RUDIN CHAPTER 4
SOLUTIONS

JOHN WANG

1. Problem 4.1

Theorem 1.1. *Suppose f is a real function defined on $\mathbb{R}^1$ which satisfies $\lim_{h\to 0}[f(x+h)-f(x-h)] = 0$ for all $x \in \mathbb{R}^1$. This does not necessarily imply that f is continuous.*

Proof. Consider the following function on $\mathbb{R}^1$:

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{else} \end{cases} \tag{1.2}$$

We can see that $f(x)$ satisfies $\lim_{h\to 0}[f(x+h)-f(x-h)] = 0$ for all $x \neq 0$. This is because we can define $g_x(h) = f(x+h) - f(x-h)$ for each $x \in \mathbb{R}^1$. It is clear that for all $x \neq 0$ we have $\lim_{h\to 0} g_x(h) = 0$ because $f(x) = 0$ is a constant. Thus we are left to show that $\lim_{h\to 0} g_0(h) = 0$. Thus, we must obtain $\lim_{h\to 0} g_0(h+)$ and $\lim_{h\to 0} g_0(h-)$. If both of them are equal to zero, then we have shown that $f(x)$ satisfies the hypothesis given in the theorem.

Indeed, we can see that $\lim_{h\to 0} g_0(h+) = \lim_{h\to 0} g_0(h-)$. This is due, first, to the symmetry of $f(x)$ about 0. Second, we know that $f(h) = f(-h)$ for all $h \neq 0$. Thus, $f(h) - f(-h) = 0$ and $f(-h) - f(h) = 0$ if $h \neq 0$. This shows that $\lim_{h\to 0} f(h) - f(-h) = \lim_{h\to 0} f(-h) - f(h) = 0$.

Finally, we can see that this function is not continuous. This is because $\lim_{x\to 0} f(x) = 0$ while $f(0) = 1$, which by a theorem in Rudin shows that $f(x)$ is not continuous at $x = 0$. □

2. Problem 4.2

Theorem 2.1. *If f is a continuous mapping of a metric space X into a metric space Y, then $f(\bar{E}) \subset f(\bar{E})$ for every set $E \subset X$. Moreover, $f(\bar{E})$ can be a proper subset of $\overline{f(E)}$.*

Proof. Let $x_0 \in \bar{E}$. Then we must have $x_0 \in E$ or $x_0 \in E'$. If we assume that $x_0 \in E$, then we have $f(x_0) \in f(E)$. Moreover, since $x_0 \in \bar{E}$, we see that $f(\bar{E}) \subset \overline{f(E)}$. If $x_0 \in E'$, then there is a sequence $\{p_n\}$ such that $\lim_{n\to\infty} f(p_n) = f(x_0)$ for every $\{p_n\} \to x_0$ since x_0 is a limit point of E. Since every $\{f(p_n)\}$ is in $f(E)$, we know that $f(x_0) \in \overline{f(E)}$.

To show that $f(\bar{E})$ can be a proper subset of $\overline{f(E)}$, take the function $f(x) = \frac{1}{x}$. Next, take the following interval $E = (1, \infty)$. We can see that $\bar{E} = [1, \infty)$ and that $f(E) = (0, 1)$. Thus, $\overline{f(E)} = [0, 1]$. However, we can see that $f(\bar{E}) = f([1, \infty]) = (0, 1]$. It is easy to see that $f(\bar{E})$ is be a proper subset of $\overline{f(E)}$. □

3. Problem 4.4

Theorem 3.1. *Let f and g be continuous mappings of a metric space X into a metric space Y and let E be a dense subset of X. Then $f(E)$ is dense in $f(X)$.*

Proof. We must show that each element of $f(X)$ is either in $f(E)$ or is a limit point of $f(E)$. Since we know E to be dense in X, we know that each point in X is either a point in E or a limit point of E. If $p \in E$, then we know $f(p) \in f(X)$. Thus, we must show that for each limit point of E not also in E, denoted by $p' \in E'$, we have a corresponding $f(p')$ that is a limit point of $f(E)$.

