, preserving the limits.

3. $f : M \to N$ is continuous if $\forall$ closed $S \subset N$, $f^{-1}(S)$ is closed in M. (Here f^{-1} is the notation for preimage of a set in a target space).

4. $f : M \to N$ is continuous if $\forall$ open $S \subset N$, $f^{-1}(S)$ is open in M.

Proof: $1 \Longrightarrow 2$: Suppose f is continuous, (x_n) is a sequence in M s.t. $x_n \to x$. Let $\epsilon > 0$ be given. By 1, we know that $\exists \delta > 0$ s.t. $d_M(x, y) < 0 \Longrightarrow d_N(f(x), f(y)) < \epsilon$. Since $x_n \to x$ $\exists L$ s.t. $n \geq L \Longrightarrow d_M(x_n, x) < \delta$. But that means that $n \geq L \Longrightarrow d_N(f(x_n), f(x)) < \epsilon$, and therefore $f(x_n) \to f(x)$.

$2 \Longrightarrow 3$: Let $S \subset N$ be closed in N, and let (x_n) be a sequence in $S^{-1} = f^{-1}(S)$, s.t. $x_n \to x \in M$. We need to show that $x \in S^{-1}$. By 2 and by the fact that $x_n \to x$, we know that $f(x_n) \to f(x)$. Since $(f(x_n))$ is a sequence in S and S is closed, we have $f(x) \in S$, and therefore $x \in S^{-1}$.

$3 \Longrightarrow 4$: Let $S \subset N$ be open in N. Then $f^{-1}(S) = (f^{-1}(S^c))^c$, and since the latter is the complement of a closed set (by 3), it is open in M.

$4 \Longrightarrow 1$: Let $x \in M$ and $\epsilon > 0$ be given. Then we know that $B_\epsilon(f(x))$ is open in N and therefore $f^{-1}(B_\epsilon(f(x))$ is open in M (by 4). Since $x \in f^{-1}(B_\epsilon(f(x))$ and $f^{-1}(B_\epsilon(f(x))$ is open, $\exists \delta > 0$ s.t. $B_\delta(x) \subset f^{-1}(B_\epsilon(f(x))$, i.e. $d_M(x, y) < \delta \Longrightarrow d_N(f(x), f(y)) < \epsilon$.

Theorem: Composite of continuous functions is continuous.

Proof: Let $f : M \to N$ and $g : N \to L$ be continuous, and let $U \subset L$ be open in L, and denote $h = g \circ f$ $M \to L$. Then $g^{-1}(U)$ is open in N and $f^{-1}(g^{-1}(U)) = h^{-1}(U)$ is open in M by definition 4 above. We conclude that h is continuous by definition 4.

It is worth pointing out that while continuous functions preserve the convergent sequences, they in general do not preserve the **non-convergent** Cauchy sequences. For example, the continuous funcion $f : (0, 1] \to \mathbb{R}$ given by $f(x) = 1/x$ maps the Cauchy sequence 1,1/2,1/3,1/4,... in $(0, 1]$ to non-Cauchy sequence 1,2,3,4,... in $\mathbb{R}$. Uniform continuity ensures the preservation of Cauchy sequences.

Theorem: Let M, N be metric spaces, $f : M \to N$ a uniformly continuous function, and (x_n) a Cauchy sequence in M. Then $(f(x_n))$ is a Cauchy sequence in N.

Proof: Let $\epsilon > 0$ be given. Then since f is uniformly continuous, $\exists \delta > 0$ s.t. $d_M(x, y) < \delta \implies d_N(f(x), f(y)) < \epsilon$. Since (x_n) is Cauchy, $\exists L$ s.t. $m, n > L \implies d_M(x_m, x_n) < \delta \implies d(f(x_n), f(x_m)) < \epsilon \implies (f(x_n))$ is Cauchy.

You can show pretty easily that **every** function defined on a discrete metric space is uniformly continuous (why?).

Example: Consistent Estimates

In statistics, we usually estimate parameters of interest from the sample we have at hand. Suppose that $\hat{\theta}_n$ is the estimate based on sample of size n, then we say that the estimate is **consistent in probability** if $\forall\ \epsilon > 0$, $P(|\hat{\theta}_n - \theta| > \epsilon) \to 0$ as $n \to \infty$. For example, the Weak Law of Large Numbers states that if we have a sequence of IID random variables $X_1, X_2, \ldots, X_n$ with expected value μ and variance σ^2, then the sample mean $\hat{\mu}_1 = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is a consistent estimate of μ. Moreover, if we consider the sequence ${X_1}^k, {X_2}^k, \ldots, {X_n}^k$, then by applying the Weak Law again, we conclude that $\hat{\mu}_k = \overline{X^k} = \frac{1}{n}\sum_{i=1}^{n} {X_i}^k$ is a consistent estimate of k-th moment of X, $\mu_k = E(X^k)$.

Theorem: Suppose $\hat{\theta}_n$ is a consistent estimate of θ, and suppose that f is continuous. Then $\hat{\rho}_n = f(\hat{\theta}_n)$ is a consistent estimate of $\rho = f(\theta)$.

Proof: Suppose not. Then for some $\epsilon^* > 0$, $P(|f(\hat{\theta}_n) - f(\theta)| > \epsilon^*) \nrightarrow 0$, i.e. for some $L > 0$ $\exists$ subsequence $(\hat{\theta}_{n_k})$ of $(\hat{\theta}_n)$ s.t. $P(|f(\hat{\theta}_{n_k}) - f(\theta)| > \epsilon^*) > L$. Now, since f is continuous, $\exists\ \delta > 0$ s.t. $|\hat{\theta}_n - \theta| \leq \delta \implies |f(\hat{\theta}_n) - f(\theta)| \leq \epsilon^*$, and therefore we have that $|f(\hat{\theta}_n) - f(\theta)| > \epsilon^* \implies |\hat{\theta}_n - \theta| > \delta$. Also, since $\hat{\theta}_n$ is a consistent estimate of θ, we know that $\exists N$ s.t. $n \geq N \implies P(|\hat{\theta_n} - \theta| > \delta) < L$, in particular this holds $\forall\ n_k \geq N$. Combining the last two results with the fact that $(A \implies B) \implies (P(A) \leq P(B))$, we have $n_k \geq N$ $\implies L < P(|f(\hat{\theta}_{n_k}) - f(\theta)| > \epsilon^*) \leq P(|\hat{\theta}_{n_k} - \theta| > \delta) < L$, which leads to a contradiction. Therefore we conclude that $\hat{\rho}_n = f(\hat{\theta}_n)$ is indeed a consistent estimate of $\rho = f(\theta)$.

