Answer all of the questions below. Be as complete, correct, and concise as possible. There are 6 questions for a total of 165 points possible; point values for each problem are in parentheses. For questions with subparts, each subpart is worth the same number of points. You have 180 minutes to complete the exam. Use the points as a guide to allocating your time. You may use any result from class with appropriate references unless you are specifically being asked to prove it.

1. (15) Define or state each of the following.
   (a) continuous function \( f : X \rightarrow Y \) from a metric space \((X, d)\) to a metric space \((Y, \rho)\)
   (b) eigenvalue of a linear transformation
   (c) Separating Hyperplane Theorem

   **Solution:** See notes.

2. (30) Prove that for every \( n \in \mathbb{N} = \{1, 2, 3, \ldots\} \),

\[
\sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{n}{n+1}
\]

   **Solution:** The proof is by induction. For the base case, let \( n = 1 \). Then

\[
\frac{1}{1(1+1)} = \frac{1}{2} = \frac{1}{1+1}
\]

Thus the claim is true for \( n = 1 \). For the induction hypothesis, assume that for some \( n \geq 1 \),

\[
\sum_{k=1}^{n-1} \frac{1}{k(k+1)} = \frac{n-1}{n-1+1} = \frac{n-1}{n}
\]

Now consider \( n \):

\[
\sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n-1} \frac{1}{k(k+1)} + \frac{1}{n(n+1)}
\]

\[
= \frac{n-1}{n} + \frac{1}{n(n+1)} \quad \text{by the induction hypothesis}
\]

\[
= \frac{(n-1)(n+1) + 1}{n(n+1)}
\]
Thus by induction, the claim holds for every $n \in \mathbb{N}$.

3. Let $X$ and $Y$ be vector spaces over the same field $F$. Let $T : X \rightarrow Y$ be a linear transformation such that $\ker T = \{0\}$. Show that if the set $V$ is a basis of $X$, then the set $T(V) = \{T(v) : v \in V\}$ is a basis of $\text{Im } T$.

**Solution:** $V$ is a basis of $X$, so $V$ is linearly independent and $V$ spans $X$. To show that $T(V)$ is a basis of $\text{Im } T$, we must show that $T(V)$ is linearly independent and that $T(V)$ spans $\text{Im } T$.

First, $T(V)$ spans $\text{Im } T$. To show this, let $y \in \text{span } T(V)$, so $y = \sum_{i=1}^{n} \alpha_i T(v_i)$ for some $\alpha_1, \ldots, \alpha_n \in F$ and some $v_1, \ldots, v_n \in V$. Then

$$y = \sum_{i=1}^{n} \alpha_i T(v_i) = T\left(\sum_{i=1}^{n} \alpha_i v_i\right)$$

since $T$ is linear. Then $\sum_{i=1}^{n} \alpha_i v_i \in X$, so $y = T\left(\sum_{i=1}^{n} \alpha_i v_i\right) \in \text{Im } T$ by definition. So span $T(V) \subseteq \text{Im } T$. Now let $y \in \text{Im } T$. Then by definition, there exists $x \in X$ such that $y = T(x)$. Since $V$ is a basis of $X$, there exist $\beta_1, \ldots, \beta_m \in F$ and $v_1, \ldots, v_m \in V$ such that $x = \sum_{i=1}^{m} \beta_i v_i$. Then

$$y = T(x) = T\left(\sum_{i=1}^{m} \beta_i v_i\right) = \sum_{i=1}^{m} \beta_i T(v_i) \quad \text{since } T \text{ is linear}$$

Thus $y \in \text{span } T(V)$. So $\text{Im } T \subseteq \text{span } T$, which implies that $\text{Im } T = \text{span } T(V)$. Then $T(V)$ is linearly independent. To see this, suppose $\sum_{i=1}^{n} \gamma_i T(v_i) = 0$ for some $\gamma_1, \ldots, \gamma_n \in F$ and some $v_1, \ldots, v_n \in V$. Then

$$\sum_{i=1}^{n} \gamma_i T(v_i) = 0$$

$$\Rightarrow T\left(\sum_{i=1}^{n} \gamma_i v_i\right) = 0 \quad \text{since } T \text{ is linear}$$

$$\Rightarrow \sum_{i=1}^{n} \gamma_i v_i = 0 \quad \text{since } \ker T = \{0\}$$

$$\Rightarrow \gamma_i = 0 \ \forall i = 1, \ldots, n \quad \text{since } V \text{ is linearly independent}$$

Thus $T(V)$ is linearly independent.
4. (30) Let \( f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n \) be continuous.

