Chapter 2. Real Limits, Continuity and Differentiation

Limits of Sequences in $\mathbb{R}$

2.1 Definition: For $p \in \mathbb{Z}$, let $\mathbb{Z}_{\geq p} = \{k \in \mathbb{Z} \mid k \geq p\}$. A **sequence** in a set A is a function of the form $x : \mathbb{Z}_{\geq p} \to A$ for some $p \in \mathbb{Z}$. Given a sequence $x : \mathbb{Z}_{\geq p} \to A$, the k^{th} **term** of the sequence is the element $x_k = x(k) \in A$, and we denote the sequence x by

$$(x_k)_{k \geq p} = (x_p, x_{p+1}, x_{p+2}, \cdots).$$

Note that the range of the sequence $(x_k)_{k\geq p}$ is the set $\{x_k\}_{k\geq p} = \{x_k | k \geq p\}$.

2.2 Definition: Let $(x_k)_{k\geq p}$ be a sequence in $\mathbb{R}$. For $a \in \mathbb{R}$ we say that the sequence $(x_k)_{k\geq p}$ **converges** to a (or that the **limit** of $(x_k)_{k\geq p}$ is equal to a), and we write $x_k \to a$ (as $k \to \infty$), or we write $\lim\limits_{k\to\infty} x_k = a$, when

$$\forall\, 0 < \epsilon \in \mathbb{R}\ \exists\, m \in \mathbb{Z}_{\geq p}\ \forall\, k \in \mathbb{Z}_{\geq p}\ \big(k \geq m \Longrightarrow |x_k - a| < \epsilon\big).$$

We say that the sequence $(x_k)_{k\geq p}$ **converges** (in $\mathbb{R}$) when there exists $a \in \mathbb{R}$ such that $(x_k)_{k\geq p}$ converges to a. We say that the sequence $(x_k)_{k\geq p}$ **diverges** (in $\mathbb{R}$) when it does not converge (to any $a \in \mathbb{R}$). We say that $(x_k)_{k\geq p}$ **diverges to infinity**, or that the limit of $(x_k)_{k\geq p}$ is equal to **infinity**, and we write $x_k \to \infty$ (as $k \to \infty$), or we write $\lim\limits_{k\to\infty} x_k = \infty$, when

$$\forall\, r \in \mathbb{R}\ \exists\, m \in \mathbb{Z}_{\geq p}\ \forall\, k \in \mathbb{Z}_{\geq p}\ \big(k \geq m \Longrightarrow x_k > r\big).$$

Similarly we say that $(x_k)_{k\geq p}$ **diverges to** $-\infty$, or that the limit of $(x_k)_{k\geq p}$ is equal to **negative infinity**, and we write $x_k \to -\infty$ (as $k \to \infty$), or we write $\lim\limits_{k\to\infty} x_k = -\infty$ when

$$\forall\, r \in \mathbb{R}\ \exists\, m \in \mathbb{Z}_{\geq p}\ \forall\, k \in \mathbb{Z}_{\geq p}\ \big(k \geq m \Longrightarrow x_k < r\big).$$

2.3 Note: We shall assume that students are familiar with sequences and limits of sequences from first-year calculus. For example, students should know that if the limit of a sequence exists then it is unique. Also, the limit does not depend on the first few terms (indeed the first finitely many terms) and so we often omit the starting value p from our notation and write the sequence $(x_k)_{k\geq p}$ as (x_k). Students should also be able to calculate limits using various limit rules, such as Operations on Limits, the Comparison Theorem and the Squeeze Theorem (which can all be found in the Appendix).

2.4 Definition: Let (x_k) be a sequence in $\mathbb{R}$. For $b \in \mathbb{R}$, we say that the sequence (x_k) is **bounded above** by b when the set $\{x_k\}$ is bounded above by b, that is when $x_k \leq b$ for all k, and we say that the sequence (x_k) is **bounded below** by b when the set $\{x_k\}$ is bounded below by b, that is when $b \leq x_k$ for all k. We say (x_k) is **bounded above** when it is bounded above by some element $b \in \mathbb{R}$, we say that (x_k) is **bounded below** when it is bounded below by some $b \in \mathbb{R}$, and we say that (x_k) is **bounded** when it is bounded above and bounded below.

2.5 Definition: Let $(x_k)_{k\geq p}$ be a sequence in $\mathbb{R}$. We say that (x_k) is **increasing** (or **nondecreasing**) when for all $k, l \in \mathbb{Z}_{\geq p}$, if $k \leq l$ then $x_k \leq x_l$. We say that (x_k) is **strictly increasing** when for all $k, l \in \mathbb{Z}_{\geq p}$, if $k < l$ then $x_k < x_l$. Similarly, we say that (x_k) is **decreasing** (or **nonincreasing**) when for all $k, l \in \mathbb{Z}_{\geq p}$, if $k \leq l$ the $x_k \geq x_l$ and we say that (x_k) is **strictly decreasing** when for all $k, l \in \mathbb{Z}_{\geq p}$, if $k < l$ the $x_k > x_l$. We say that (x_k) is **monotonic** when it is either increasing or decreasing.

2.6 Theorem: *(Monotonic Convergence Theorem) Let* (x_k) *be a sequence in* $\mathbb{R}$.

(1) Suppose (x_k) *is increasing. If* (x_k) *is bounded above then* $x_k \to \sup\{x_k\}$*, and if* (x_k) *is not bounded above then* $x_k \to \infty$.
(2) Suppose (x_k) *is decreasing. If* (x_k) *is bounded below then* $x_k \to \inf\{x_k\}$*, and if* (x_k) *is not bounded below then* $x_k \to -\infty$.

Proof: We prove Part 1 in the case that $(x_k)_{k\geq p}$ is increasing and bounded above, say by $b \in \mathbb{R}$. Let $A = \{x_k | k \geq p\}$ (so A is the range of the sequence (x_k)). Note that A is nonempty and bounded above (indeed b is an upper bound for A). By the Least Upper Bound Property of $\mathbb{R}$, A has a supremum in $\mathbb{R}$. Let $a = \sup\{x_k | k \geq p\}$. Note that $a \geq x_k$ for all $k \geq p$ and $a \leq b$, by the definition of the supremum. Let $\epsilon > 0$. By the Approximation Property of the supremum, we can choose an index $m \geq p$ so that the element $x_m \in A$ satisfies $a - \epsilon < x_m \leq a$. Since (x_k) is increasing, for all $k \geq m$ we have $x_k \geq x_m$, so we have $a - \epsilon < x_m \leq x_k \leq a$ and hence $|x_k - a| < \epsilon$. Thus $\lim\limits_{k\to\infty} x_k = a \leq b$.