This result is very useful in general and is the foundation of estimation technique known as **Method of Moments**. Once again, suppose we have a sequence of IID random variables $X_1, X_2, \ldots, X_n$ with expected value μ and variance σ^2. We know that $\bar{X}$ is a consistent estimate of μ and by above result we have $\bar{X}^2$ is a consistent estimate of μ^2. Then you can prove yourself that $\overline{X^2} - \bar{X}^2$ is a consistent estimate of $\sigma^2 = Var(X)$ (do it!). In general, if there is a continuous function $\theta = f(\mu_1, \mu_2, \ldots, \mu_k)$, then $\hat{\theta} = f(\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k)$ is a consistent estimate of θ. Therefore if we can express a parameter of interest as a continuous function of the moments of distribution, then applying function to sample moments will give us a consistent estimate of the parameter. For example, if we are sampling from $N(\mu, \sigma^2)$ distribution and we have a sample of size n, then $\mu = \mu_1$ and $\sigma^2 = E(X^2) - \mu^2 = \mu_2 - {\mu_1}^2$, and therefore $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = \left(\frac{1}{n} \sum_{i=1}^{n} X_i^2 \right) - \bar{X}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$ are consistent estimates of μ and σ^2, respectively.

3.5 Product Metrics

Let $M = M_1 \times M_2$ be the Cartesian product of metric spaces M_1, M_2, i.e. M is the set of all points $x = (x_1, x_2)$ s.t. $x_1 \in M_1$ and $x_2 \in M_2$. For example, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is the set of all points in the plane. How would one define a useful metric on this product space? Three of the possibilities are listed below:

1. $d_M(x, y) = \sqrt{d_{M_1}(x_1, y_1)^2 + d_{M_2}(x_2, y_2)^2} = d_E(x, y)$, a Euclidean metric

2. $d_M(x, y) = max\{d_{M_1}(x_1, y_1), d_{M_2}(x_2, y_2)\} = d_{max}(x, y)$

3. $d_M(x, y) = d_{M_1}(x_1, y_1) + d_{M_2}(x_2, y_2) = d_{sum}(x, y)$

Continuing with our example, if we took points $x = (-5, 2)$ and $y = (7, 11)$ in $\mathbb{R}^2$, then $d_E(x, y) = \sqrt{12^2 + 9^2} = 15$, $d_{max}(x, y) = max\{12, 9\} = 12$, and $d_{sum}(x, y) = 12 + 9 = 21$, where we use the usual distance metric on $\mathbb{R}$. It turns out that in a way, all these metrics are equivalent.

Theorem: $d_{max}(x, y) \leq d_E(x, y) \leq d_{sum}(x, y) \leq 2d_{max}(x, y)$.

Proof Some basic arithmetic.

Theorem: Let $M = M_1 \times M_2$, and let $(x_n) = ((x_{1,n}, x_{2,n}))$ be a sequence in M. Then (x_n) converges with respect to (wrt) d_{max} iff it converges wrt d_E iff it converges wrt d_{sum} iff both $(x_{1,n})$ and $(x_{2,n})$ converge in M_1 and M_2, respectively.

Proof: The equivalence of the first 3 convergences is obvious from previous theorem. For example, suppose (x_n) converges to some $x \in M$wrt d_{max}. Let $\epsilon > 0$ be given, then $\exists N$ s.t. $n \geq N \implies d_{max}(x_n, x) < \epsilon/2 \implies d_E(x_n, x) \leq d_{sum}(x_n, x) \leq \epsilon$. The convergence of component sequences is obviously equivalent to convergence wrt d_{max}.

In particular, if $x_n \to x$ and $y_n \to y$ in $\mathbb{R}$, then $z_n = (x_n, y_n) \to z = (x, y)$ in $\mathbb{R}^2$.

Theorem: The above results extend to the Cartesian product of $n > 2$ metric spaces.

Proof: By induction.

Theorem: $\mathbb{R}^m$ is complete.

Proof: Let (x_n) be a Cauchy sequence in $\mathbb{R}^m$ wrt any of the above metrics. Then each component sequence $(x_{k,n})$ for $k = 1, 2, 3, \ldots, m$ is Cauchy in $\mathbb{R}$ (by same reasoning as convergence equivalence theorem above), and since $\mathbb{R}$ is complete, all component sequences converge, $x_{k,n}$ converges to some x_k $\forall k$, therefore $x_n \to (x_1, x_2, \ldots, x_m)$.

3.6 Connectedness

A metric space M is said to be **disconnected** if it can be written as a disjoint union of two non-empty clopen sets. Note that it is sufficient to discover one proper (i.e $\neq \varnothing$ or M itself) clopen subset S of M, since its complement would also have to be clopen and proper (hence non-empty), and we'd have $M = S \sqcup S^c$, where symbol $\sqcup$ indicates disjoint union. M is said to be **connected** if it is not disconnected. $S \subset M$ is said to be connected if it's connected when considered as metric space in its own right (with metric inherited from M), and is disconnected otherwise. For example, interval $(2, 5)$ is a connected subset of $\mathbb{R}$, but $(2, 5) \cup (5, 8)$ is not. $\mathbb{Q}$ is also disconnected, since $\mathbb{Q} = (\mathbb{Q} \cap (-\infty, \pi)) \sqcup (\mathbb{Q} \cap (\pi, \infty))$.

Theorem: Suppose M is connected, and $f : M \to N$ is a continuous function onto N. Then N is connected.

Proof: We note here first that $f : M \to N$ is **onto** if $\forall y \in N$ $\exists x \in M$ s.t. $y = f(x)$, i.e. every point in N is image under f of some point in M (there could be many points in M all mapping to same point in N). Now, suppose N is not connected, and let $S \subset N$ be a proper clopen subset, then $f^{-1}(S)$ is a non-empty (since f is onto), clopen (since f is continuous) subset of M and so is $f^{-1}(S^c)$, resulting in $M = (f^{-1}(S)) \sqcup (f^{-1}(S^c))$ is a disjoint union of proper clopen subsets, contradicting the connectedness of M. We conclude that N is connected.