To show this, we note that f is continuous, which implies that for all $\epsilon > 0$, there exists a $\delta > 0$ such that $d(f(p), f(p')) < \epsilon$ if $d(p, p') < \delta$, where $p \in E$ and $p' \in E'$. In other words, $\forall \epsilon > 0$, there is a corresponding neighborhood $N_\epsilon(f(p'))$ with $f(p) \in f(E)$ and $f(p) \in N_\epsilon(f(p'))$. This shows that $f(p')$ is a limit point of $f(E)$. Since each $f(p')$ is a limit point of $f(E)$, and we know that X consists entirely of p and p', we can see that $f(X)$ consits entirely of $f(p')$ and $f(p)$, which completes the proof. □

Theorem 3.2. *If $g(p) = f(p)$ for all $p \in E$, then $g(p) = f(p)$ for all $p \in X$.*

Proof. Since E is dense in X, we know that $X = E \cup E^c$ and that each $p' \in E^c$ is a limit point of E. Since we assumed $g(p) = f(p)$ for all $p \in E$, we only need to show $g(p') = f(p')$ for all $p' \in E^c$. We know that there exists a sequence $\{p_n\}$ in E which converges to each $p' \in E^c$ because p' is a limit point of E. Thus we know that $f(p_n) = g(p_n)$, because $p_n \in E$. Moreover by continuity of f and g, we know that $\lim_{p_n \to p'} f(p_n) = f(p')$ and $\lim_{p_n \to p'} g(p_n) = g(p')$. Since $f(p_n) = g(p_n)$, we know that the limits are the same as well so that $f(p') = g(p')$. This completes the proof. □

4. Problem 4.5

Theorem 4.1. *If f is a real continuous function defined on a closed set $E \subset \mathbb{R}^1$, then there exist continuous real functions g on $\mathbb{R}^1$ such that $g(x) = f(x)$ for all $x \in E$. However, the result fails when the word "closed" is omitted from the hypothesis.*

Proof. Since a closed set on $\mathbb{R}^1$ must be a closed interval, we will let $E = [a, b]$. Next, consider the following function:

$$g(x) = \begin{cases} f(a) & x \leq a \\ f(x) & a < x < b \\ f(b) & x \geq b \end{cases} \tag{4.2}$$

We see that $g(x) = f(x)$ for all $x \in [a, b]$. Now, we will show that $g(x)$ is continuous on $\mathbb{R}^1$. First, we know that $g(x)$ is continuous on (a, b) because $f(x)$ is a continuous function by assumption. We also know that $g(x)$ is continuous on $(-\infty, a) \cup (b, \infty)$ because $f(a)$ and $f(b)$ are constants, and constants are continuous functions. Thus, we must show that $g(x)$ is also continuous at points $x = a$ and $x = b$. To do this, we shall look at the left-hand and right-hand limits of $f(a)$. We see that $\lim_{x \to a^-} f(x) = f(a)$ because $f(x) = f(a)$ for $x \leq a$. We also know that $\lim_{x \to a^+} f(x) = f(a)$ because the continuity of $f(x)$ on $[a, b]$ implies that $f(a) = lim_{x \to a} f(x) = \lim_{x \to a^+} f(x)$. Thus, $g(x)$ is a continuous function on $\mathbb{R}^1$.

This result fails when the set E is not closed. Consider for instance $E = (-\infty, 0) \cup (0, \infty)$ and $f(x) = \frac{1}{x}$. Then we see that $\lim_{x \to 0^-} f(x) = -\infty$ and $\lim_{x \to 0^+} f(x) = +\infty$. This means that no value for $g(0)$ will be continuous if $g(x) = f(x)$ for all $x \in E$. This is because to be continuous, $g(0-) = -\infty$ must equal $g(0+) = +\infty$, which is impossible. □

Theorem 4.3. *Let $\vec{f} = (f_1, f_2, \ldots, f_k)$ be a real continuous vector valued function defined on a closed set $E \subset \mathbb{R}^k$. Then there exist continuous real functions $\vec{g}$ on $\mathbb{R}^k$ such that $\vec{g}(x) = \vec{f}(x)$ for all $\vec{x} \in E$.*

Proof. If E is a closed set in $\mathbb{R}^k$, then it must be composed of a k-cell of closed intervals $[a_1, b_1], [a_2, b_2], \ldots, [a_k, b_k]$. We can define $\vec{g}(x) = (g_1(x), g_2(x), \ldots, g_k(x))$ by defining the individual functions $g_n(x)$, $n = 1, 2, \ldots, k$ as the following:

$$g_n(x) = \begin{cases} f_n(a_n) & x \leq a_n \\ f_n(x) & a_n < x < b_n \\ f_n(b_n) & x \geq b_n \end{cases} \tag{4.4}$$

We know that $f_1(x), f_2(x), \ldots, f_k(x)$ are all continuous on $[a_1, b_1], [a_2, b_2], \ldots, [a_k, b_k]$ respectively by the fact that $\vec{f}$ is continuous. Thus, using the argument from the above theorem, we can see that each $g_n(x)$ is continuous. Thus, $\vec{g}$ is continuous because each of its component functions is also continuous. □