2.7 Definition: For $a, b \in \mathbb{R}$ with $a \leq b$ we write

$$\begin{gathered}(a,b) = \big\{x \in \mathbb{R}\big|a < x < b\big\}\ ,\ [a,b] = \big\{x \in \mathbb{R}\big|a \leq x \leq b\big\}\,,\\ (a,b] = \big\{x \in \mathbb{R}\big|a < x \leq b\big\}\ ,\ [a,b) = \big\{x \in \mathbb{R}\big|a \leq x < b\big\}\,,\\ (a,\infty) = \big\{x \in \mathbb{R}\big|a < x\big\}\ ,\ [a,\infty) = \big\{x \in \mathbb{R}\big|a \leq x\big\}\,,\\ (-\infty,b) = \big\{x \in \mathbb{R}\big|x < b\big\}\ ,\ (-\infty,b] = \big\{x \in \mathbb{R}\big|x \leq b\big\}\,,\\ (-\infty,\infty) = \mathbb{R}\,.\end{gathered}$$

An **interval** in $\mathbb{R}$ is any set of one of the above forms. In the case that $a = b$ we have $(a,b) = [a,b) = (a,b] = \emptyset$ and $[a,b] = \{a\}$, and these intervals are called **degenerate** intervals. The **nondegenerate** intervals contain at least two points. The intervals $\emptyset$, (a,b), (a,∞), $(-\infty,b)$ and $(-\infty,\infty)$ are called **open** intervals. The intervals $\emptyset$, $[a,b]$, $[a,\infty)$, $(-\infty,b]$ and $(-\infty,\infty)$ are called **closed** intervals. The intervals $\emptyset$, (a,b), $(a,b]$, $[a,b)$ and $[a,b]$ are **bounded** and the intervals (a,∞), $[a,\infty)$, $(-\infty,b)$, $(-\infty,b]$ and $(-\infty,\infty)$ are **unbounded**.

2.8 Theorem: *(Nested Interval Theorem) Let* $I_1, I_2, I_3, \cdots$ *be nonempty, closed bounded intervals in* $\mathbb{R}$*. Suppose that* $I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots$*. Then* $\bigcap\limits_{k=1}^{\infty} I_k \neq \emptyset$.

Proof: For each $k \geq 1$, let $I_k = [a_k, b_k]$ with $a_k \leq b_k$. For each k, since $I_{k+1} \subseteq I_k$ we have $a_k \leq a_{k+1} \leq b_{k+1} \leq b_k$. Since $a_k \leq a_{k+1}$ for all k, the sequence (a_k) is increasing. Since $a_k \leq b_k \leq b_{k-1} \leq \cdots \leq b_1$ for all k, the sequence (a_k) is bounded above by b_1. Since (a_k) is increasing and bounded above, it converges. Let $a = \sup\{a_k\} = \lim\limits_{k\to\infty} a_k$. Similarly, (b_k) is decreasing and bounded below by a_1, and so it converges. Let $b = \inf\{b_k\} = \lim\limits_{k\to\infty} b_k$. Since $a_k \leq b_k$ for all k, by the Comparison Theorem we have $a \leq b$, and so the interval $[a,b]$ is not empty. Since (a_k) is increasing with $a_k \to a$, it follows (we leave the proof as an exercise) that $a_k \leq a$ for all $k \geq 1$. Similarly, we have $b_k \geq b$ for all $k \geq 1$ and so $[a,b] \subseteq [a_k, b_k] = I_k$. Thus $[a,b] \subseteq \bigcap\limits_{k=1}^{\infty} I_k$, and so $\bigcap\limits_{k=1}^{\infty} I_k \neq \emptyset$.

2.9 Note: The above theorem does not hold for bounded open intervals. For example, for $I_k = \left(0, \frac{1}{k}\right)$ we have $I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots$ but $\bigcap\limits_{k=1}^{\infty} I_k = \emptyset$. The theorem also does not hold for unbounded closed intervals. For example, consider $I_k = [k, \infty)$.

2.10 Definition: Let $(x_k)_{k\geq p}$ be a sequence in a set A. Given a strictly increasing function $f:\mathbb{Z}_{\geq q}\to\mathbb{Z}_{\geq p}$, write $k_l=f(l)$ and let $y_l=x_{k_l}$ for all $l\geq q$. Then the sequence $(y_l)_{l\geq q}$ is called a **subsequence** of the sequence $(x_k)_{k\geq p}$. In other words, a subsequence of $(x_k)_{k\geq p}$ is a sequence of the form

$$\left(x_{k_q},x_{k_{q+1}},x_{k_{q+2}},\cdots\right)\text{ with }p\leq k_q<k_{q+1}<k_{q+2}<\cdots.$$

Given a bijective function $f:\mathbb{Z}_{\geq q}\to\mathbb{Z}_{\geq p}$, write $k_l=f(l)$ and let $y_l=x_{k_l}$ for $l\geq q$. Then the sequence $(y_l)_{l\geq q}$ is called a **rearrangement** of the sequence (x_k).

2.11 Theorem: *(Subsequences and Rearrangements) Let (x_k) be a convergent sequence in $\mathbb{R}$ with $x_k\to a$. Then*

(1) every subsequence of (x_k) converges to a, and
(2) every rearrangement of (x_k) converges to a.

Proof: We shall prove Parts 1 and 2 simultaneously. Let $f:\mathbb{Z}_{\geq q}\to\mathbb{Z}_{\geq p}$ be an injective map. Write $k_l=f(l)$ and let $y_l=x_{k_l}$ for $k\geq l$. Let $\epsilon>0$. Choose $m_1\in\mathbb{Z}$ so that $k\geq m_1\implies|x_k-a|<\epsilon$. Since f is injective, there are only finitely many indices l with $p\leq f(l)<m_1$. Choose $m\in\mathbb{Z}$ with m larger than every such index l. Then for $l\geq m$ we have $k_l=f(l)\geq m_1$ and so $|y_l-a|=|x_{k_l}-a|<\epsilon$.

2.12 Theorem: *(Bolzano-Weirstrass Theorem in $\mathbb{R}$) Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

Proof: Let (x_k) be a bounded sequence in $\mathbb{R}$. Choose $a,b\in\mathbb{R}$ with $a\leq x_k$ for all k and $x_k\leq b$ for all k. Then we have $x_k\in[a,b]$ for all k. We define a sequence of nonempty closed intervals recursively as follows. Let $I_0=[a_0,b_0]=[a,b]$. Note that $I_0=\left[a,\frac{a+b}{2}\right]\cup\left[\frac{a+b}{2},b\right]$. Let $I_1=[a_1,b_1]$ be equal to one of the two intervals $\left[a,\frac{a+b}{2}\right]$ and $\left[\frac{a+b}{2},b\right]$, chosen in such a way that there are infinitely many indices k with $x_k\in I_1$. Suppose we have chosen intervals $I_j=[a_j,b_j]$ with $b_j-a_j=\frac{1}{2^j}(b-a)$ for $1\leq j\leq n$, such that $I_0\supseteq I_1\supseteq I_2\supseteq\cdots\supseteq I_n$ and such that for each index j, there are infinitely many indices k with $x_k\in I_j$. Note that $I_n=[a_n,b_n]=\left[a_n,\frac{a_n+b_n}{2}\right]\cup\left[\frac{a_n+b_n}{2},b_n\right]$. Let I_{n+1} be equal to one of the two intervals $\left[a_n,\frac{a_n+b_n}{2}\right]$ and $\left[\frac{a_n+b_n}{2},b_n\right]$, chosen in such a way that there are infinitely many indices k with $x_k\in I_{n+1}$. In this way, we obtain a sequence $(I_j)_{j\geq 0}$ of nonempty closed intervals. By the Nested Interval Theorem, $\bigcap\limits_{j=0}^{\infty}I_j$ is not empty. Choose a point c with $c\in I_n$ for every $n\geq 0$.