Thus a continuous image of a connected set is connected. A discontinuous image of a connected set need not be connected, however. For example if $f : \mathbb{R} \to \mathbb{R}$ sends all negative numbers to 0 and all non-negative numbers to 1, then the image is $\{0, 1\}$ a finite set, which is clearly disconnected. Of course this example only works if $\mathbb{R}$ was connected to begin with.

Theorem: $\mathbb{R}$ is connected.

Proof: Suppose $\mathbb{R}$ is disconnected. Then $\exists S \subset \mathbb{R}$ s.t. S is proper and clopen. Since S is opein in $\mathbb{R}$ we know that S is a countable disjoint union of open intervals. Let (a, b) be one of these intervals (we know that S is non-empty since it's proper, so (a, b) exists). We have seen in the proof of the theorem about every open set U in $\mathbb{R}$ being the countable disjoint union of open intervals that the endpoints of hte intervals do not belong to U, thus $a, b \notin S$. Suppose $b < \infty$, then b is clearly a limit point of of S and $b \notin S$ contradicts S being closed. Therefore $b = \infty$, and similarly $a = -\infty$. But then $S = \mathbb{R}$ and S is not proper and we arrive at contradiction. Thus, $\mathbb{R}$ is connected.

Theorem: Open and closed intervals in $\mathbb{R}$ are connected.

Proof: Let $(a, b) \subset \mathbb{R}$. We know that $f : \mathbb{R} \to (-\pi/2, \pi/2)$ given by $f(x) = tan^{-1}(x)$ is continuous (you can verify that for yourself). You are also invited to find continuous function $g : (-\pi/2, \pi/2) \to (a, b)$ (which is easy). Then $g \circ f$ is continuous as a composition of continuous functions and therefore (a, b) is connected as a continuous image of the connected set $\mathbb{R}$.

Now, let $[a, b] \subset \mathbb{R}$. Then define $f : \mathbb{R} \to [a, b]$ as follows:

$$f(x) = \begin{cases} a & \text{if } x < a \\ x & \text{if } a \leq x \leq b \\ b & \text{if } x > b \end{cases}$$

Clearly f is continuous and therefore $[a, b]$ is connected.

Theorem: Suppose that $\{S_\alpha\}$ is a collection of connected sets and that $\cap S_\alpha \neq \varnothing$. Then $S = \cup S_\alpha$ is connected.

Proof: Let $x \in \cap S_\alpha$. Now suppose that S is disconnected, $S = A \sqcup A^c$, where A and A^c are disjoint non-empty clopen subsets of S. Now, $x \in \cap S_\alpha \implies x \in S_\alpha$ for some $\alpha \implies$

$x \in S = \cup S_\alpha \Longrightarrow x \in A$ or $x \in A^c$. Without loss of generality, assume $x \in A$. Since each S_α is a subset of S, S_α inherits its closed and open sets from S $\forall \alpha$ by the inheritance theorem, and since $x \in S_\alpha$ $\forall \alpha$ we have $A \cap S_\alpha$ is a non-empty clopen subset of S_α $\forall \alpha$. Since each S_α is connected, we conclude that $A \cap S_\alpha = S_\alpha$ $\forall \alpha$ and therefore $A = \cup(A \cap S_\alpha) = \cup S_\alpha = S$, contradicting A being a proper subset of S. We conclude that S is connected.

Theorem: Let $S \subset M$ be connected. Then $S \subset T \subset \bar{S} \Longrightarrow T$ is connected. In particular, the closure of a connected set is connected.

Proof: Suppose T is disconnected, $T = A \sqcup A^c$, where A and A^c are proper clopen subsets of T. By intheritance principle, $A \cap S$ is a clopen subset of S. Since S is connected, $B = A \cap S$ cannot be proper, therefore either $B = S$ or $B = \varnothing$. Without loss of generality, suppose $B = \varnothing$ (if $B = S$, just look at $C = A^c \cap S$, the complement of B in S, which leads to $C = \varnothing$). Since $A \cap S = \varnothing$, we have $S \subset A^c$ in T. But $A \neq \varnothing$ and $A \subset T \subset \bar{S} \Longrightarrow \exists x \in A$ s.t. x is a limit point of S. Since A is open, $\exists r > 0$, s.t. $B_r(x) \subset A$. But x is a limit point of S and therefore $B_r(x) \cap S \neq \varnothing \Longrightarrow A \cap A^c \neq \varnothing$, and we arrive at a contradiction. We conclude that T is connected.

Now we introduce the **intermediate value property**. Let $f : M \to \mathbb{R}$ be a function. Then f is said to have intermediate value property if $\forall x, y \in M$ it is true that if

$f(x) = a < b = f(y)$ then $\forall c \in (a,b)$ $\exists z \in M$ s.t. $f(z) = c$. In other words, if function assumes two distinct values in $\mathbb{R}$ then it also has to assume all the values in-between to have the intermediate value property.

Theorem: Let M be connected, and $f : M \to \mathbb{R}$ be continuous. Then f has the intermediate value property. In particular, every continuous function $f : \mathbb{R} \to \mathbb{R}$ has the intermediate value property.

Proof: Suppose not. Then $\exists x, y \in M$ s.t. $f(x) = a < b = f(y)$ and $\exists c \in (a,b)$ s.t $\forall z \in M$ $f(z) \neq c$. Let $A = (-\infty, c)$, and $A^* = (c, \infty)$. Then $M = B \sqcup B^c$, where $B = f^{-1}(A)$ and $B^c = f^{-1}(A^*)$ is the complement of B in M. Note that $x \in B$ and $y \in B^c$, so $B \neq \varnothing \neq B^c$. Since f is continuous, the preimage of an open set is open and since both A and A^* are open subsets of $\mathbb{R}$, we have that B and B^c are both open, and therefore also both closed as complements of each other. We conclude that M is a disjoint union of non-empty clopen sets and is therefore disconnected, which contradicts our assumption. Thus, we conclude that f has the intermediate value property.