5. Problem 4.8

Theorem 5.1. *If f is a real uniformly continuous function on the bounded set $E \subset \mathbb{R}^1$, then f is bounded in E.*

Proof. If E is bounded, then we know that $\bar{E}$ is also bounded. This is because for each limit point $p \in E'$, there is a point $x \in E$ such that $d(x, p) < \epsilon$ for all $\epsilon > 0$. Since E is bounded, there exists an $M > 0$ such that $d(x, q) < M$ for all $x \in E$ and for some $q \in \mathbb{R}$. By the triangle inequality, $d(p, q) \leq d(x, p) + d(x, q)$ for some $p \in \bar{E}$. Thus we see that $d(p, q) \leq \epsilon + M$, which shows that $\bar{E}$ is bounded.

Thus, we see that $\bar{E}$ is compact by Heine-Borel because it is both closed and bounded. Since f is uniformly continuous on E, its extension to $\bar{E}$, let us call it $\bar{f}$, must also be continuous on $\bar{E}$. This is because for some limit point $p \in E'$, we have $d(p, x) < \rho$ for all $\rho > 0$ and $x \in E$ by the definition of limit point. Thus, we can pick a $\rho < \delta$ such that $d(p, x) < \rho < \delta$, implying that $d(f(p), f(x)) < \epsilon$ for all $\epsilon > 0$ by uniform continuity.

Since $\bar{f}$ is a continuous function on a compact set $\bar{E}$, we see that its image $\bar{f}(\bar{E})$ is also a compact set in $\mathbb{R}^1$. In particular, we know that it is bounded. Moreover, since $f(E) \subset \bar{f}(\bar{E})$, we know that $f(E)$ is bounded. □

Theorem 5.2. *If $E \subset \mathbb{R}^1$ is not bounded, then f is not bounded in E.*

Proof. Take the function $f(x) = x$ and the set $E = \mathbb{R}^1$. We can see that $f(x) = x$ is a real function that is uniformly continuous because for $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ if $|x - y| < \delta$ for all $x, y \in \mathbb{R}^1$. This is because for any $\epsilon > 0$, we can pick $\delta = \epsilon$ to satisfy the inequalities, since $f(x) = x$ and $f(y) = y$.

Moreover, it is clear that f is not bounded in E, because E itself is not bounded and the range of f is given by E. □

6. Problem 4.14

Theorem 6.1. *Let $I = [0, 1]$ be the closed unit interval. Suppose f is a continuous mapping of I onto I. Then $f(x) = x$ for at least one $x \in I$.*

Proof. Define $g(x) = f(x) - x$. Then we can see that $g(x)$ is continuous on $[0, 1]$ because $f(x)$ is continuous by assumption, x is continuous as a polynomial, and $(h - j)(x)$ of two continuous functions h and j is also continuous. Notice that if we have $g(0) = 0$ or $g(1) = 0$, then the proof is completed.

Otherwise, we must have $g(0) = f(0) - 0 > 0$ and $g(1) = f(1) - 1 < 0$ because $f(x)$ is mapped onto $I = [0, 1]$. Since $g(0) > 0$ and $g(1) < 0$, we can use the intermediate value theorem to show that there must be some z in the interval $[0, 1]$ such that $g(z) = 0$ because g is continuous. Thus, we see that $f(z) - z = 0$ or $f(z) = z$ for some $z \in I$. This completes the proof. □

7. Problem 4.15

Theorem 7.1. *Call a mapping of X into Y open if $f(V)$ is an open set in Y whenever V is an open set in X. Then every continuous open mapping of $\mathbb{R}^1$ into $\mathbb{R}^1$ is monotonic.*

Proof. First we will prove two lemmas for a continuous open mapping $f : \mathbb{R}^1 \to \mathbb{R}^1$.

Lemma 7.2. *If $a \neq b$, then $f(a) \neq f(b)$.*

Proof. Let f be a mapping in $[a, b]$. Then we know that $[a, b]$ is compact because it is real, closed, and bounded. Also, we know that for every continuous function on a compact space, $M = \sup_{x \in [a,b]} f(x)$ and $m = \inf_{x \in [a,b]} f(x)$ exist. There must be some values $x_1, x_2 \in [a, b]$ such that $f(x_1)$ is a maximum and $f(x_2)$ is a minimum. The case where $M = m$ implies that $f(x)$ is constant on $[a, b]$. Since a finite point is not open, then $f(x)$ does not satisfy the hypothesis of an open mapping for $M = m$. We must assume then that $M > m$.