We shall now construct a subsequence of (x_k) which converges to c. Since for each $j\geq 0$ there exist infinitely many indices k with $x_k\in I_j$, we can construct a subsequence of (x_k) as follows. Choose k_0 so that $x_{k_0}\in I_0$, then choose $k_1>k_0$ so that $x_{k_1}\in I_1$, then choose $k_2>k_1$ with $x_{k_2}\in I_2$, and so on. In this way, we obtain a subsequence $(x_{k_j})_{j\geq 0}$ of (x_k) with $x_{k_j}\in I_j$ for all $j\geq 0$. We claim that $x_{k_j}\to c$ as $j\to\infty$. Let $\epsilon>0$ Choose $m\in\mathbb{Z}$ so that $\frac{1}{2^m}(b-a)<\epsilon$. For $j\geq m$, since $c\in[a_j,b_j]\subseteq[a,b]$ and $x_{k_j}\in[a_j,b_j]$, it follows that

$$|x_{k_j}-c|=\max\{x_{k_j},c\}-\min\{x_{k_j},c\}\leq b_j-a_j=\frac{1}{2^j}(b-a)\leq\frac{1}{2^m}(b-a)<\epsilon.$$

Thus $x_{k_j}\to c$ as $j\to\infty$, as claimed.

2.13 Definition: Let $(x_k)_{k\geq p}$ be a sequence in $\mathbb{R}$. We say that (x_k) is **Cauchy** when

$$\forall\,\epsilon{>}0\ \ \exists\, m{\in}\mathbb{Z}_{\geq p}\ \forall\, k,l{\in}\mathbb{Z}_{\geq p}\ \big(k,l\geq m \Longrightarrow |x_k-x_l|<\epsilon\big).$$

2.14 Theorem: *(The Completeness of $\mathbb{R}$, or The Cauchy Criterion for Convergence in $\mathbb{R}$) For a sequence (x_k) in $\mathbb{R}$, the sequence (x_k) converges if and only if it is Cauchy.*

Proof: Let (x_k) be a sequence in $\mathbb{R}$. Suppose that (x_k) converges, say $x_k\to a$. Let $\epsilon>0$ and choose $m\in\mathbb{Z}$ so that $k\geq m\Longrightarrow |x_k-a|<\frac{\epsilon}{2}$. Then for $k,l\geq m$ we have

$$|x_k-x_l|=|x_k-a+a-x_l|\leq|x_k-a|+|a-x_l|<\tfrac{\epsilon}{2}+\tfrac{\epsilon}{2}=\epsilon.$$

Thus (x_k) is Cauchy.

Now suppose that (x_k) is Cauchy. We claim that (x_k) is bounded. Since (x_k) is Cauchy, we can choose $m\in\mathbb{Z}$ so that $k,l\geq m\Longrightarrow|x_k-x_l|<1$. In particular, for all $k\geq m$ we have $|x_k-x_m|<1$ and so $|x_k|=|x_k-x_m+x_m|\leq|x_k-x_m|+|x_m|<1+|x_m|$. It follows that (x_k) is bounded by $b=\max\big\{|x_p|,|x_{p+1}|,\cdots,|x_{m-1}|\,,\,1+|x_m|\big\}$.

Because (x_k) is bounded, it has a convergent subsequence, by the Bolzano Weierstrass Theorem. Let (x_{k_j}) be a convergent subsequence of (x_k) and let $a=\lim\limits_{j\to\infty}x_{k_j}$. We claim that the original sequence (x_k) converges with $\lim\limits_{k\to\infty}x_k=a$. Let $\epsilon>0$. Since (x_k) is Cauchy, we can choose $m\in\mathbb{Z}$ so that $k,l\geq m\Longrightarrow|x_k-x_l|<\frac{\epsilon}{2}$. Since $x_{k_j}\to a$ we can choose $m_0\in\mathbb{Z}$ so that $j\geq m_0\Longrightarrow|x_{k_j}-a|<\frac{\epsilon}{2}$. Choose an index $j\geq m_0$ so that $k_j\geq m$. Then for all $k\geq m$ we have

$$|x_k-a|=|x_k-x_{k_j}+x_{k_j}-a|\leq|x_k-x_{k_j}|+|x_{k_j}-a|<\tfrac{\epsilon}{2}+\tfrac{\epsilon}{2}=\epsilon.$$

Thus $x_k\to a$, as claimed.

2.15 Definition: Let $(x_k)_{k\geq p}$ be a sequence in $\mathbb{R}$. The **series** $\sum\limits_{k\geq p}x_k$ is defined to be the sequence $(S_n)_{n\geq k}$ where $S_n=\sum\limits_{k=p}^{n}x_k=x_p+x_{p+1}+\cdots+x_n$. The term S_n is called the n^{th} **partial sum** of the series $\sum\limits_{k\geq p}x_k$. The **sum** of the series, denoted by $S=\sum\limits_{k=p}^{\infty}x_k=x_p+x_{p+1}+x_{p+2}+\cdots$, is the limit of the sequence of partial sums, if it exists, and we say the series **converges** when the sum exists and is finite.

2.16 Note: As with sequences, we assume students are familiar with series and various tests for convergence.

2.17 Theorem: *(Cauchy Criterion for Series) Let $(x_k)_{k\geq p}$ be a sequence. Then the series $\sum x_k$ converges if and only if*

$$\forall\,\epsilon{>}0\ \ \exists\,\ell{\in}\mathbb{Z}_{\geq p}\ \forall\, m,n{\in}\mathbb{Z}_{\geq p}\left(m{>}n{\geq}\ell\Longrightarrow\Big|\sum_{k=n+1}^{m}x_k\Big|<\epsilon\right).$$

Proof: This follows from the Cauchy Criterion for the convergence of the sequence of partial sums. Indeed (S_n) converges if and only if for all $\epsilon>0$ there exists $\ell\geq p$ such that $m>n\geq\ell\Longrightarrow|S_m-S_n|<\epsilon$, and we have $|S_m-S_n|=\Big|\sum\limits_{k=p}^{m}x_k-\sum\limits_{k=p}^{n}x_k\Big|=\Big|\sum\limits_{k=n+1}^{m}x_k\Big|$.

Limit Inferior and Limit Superior

I may include a discussion of the limit supremum and limit infimum.

Limits of Functions and Continuity in $\mathbb{R}$

2.18 Definition: Let $A \subseteq \mathbb{R}$. For $a \in \mathbb{R}$, we say that a is a **limit point** of A when

$$\forall\,\delta>0\ \exists\,x\in A\ \ 0<|x-a|<\delta.$$

When $a \in A$ and a is not a limit point of A we say that a is an **isolated point** of A.

2.19 Definition: Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$. When a is a limit point of A, we make the following definitions.