(a) Let \( C \subseteq \mathbb{R}^n \) be compact. Let \( \Psi : \mathbb{R}^p \to 2^C \) be given by

\[
\Psi(a) = \{ x \in C : f(x, a) = 0 \}
\]

Show that \( \Psi \) is upper hemicontinuous.

\textbf{Solution:} First note that since \( C \) is compact and \( \Psi(a) \subseteq C \) for each \( a \), it suffices to show that \( \Psi \) has closed graph. To that end, suppose \((a_n, x_n) \in \text{graph } \Psi\) for each \( n \) and \((a_n, x_n) \to (a, x) \in \mathbb{R}^p \times C\). Then since \((a_n, x_n) \in \text{graph } \Psi\) for each \( n \), \( f(x_n, a_n) = 0 \) for each \( n \). Since \( f \) is continuous,

\[
f(x, a) = \lim_{n} f(x_n, a_n) = 0
\]

Thus \( x \in \Psi(a) \), that is, \((a, x) \in \text{graph } \Psi\). This shows that \( \Psi \) has closed graph, which implies that \( \Psi \) is uhc.

Here is an argument using the definition of uhc. Fix \( a \in \mathbb{R}^p \) and let \( V \subseteq \mathbb{R}^n \) be an open set such that \( \Psi(a) \subseteq V \). We must show that there exists an open set \( U \subseteq \mathbb{R}^p \) such that \( a \in U \) and for all \( a' \in U \), \( \Psi(a') \subseteq V \). If not, then for each \( n \) there exists \( a_n \in B_{\frac{1}{n}}(a) \) and \( x_n \in \Psi(a_n) \) such that \( x_n \notin V \). By construction, \( a_n \to a \) and \( \{x_n\} \subseteq C \). Since \( C \) is compact, \( \{x_n\} \) has a subsequence \( \{x_{n_k}\} \) converging to an element \( x \in C \). Thus \( x_{n_k} \to x \) and \( a_{n_k} \to a \), and by construction \( x_{n_k} \in \Psi(a_{n_k}) \) for each \( k \). Then by definition, \( f(x_{n_k}, a_{n_k}) = 0 \) for each \( k \). Since \( f \) is continuous,

\[
f(x, a) = \lim_{k} f(x_{n_k}, a_{n_k}) = 0
\]

Thus \( x \in \Psi(a) \subseteq V \). But \( V \) is open and \( x_{n_k} \to x \), so there exists \( K \) sufficiently large so that for all \( k \geq K \), \( x_{n_k} \in V \). This is a contradiction, since \( x_{n_k} \notin V \) for all \( k \) by construction. Thus there exists an open set \( U \subseteq \mathbb{R}^p \) with \( a \in U \) such that for all \( a' \in U \), \( \Psi(a') \subseteq V \). Thus \( \Psi \) is uhc at \( a \). Since \( a \) was arbitrary, \( \Psi \) is uhc.

(b) Let \( X \subseteq \mathbb{R}^n \) and \( A \subseteq \mathbb{R}^p \) be open sets, and let \( f : X \times A \to \mathbb{R}^n \) be \( C^1 \). Let \( \Psi : A \to 2^X \) be given by

\[
\Psi(a) = \{ x \in X : f(x, a) = 0 \}
\]

Suppose \( 0 \) is a regular value of \( f(\cdot, a^\ast) \), that is, for each \( x^\ast \) such that \( f(x^\ast, a^\ast) = 0 \), \( \det(D_xf(x^\ast, a^\ast)) \neq 0 \). Show that \( \Psi \) is lower hemicontinuous at \( a^\ast \).