Now, if $f(x_1) = m$ for some $x_1 \in (a, b)$, then f is not open on (a, b). This is because every neighborhood $N_r(f(x_1))$ contains values $p < m$ such that $p \notin f((a, b))$. Thus, for no value of r is $N_r(f(x_1)) \subset f((a, b))$, which means that $f((a, b))$ is not open for this case.

The same applies for $f(x_2) = M$ for some $x_2 \in (a, b)$. Every neighborhood $N_r(f(x_2))$ contains values $p > M$ such that $p \notin f((a, b))$. Thus, for no value of r is $N_r(f(x_2)) \subset f((a, b))$, which means that $f((a, b))$ is not open for this case.

Thus, there are only two cases left. First, we could have $f(a) = m$ and $f(b) = M$, which would imply that $f(a) < f(b)$ and thus that $f(a) \neq f(b)$. Second, we could have $f(a) = M$ and $f(b) = m$, which would imply that $f(a) > f(b)$ and thus that $f(a) \neq f(b)$. This completes the proof of the lemma. □

Now, we will proceed to the prove the second lemma.

Lemma 7.3. *If $a < b < c$ and $f(a) < f(b)$, then $f(b) < f(c)$.*

Proof. First, notice that if $f(c) = f(a)$ or $f(c) = f(b)$, this would contradict the previous lemma. Thus, there are three cases left that could possibly occur. The first case is where $f(a) < f(b)$ and $f(a) < f(c) < f(b)$. Using the intermediate value theorem, there must exist some $x \in (a, b)$ such that $f(x) = f(c)$. Since $c \notin (a, b)$, this would mean that $x \neq c$, but that $f(x) = f(c)$, which is a contradiction of our previous lemma.

The second case is if $f(c) < f(a)$. In this case, we have $f(c) < f(a) < f(b)$. Using the intermediate value theorem, there must exist some $x \in (b, c)$ such that $f(x) = f(a)$. Since $a \notin (b, c)$, we must have $x \neq a$ and $f(x) = f(a)$, which is a contradiction of our previous lemma.

Thus, the only possibility left is $f(a) < f(b) < f(c)$. Since the real field R is ordered, we know that this last possibility must always hold. □

Now, we can proceed to prove the theorem. Assume by contradiction that an open mapping f from $\mathbb{R}^1$ to $\mathbb{R}^1$ is not monotonically increasing or decreasing. Then for points $x_1 < x_2 < x_3$ in $\mathbb{R}^1$, we must have either $f(x_1) < f(x_2)$ and $f(x_3) < f(x_2)$ or $f(x_1) > f(x_2)$ and $f(x_3) > f(x_2)$. The first case is clearly a contradiction of our second lemma. For the second case, let us use $g(x) = -f(x)$ to obtain the following relations: $g(x_1) < g(x_2)$ and $g(x_3) < g(x_2)$. We know that $g(x)$ is still continuous because we have simply multiplied a continuous function by a constant. We also know that $g(x)$ is an open mapping. This is because $f(V)$ is open for every open set $V \in \mathbb{R}^1$, implying that $-f(V)$ is also open. Thus, we have found a continuous open mapping $g(x)$ which is a contradiction of our second lemma. Therefore, our assumption must be incorrect, and each open mapping from $\mathbb{R}^1$ to $\mathbb{R}^1$ is monotonic. □

8. Problem 4.16

Theorem 8.1. *Let $[x]$ denote the largest integer contained in x, that is $[x]$ is the integer suh that $x-1 < [x] \leq x$; and let $(x) = x - [x]$ denote the fractional part of x. Then $f(x) = [x]$ and $g(x) = (x)$ are discontinuous for all $x \in \mathbb{Z}$ and continuous everywhere else.*

Proof. To begin, we shall show that $f(x)$ and $g(x)$ are continuous at all points $x \notin \mathbb{Z}$. We know that there exists an integer n such that $x \in (n, n+1)$ if x is not an integer. Thus, we see that $f(x) = n$ and $g(x) = x - n$. Since n is a constant, $f(x)$ is continuous. Moreover, since x is a polynomial, which is continuous, and $(h + j)(x)$ is continuous if h and j are continuous, we can see that $g(x) = x - n$ is also continuous. This shows that for non-integer values of x, $f(x)$ and $g(x)$ are continuous.