(1) For $b \in \mathbb{R}$, we say that the **limit** of $f(x)$ as x tends to a is equal to b, and we write $\lim\limits_{x\to a} f(x) = b$ or we write $f(x) \to b$ as $x \to a$, when

$$\forall\,\epsilon>0\ \exists\,\delta>0\ \forall\,x\in A\ \big(0<|x-a|<\delta \Longrightarrow |f(x)-b|<\epsilon\big).$$

(2) We say the limit of $f(x)$ as x tends to a is equal to **infinity**, and we write $\lim\limits_{x\to a} f(x) = \infty$, or we write $f(x) \to \infty$ as $x \to a$, when

$$\forall\,r\in\mathbb{R}\ \exists\,\delta>0\ \forall\,x\in A\ \big(0<|x-a|<\delta \Longrightarrow f(x)>r\big).$$

(3) We say that the limit of $f(x)$ as x tends to a is equal to **negative infinity**, and we write $\lim\limits_{x\to a} f(x) = -\infty$, or we write $f(x) \to -\infty$ as $x \to a$, when

$$\forall\,r\in\mathbb{R}\ \exists\,\delta>0\ \forall\,x\in A\ \big(0<|x-a|<\delta \Longrightarrow f(x)<r\big).$$

2.20 Note: We assume that students are familiar with limits of functions and are able to calculate limits using various limit rules (such as Operations on Limits, and the Comparison and the Squeeze Theorems) We also assume familiarity with one-sided limits $\lim\limits_{x\to a^+} f(x)$ and $\lim\limits_{x\to a^-} f(x)$ as well as asymptotic limits $\lim\limits_{x\to -\infty} f(x)$ and $\lim\limits_{x\to\infty} f(x)$. Here is one theorem that relates limits of functions and limits of sequences which students may not have seen.

2.21 Theorem: *(Sequential Characterization of Limits of Functions) Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$, let $a \in \mathbb{R}$ be a limit point of A, and let $b \in \mathbb{R}$. Then $\lim\limits_{x\to a} f(x) = b$ if and only if for every sequence (x_k) in $A \setminus \{a\}$ with $x_k \to a$ we have $f(x_k) \to b$.*

Proof: Suppose that $\lim\limits_{x\to a} f(x) = b$. Let (x_k) be a sequence in $A \setminus \{a\}$ with $x_k \to a$. Let $\epsilon > 0$. Since $\lim\limits_{x\to a} f(x) = b$, we can choose $\delta > 0$ so that $0 < |x - a| < \delta \Longrightarrow |f(x) - b| < \epsilon$. Since $x_k \to a$ we can choose $m \in \mathbb{Z}$ so that $k \ge m \Longrightarrow |x_k - a| < \delta$. Then for $k \ge m$, we have $|x_k - a| < \delta$ and we have $x_k \ne a$ (since the sequence (x_k) is in the set $A \setminus \{a\}$) so that $0 < |x_k - a| < \delta$ and hence $|f(x_k) - b| < \epsilon$. This shows that $f(x_k) \to b$.

Conversely, suppose that $\lim\limits_{x\to a} f(x) \ne b$. Choose $\epsilon_0 > 0$ so that for all $\delta > 0$ there exists $x \in A$ with $0 < |x - a| < \delta$ and $|f(x) - b| \ge \epsilon_0$. For each $k \in \mathbb{Z}^+$, choose $x_k \in A$ with $0 < |x_k - a| \le \frac{1}{k}$ and $|f(x_k) - b| \ge \epsilon_0$. In this way we obtain a sequence $(x_k)_{k\ge 1}$ in $A \setminus \{a\}$. Since $|x_k - a| \le \frac{1}{k}$ for all $k \in \mathbb{Z}^+$, it follows that $x_k \to a$ (indeed, given $\epsilon > 0$ we can choose $m \in \mathbb{Z}$ with $m > \frac{1}{\epsilon}$ and then $k \ge m \Longrightarrow |x_k - a| \le \frac{1}{k} \le \frac{1}{m} < \epsilon$). Since $|f(x_k) - b| \ge \epsilon_0$ for all k, it follows that $f(x_k) \not\to b$ (indeed if we had $f(x_k) \to b$ we could choose $m \in \mathbb{Z}$ so that $k \ge m \Longrightarrow |f(x_k) - b| < \epsilon_0$ and then we could choose $k = m$ to get $|f(x_k) - b| < \epsilon_0$).

2.22 Definition: Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$. For $a \in A$, we say that f is **continuous** at a when

$$\forall \epsilon{>}0 \ \exists \delta{>}0 \ \forall x{\in}A \ \big(|x-a|{<}\delta \Longrightarrow |f(x)-f(a)|{<}\epsilon\big).$$

We say that f is **continuous** (on A) when f is continuous at every point $a \in A$.

2.23 Note: Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$ and let $a \in A$. Verify, as an exercise, that
(1) if a is an isolated point of A then f is continuous at a, and
(2) if a is a limit point of A then f is continuous at a if and only if $\lim\limits_{x\to a} f(x) = f(a)$.

2.24 Note: We assume the reader is familiar with continuity. In particular, we assume the reader knows that every elementary function is continuous in its domain (an **elementary** function is any function which can be obtained from the basic elementary functions x, $\sqrt[n]{x}$, e^x, $\ln x$, $\sin x$ and $\sin^{-1} x$ using addition, subtraction, multiplication, division, and composition of functions).

2.25 Theorem: *(The Sequential Characterization of Continuity) Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$ and let $a \in A$. Then f is continuous at a if and only if for every sequence (x_k) in A with $x_k \to a$ we have $f(x_k) \to f(a)$.*

Proof: Suppose that f is continuous at a. Let (x_k) be a sequence in A with $x_k \to a$. Let $\epsilon > 0$. Choose $\delta > 0$ so that for all $x \in A$ we have $|x-a| < \delta \Longrightarrow |f(x)-f(a)| < \epsilon$. Choose $m \in \mathbb{Z}$ so that for all indices k we have $k \geq m \Longrightarrow |x_k - a| < \delta$. Then when $k \geq m$ we have $|x_k - a| < \delta$ and hence $|f(x_k) - f(a)| < \epsilon$. Thus we have $f(x_k) \to f(a)$.

Conversely, suppose that f is not continuous at a. Choose $\epsilon_0 > 0$ so that for all $\delta > 0$ there exists $x \in A$ with $|x-a| < \delta$ and $|f(x)-f(a)| \geq \epsilon_0$. For each $k \in \mathbb{Z}^+$, choose $x_k \in A$ with $|x_k - a| \leq \frac{1}{k}$ and $|f(x_k) - f(a)| \geq \epsilon_0$. Consider the sequence (x_k) in A (we remark that the Axiom of Choice is being used here). Since $|x_k - a| \leq \frac{1}{k}$ for all $k \in \mathbb{Z}^+$, it follows that $x_k \to a$. Since $|f(x_k) - f(a)| \geq \epsilon_0$ for all $k \in \mathbb{Z}^+$, it follows that $f(x_k) \not\to f(a)$.

2.26 Theorem: *(Intermediate Value Theorem) Let I be an interval in $\mathbb{R}$ and let $f : I \to \mathbb{R}$ be continuous. Let $a, b \in I$ with $a \leq b$ and let $y \in \mathbb{R}$. Suppose that either $f(a) \leq y \leq f(b)$ or $f(b) \leq y \leq f(a)$. Then there exists $x \in [a,b]$ with $f(x) = y$.*

Proof: We prove the theorem in the case that $f(a) \leq y \leq f(b)$. If $y = f(a)$ then we can take $x = a$ and if $y = f(b)$ then we can take $x = b$. Suppose that $f(a) < y < f(b)$. Let $A = \big\{t \in [a,b] \big| f(t) \leq y\big\}$. Note that $A \neq \emptyset$ (since $a \in A$) and A is bounded above (by b) and so A has a supremum in $\mathbb{R}$. Let $x = \sup A$. Since $a \in A$ and $x = \sup A$ we have $x \geq a$. Since b is an upper bound for A and $x = \sup A$ we have $x \leq b$. Thus $x \in [a,b]$.

