ME 550 Foundations of Engineering Systems Analysis

Hahn-Banach Theorem

The Hahn-Banach theorem is an extension for linear functionals and has many other applications. It allows manipulation of normed spaces and associated bounded linear functionals, and also provides an adequate theory of dual spaces. Specifically, it states that a bounded linear functional on a subspace of a normed vector space can be extended to a bounded linear functional on the entire space with the same norm.

I. Zorn's Lemma

Zorn's Lemma is necessary in the proof of Hahn-Banach Theorem and has also other applications. We introduce introduce the concept of Zorn's Lemma.

Definition 1.1: A partially ordered set, abbreviated as poset, is a set S on which a binary relation, known as partial ordering and denoted as $\preccurlyeq$, satisfies the following conditions for every $\alpha, \beta, \gamma \in S$:

$$\alpha \preccurlyeq \alpha \quad (Reflexivity)$$

$$If\ \alpha \preccurlyeq \beta \text{ and } \beta \preccurlyeq \alpha, \text{ then } \alpha = \beta \quad (Antisymmetry)$$

$$If \alpha \preccurlyeq \beta \text{ and } \beta \preccurlyeq \gamma, \text{ then } \alpha \preccurlyeq \gamma \quad (Transitivity)$$

Definition 1.2: Two elements α and β of a partially ordered set are called *incomparable* for which neither $\alpha \preccurlyeq \beta$ nor $\beta \preccurlyeq \alpha$ holds; two elements are called *comparable* if they satisfy the condition $\alpha \preccurlyeq \beta$ or $\beta \preccurlyeq \alpha$ or both.

Definition 1.3: A totally ordered (also called linearly ordered) set or a chain is a partially ordered set such that every pair of elements in the set are comparable. In other words, a chain is a a partially ordered set having no incomparable elements.

Definition 1.4: Let $(\mathcal{P}, \preccurlyeq)$ be a partially ordered set. Then, $\mathcal{Q}$ is a maximally totally ordered subset of $\mathcal{P}$ if (i) $\mathcal{Q} \subseteq \mathcal{P}$, (ii) $(\mathcal{Q}, \preccurlyeq)$ is totally ordered, and (iii) if any member of $\mathcal{P}$ not in $\mathcal{Q}$ is adjoined to $\mathcal{Q}$, then the resulting collection of sets is no longer totally ordered by $\preccurlyeq$.

Remark 1.1: Every subset of a nonempty set, which consists of a single element, is totally ordered.

Definition 1.5: Let S be a partially ordered set. An upper bound of $W \subseteq S$ is an element $\alpha \in S$ such that

$$\theta \preccurlyeq \alpha \quad \forall\ \theta \in W \tag{1}$$

A lower bound of $W \subseteq S$ is an element $\beta \in S$ such that

$$\beta \preccurlyeq \theta \quad \forall\ \theta \in W \tag{2}$$

Depending on S and W, an upper bound or a lower bound of W may or may not exist.

Definition 1.6: Let $(S, \preccurlyeq)$ be a partially ordered set. An element $\alpha \in S$ is called a maximal element of S if $\theta \preccurlyeq \alpha$ for every $\theta \in S$ which is comparable to α. In other words,

$$\forall \theta \in S, (\alpha \preccurlyeq \theta) \ \Rightarrow \ (\alpha = \theta) \tag{3}$$

Similarly, a minimal element of S is an element $\beta \in S$ such that

$$\forall \theta \in S, (\theta \preccurlyeq \beta) \ \Rightarrow \ (\beta = \theta) \tag{4}$$

A partially ordered set S may or may not have a maximal element or a minimal element. Furthermore, a maximal element need not be an upper bound. Similarly, a minimal element need not be a lower bound.

Example 1.1: Let $S = (0, 1) \subset \mathbb{R}$; then, $(S, \leq)$ is a totally ordered set that has no maximal element and no minimal element. However, $1 \in \mathbb{R}$ is an upper bound of S; similarly, $0 \in \mathbb{R}$ is a lower bound of S. As a matter of fact, 1 is the least upper bound of S and 0 is the greatest lower bound of S.

Example 1.2: Let S be the set of all points (x, y) in the plane $\mathbb{R}^2$ with $y \leq 0$. Let us define an ordering $\preccurlyeq$ on S as

$$\Big((x, y) \preccurlyeq (\tilde{x}, \tilde{y})\Big) \ \Rightarrow \ \Big((x = \tilde{x}) \ \bigwedge \ (y \leq \tilde{y})\Big)$$

Then, the partially ordered set $(S, \preccurlyeq)$ has infinitely many maximal elements.