To show that $f(x)$ and $g(x)$ are discontinuous for $x \in \mathbb{Z}$, we will compute the right-hand and left-hand limits of both functions. We see that as x approaches $n \in \mathbb{Z}$, we have the following limits:

$$\lim_{x \to n^+} [x] = n \tag{8.2}$$

$$\lim_{x \to n^-} [x] = n - 1 \tag{8.3}$$

$$\lim_{x \to n^+} (x) = 0 \tag{8.4}$$

$$\lim_{x \to n^-} (x) = 1 \tag{8.5}$$

We see that $f(n+) \neq f(n-)$ and $g(n+) \neq g(n-)$. This implies that $[x]$ and (x) are not continuous when $x \in \mathbb{Z}$. □

9. Problem 4.22

Theorem 9.1. *Let A and B be disjoint nonempty closed sets in a metric space X, and define $f(p) = \frac{\rho_A(p)}{\rho_A(p)+\rho_B(p)}$, where $p \in X$ and $\rho_E(p) = \inf_{z \in E} d(p, z)$. Then f is a continuous function on X whose range lies in $[0, 1]$.*

Proof. We will show that $\rho_E(p)$ is continuous and that $\rho_E(p) = 0$ if and only if $p \in \bar{E}$. This will allow us to prove that f is a continuous function by the rules for the addition of continuous functions and because A and B are disjoint, nonempty, and closed. Thus, we will need two lemmas.

Lemma 9.2. *If ρ_E is defined as above, then $\rho_E(p) = 0$ if and only if $p \in \bar{E}$.*

Proof. Suppose $\rho_E(p) = 0$, then $0 = \inf_{z \in E} d(p, z)$. Suppose by contradiction that $p \notin \bar{E}$ so that p is neither a limit point of E nor a point in E. Then for some $r > 0$, there does not exist a neighborhood $N_r(p)$ such that $z \in E$ and $z \in N_r(p)$. This would imply that $d(p, z) > r$, and that $r = \inf_{z \in E} d(p, z)$. Since $r > 0$ implies $r \neq 0$, we obtain a contradiction with the hypothesis that $\inf_{z \in E} d(p, z) = 0$.

For the converse, assume that $p \in \bar{E}$. Then we have either $p \in E$ or $p \in E'$. The first case is trivial because for some $z \in E$, we must have $d(p, z) = 0$. If $p \in E'$, then for every $\epsilon > 0$, we have $d(p, z) < \epsilon$. Since ϵ is arbitrary, we have $\inf_{z \in E} d(p, z) = 0$. □

Lemma 9.3. *If ρ_E is defined as above, then ρ_E is uniformly continuous.*

Proof. Since we have defined $\rho_E(p) = \inf_{z \in E} d(p, z)$, we can use the triangle inequality to show that $\inf_{z \in E} d(p, z) \leq \inf_{z \in E} d(p, y) + d(y, z)$ where $y \in X$. Moreover, since we only obtain the infimum of $d(p, y) + d(y, z)$ when we have the infimum of both $d(p, y)$ and $d(y, z)$, then we can see that $\rho_E(p) \leq d(p, y) + \rho_E(y)$. We can rearrange this to show that $|\rho_E(p) - \rho_E(y)| \leq d(p, y)$. Thus, for any $\epsilon > 0$, we can always find $\delta = \epsilon$ such that $|\rho_E(p) - \rho_E(y)| \leq d(x, y) < \delta = \epsilon$. This shows that ρ_E is uniformly continuous. □

Thus, we have shown that ρ_E is uniformly continuous, and thus continuous by the second lemma. Since ρ_A and ρ_B are both continuous, we know that $\rho_A+\rho_B$ is continuous. Moreover, we know that f is continuous at all points except possibly where $\rho_A(p)+\rho_B(p)=0$, by the rule for the division of continuous functions. However, we know that A and B are disjoint, nonempty closed sets so that $\bar{A}=A$ and $\bar{B}=B$. This means that if $x\in A$, then $x\in\bar{A}$. Since the two sets are disjoint, then $x\notin B$ implies $x\notin\bar{B}$. By the first lemma, if $\rho_A(p)=0$, then $\rho_B(p)\neq 0$ and vice versa. Thus, $\rho_A(p)+\rho_B(p)\neq 0$ for any value of $p\in X$. This means that f is continuous everywhere.

Next, we will show that f has a range $[0,1]$. First, we know that $\rho_A(p)\leq\rho_A(p)+\rho_B(p)$, since $\rho_B(p)\geq 0$ by the fact that a metric must be non-negative. This gives an upper bound of $f(p)\leq 1$. Next, since $\rho_A(p)\geq 0$ by the same fact that metrics must be non-negative, we know that $\rho_A(p)+\rho_B(p)\geq 0$ and so that $\frac{\rho_A(p)}{\rho_A(p)+\rho_B(p)}\geq 0$. This gives a lower bound $f(p)\geq 0$, and since f is continuous, it has a range of $[0,1]$. ☐

Theorem 9.4. *Let f be defined as above, then $f(p)=0$ precisely on A and $f(p)=1$ precisely on B.*

Proof. We have shown above in the first lemma that $\rho_E(p)=0$ if and only if $p\in\bar{E}$. Thus, the only time that $\rho_A(p)=0$ is when $p\in\bar{A}$ and since A is closed, this is equivalent to $p\in A$. Since $f(p)=\frac{\rho_A(p)}{\rho_A(p)+\rho_B(p)}$, the only time that $f(p)=0$ is when $\rho_A(p)=0$, which is when $p\in A$.