We claim that $f(x) = y$. Suppose, for a contradiction, that $f(x) > y$. Since $x \neq a$ (because $f(a) < y$ but $f(x) > y$) we can choose $\delta_1 > 0$ so that $(x - \delta_1, x] \subseteq [a,b]$. Since f is continuous at x with $f(x) > y$, we can choose δ_2 so that for all $t \in [a,b]$ we have $|t-x| < \delta_2 \Longrightarrow f(t) > y$. Let $\delta = \min\{\delta_1, \delta_2\}$. Since $x = \sup A$, by the Approximation Property we can choose $t \in A$ with $x - \delta < t \leq x$. Since $t \in A$ we have $f(t) \leq y$, but since $t \in (x-\delta, x]$ we have $f(t) > y$, so we have obtained the desired contradiction. Now suppose, for a contradiction, that $f(x) < y$. Since $x \neq b$ (because $f(b) > y$ but $f(x) < y$) we can choose $\delta_1 > 0$ so that $[x, x+\delta_1) \subseteq [a,b]$. Since f is continuous at x with $f(x) < y$ we can choose $\delta_2 > 0$ so that for all $t \in [a,b]$ we have $|t - x| < \delta_2 \Longrightarrow f(t) < y$. Let $\delta = \min\{\delta_1, \delta_2\}$ so that $[x, x+\delta) \subseteq [a,b]$ and for all $t \in [x, x+\delta)$ we have $f(t) < y$. But then $x + \delta \in A$ so we cannot have $x = \sup A$, and we have obtained the desired contradiction.

2.27 Exercise: Let $A \subseteq \mathbb{R}$. Show that A is an interval if and only if A has the intermediate value property that for all $a, b, x \in \mathbb{R}$ with $a < x < b$, if $a \in A$ and $b \in A$ then $x \in A$.

2.28 Theorem: *(Continuous Functions and Intervals) Let A be an interval in $\mathbb{R}$ and let $f : A \to \mathbb{R}$ be continuous. Then the range $B = f(A)$ is an interval in $\mathbb{R}$.*

Proof: If $B = f(A)$ contains less than 2 points then it is a (degenerate) interval. Suppose that B contains at least two points. Let $u, v \in B$ and let $y \in \mathbb{R}$ with $u < y < v$. Since $B = f(A)$ we can choose $a, b \in A$ with $f(a) = u$ and $f(b) = v$. Since $f(a) = u \neq v = f(b)$ we have $a \neq b$. Since y lies between $f(a) = u$ and $f(b) = v$, and since f is continuous, it follows from the Intermediate Value Theorem that we can choose x between a and b with $f(x) = y$. Since A is an interval in $\mathbb{R}$, it has the intermediate value property (by Exercise 2.27), and so we have $x \in A$. Since $x \in A$ and $y = f(x)$ we have $y \in f(A) = B$. This proves that B has the intermediate value property, and so (by Exercise 2.27) B is an interval, as required.

2.29 Definition: Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$. For $a \in A$, if $f(a) \geq f(x)$ for every $x \in A$, then we say that $f(a)$ is the **maximum value** of f and that f attains its maximum value at a. Similarly for $b \in A$, if $f(b) \leq f(x)$ for every $x \in A$ then we say that $f(b)$ is the **minimum value** of f (in A) and that f attains its minimum value at b. We say that f attains its **extreme values** in A when f attains its maximum value at some point $a \in A$ and f attains its minimum value at some point $b \in A$.

2.30 Theorem: *(Extreme Value Theorem) Let $a, b \in \mathbb{R}$ with $a < b$, and let $f : [a, b] \to \mathbb{R}$ be continuous. Then f attains its extreme values in $[a, b]$.*

Proof: We prove that f attains its maximum. First we claim that f is bounded above. Suppose, for a contradiction, that it is not. For each $k \in \mathbb{Z}^+$, choose $x_k \in [a, b]$ such that $f(x_k) \geq k$. By the Bolzano Weierstrass Theorem, we can choose a convergent subsequence (x_{k_j}). Let $p = \lim\limits_{j\to\infty} x_{k_j}$. Note that $p \in [a, b]$ by Comparison (since $x_{k_j} \geq a$ for all j we have $p \geq a$, and since $x_{k_j} \leq b$ for all j we have $p \leq b$). Since $f(x_{k_j}) \geq k_j$ and $k_j \to \infty$ we must have $f(x_{k_j}) \to \infty$ as $j \to \infty$. But by the Sequential Characterization of Continuity, we should have $f(x_{k_j}) \to f(p) \in \mathbb{R}$, so we have obtained the desired contradiction. Thus f is bounded above, as claimed.

Since the range $f([a, b])$ is nonempty and bounded above, it has a supremum. Let $m = \sup f([a, b])$. By the Approximation Property of the supremum, for each $k \in \mathbb{Z}^+$ we can choose $y_k \in [a, b]$ such that $m - \frac{1}{k} \leq f(y_k) \leq m$. By the Bolzano Weierstrass Theorem, we can choose a convergent subsequence (y_{k_j}). Let $c = \lim\limits_{j\to\infty} y_{k_j}$. Since we have $m - \frac{1}{k_j} \leq f(y_{k_j}) \leq m$ and $\frac{1}{k_j} \to 0$, we have $f(y_{k_j}) \to m$ as $j \to \infty$ by the Squeeze Theorem. Since f is continuous at c, by the Sequential Characterization of Continuity we have $f(y_{k_j}) \to f(c)$, and so by the Uniqueness of Limits, we have $f(c) = m$. Thus f attains its maximum value at c.

2.31 Definition: Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$. We say that f is **uniformly continuous** (on A) when

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall a \in A\ \forall x \in A\ \big(|x - a| < \delta \Longrightarrow |f(x) - f(a)| < \epsilon\big).$$

2.32 Example: Define $f : (0, \infty) \to (0, \infty)$ by $f(x) = \frac{1}{x}$. Let $\epsilon = 1$. Let $\delta > 0$. If $\delta \geq 1$ then for $x = \frac{1}{3}$ and $a = 1$ we have $|x - a| = \frac{2}{3} < \delta$ but $|f(x) - f(a)| = 2 > \epsilon$. If $0 < \delta < 1$ then for $x = \frac{\delta}{3}$ and $a = \delta$ we have $|x - a| = \frac{2}{3}\delta < \delta$ but $|f(x) - f(a)| = \frac{2}{\delta} > 2 > \epsilon$. This proves that f is not uniformly continuous (but f is continuous because it is elementary).