Zorn's lemma: Let $S \neq \emptyset$ be a partially ordered set such that every chain $T \subseteq S$ has an upper bound. Then, S has at least one maximal element.

Hausdorff Maximality Theorem: Every (nonempty) partially ordered set contains a maximal totally ordered subset. In other words, if S is a maximal totally ordered subset of a (nonempty) partially ordered set X and if T is a totally ordered subset of X, then $\big(S \subseteq T \subseteq X\big) \Rightarrow \big(S = T\big)$.

Axiom of Choice: Let $S \neq \emptyset$ be a set and $\mathcal{I} \neq \emptyset$ be an index set. Then, there exists a mapping, called the *choice function*, $f : \mathcal{I} \to S$ such that $f(\alpha) \in S_\alpha \subseteq S$ and $S_\alpha \neq \emptyset$. That is, for every nonempty set, there exists a choice function.

The axiom of choice can also be stated as: *The product of a family of nonempty sets indexed by a nonempty set is nonempty*.

Remark 1.2: Zorn's Lemma and Hausdorff Maximality Theorem are equivalent and they are also equivalent to Axiom of Choice. For details, see Appendix, pp. 392-393, on Hausdorff Maximality Theorem in *Real and Complex Analysis* by Rudin and p. 13 in *Algebra* by Thomas Hungerford.

Let us illustrate a simple application of Zorn's lemma. We first make the following assertions:

- V is a vector space and A is a set of linearly independent vectors belonging to V.
- $\mathbb{X}$ is the collection of all linearly independent sets of vectors in V such that A is a subset of each member in $\mathbb{X}$.
- $\subseteq$ is a partial ordering on $\mathbb{X}$.
- H is a Hamel basis of V such that $A \subseteq H$.
- $\mathcal{I}$ is a non-empty index set and $\mathbb{Y} = \{B_i : i \in \mathcal{I}\}$ is a chain of $\mathbb{X}$.
- $B = \bigcup_{i \in \mathcal{I}} B_i$

It follows that the sets in the chain $\mathbb{Y}$ can be ordered as: $B_{i_1} \subseteq B_{i_2} \subseteq \cdots \subseteq B_{i_n} \subseteq \cdots$ and $\mathbb{Y}$ has an upper bound B. Since $\mathbb{Y}$ can be arbitrarily chosen, $\mathbb{X}$ has a maximal element H by Zorn's lemma.

II. Extension of Linear Functionals

In Hahn-Banach theorem, the objective is to extend a linear functional f, defined on a subspace U of a vector space V, which has a certain boundedness property.

Definition 2.1: Let f be a linear functional on a subspace U of a vector space V over the real field $\mathbb{R}$. A linear functional f_{ext}, on another subspace $W \subseteq V$, is called an extension of f from U to W if

- U is a proper subspace of W, i.e., $U \subset W$.
- $f_{ext}(x) = f(x) \quad \forall x \in U$

Definition 2.2: Let V be a vector space over the real field $\mathbb{R}$, and let $p : V \to \mathbb{R}$. Then, p is called a sublinear functional on V if it has the following two properties:

- *Subadditive*: $p(x + y) \leq p(x) + p(y) \quad \forall x, y \in V$
- *Positive homogeneous*: $p(\alpha x) = \alpha p(x) \quad \forall \alpha \in [0, \infty) \quad \forall x \in V$

Remark 2.1: A norm on a vector space is a sublinear functional.

Theorem 2.1: (Hahn-Banach Theorem: Extension of Linear Functionals)

Let V be a vector space over the real field $\mathbb{R}$, and let p be a sublinear functional on V. Let f be a linear functional on a subspace $U \subset V$. If $f(x) \leq p(x) \quad \forall x \in U$, then there exists an extension f_{ext} of f from U to V such that $f_{ext}(x) \leq p(x) \quad \forall x \in V$.

Proof: The theorem is proved in the following three steps:

- $Step$ 1: Let us construct the set E consisting of the linear functional f and all linear extensions g of f, which satisfy the relation: $g(x) \leq p(x)$ on the domain $\mathcal{D}(g)$. The set E is partially ordered and Zorn's lemma yields a maximal element $f_{ext} \in E$.
- $Step$ 2: The linear functional f_{ext} is defined on the entire space V.
- $Step$ 3: The relation $f_{ext} \leq p(x) \; \forall x \in V$ is established.