3.7 Compactness

We say that $S \subset M$ is **sequentially compact** if every sequence (x_n) in S, has a convergent subsequence (x_{n_k}) s.t. $x_{n_k} \to x$ for some $x \in S$. For example, every finite subset

$S = \{x_1, x_2, \ldots, x_n\}$ of a metric space M is compact, since any sequence in S has to repeat at least one x_k infinitely many times, and that results in existence of at least one convergent subsequence $x_k, x_k, x_k, \ldots$ Notice also that the convergent subsequence or indeed the limit need not be unique, as the example of a sequence 1,2,1,2,1,2,1,2, ... in $S = \{1, 2\} \subset \mathbb{N}$ shows. Also, $\mathbb{N}$ is not a compact subset of $\mathbb{R}$ since the sequence 1,2,3,4,5,... does not converge to a point in $\mathbb{N}$. Other non-compact subsets of $\mathbb{R}$ include $(0, 1]$ (sequence 1, 1/2, 1/3, ... does not have a subsequence that would converge to a point in $(0, 1]$) and $\mathbb{Q}$ (every subsequence of the sequence 3, 3.1, 3.14, 3.141, ... converges to $\pi \notin \mathbb{Q}$).

We say that a collection $\{U_\alpha\}$ of open subsets of M is an **(open) cover** for $S \subset M$, if $\forall x \in S\ \exists \alpha$ s.t $x \in U_\alpha$. We say that a collection $\{V_\beta\}$ is a **subcover** of $\{U_\alpha\}$ if $\forall \beta\ V_\beta = U_\alpha$ for some α, i.e. $\{V_\beta\} \subset \{U_\alpha\}$. We then also say that $\{U_\alpha\}$ **reduces to** $\{V_\beta\}$. We say that $S \subset M$ is **covering compact** if **every** open cover $\{U_\alpha\}$ reduces to finite subcover. In other words, if we're given any open cover for S whatsoever, and we can always throw away enough members of it, so that we're left with only finitely many and they still cover all of S then S is covering compact. Note that every subset S of M has at least one open cover, namely $\{M\}$, since M is open and definitely covers S (this particular open cover is already finite). Also, in order to see that not all subsets are covering compact, consider the cover of $(0, 1] \subset \mathbb{R}$ by open intervals $U_n = (1/n, 1 + 1/n)$. Clearly, $\{U_n : n \in \mathbb{N}\}$ is an open cover for $(0, 1]$, but you can easily see that it can't be reduced to a finite subcover, since in that case we'd be left with $\{U_{n_1}, U_{n_2}, \ldots, U_{n_k}\}$ and letting $N = max\{n_1, n_2, \ldots, n_k\}$, we observe that $\forall x \in (0, 1/N)$ x is not contained in any of $U_{n_1}, U_{n_2}, \ldots, U_{n_k}$.

Now, let $S \subset M$ and let $\{U_\alpha\}$ be some open cover for S. Then if $\exists \lambda > 0$ s.t. $\forall x \in S$ $B_\lambda(x) \subset U_\alpha$ for some α we say that λ is a **Lebesgue number** for $\{U_\alpha\}$. In other words, a Lebesgue number for an open cover of S is some small radius, s.t. neighborhood of every point in S of that radius is contained in some member of the cover (obviously, which member it is, depends on the particular point). Notice that a given cover for a given set might not have a Lebesgue number. For example, $(0, 1) \subset \mathbb{R}$ has as one possible cover $\{(0, 1)\}$, i.e. it's covered by itself. Suppose it had Lebesgue number λ, then pick $x \in (0, \lambda)$. Since $B_\lambda(x) = (x - \lambda, x + \lambda)$ contains negative points, we have $B_\lambda(x) \not\subseteq (0, 1)$, and thus this neighborhood is not contained in any member of the cover, leading to contradiction.

Theorem Every open covering of a sequentially compact $S \subset M$ has a Lebesgue number.

Proof: Suppose not. Let $\{U_\alpha\}$ be an open cover for S s.t. $\forall \lambda_n$ $\exists x_n \in S$ s.t. $B_{\lambda_n}(x_n) \not\subseteq U_\alpha$ $\forall \alpha$. Let $\lambda_n = 1/n$, and x_n be as above. Then since S is sequentially compact, (x_n) has some convergent subsequence $(x_{n_k}) \to x \in S$. Since $\{U_\alpha\}$ is a cover for S, $x \in U_\alpha$ for some α. Since U_α is open, $\exists r > 0$ s.t. $B_r(x) \subset U_\alpha$. Now, since $x_{n_k} \to x$, $\exists N_1$ s.t. $n_k \geq N_1 \implies d(x, x_{n_k}) < r/2$. Moreover, $\exists N_2 \geq N_1$ s.t $n_k \geq N_2 \implies \lambda_{n_k} < r/2$. Now pick some x_{n_N} s.t. $n_N > N_2$. Let $y \in B_{\lambda_{n_N}}(x_{n_N})$, then:

$$d(x,y) \leq d(x,x_{n_N}) + d(x_{n_N},y) < r/2 + \lambda_{n_N} < r/2 + r/2 = r$$

We conclude that $B_{\lambda_{n_N}}(x_{n_N}) \subset B_r(x) \subset U_\alpha$, contrary to our assumption that $\forall n\ B_{\lambda_n}(x_n) \not\subseteq U_\alpha\ \forall\alpha$. Thus, every open covering of a sequentically compact set does indeed have a Lebesgue number.

Theorem: $S \subset M$ is sequentially compact iff it is covering compact.

Proof: Suppose S is covering compact, but not sequentially compact. Then let (x_n) be a sequence in S s.t. no subsequence of (x_n) converges to a point in S. That implies that $\forall x \in S\ \exists r_x > 0$ s.t. $B_{r_x}(x)$ contains only finitely many terms of (x_n) (otherwise there'd be a subsequence converging to x). Now, $\{B_{r_x}(x)\}$ is an open cover for S and since S is covering compact, it reduces to a finite subcover $\{B_{r_{x_1}}(x_1), B_{r_{x_2}}(x_2), \ldots, B_{r_{x_k}}(x_k)$ and since each member of this subcover contains only finitely many terms of (x_n), we conclude that S contains finitely many terms of (x_n), an obvious contradiction. Thus covering compactness implies sequential compactness.