2.33 Theorem: *(Closed Bounded Intervals and Uniform Continuity) Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$. If f is continuous then f is uniformly continuous (on $[a, b]$).*

Proof: Suppose, for a contradiction, that $f : [a, b] \to \mathbb{R}$ is continuous but not uniformly continuous on $[a, b]$. Choose $\epsilon > 0$ so that for all $\delta > 0$ there exist $x, y \in [a, b]$ such that $|x - y| < \delta$ but $|f(x) - f(y)| \geq \epsilon$. For each $k \in \mathbb{Z}^+$ choose x_k and y_k in $[a, b]$ with $|x_k - y_k| \leq \frac{1}{k}$ and $|f(x_k) - f(y_k)| \geq \epsilon$. By the Bolzano Weierstrass Theorem, we can choose a convergent subsequence (y_{k_j}) of (y_k). Let $c = \lim_{j\to\infty} y_{k_j}$. For all j we have $|x_{k_j} - y_{k_j}| \leq \frac{1}{k_j}$ hence $y_{k_j} - \frac{1}{k_j} \leq x_{k_j} \leq y_{k_j} + \frac{1}{k_j}$. Since $y_{k_j} \to c$ and $\frac{1}{k_j} \to 0$ we have $y_{k_j} \pm \frac{1}{k_j} \to c$ and hence $x_{k_j} \to c$ by the Squeeze Theorem. Since f is continuous at c and $x_{k_j} \to c$ and $y_{k_j} \to c$, we have $f(x_{k_j}) \to f(c)$ and $f(y_{k_j}) \to f(c)$ by the Sequential Characterization of Continuity. Since $f(x_{k_j}) \to c$ and $f(y_{k_j}) \to c$ we have $f(x_{k_j}) - f(y_{k_j}) \to 0$. But this implies that we can choose j so that $|f(x_{k_j}) - f(y_{k_j})| < \epsilon$, giving the desired contradiction.

Differentiation in $\mathbb{R}$

2.34 Definition: For a subset $A \subseteq \mathbb{R}$, we say that A is **open** when it is a union of open intervals. Let $A \subseteq \mathbb{R}$ be open, let $f : A \to \mathbb{R}$. For $a \in A$, we say that f is **differentiable** at a when the limit

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a}$$

exists in $\mathbb{R}$. In this case we call the limit the **derivative** of f at a, and we denote to by $f'(a)$, so we have

$$f'(a) = \lim_{x\to a} \frac{f(x) - f(a)}{x - a}.$$

We say that f is **differentiable** (on A) when f is differentiable at every point $a \in A$. In this case, the **derivative** of f is the function $f' : A \to \mathbb{R}$ defined by

$$f'(x) = \lim_{u\to x} \frac{f(u) - f(x)}{u - x}.$$

When f' is differentiable at a, denote the derivative of f' at a by $f''(a)$, and we call $f''(a)$ the **second derivative** of f at a. When $f''(a)$ exists for every $a \in A$, we say that f is **twice differentiable** (on A), and the function $f'' : A \to \mathbb{R}$ is called the **second derivative** of f. Similarly, $f'''(a)$ is the derivative of f'' at a and so on.

2.35 Remark: Note that

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}.$$

To be precise, the limit on the left exists in $\mathbb{R}$ if and only if the limit on the right exists in $\mathbb{R}$, and in this case the two limits are equal.

2.36 Note: The student should be familiar with derivatives from first year calculus, and should be able to calculate the derivatives of elementary functions using differentiation rules including the Product Rule, the Quotient Rule and the Chain Rule. We shall provide proofs of some of the theorems whose proofs are often omitted in calculus courses.

2.37 Exercise: Let $A \subseteq \mathbb{R}$ be open, let $f : A \to \mathbb{R}$, and let $a \in A$. Show that f is differentiable at a with derivative $f'(a)$ if and only if

$$\forall\, \epsilon > 0\ \exists\, \delta > 0\ \forall\, x \in A\ \Big(|x-a| \le \delta \implies \big|f(x) - f(a) - f'(a)(x-a)\big| \le \epsilon\,|x-a| \Big).$$

2.38 Theorem: *(Differentiability Implies Continuity) Let $A \subseteq \mathbb{R}$ be open, let $f : A \to \mathbb{R}$ and let $a \in A$. If f is differentiable at a then f is continuous at a.*

Proof: The proof is left as an exercise (the proof is often given in first year calculus).

2.39 Theorem: *(Chain Rule) Let $A, B \subseteq \mathbb{R}$ be open, let $f : A \to B$, let $g : B \to \mathbb{R}$ and let $h = g \circ f : A \to \mathbb{R}$. Let $a \in A$ and let $b = f(a) \in B$. Suppose that f is differentiable at a and g is differentiable at b. Then h is differentiable at a with*

$$h'(a) = g'\big(f(a)\big)\, f'(a).$$

Proof: We shall use the ϵ-δ formulation of the derivative given in Exercise 1.48. Note first that for $x \in A$ and $y = f(x) \in B$ we have

$$\begin{aligned}
&\big|h(x) - h(a) - g'(f(a))f'(a)(x-a)\big| \\
&\quad = \big|g(f(x)) - g(f(a)) - g'(f(a))f'(a)(x-a)\big| \\
&\quad = \big|g(y) - g(b) - g'(b)f'(a)(x-a)\big| \\
&\quad = \big|g(y) - g(b) - g'(b)(y-b) + g'(b)(y-b) - g'(b)f'(a)(x-a)\big| \\
&\quad \le \big|g(y) - g(b) - g'(b)(y-b)\big| + |g'(b)|\,|y - b - f'(a)(x-a)| \\
&\quad = \big|g(y) - g(b) - g'(b)(y-b)\big| + |g'(b)|\,\big|f(x) - f(a) - f'(a)(x-a)\big|
\end{aligned}$$

and also

$$\begin{aligned}
|y-b| &= \big|f(x) - f(a)\big| = \big|f(x) - f(a) - f'(a)(x-a) + f'(a)(x-a)\big| \\
&\le \big|f(x) - f(a) - f'(a)(x-a)\big| + |f'(a)|\,|x-a|.
\end{aligned}$$

Let $\epsilon > 0$. Since g is differentiable at b, we can choose $\delta_0 > 0$ so that

$$|y-b| \le \delta_0 \implies \big|g(y) - g(b) - g'(b)(y-b)\big| \le \tfrac{\epsilon}{2(1+|f'(a)|)}\,|y-b|.$$

Since f is continuous at a (by Theorem 1.49), we can choose δ_1 so that

$$|x-a| \le \delta_1 \implies |f(x) - f(a)| \le \delta_0 \implies |y-b| \le \delta_0.$$

Since f is differentiable at a we can choose $\delta_2 > 0$ and $\delta_3 > 0$ so that

$$\begin{aligned}
&|x-a| \le \delta_2 \implies \big|f(x) - f(a) - f'(a)(x-a)\big| \le |x-a| \text{ and} \\
&|x-a| \le \delta_3 \implies \big|f(x) - f(a) - f'(a)(x-a)\big| \le \tfrac{\epsilon}{2(1+|g'(b)|)}\,|x-a|.
\end{aligned}$$

Let $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Let $x \in A$ and let $y = f(x) \in B$. Then when $|x-a| \le \delta$ we have

$$\begin{aligned}
&\big|h(x) - h(a) - g'(f(a))f'(a)(x-a)\big| \\
&\quad \le \big|g(y) - g(b) - g'(b)(y-b)\big| + |g'(b)|\,\big|f(x) - f(a) - f'(a)(x-a)\big| \\
&\quad \le \tfrac{\epsilon}{2(1+|f'(a)|)}\,|y-b| + (1 + |g'(b)|) \cdot \tfrac{\epsilon}{2(1+|g'(b)|)}\,|x-a| \\
&\quad \le \tfrac{\epsilon}{2(1+|f'(a)|)}\Big(\big|f(x) - f(a) - f'(a)(x-a)\big| + |f'(a)|\,|x-a|\Big) + \tfrac{\epsilon}{2}\,|x-a| \\
&\quad \le \tfrac{\epsilon}{2(1+|f'(a)|)}\Big(|x-a| + |f'(a)|\,|x-a|\Big) + \tfrac{\epsilon}{2}\,|x-a| \\
&\quad = \tfrac{\epsilon}{2}\,|x-a| + \tfrac{\epsilon}{2}\,|x-a| = \epsilon|x-a|.
\end{aligned}$$

Thus h is differentiable at a with $h'(a) = g'(f(a))f'(a)$, as required.