Step 1: It is obvious that E is nonempty because $f \in E$. Let us define a partial ordering on E as: $g \leq h \Rightarrow h$ is an extension of g. That is, $\mathcal{D}(g) \subseteq \mathcal{D}(h)$ and $h(x) = g(x) \quad \forall x \in \mathcal{D}(g)$.

For any chain $H \subseteq E$, let us define a linear functional $\widetilde{g} \in E$ as:

$$\mathcal{D}(\widetilde{g}) = \bigcup_{g \in H} \mathcal{D}(g) \text{ and } \widetilde{g}(x) = g(x) \text{ if } x \in \mathcal{D}(g) \tag{5}$$

Note that, for an $x \in \mathcal{D}(g_1) \bigcap \mathcal{D}(g_2)$ with $g_1, g_2 \in H$, we have $g_1(x) = g_2(x)$ because H is a chain so that $g_1 \leq g_2$ or

$g_2 \leq g_1$. Then, $g \leq \widetilde{g}$ for all $g \in H$. Hence, H has an upper bound. Since selection of $H \subseteq E$ is arbitrary, Zorn's lemma implies that E has a maximal element; let us call this maximal element as f_{ext}. By definition, f_{ext} is a linear extension of f that satisfies the condition:

$$f_{ext}(x) \leq p(x) \quad \forall x \in \mathcal{D}(f_{ext}) \tag{6}$$

Step 2: Now we prove, by contradiction, that $\mathcal{D}(f_{ext})$ spans the entire vector space V. Let us assume that the assertion is false, i.e., $\mathcal{D}(f_{ext})$ is a proper subset of V. Then, there exists $z \in \big(V - \mathcal{D}(f_{ext})\big)$ and $z \neq \underline{0}$ because $\underline{0} \in \mathcal{D}(f_{ext})$. Let the subspace W be spanned by $\mathcal{D}(f_{ext})$ and the vector z. Thus, any $x \in W$ can be expressed as:

$$x = y + \alpha z \ \text{ where } y \in \mathcal{D}(f_{ext}) \ \text{ and } \alpha \text{ is a scalar} \tag{7}$$

The above representation is unique because $y \in \mathcal{D}(f_{ext})$ and $z \in \big(V - \mathcal{D}(f_{ext})\big)$.

A linear functional g on W is defined by

$$g(y + \alpha z) = f_{ext}(y) + \alpha c \ \text{ where } \ g(z) = c \in \mathbb{R} \tag{8}$$

Note that g is a proper extension of f_{ext}, i.e., $\mathcal{D}(f_{ext})$ is a proper subset of $\mathcal{D}(g)$, because if $\alpha = 0$, then $g(y) = f_{ext}(y) \quad \forall y \in \mathcal{D}(f_{ext})$. Consequently, if it is proven that $g \in E$ by showing that $g(x) \leq p(x) \ \forall x \in \mathcal{D}(g)$, then this will contradict maximality of f_{ext} so that the assertion $\mathcal{D}(f_{ext}) \neq V$ is false, i.e., the truth of the statement $\mathcal{D}(f_{ext}) = V$ is established.

Step 3: We will show that g with a real constant value of c in Eq. (8) satisfies the condition $g(x) \leq p(x) \ \ \forall x \in \mathcal{D}(g)$.

Let $y, z \in \mathcal{D}(f_{ext})$ and let $w \in \mathcal{D}(f_{ext})$ be fixed. Since p is a subadditive functional and the linear functional $f_{ext} \leq p$,

$$\begin{aligned} f_{ext}(y) - f_{ext}(z) = f_{ext}(y - z) &\leq p(y - z) \\ &= p(y + w - w - z) \\ &\leq p(y + w) + p(-w - z) \end{aligned} \tag{9}$$