Moreover, we know that the only time $f(p)=1$ is when $\rho_B(p)=0$ and $\rho_A(p)\neq 0$ because then $f(p)=\frac{\rho_A(p)}{\rho_A(p)}=1$. This occurs when $p\in\bar{B}$, and since $\bar{A}\cap\bar{B}=\emptyset$, we can be sure that $p\notin\bar{A}$. Thus, $f(p)=1$ when $p\in B$, since B is closed already. ☐

Theorem 9.5. *Set $V=f^{-1}([0,\frac{1}{2}))$ and $W=f^{-1}((\frac{1}{2},1])$, then V and W are open and disjoint, and $A\subset V$, $B\subset W$.*

Proof. First, we can see that $[0,\frac{1}{2})$ and $(\frac{1}{2},1]$ are open sets in $f(X)$ because the range of $f(X)$ is $[0,1]$. Since f is continuous, we know that the sets $f^{-1}([0,\frac{1}{2}))$ and $f^{-1}((\frac{1}{2},1])$ are also open. Moreover, these two sets are disjoint because each $f(p)$ maps to a single value, so that $f^{-1}(f(p))$ belongs to only one of V or W, which ensures that $V\cap W=\emptyset$. We know that $f(p)=0$ when $p\in A$ and that $A=f^{-1}(0)\subset f^{-1}([0,\frac{1}{2}))=V$ and since $f(p)=1$ when $p\in B$, we see that $B=f^{-1}(1)\subset f^{-1}((\frac{1}{2},1])=W$. Thus, we have $A\subset V$ and $B\subset W$. ☐

10. Problem 4.23

Theorem 10.1. *A real valued function f defined in (a,b) is said to be convex if $f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y)$ whenever $a<x<b, a<y<b, 0<\lambda<1$. We will prove the last proposition first, namely that if f is convex in (a,b) and if $a<s<t<u<b$, then*

$$\frac{f(t)-f(s)}{t-s}\leq\frac{f(u)-f(s)}{u-s}\leq\frac{f(u)-f(t)}{u-t} \tag{10.2}$$

Proof. We can let $t=\lambda s+(1-\lambda)u$ where $\lambda=\frac{u-t}{u-s}$. Using this definition, it is easy to see that $0<\lambda<1$. Thus, we have the following:

$$f(t) = f(\lambda s+(1-\lambda)u)\leq\lambda f(s)+(1-\lambda)f(u) \tag{10.3}$$

$$0 \leq -f(t)+\frac{u-t}{u-s}f(s)+\left(1-\frac{u-t}{u-s}\right)f(u) \tag{10.4}$$

$$0 \leq -(u-s)f(t)+(u-t)f(s)+(t-s)f(u). \tag{10.5}$$

We can see that if we rearrange the terms in the above inequality and add $sf(s)-sf(s)$ to the right side, we obtain:

$$0 \leq (u-s)(f(s)-f(t))+sf(s)-tf(s)+(t-s)f(u) \tag{10.6}$$

$$0 \leq -(u-s)(f(t)-f(s))+(t-s)(f(u)-f(s)) \tag{10.7}$$

$$\frac{f(t)-f(s)}{t-s}\leq\frac{f(u)-f(s)}{u-s} \tag{10.8}$$

Which completes the first inequality that we wanted to show. To show the second inequality, we return to inequality (10.5) and add $uf(u)-uf(u)$ to the right hand side.

$$0 \leq (u-s)(f(u)-f(t)) - uf(u) + (u-t)f(s) + tf(u) \tag{10.9}$$
$$0 \leq (u-s)(f(u)-f(t)) - (u-t)(f(u)-f(s)) \tag{10.10}$$

$$\frac{f(u)-f(s)}{u-s} \leq \frac{f(u)-f(t)}{u-t} \tag{10.11}$$

Putting these two inequalities together, we obtain what we wanted to prove. □

Theorem 10.12. *If f is a convex function, then it is continuous.*

Proof. Let f be defined on the interval (a,b) and let $x \in (a,b)$ be an element of f. Define $\delta > 0$ such that $(x-\delta, x+\delta) \subset (a,b)$. Now consider a point $z \in (x-\delta, x)$. Then we have $a < z < x < x+\delta < b$. Thus, using the above inequalities, we find:

$$\frac{f(x)-f(z)}{x-z} \leq \frac{f(x+\delta)-f(z)}{(x+\delta)-z} \leq \frac{f(x+\delta)-f(x)}{(x+\delta)-x} \tag{10.13}$$