2.40 Exercise: Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$. Show that f is strictly monotonic if and only if f has the property that for all $a, b, c \in A$, if b lies strictly between a and c then $f(b)$ lies strictly between $f(a)$ and $f(c)$.

2.41 Theorem: *(The Inverse Function Theorem) Let A be an interval in $\mathbb{R}$, let $f : A \to \mathbb{R}$ be injective and continuous, let $B = f(A)$ (and note that $f : A \to B$ is bijective) and let $g = f^{-1} : B \to A$ be the inverse function. Then*

(1) the functions f and g are strictly monotonic and g is continuous, and

(2) if A is an open interval then so is B, and if f is differentiable at $a \in A$ with $f'(a) \neq 0$, then g is differentiable at $b = f(a)$ with $g'(b) = \frac{1}{f'(a)}$.

Proof: To prove Part 1, suppose that f is injective and continuous. Let $a, b, c \in A$ with $a < b < c$. Since f is injective and $a \neq c$, we have $f(a) \neq f(c)$, so either $f(a) < f(c)$ or $f(a) > f(c)$. Consider the case that $f(a) < f(c)$. Suppose, for a contradiction, that $f(b) \geq f(c)$. Note that since f is injective and $b \neq c$ we have $f(b) \neq f(c)$ and so $f(b) > f(c)$. Choose y with $f(c) < y < f(b)$. Since f is continuous on $[a, b]$ and on $[b, c]$, by the Intermediate Value Theorem, we can choose $x_1 \in [a, b]$ and $x_2 \in [b, c]$ with $f(x_1) = y = f(x_2)$. Since $y \neq f(b)$ we cannot have $x_1 = b$ or $x_2 = b$ so we have $x_1 < b < x_2$ with $f(x_1) = f(x_2)$, which contradicts the fact that f is injective. Thus we cannot have $f(b) \geq f(c)$ so we have $f(b) < f(c)$. A similar argument shows that we cannot have $f(b) \leq f(a)$ so we must have $f(b) > f(a)$. This proves that in the case that $f(a) < f(c)$ we have $f(a) < f(b) < f(c)$. A similar argument shows that in the case that $f(a) > f(c)$ we have $f(a) > f(b) > f(c)$. It follows that f is strictly monotonic, by Exercise 1.51. It is easy to see that if f is strictly increasing then g is strictly increasing (indeed when $u, v \in B$ with $u < v$ and $a = g(u)$ and $b = g(v)$, we must have $a < b$ because if $a = b$ then $u = v$ and if $a > b$ then $u > v$ since f is strictly increasing) and if f is strictly decreasing then g is strictly decreasing.

To complete the proof of Part 1, it remains to show that g is continuous. Suppose that f and g are strictly increasing (the case that f and g are strictly decreasing is similar). Let $b \in B$ and let $a = g(b)$ so that $f(a) = b$. Since f and g are strictly increasing, it follows that b is the left (or right) endpoint of B if and only if a is the left (or right) endpoint of A. To show that g is continuous at b, it suffices to show that if b is not the right endpoint of B then $\lim_{y \to b^+} g(y) = g(b)$ and that if b is not the left endpoint then $\lim_{y \to b^-} g(y) = g(b)$. We shall prove the first of these two statements (the proof of the second is similar). Suppose that b is not the right endpoint of B and hence a is not the right endpoint of A. Let $\epsilon > 0$ be small enough that $a + \epsilon \in A$. Choose $\delta = f(a + \epsilon) - b = f(a + \epsilon) - f(a)$ and note that $\delta > 0$ since f is strictly increasing. Then for all $y \in B$, if $b < y < b + \delta$ then $a = g(b) < g(y) < g(b + \delta) = g(f(a + \epsilon)) = a + \epsilon$. Thus $\lim_{y \to b^+} g(y) = g(b)$, as required.

To prove Part 2, suppose that A is an open interval and that f is differentiable at $a \in A$ with $f'(a) \neq 0$. Note that B is an interval by Theorem 1.39 and B is open because, as mentioned above, if $u \in B$ was a right or left endpoint of B then $g(u)$ would be a right or left endpoint of A. By Part 1, we know that g is continuous at $b = f(a)$, and so as $y \to b$ in B we have $g(y) \to g(b)$ in A, and so for $x = g(y)$ we have

$$\frac{g(y) - g(b)}{y - b} = \frac{x - a}{f(x) - f(a)} = \frac{1}{\frac{f(x) - f(a)}{x - a}} \longrightarrow \frac{1}{f'(a)} \quad \text{as } y \to b.$$

2.42 Definition: Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$ and let $a \in A$. We say that f has a **local maximum** value at a when $\exists\, \delta > 0\ \forall x \in A\ \big(|x-a| \leq \delta \implies f(x) \leq f(a)\big)$. Similarly, we say that f has a **local minimum** value at a when $\exists\, \delta > 0\ \forall x \in A\ \big(|x-a| \leq \delta \implies f(x) \geq f(a)\big)$.

2.43 Theorem: *(Fermat's Theorem) Let $A \subseteq \mathbb{R}$ be open, let $f : A \to \mathbb{R}$, and let $a \in A$. Suppose that f is differentiable at a and that f has a local maximum or minimum value at a. Then $f'(a) = 0$.*

Proof: The proof is left as an exercise (you probably saw the proof in first year calculus).

2.44 Theorem: *(Mean Value Theorems) Let $a, b \in \mathbb{R}$ with $a < b$.*

(1) (Rolle's Theorem) If $f : [a,b] \to \mathbb{R}$ is differentiable in (a,b) and continuous at a and b with $f(a) = 0 = f(b)$ then there exists a point $c \in (a,b)$ such that $f'(c) = 0$.

(2) (The Mean Value Theorem) If $f : [a,b] \to \mathbb{R}$ is differentiable in (a,b) and continuous at a and b then there exists a point $c \in (a,b)$ with

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

(3) (Cauchy's Mean Value Theorem) If $f, g : [a,b] \to \mathbb{R}$ are differentiable in (a,b) and continuous at a and b, then there exists a point $c \in (a,b)$ such that

$$f'(c)\big(g(b) - g(a)\big) = g'(c)\big(f(b) - f(a)\big).$$

Proof: To Prove Rolle's Theorem, let $f : [a,b] \to \mathbb{R}$ be differentiable in (a,b) and continuous at a and b with $f(a) = 0 = f(b)$. If f is constant, then $f'(x) = 0$ for all $x \in [a,b]$, so we can choose any $c \in (a,b)$ and we have $f'(c) = 0$. Suppose that f is not constant. Either $f(x) > 0$ for some $x \in (a,b)$ or $f(x) < 0$ for some $x \in (a,b)$. Suppose that $f(x) > 0$ for some $x \in (a,b)$ (the case that $f(x) < 0$ for some $x \in (a,b)$ is similar). By the Extreme Value Theorem, f attains its maximum value at some point, say $c \in [a,b]$. Since $f(x) > 0$ for some $x \in (a,b)$, we must have $f(c) > 0$. Since $f(a) = f(b) = 0$ and $f(c) > 0$, we have $c \in (a,b)$. By Fermat's Theorem, we have $f'(c) = 0$. This completes the proof of Rolle's Theorem.