Now suppose S is sequentially compact, and let $\{U_\alpha\}$ be some open cover for S. We know that $\{U_\alpha\}$ has some Lebesgue number $\lambda > 0$. Pick $x_1 \in S$ and some $U_1 \in \{U_\alpha\}$ s.t. $B_\lambda(x_1) \subset U_1$. If $S \subset U_1$ then we have succeded in reducing $\{U_\alpha\}$ to a finite subcover. If not, then pick uncovered point $x_2 \in S$ ($x \in S \cap {U_1}^c$) and U_2 s.t. $B_\lambda(x_2) \subset U_2$. If $S \subset U_1 \cup U_2$, we're done, if not we continue picking uncovered points to obtain a sequence of points (x_n) in S and a sequence (U_n) of members of $\{U_\alpha\}$ s.t. $B_\lambda(x_n) \subset U_n$ and $x_{n+1} \in A \cap {U_1}^c \cap {U_2}^c \cap \ldots \cap {U_n}^c$. If at some point the sequences terminate (if for some $N, A = (U_1 \cup U_2 \cup \ldots \cup U_N)$) then we have reduced $\{U_\alpha\}$ to a finite subcover. Now suppose, the sequences never terminate. Then since S is sequentially compact, there is some subsequence (x_{n_k}) of (x_n) s.t. $x_{n_k} \to x \in S$. Therefore, $\exists N$ s.t $n_k \geq N \implies d(x_{n_k}, x) < \lambda$. In particular, $d(x_N, x) < \lambda \implies x \in B_\lambda(x_N) \subset U_N$. Since U_N is open, $\exists r > 0$, s.t. $B_r(x) \subset U_N$. But $n_k > N \implies x_{n_k} \notin U_N$, and therefore $B_r(x)$ contains only finitely many terms of (x_{n_k}), a clear contradiction to convergence. We conclude that in fact, $\{U_\alpha\}$ reduces to a finite subcover, and sequential compactness implies covering compactness.

$S \subset M$ is said to be **compact** if it is sequentially/covering compact. In the proofs of the theorems about compactness that follow, we will use whichever definition of compactness leads to a more direct proof and we will supply two proofs for some of the theorems to give you better understanding of both kinds of compactness.

Theorem: Let $S_1 \subset M_1$ and $S_2 \subset M_2$ be compact. Then $S = S_1 \times S_2 \subset M_1 \times M_2 = M$ is compact.

Proof: Let (x_n, y_n) be a sequence in S (we avoid here cumbersome notation $((x_n, y_n))$).

Then (x_n) is a sequence in S_1 and therefore has some subsequence (x_{n_k}), that converges in S_1. Now (y_{n_k}) is then a sequence in S_2 and therefore has some subsequence $(y_{n_{k_l}})$ that converges in S_2. Moreover, since (x_{n_k}) was a convergent sequence in S_1, the subsequence $(x_{n_{k_l}})$ is also a convergent sequence in S_1. Therefore, the sequence (x_n, y_n) has a convergent subsequence $(x_{n_{k_l}}, y_{n_{k_l}})$ with respect to all product space metrics discussed above (since both component sequences converge), and we conclude that Cartesian product of two compact sets is compact.

Theorem: The Cartesian product of n compact sets is compact.

Proof: By induction and above theorem.

Theorem: Closed interval $[a, b] \subset \mathbb{R}$ is compact.

Proof 1: Let (x_n) be a sequence in $[a, b]$. Consider set $C = \{x \in [a, b] : x_n < x$ only finitely often $\}$. Clearly, $a \in C$, and b is an upper bound for C, therefore C has a least upper bound x^*. We will show that there is some subsequence (x_{n_k}) of (x_n) converging to x^*. Suppose $x^* = b$, then $\exists N$ s.t $n \geq N \implies x_n = b$, and clearly we have a subsequence converging to $b = x^*$. Now, suppose $x^* < b$ and no subsequence converges to x^*, then $\exists$ $b - x^* > r > 0$ s.t. $x_n \in B_r(x^*) = (x^* - r, x^* + r)$ only finitely often. But then $x^* + r \in C$ (note that $x^* + r \in [a, b]$), contradiction to x^* being the l.u.b for C. We conclude that in fact there is some subsequence converging to x^* and therefore $[a, b]$ is a compact subset of $\mathbb{R}$.

Proof 2: Let $\{U_\alpha\}$ be some open cover for $[a, b]$ and consider set $C = \{x \in [a, b]$: finitely many U_α would suffice to cover the interval $[a, x]\}$. Clearly $a \in C$ and b is an upper bound, therefore C has the l.u.b. x^*. Suppose $x^* < b$, and let $U_{\alpha_1}, U_{\alpha_2}, \ldots, U_{\alpha_n}$ be those finitely many members of $\{U_\alpha\}$ that suffice to cover $[a, x^*]$. Then $x^* \in U_{\alpha_k}$ for some k (not necessarily unique), and since U_{α_k} is open, $\exists r > 0$ s.t. $B_r(x^*) = (x^* - r, x^* + r) \subset U_{\alpha_k}$. Now $\forall y \in (x^*, x^* + r)$, y is covered by same $U_{\alpha_1}, \ldots, U_{\alpha_n}$ as x^*, and therefore picking some specific y we have that finitely many members of $\{U_{alpha}\}$ suffice to cover $[a, y]$, which contradicts x^* being the l.u.b. for C. We conclude that $x^* = b$ and therefore finitely many U_α cover $[a, b]$, and thus $[a, b]$ is compact.

Theorem: Closed box $[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$ is compact.

Proof: It is a finite Cartesian product of compact sets.

Theorem (Bolzano-Weierstrass): Any bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.

Proof: Any bounded sequence in $\mathbb{R}^n$ is contained in some closed box, and therefore has a subsequence that converges to some point in that box.

Theorem: Suppose M is compact and $S \subset M$ is closed. Then S is compact.

Proof 1: Let (x_n) be a sequence in S, then (x_n) is also a sequence in M and therefore has a convergent subsequence $x_{n_k} \to x \in M$. Since S is closed in M, we have $x \in S$, and therefore (x_n) has a subsequence that converges in S. We conclude that S is compact.

Proof 2: Let $\{U_\alpha\}$ be some open cover for S, then since S^c is open, $\{U_\alpha\} \cup \{S^c\}$ is an open cover for M and since M is compact, it reduces to a finite subcover for M and therefore it is a finite cover for S. Now there are two possibilities. First, S^c might not be a member of this finite subcover, then we have reduced $\{U_\alpha\}$ to a finite subcover for S. Now, suppose S^c is a member of the finite subcover we found ourselves with, then it is obvious that the rest of the finite subcover still covers S (since no element of S is contained in S^c) and by excluding S^c we are left with the finite subcover for S to which the original cover $\{U_\alpha\}$ is now reduced. We conclude that S is compact.