Since we also have $a < x-\delta < z < x < b$, we can use the previously derived inequalities to obtain:

$$\frac{f(z)-f(x-\delta)}{z-(x-\delta)} \leq \frac{f(x)-f(x-\delta)}{x-(x-\delta)} \leq \frac{f(x)-f(z)}{x-z} \tag{10.14}$$

Combining the above inequalities, we can derive:

$$\frac{f(x)-f(x-\delta)}{\delta} \leq \frac{f(x)-f(z)}{x-z} \leq \frac{f(x+\delta)-f(x)}{\delta} \tag{10.15}$$

These inequalities hold similarly for $z \in (x, x+\delta)$ by symmetric arguments. Thus, since we have fixed x and δ, we see that $c_1(x-z) \leq f(x)-f(z) \leq c_2(x-z)$, for constants c_1 and c_2. This further implies, for all $z \in (x-\delta, x+\delta)$:

$$|f(x)-f(z)| \leq C(x-z) \tag{10.16}$$

Where C is a constant. We have shown that if $|x-z| < B$, then there exists some number A where $|f(x)-f(z)| < A$, which shows that a convex function is continuous. □

Theorem 10.17. *Every increasing convex function of a convex function is convex.*

Proof. Define $f(x)$ be a convex function defined on (a,b) and $g(x)$ be an increasing convex function on (c,d) where $(a,b) \subset (c,d)$. We must show that $h(x) = g(f(x))$ is convex. Since $f(x)$ is convex, we know that $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ for $x, y \in (a,b)$ and $0 < \lambda < 1$. Since $g(x)$ is increasing, we know that $g(f(\lambda x + (1-\lambda)y)) \leq g(\lambda f(x) + (1-\lambda)f(y))$. Thus, since $g(x)$ is also convex, we have:

$$h(\lambda x + (1-\lambda)y) = g(f(\lambda x + (1-\lambda)y)) \tag{10.18}$$
$$\leq g(\lambda f(x) + (1-\lambda)f(y)) \tag{10.19}$$
$$\leq \lambda g(f(x)) + (1-\lambda)g(f(y)) \tag{10.20}$$
$$\leq \lambda h(x) + (1-\lambda)h(y) \tag{10.21}$$

This proves the convexity of $h(x)$. □

11. Problem 4.24

Theorem 11.1. *Assume that f is a continuous real function defined in (a,b) such that $f(\frac{x+y}{2}) \leq \frac{f(x)+f(y)}{2}$ for all $x, y \in (a,b)$, then f is convex.*

Proof. First, we will use mathematical induction to show the convexity of f for every λ of the form $\frac{m}{2^n}$ where $0 \leq m \leq 2^n$ and $m, n \in \mathbb{Z}_+$. Then, we will show that every $\lambda \in \mathbb{R}$ has a sequence of numbers of the form $\frac{m}{2^n}$ converging to it.

So, let $\lambda = \frac{m}{2^n}$, $0 \leq m \leq 2^n$, and $m, n \in \mathbb{Z}_+$. Now start with the base case of $n = 1$. The possible values of m are given by $m = 0, 1, 2$. It is trivial to show convexity for $m = 0$ and $m = 2$, since $\lambda = 0$ or $\lambda = 1$ respectively. For $m = 1$, we have:

$$f(\lambda x+(1-\lambda)y)=f\left(\frac{x+y}{2}\right)\leq\frac{f(x)+f(y)}{2}=\lambda f(x)+(1-\lambda)f(y) \tag{11.2}$$

Thus, we have established convexity of f for $n=1$. Now assume we have shown convexity of f for $n=k$. We shall proceed to show that f is convex for $n=k+1$. We can rewrite λ as the following:

$$\lambda=\frac{m}{2^{k+1}}=\frac{1}{2}\left(\frac{m-1}{2^k}+\frac{1}{2^k}\right) \tag{11.3}$$