To prove the Mean Value Theorem, suppose that $f : [a,b] \to \mathbb{R}$ is differentiable in (a,b) and continuous at a and b. Let $g(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a)$. Then g is differentiable in (a,b) with $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$ and g is continuous at a and b with $g(a) = 0 = g(b)$. By Rolle's Theorem, we can choose $c \in (a,b)$ so that $g'(c) = 0$, and then $g'(c) = \frac{f(b)-f(a)}{b-a}$, as required.

Finally, we use the Mean Value Theorem to Prove Cauchy's Mean Value Theorem. Suppose $f, g : [a,b] \to \mathbb{R}$ are both differentiable in (a,b) and continuous at a and b. Let $h(x) = f(x)\big(g(b) - g(a)\big) - g(x)\big(f(b) - f(a)\big)$. Then h is differentiable in (a,b) and continuous at a and b with $h(a) = f(a)g(b) - g(a)f(b) = h(b)$. By the Mean Value Theorem, we can choose $c \in (a,b)$ so that $h'(c) = \frac{h(b)-h(a)}{b-a} = 0$, and then we have $f'(c)\big(g(b) - g(a)\big) - g'(c)\big(f(b) - f(a)\big) = 0$, as required.

2.45 Corollary: *Let $a, b \in \mathbb{R}$ with $a < b$. Let $f : [a,b] \to \mathbb{R}$. Suppose that f is differentiable in (a,b) and continuous at a and b. If $f'(x) > 0$ for all $x \in (a,b)$ then f is strictly increasing on $[a,b]$.*

Proof: The proof is left as an exercise (the proof is often given in first year calculus).

2.46 Theorem: *(l'Hôpital's Rule) Let A be a nonempty open interval in $\mathbb{R}$. Let $a \in A$, or let a be an endpoint of A. Let $f, g : A \setminus \{a\} \to \mathbb{R}$. Suppose that f and g are differentiable in $A \setminus \{a\}$ with $g'(x) \neq 0$ for all $x \in A \setminus \{a\}$. Suppose either that $\lim\limits_{x\to a} f(x) = 0 = \lim\limits_{x\to a} g(x)$ or that $\lim\limits_{x\to a} g(x) = \pm\infty$. Suppose that $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)} = u \in \mathbb{R}$. Then $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = u$.*

Similar results hold for limits $x \to a^+$, $x \to a^-$, $x \to \infty$ and $x \to -\infty$ and also when the limit is $u = \pm\infty$.

Proof: We give the proof for $x \to a^+$ (assuming that $a \in A$ or a is the left endpoint of A) with $u \in \mathbb{R}$. Suppose first that $\lim\limits_{x\to a^+} f(x) = 0 = \lim\limits_{x\to a^+} g(x)$. Choose $b \in A$ with $a < b$. Extend the maps f and g to obtain maps $f, g : [a, b] \to \mathbb{R}$ by defining $f(a) = 0 = g(a)$. Note that f and g are continuous at a since $\lim\limits_{x\to a^+} f(x) = 0$ and $\lim\limits_{x\to a^+} g(x) = 0$. Let (x_k) be a sequence in $(a, b\,]$ with $x_k \to a$. For each index k, by Cauchy's Mean Value Theorem we can choose $c_k \in (a, x_k)$ so that $f'(c_k)\big(g(x_k) - g(a)\big) = g'(c_k)\big(f(x_k) - f(a)\big)$. Since $f(a) = 0 = g(a)$, this simplifies to $f'(c_k)g(x_k) = g'(c_k)f(x_k)$ and so we have $\frac{f(x_k)}{g(x_k)} = \frac{f'(c_k)}{g'(c_k)}$. Since $a < c_k < x_k$ and $x_k \to a$, we have $c_k \to a$ by the Squeeze Theorem. Since $\lim\limits_{x\to a^+} \dfrac{f'(x)}{g'(x)} = u$ and $c_k \to a$, we have $\frac{f(x_k)}{g(x_k)} = \frac{f'(c_k)}{g'(c_k)} \to u$ by the Sequential Characterization of Limits. We have shown that for every sequence (x_k) in $(a, b\,]$ with $x_k \to a$ we have $\frac{f(x_k)}{g(x_k)} \to u$. By the Sequential Characterization of limits, it follows that $\lim\limits_{x\to a^+} \dfrac{f(x)}{g(x)} = u$.

Now suppose that $\lim\limits_{x\to a^+} g(x) = \infty$. Since $\lim\limits_{x\to a^+} g(x) = \infty$ we can choose $b \in A$ with $b > a$ so that $g(x) > 0$ for all $x \in (a, b\,]$. Let (x_k) be a sequence in $(a, b\,]$ with $x_k \to a$. For each pair of indices k, l, by Cauchy's Mean Value Theorem we can choose $c_{kl} \in (a, x_k)$ so that $f'(c_{kl})\big(g(x_k) - g(x_l)\big) = g'(c_{kl})\big(f(x_k) - f(x_l)\big)$. Divide both sides by $g'(c_{kl})g(x_l)$ to get

$$\frac{f'(c_{kl})}{g'(c_{kl})}\,\frac{g(x_k)}{g(x_l)} - \frac{f'(c_{kl})}{g'(c_{kl})} = \frac{f(x_k)}{g(x_l)} - \frac{f(x_l)}{g(x_l)}.$$

so we have

$$\frac{f(x_l)}{g(x_l)} = \frac{f'(c_{kl})}{g'(c_{kl})} + \frac{f(x_k)}{g(x_l)} - \frac{f'(c_{kl})}{g'(c_{kl})}\,\frac{g(x_k)}{g(x_l)}.$$

Let $\epsilon > 0$. Since $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)} = u$ we can choose $\delta > 0$ so that $|x - a| \leq \delta \Longrightarrow \left|\frac{f'(x)}{g'(x)} - u\right| \leq \frac{\epsilon}{3}$. Since $x_k \to a$ we can choose $m \in \mathbb{Z}^+$ so $k \geq m \Longrightarrow |x_k - a| \leq \delta$. Note that when $k, l \geq m$, since c_{kl} lies between x_k and x_l we also have $|c_{kl} - a| \leq \delta$ so $\left|\frac{f'(c_{kl})}{g'(c_{kl})} - u\right| \leq \min\left\{1, \frac{\epsilon}{3}\right\}$. Fix $k \geq m$. Choose l large enough so that $\left|\frac{f(x_k)}{g(x_l)}\right| \leq \frac{\epsilon}{3}$ and $\left|\frac{f'(c_{kl})}{g'(c_{kl})}\,\frac{g(x_k)}{g(x_l)}\right| \leq \frac{\epsilon}{3}$. Then we have

$$\left|\frac{f(x_l)}{g(x_l)} - u\right| \leq \left|\frac{f'(c_{kl})}{g'(c_{kl})} - u\right| + \left|\frac{f(x_k)}{g(x_l)}\right| + \left|\frac{f'(c_{kl})}{g'(c_{kl})}\,\frac{g(x_k)}{g(x_l)}\right| \leq \epsilon.$$