\textbf{Solution:} Let \( V \subseteq X \) be open and \( \Psi(a^\ast) \cap V \neq \emptyset \). Then \( \exists x^\ast \in \Psi(a^\ast) \cap V \), so \( x^\ast \in V \) and \( f(x^\ast, a^\ast) = 0 \). By assumption, \( \det(D_xf(x^\ast, a^\ast)) \neq 0 \). Then by the
Implicit Function Theorem, \( \exists \) open sets \( U \ni x^* \) and \( W \ni a^* \) and a \( C^1 \) function \( g : W \to U \) such that
\[
\forall (x, a) \in U \times W \ f(x, a) = 0 \iff x = g(a)
\]
Then \( g(a^*) = x^* \in U \cap V \). Since \( U \cap V \) is an open set and \( g \) is continuous, there exists an open set \( W' \ni a^* \) such that \( a \in W' \Rightarrow g(a) \in U \cap V \). By definition, \( f(g(a), a) = 0 \) for each \( a \in W' \), that is, \( g(a) \in \Psi(a) \) for each \( a \in W' \). So
\[
\Psi(a) \cap V \neq \emptyset \quad \forall a \in W'
\]
Therefore \( \Psi \) is lhc at \( a^* \).

5. (30) Let \( X \subseteq \mathbb{R}^n \) be open and \( f : X \to \mathbb{R} \) be differentiable. Let \( C \subseteq X \) be a compact, convex set such that for each \( x \in C \), there exists \( \varepsilon_x > 0 \) and \( M_x \geq 0 \) such that \( \|Df(y)\| \leq M_x \) for all \( y \in B_{\varepsilon_x}(x) \). Show that \( f \) is Lipschitz continuous on \( C \).

(Hint: Use the open cover definition of compactness.)

**Solution:** The collection \( \{B_{\varepsilon_x}(x) : x \in C\} \) is an open cover of \( C \), and \( C \) is compact, so there exists \( x_1, \ldots, x_n \in C \) such that
\[
C \subseteq B_{\varepsilon_{x_1}}(x_1) \cup \cdots \cup B_{\varepsilon_{x_n}}(x_n)
\]
Then let \( x, y \in C \). By the Mean Value Theorem, there exists \( z \in \ell(x, y) = \{\alpha x + (1 - \alpha)y : \alpha \in [0, 1]\} \) such that
\[
f(x) - f(y) = Df(z)(x - y)
\]
Since \( C \) is convex and \( x, y \in C, z \in C \). Thus \( z \in B_{\varepsilon_{x_i}}(x_i) \) for some \( i = 1, \ldots, n \), which implies
\[
\|Df(z)\| \leq M_{x_i}
\]
Set \( M = \max\{1, M_{x_1}, \ldots, M_{x_n}\} \). Then \( M \) is defined (that is, \( M \) is finite), and \( M > 0 \). Then
\[
\|f(x) - f(y)\| = \|Df(z)(x - y)\|
\leq \|Df(z)\|\|x - y\|
\leq M_{x_i}\|x - y\|
\leq M\|x - y\|
\]
Since \( x, y \in C \) were arbitrary, \( f \) is Lipschitz continuous on \( C \).
6. (30) Let \((X, d)\) be a metric space, and \(C \subseteq X\) be a nonempty compact set. Let 
\(f : C \to C\) be a function such that 
\[
d(f(x), f(y)) < d(x, y) \quad \forall x, y \in C \text{ such that } x \neq y
\]
Show that \(f\) has a unique fixed point in \(C\).

(Hint: No theorem from class will immediately imply this result.)

Solution: First, note that \(f\) is continuous, as for \(\varepsilon > 0\), set \(\delta = \varepsilon\). Then 
\[
d(x, y) < \delta \Rightarrow \text{either } x = y, \text{ in which case } d(f(x), f(y)) = 0, \text{ or } x \neq y \text{ and}
\]
\[
d(f(x), f(y)) < d(x, y) < \delta = \varepsilon
\]
Thus 
\[
d(x, y) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon
\]
Now note that \(f\) can have at most one fixed point in \(C\), as if \(x^*, y^* \in C\) are both fixed 
points of \(f\) but \(x^* \neq y^*\), then \(f(x^*) = x^*\) and \(f(y^*) = y^*\), so 
\[
0 < d(x^*, y^*) = d(f(x^*), f(y^*)) < d(x^*, y^*)
\]
which is a contradiction.