Letting $\lambda_1=\frac{m-1}{2^k}$ and $\lambda_2=\frac{1}{2^k}$, we see that $\lambda=\frac{1}{2}(\lambda_1+\lambda_2)$. We have assumed f is convex for $\lambda_1=\frac{m-1}{2^k}$ and $\lambda_2=\frac{1}{2^k}$ because we have assumed convexity for all $n=k$. Therefore, we know that $f(\lambda_1x+(1-\lambda_1)y)\leq\lambda_1f(x)+(1-\lambda_1)f(y)$ and $f(\lambda_2x+(1-\lambda_2)y)\leq\lambda_2f(x)+(1-\lambda_2)f(y)$. Thus, we have:

$$\begin{aligned}
f(\lambda x+(1-\lambda)y) &= f\left(\frac{\lambda_1x+(1-\lambda_1)y+\lambda_2x+(1-\lambda_2)y}{2}\right) && (11.4)\\
&\leq \frac{1}{2}f(\lambda_1x+(1-\lambda_1)y)+\frac{1}{2}f(\lambda_2x+(1-\lambda_2)y) && (11.5)\\
&\leq \frac{1}{2}\lambda_1f(x)+\frac{1}{2}(1-\lambda_1)f(y)+\frac{1}{2}\lambda_2f(x)+\frac{1}{2}(1-\lambda_2)f(y) && (11.6)\\
&= \frac{1}{2}(\lambda_1+\lambda_2)f(x)+(1+\frac{1}{2}(\lambda_1+\lambda_2))f(y) && (11.7)\\
&= \lambda f(x)+(1-\lambda)f(y) && (11.8)
\end{aligned}$$

Thus, we have shown by induction that if $\lambda=\frac{m}{2^n}$ for $m,n\in\mathbb{Z}_+$ and $0\leq m\leq 2^n$, then f is convex. Now, we must show that f is convex for any $\lambda\in\mathbb{R}$ with $0<\lambda<1$. We will show that every $\lambda\in\mathbb{R}$ in $0<\lambda<1$ has a sequence of $\{\lambda_k\}$ that converges to λ. This is because each real number has a binary expansion (see the lemma at the bottom). This implies that there exists a sequence $\{\lambda_k\}$ of the form $\frac{m}{2^n}$ such that $\{\lambda_k\}\to\lambda$ as $k\to\infty$. This implies the that $f(\lambda x+(1-\lambda)y)=\lim_{k\to\infty}f(\lambda_kx+(1-\lambda_k)y)\leq\lim_{k\to\infty}\lambda_kf(x)+(1-\lambda_k)f(y)=\lambda f(x)+(1-\lambda)f(x)$ for $\lambda\in(0,1)$. Thus, we have shown f is convex for all $x,y\in(a,b)$ and $\lambda\in(0,1)$. ☐

Lemma 11.9. *Each real number $x\in(0,1)$ has a binary expansion.*

Proof. Since we have $x\in(0,1)$, we will construct $x_n\leq x<x_n+2^{-n}$ and $a_n\in\{0,1\}$ such that $x_{n+1}=x_n+a_{n+1}2^{-(n+1)}$. Thus we will have

$$x_n=\sum_{i=1}^{n}a_i2^{-i} \tag{11.10}$$

First, let $a_0=x_0=[x]$. Then since $x\in(0,1)$ and $a_0\in\{0,1\}$, we have $x_0\leq x<x_0+\frac{1}{2}$. Now we use mathematical induction to construct the binary expansion. Assume we already have constructed x_n so that $x_n\leq x<x_n+2^{-n}$. Then we want to construct x_{n+1}. We know the interval $[x_n,x_n+2^{-n})$ has a length 2^{-n} between its two endpoints. Let $x_{n+1}\in[x_n,x_n+2^{-n})$ be the largest number of the form $x_n+a_{n+1}2^{-(n+1)}$ which does not exceed x and $a_{n+1}\in\{0,1\}$. Therefore, we know that $x_{n+1}=x_n+a_{n+1}2^{-(n+1)}$ and that $x_{n+1}\leq x<x_n+(a_{n+1}+1)2^{-(n+1)}=x_{n+1}+2^{-(n+1)}$.

Therefore, we have shown that $x\in[x_{n+1},x_{n+1}+2^{-(n+1)})$. Therefore, by induction, we have constructed x_n. To show that $\{x_n\}$ converges to x, we see that $\{x_n\}$ is monotonically increasing by construction and bounded from above by x. Therefore, there is a limit $p\in\mathbb{R}$. Also, we know that:

$$\lim_{n\to\infty}x_n\leq x\leq\lim_{n\to\infty}x+2^{-n} \tag{11.11}$$

And since $\lim_{n\to\infty}x_n=p$ and $\lim_{n\to\infty}x+2^{-n}=p$, we have $p\leq x\leq p$ so that by the squeeze theorem, we must have $p=x$. This shows that each real number $x\in(0,1)$ has a binary expansion. ☐