Theorem: Compact set $S \subset M$ is closed and bounded.

Proof: Suppose S is not closed, then $\exists (x_n)$, a sequence in S s.t. $x_n \to x \notin S$. Since S is compact, $\exists$ some subsequence $x_{n_k} \to x^* \in S$, but since (x_n) is a convergent sequence in M, all of its subsequences converge to the same limit in M, and we have that $x_{n_k} \to x$, and by uniqueness of limits we have $x = x^* \in S$, a contradiction to x not being in S. So S is closed. Now let $x \in M$ and suppose S is not bounded. Then $\forall n \in \mathbb{N}\ \exists x_n \in S$ s.t. $d(x, x_n) > n$. Since S is compact, $\exists$ some subsequence (x_{n_k}) of a sequence (x_n) in S, s.t. $x_{n_k} \to x^* \in S$. Since all convergent sequences are bounded, for some $r > 0$ and for some $y \in M$ (for example x^*), we have $x_{n_k} \in B_r(y)\ \forall n_k$. Now, let $r^* = d(x, y)$. We observe that by the property 3 of the metric, we have $d(x, x_{n_k}) \leq d(x, y) + d(y, x_{n_k}) \leq r^* + r$, contradictory to the fact that $d(x, x_{n_k}) \to \infty$ as $n_k \to \infty$. We conclude that S is in fact bounded.

The converse to that statement (every closed and bounded subset of a metric space is compact) is not true in general. For example, let $M = \mathbb{N}$ with discrete metric. Then M is a closed and bounded subset of itself ($x \in B_2(1) \forall x \in M$), but the sequence 1,2,3,4,... in M has no convergent subsequence, and therefore M is not compact. The converse is, however, true in $\mathbb{R}^n$ and a more general converse will be provided again later.

Theorem (Heine-Borel): Every closed and bounded $S \subset \mathbb{R}^n$ is compact.

Proof: Since S is bounded, it is contained in some closed box in $\mathbb{R}^n$, and is therefore a closed subset of a compact set. We conclude that S is compact.

We will now proceed to construct a more general converse statement. We start by introducing a new notion: $S \subset M$ is said to be **totally bounded** if $\forall \epsilon > 0$ there exists a finite covering of S by ϵ-neighborhoods. Notice that this definition is different from the definition of (non-total) boundedness given earlier.

Theorem: Let M be a complete metric space. Then $S \subset M$ is compact iff S is closed and totally bounded.

Proof: Suppose $S \subset M$ is compact. We have already shown that S has to be closed. Now let $\epsilon > 0$ be given, and consider the following open cover for S: $\{B_\epsilon(x) : x \in S\}$. Since S is compact, the cover reduces to a finite subcover and we have a finite covering of S by ϵ-neighborhoods, and we conclude that S is totally bounded.

Now suppose $S \subset M$ is closed and totally bounded. Let (x_n) be a sequence in S. Let $\epsilon_n = 1/n \ \forall n$. Since S is totally bounded, we have a finite covering for S by ϵ_1-neighborhoods $B_{\epsilon_1}(y_{1,1}), B_{\epsilon_1}(y_{1,2}), \ldots, B_{\epsilon_1}(y_{1,m_1})$ for some $y_{1,1}, y_{1,2}, \ldots, y_{1,m_1} \in M$. Then at least one of these neighborhoods contains infinitely many terms of the sequence (x_n), suppose $B_{\epsilon_1}(y_{1,k_1})$, and let N_1 be s.t. $x_{N_1} \in B_{\epsilon_1}(y_{1,k_1})$. Now, since every subset of a totally bounded set is totally bounded (you can easily show it), we have a finite covering of $B_{\epsilon_1}(y_{1,k_1})$ by ϵ_2-neighborhoods $B_{\epsilon_2}(y_{2,1}), B_{\epsilon_2}(y_{2,2}), \ldots, B_{\epsilon_2}(y_{2,m_2})$ and once again one of these neighborhoods contains infinitely many terms of the sequence (x_n), suppose $B_{\epsilon_2}(y_{2,k_2})$, and let $N_2 > N_1$ be s.t. $x_{N_2} \in B_{\epsilon_2}(y_{2,k_2})$. Continuing in this manner we obtain a subsequence (x_{N_k}) which is

Cauchy, since $n, m > N^* \implies d(x_{N_n}, x_{N_n}) < \epsilon_{N^*} = 1/N^*$. Since M is complete, $x_{N_k} \to x \in M$. Since S is closed in M, $x \in S$. Thus (x_n) has a subsequence that converges to a limit in S and we conclude that S is compact.

We note here that the conditions specified are indeed necessary. Take completeness, for example. Consider $S = \mathbb{Q} \cap [-\pi, \pi]$, a subset of a metric space $\mathbb{Q}$ with the usual distance metric. Then S is closed and totally bounded subset of $\mathbb{Q}$ but it's clearly not compact since sequence 3, 3.1, 3.14, 3.15, $\ldots$ in S doesn't have a subsequence that converges to a point in S or even in $\mathbb{Q}$ for that matter. Also, substituting boundedness for total boundedness in the theorem above would not suffice, since we have already seen that a complete metric space $\mathbb{N}$ with discrete metric is a closed and bounded non-compact subset of itself.

Below are some important results on the behavior of continuous functions on compact sets.

Theorem: The continuous image of a compact set is compact.

Proof 1: Let $f : M \to N$ be continuous, let $S \subset M$ be compact and let $f(S) \subset N$ denote the image of S under f. Let (y_n) be a sequence in $f(S)$, then $\forall n\ \exists x_n \in S$ s.t. $f(x_n) = y_n$. Since S is compact, the sequence (x_n) in S has some convergent subsequence $x_{n_k} \to x \in S$.

By continuity of f, we then have that $y_{n_k} = f(x_{n_k}) \to f(x) \in f(S)$, and therefore (y_n) has a subsequence that converges to the limit in $f(S)$, and we conclude that $f(S)$ is compact.