Then note that \(x^* = f(x^*) \iff d(x^*, f(x^*)) = 0\). Since \(f\) is continuous, the function 
\(g : C \to \mathbb{R}\) given by 
\[
g(x) = d(x, f(x))
\]
is continuous. Since \(C\) is compact, by the Extreme Value Theorem \(g\) achieves its 
minimum on \(C\), that is, \(\exists x^* \in C\) such that 
\[
g(x^*) = d(x^*, f(x^*)) \leq d(x, f(x)) = g(x) \quad \forall x \in C
\]
Now claim \(d(x^*, f(x^*)) = 0\) (so that \(x^*\) is the unique fixed point of \(f\)). To see this, 
suppose not. Then \(d(x^*, f(x^*)) > 0\), so \(x^* \neq f(x^*)\). But then 
\[
d(f(x^*), f(f(x^*))) < d(x^*, f(x^*)) \text{ and } f(x^*) \in C
\]
This is a contradiction. Therefore \(d(x^*, f(x^*)) = 0\), that is, \(x^* = f(x^*)\).

Here is a second argument that is constructive, along the lines of the proof of the 
Contraction Mapping Theorem. The Contraction Mapping Theorem does not apply, 
and the conditions of the problem are not sufficient to imply that \(f\) is a contraction 
(these were both common mistakes). But the same algorithm used in the proof of 
the Contraction Mapping Theorem yields a sequence that converges to the fixed point 
from any initial point.

Fix \(x_0 \in C\). Let \(x_n = f^n(x_0)\) for each \(n \geq 1\). Then \(\{x_n\} \subseteq C\), and \(C\) is compact, so 
there is a subsequence \(\{x_{n_k}\}\) such that \(x_{n_k} \to x^*\) for some \(x^* \in C\). Note by definition 
\[
x_{n_k+1} = f(x_{n_k}) = f^{n_k+1}(x_0) \quad \forall k
\]
and since $f$ is continuous,

$$x_{n_k+1} = f(x_{n_k}) = f^{n_k+1}(x_0) \to f(x^*)$$

Now claim $x^* = f(x^*)$. To see this, suppose not. Then $d(x^*, f(x^*)) = c > 0$, and since $x^* \neq f(x^*)$,

$$d(f(x^*), f(f(x^*))) < d(x^*, f(x^*))$$

So

$$0 \leq \frac{d(f(x^*), f(f(x^*)))}{d(x^*, f(x^*)))} < r < 1 \quad \text{for some } r \in (0, 1)$$

Let $D = \{(x, y) \in C \times C : x = y\}$; note that $D$ is a closed set, and $x^* \neq f(x^*) \Rightarrow (x^*, f(x^*)) \notin D$. Since $f$ is continuous, there is an open set $U \subseteq (C \times C) \setminus D$ with $(x^*, f(x^*)) \in U$ and $R \in (0, 1)$ such that $\forall (x, y) \in U$,

$$0 \leq \frac{d(f(x), f(y))}{d(x, y)} < R < 1$$

Then choose $\varepsilon > 0$ sufficiently small so that $\varepsilon < \frac{c}{4}$ and such that

$$B_\varepsilon(x^*) \times B_\varepsilon(f(x^*)) \subseteq U$$

Since $x_{n_k} \to x^*$ and $f(x_{n_k}) \to f(x^*)$, there exists $K > 0$ such that $k \geq K \Rightarrow$

$$f^{n_k}(x_0) = x_{n_k} \in B_\varepsilon(x^*) \quad \text{and} \quad f^{n_k+1}(x_0) = f(x_{n_k}) \in B_\varepsilon(f(x^*))$$

Thus for all $k \geq K$,

$$d(f^{n_k}(x_0), f^{n_k+1}(x_0)) \geq \frac{c}{2} > 0$$

Then note that $\forall \ell \geq k \geq K$,

$$d(f^{n_\ell}(x_0), f^{n_\ell+1}(x_0)) \leq d(f^{n_{\ell-1}}(x_0), f^{n_{\ell-1}+1}(x_0)) \leq Rd(f^{n_{\ell-1}}(x_0), f^{n_{\ell-1}+1}(x_0)) \leq \cdots < R^{\ell-k}d(f^{n_k}(x_0), f^{n_k+1}(x_0))$$

This holds $\forall \ell \geq K$, so

$$\lim_{\ell} d(f^{n_\ell}(x_0), f^{n_\ell+1}(x_0)) = 0$$

But this is a contradiction, as for all $k \geq K$,

$$d(f^{n_k}(x_0), f^{n_k+1}(x_0)) \geq \frac{c}{2} > 0$$

Therefore $x^* = f(x^*)$.

From this it can also be shown that $x_n \to x^*$ (that is, the original sequence $\{x_n\}$ converges to $x^*$, not just some subsequence $\{x_{n_k}\}$). This is a good exercise.