Proof 2: Let $f : M \to N$ be continuous, let $S \subset M$ be compact and let $f(S) \subset N$ denote the image of S under f. Let $\{U_\alpha\}$ be an open cover for $f(S)$ in N. Then since f is continuous, $f^{-1}(U_\alpha)$ is an open set in M $\forall \alpha$, and therefore $\{f^{-1}(U_\alpha)\}$ is an open cover for S in M. Since S is compact, $\{f^{-1}(U_\alpha)\}$ reduces to a finite subscover $f^{-1}(U_1), f^{-1}(U_2), \ldots, f^{-1}(U_n)$, and clearly $U_1, U_2, \ldots, U_n$ is an open cover for $f(S)$ (you can easily see it yourself), which is a finite subcover of the original cover. We conclude that $f(S)$ is compact.

Theorem: Continuous real-valued funciton defined on a compact set assumes its maximum and minimum.

Proof: Let $f : M \to \mathbb{R}$ be continuous and let $S \subset M$ be compact. Then $f(S) \subset \mathbb{R}$ is closed and bounded, and we have seen that a closed and bounded subset of $\mathbb{R}$ contains its g.l.b and its l.u.b.

Theorem: Every continous function defined on a compact set is uniformly continuous.

Proof: Let M be compact and $f : M \to N$ continuous. Suppose f is not uniformly continuous. Then $\exists \epsilon > 0$ s.t. $\forall \delta_n > 0$ $\exists x_n, y_n \in M$ s.t. $d(x_n, y_n) < \delta_n$ but $d(f(x_n), f(y_n)) > \epsilon$.

Let $\delta_n = 1/n$ and for each n, let $x_n, y_n \in M$ be as above. Since M is compact, the sequence (x_n) in M has some convergent subsequence (x_{n_k}). Moreover, the sequence (y_{n_k}) also has some convergent subsequence $y_{n_{k_l}} \to y \in M$. Since $(x_{n_{k_l}})$ is a subsequence of a convergernt sequence, it also converges and it converges to the same limit as the mother sequence, and since $d(x_{n_{k_l}}, y_{n_{k_l}}) < 1/n_{k_l}$, we can easily see that it also converges to y. The continuity of f implies that $f(x_{n_{k_l}}) \to y$ and $f(y_{n_{k_l}}) \to y$. Now, let N be s.t. $n_{k_l} \geq N \implies d(f(x_{n_{k_l}}), f(y)) < \epsilon/2$ and $d(f(y_{n_{k_l}}), f(y)) < \epsilon/2$, then we have $n_{k_l} \geq N \implies$

$$d(f(x_{n_{k_l}}), f(y_{n_{k_l}})) \leq d(f(x_{n_{k_l}}), f(y)) + d(f(y), f(y_{n_{k_l}}) < \epsilon/2 + \epsilon/2 = \epsilon$$

which contradicts our assumption that $d(f(x_n), f(y_n)) > \epsilon \ \forall n$. We then conclude that f indeed is uniformly continuous.

3.8 More Results on $\mathbb{R}$

We introduce some notation first. Given sequence (x_n) in $\mathbb{R}$ we say that $\lim(x_n) = x \in \mathbb{R}$ if $x_n \to x$ in $\mathbb{R}$. We say that $\lim(x_n) = \infty$ if $\forall M > 0 \ \exists N$ s.t. $n \geq N \implies x_n > M$. And similarly, $\lim(x_n) = -\infty$ if $\forall M < 0 \ \exists N$ s.t. $n \geq N \implies x_n < M$. Here are some examples:

1. $(x_n) =$ 3, 3.1, 3.14, 3.141, ... Then $\lim(x_n) = \pi$

2. $(x_n) =$ 2, 3, 5, 7, 11, ... Then $\lim(x_n) = \infty$

3. $(x_n) =$ -1, -4, -9, -16, ... Then $\lim(x_n) = -\infty$

4. (x_n) = -1, 1, -2, 2, -3, 3, ... Then $\lim(x_n)$ does not exist.

Given a sequence (x_n) in $\mathbb{R}$, we say that M =lim sup(x_n) if $M = \lim_{n\to\infty}(\sup_{k\geq n} x_k)$. We define $m = $ lim inf(x_n) similarly. Here are some examples:

1. $(x_n) = 1, 2, 3, 4, 5, \ldots$, then lim sup$(x_n)$ = lim inf $(x_n) = \infty$

2. $(x_n) = 1, 1, 2, 1, 2, 3, \ldots$, then lim sup$(x_n) = \infty$, lim inf$(x_n = 1$

3. $(x_n) = 3, 3.1, 3.14, 3.141, \ldots$, then lim sup$(x_n)$ = lim inf(x_n)= π

We say that sequence (x_n) is **monotone increasing** if $n > m \Longrightarrow x_n > x_m$. The sequence is **monotone non-decreasing** if $n > m \Longrightarrow x_n \geq x_m$. The monotone decreasing and monotone non-increasing sequences are defined similarly. A sequence is said to be **monotone** if it's monotone non-increasing or monotone non-decreasing (note that every monotone increasing sequence is also monotone non-decreasing, and every monotone decreasing sequences is monotone non-increasing, so these two classes are also covered by the definition).

1. (x_n) = 3, 3.1, 3.14, 3.141, ... is monotone increasing

2. (x_n) = -1, -4, -9, -16, ... is monotone decreasing

3. (x_n) = 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, ... is monotone non-decreasing

4. (x_n) = -1, -1, -1, -1, ... is monotone non-decreasing and monotone non-increasing

5. $(x_n) = 1, 1, 2, 1, 2, 3, 1, 2, 3, 4, \ldots$ is not monotone but it has many possible monotone subsequences

The following are some basic results about sequences in $\mathbb{R}$:

1. Let $x_n \to x \in \mathbb{R}$, and let $k \in \mathbb{R}$. Then $y_n = kx_n \to kx$.

2. Let $x_n \to x$ and $y_n \to y$ in $\mathbb{R}$. Then $z_n = x_n + y_n \to x + y$.

3. Let $x_n \to x$ and $y_n \to y$ in $\mathbb{R}$. Then $z_n = x_n y_n \to xy$.

4. Let $x_n \to x \neq 0$ in $\mathbb{R}$, s.t. $x_n \neq 0$ $\forall n$. Then $y_n = 1/x_n \to 1/x$.

5. Let $x_n \to x \neq 0$ in $\mathbb{R}$, s.t. $x_n \neq 0$ $\forall n$, and let $y_n \to y$ in $\mathbb{R}$. Then $z_n = y_n/x_n \to y/x$

6. Let $\lim(x_n) = \infty$, $\lim\,(y_n) > 0$. Then $\lim(z_n = x_n y_n) = \infty$

7. Let (x_n) be monotone and bounded. Then (x_n) converges in $\mathbb{R}$.

8. Let (x_n) be any sequence in $\mathbb{R}$. Then (x_n) has a monotone subsequence. Together this and the previous result imply that every bounded sequence in $\mathbb{R}$ has a convergent subsequence.

9. Let (x_n) be a monotone sequence. Then $\lim(x_n)$ exists (it could of course take on values of ∞ and $-\infty$)

10. Let f, g be real-valued funcitons $M \to \mathbb{R}$, continuous at some $x \in M$. Then $f + g$, fg, and f/g (asusming $g(x) \neq 0$) are all continuous at x.

11. Given a sequence (x_n), lim sup(x_n) and lim inf(x_n) are well defined (either they belong to $\mathbb{R}$ or take values of $\pm\infty$

12. $x_n \to x$ iff lim sup(x_n) = lim inf$(x_n) = x$.

The proofs of these results are left as exersises.

3.9 Calculus and Function Convergence

. We say that a real-valued funciton f is **differentiable** at x, if one of the following equivalent conditions holds:

1. $\lim_{t \to x} \dfrac{f(t) - f(x)}{t - x} = L$ exists

2. $\lim_{\Delta x \to 0} \dfrac{f(x + \Delta x) - f(x)}{\Delta x} = L$ exists

3. $\forall\ \epsilon > 0\ \exists\ \delta > 0$ s.t. $|t - x| < \delta \implies |\frac{f(t)-f(x)}{t-x} - L| < \epsilon$

We call L the **derivative** of f at x and write $f'(x) = L$. We say that f is differentiable on the interval (a, b) if it is differentialbe at every $x \in (a, b)$. The following are familiar results from calculus:

1. If f is differentiable at x, then f is continuous at x.

2. If f and g are differentiable at x, then so is $f + g$ and $(f + g)'(x) = f'(x) + g'(x)$

3. If f and g are differentiable at x, then so is $(f \cdot g)$, and $(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x)$

4. If $f(x) = c \ \forall \ x$, then $f'(x) = 0 \ \forall \ x$

5. If f and g are differentiable at x and $g(x) \neq 0$, then f/g is differentiable at x and $(f/g)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$

6. Chain Rule: If f is differentiable at x, and g is differentiable at $f(x)$, then $g \circ f$ is differentiable at x and $(g \circ f)'(x) = g'(f(x))f'(x)$

Theorem Let f be differentiable on (a, b) and suppose it achieves a maximum or minimum at some $c \in (a, b)$. Then $f'(c) = 0$.

Proof: We will prove the theorem for maximum, and proof for minimum is analogous. Let t approach c from above. Then $\frac{f(t)-f(c)}{t-c} \leq 0 \ \forall t$. If we let t approach c from below, then $\frac{f(t)-f(c)}{t-c} \geq 0 \ \forall t$. Since both expressions have to tend to the same limit $L = f'(c)$, we conclude that $f'(c) = 0$.

Theorem (Mean Value): Suppose f is continuous on $[a, b]$ and differentiable on (a, b). Then $\exists \ c \in (a, b)$ s.t. $f(b) - f(a) = f'(c)(b - a)$. In particular, if $|f'(x)| \leq M \ \forall \ x \in (a, b)$, then $\forall \ t, x \in (a, b)$ we have $|f(t) - f(x)| \leq M|t - x|$.

Proof: Let $S = \frac{f(b)-f(a)}{b-a}$ and let $g(x) = f(x) - Sx$. Clearly g is differentiable on (a, b) and continuous on $[a, b]$ (why?). Moreover, $g(a) = g(b)$. Since g is continuous on a compact set, it achieves maximum and minimum and since $g(a) = g(b)$ it achieves at least one of them at some $c \in (a, b)$. Then $0 = g'(c) = f'(c) - S \implies f(b) - f(a) = f'(c)(b - a)$.

We state the following result without proof (which can be found, for example in Pugh, p.145): If f is differentiable on (a, b), then f' has the intermediate value property. Also, all familiar results from calculus (derivative and integral formulas, L'Hopital's Rule, results on series convergence, etc...) hold.

We say that a sequence of functions $f_n : [a, b] \to \mathbb{R}$ **converges pointwise** to f ($f_n \to f$), if $\forall\ x \in [a, b]\ f_n(x) \to f(x)$ as $n \to \infty$. We say that a sequence $f_n : [a, b] \to \mathbb{R}$ **converges uniformly** to f ($f_n \rightrightarrows f$) if $\forall\ \epsilon > 0$, $\exists\ N$ s.t. $n \geq N \Longrightarrow |f_n(x) - f(x)| < \epsilon\ \forall\ x \in [a, b]$. You should convince yourself that the sequence of functions $f_n(x) = x^n$ on open interval $[0, 1]$ converges pointwise to the function

$$f(x) = \begin{cases} 0 & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x = 1 \end{cases}$$

but does not converge uniformly.

Theorem: If $f_n \rightrightarrows f$ and each f_n is continuous at x, then f is continuous at x.

Proof: Let $\epsilon > 0$ be given, and let N be s.t. $n \geq N \Longrightarrow |f_n(y) - f(y)| < \epsilon/3\ \forall\ y$. Since f_N is continuous at x, $\exists\ \delta > 0$ s.t. $|x - y| < \delta \Longrightarrow |f_N(x) - f_N(y)| < \epsilon/3$. Then $|x - y| < \delta$ $\Longrightarrow |f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$, and we conclude that f is continuous at x.

Note that our example above shows that pointwise convergence of continuous functions is not sufficient to ensure continuity of the limit.

We state two more results without proving them. The proofs can be found in Pugh.

1. Suppose $f_n \rightrightarrows f$, then $\int_a^b f_n(x)dx \to \int_a^b f(x)dx$ as $n \to \infty$.
2. Suppose $f_n \rightrightarrows f$ and $f_n' \rightrightarrows g$, then $g = f'$.

4 References

1. Pugh, C.C., 'Real Mathematical Analysis', Springer, 2002
2. Rice, John A., 'Mathematical Statistics and Data Analysis', 2nd ed., Duxbury Press, 1995
3. Ross, Kenneth A., 'Elementary Analysis: The Theory of Calculus', Springer, 2000