Taking the last term to the left and the term $f_{ext}(y)$ to the right in Eq. (9), we have

$$-p(-w - z) - f_{ext}(z) \leq p(y + w) - f_{ext}(y) \tag{10}$$

Since y does not appear on the left and z does not appear on the right, the inequality in Eq. (10) continues to hold if the supremum, m, is taken over $z \in \mathcal{D}(f_{ext})$ on the left and the infimum, M, over $y \in \mathcal{D}(f_{ext})$ on the right. Therefore, with the constant c in Eq. (8) being in the closed interval $[m, M]$, it follows from Eq. (10) that

$$-p(-w - z) - f_{ext}(z) \leq c \ \ \forall z \in \mathcal{D}(f_{ext}) \tag{11}$$

$$c \leq p(y + w) - f_{ext}(y) \ \ \forall y \in \mathcal{D}(f_{ext}) \tag{12}$$

For $\alpha = 0$, we already have $x \in \mathcal{D}(f_{ext})$. Let us first prove $g(x) \leq p(x) \ \ \forall x \in \mathcal{D}(g)$ for $\alpha < 0$ in Eq. (8). Replacing z in Eq. (11) by $\alpha^{-1} y$ and multiplying both sides by the positive quantity $-\alpha$ yields:

$$\alpha p(-w - \alpha^{-1} y) + f_{ext}(y) \leq -\alpha c \tag{13}$$

From Eqs. (8) and (11), using $x = y + \alpha w$ yields:

$$g(x) = f_{ext}(y) + \alpha c \leq -\alpha p(-w - \alpha^{-1} y) = p(\alpha w + y) = p(x) \tag{14}$$

For $\alpha > 0$, let us replace y in Eq. (12) by $\alpha^{-1} y$ to obtain:

$$c \leq p(\alpha^{-1} y + w) - f_{ext}(\alpha^{-1} y) \tag{15}$$

Multiplication of Eq. (15) by α yields

$$\alpha c \leq \alpha p(\alpha^{-1} y + w) - \alpha f_{ext}(\alpha^{-1} y) = p(x) - f_{ext}(y) \tag{16}$$

A combination of Eq. (16) with Eq. (8) yields:

$$g(x) = f_{ext}(y) + \alpha c \leq p(x) \tag{17}$$

■

Remark 2.2: In some cases (e.g., finite-dimensional and separable Hilbert spaces), it is possible to prove Hahn-Banach Theorem without using Zorn's Lemma (see Chapter 5, p. 111 in *Optimization by Vector Space Methods* by Luenberger).

Theorem 2.2: (Hahn-Banach Theorem: Generalization)

Let V be a vector space over the real field $\mathbb{R}$ or the complex field $\mathbb{C}$, and let p be a real-valued functional on V, having the following two properties:

- *Subadditive*: $p(x + y) \leq p(x) + p(y) \ \ \forall x, y \in V$
- *Absolute homogeneous*: $p(\alpha x) = |\alpha| p \ \ \forall \alpha \in \mathbb{C} \ \forall x \in V$

Let f be a linear functional on a subspace $U \subset V$. If $f(x) \leq p(x) \ \ \forall x \in U$, then there exists an extension f_{ext} of f from U to V such that $|f_{ext}(x)| \leq p(x) \ \ \forall x \in V$.

Proof: If V is a vector space over $\mathbb{R}$, the proof is identical to that of Theorem 2.1. If V is a vector space over $\mathbb{C}$, then the functional f is also complex-valued and is split into real and complex parts as:

$$f(x) = f^{real}(x) + if^{imag}(x) \tag{18}$$

where both f^{real} and f^{imag} are real-valued. We note that

$$f^{real}(x) \leq |f(x)| \quad \forall x \in U$$

It follows from Theorem 2.1 that a linear extension of f^{real}_{ext} of f^{real} from U to V satisfies the following condition:

$$f^{real}_{ext}(x) \leq p(x) \quad \forall x \in V$$

Equating the real and imaginary parts of the following equation:

$$\begin{aligned} i[f^{real}(x) + if^{imag}(x)] &= if(x) = f(ix) \\ &= f^{real}(ix) + if^{imag}(ix) \quad \forall x \in U \end{aligned}$$

we have

$$f^{imag}(x) = -f^{real}(ix) \quad \forall x \in U \tag{19}$$

$$f_{ext}(x) = f^{real}_{ext}(x) - if^{real}_{ext}(ix) \quad \forall x \in V \tag{20}$$

It follows from Eqs. (19) and (20) that $f_{ext}(x) = f(x)$ on U, i.e., f_{ext} is an extension of f from U to V. It remains to prove the following tasks:

1. f_{ext} is a linear functional on the complex vector space
2. $|f_{ext}(x)| \leq p(x) \quad \forall x \in V$

Task 1 holds from the fact that, for any complex scalar $a + ib$, the following relation holds based on Eq. (20):

$$\begin{aligned} f_{ext}((a+ib)x) &= f^{real}_{ext}(ax + ibx) - if^{real}_{ext}(iax - bx) \\ &= af^{real}_{ext}(x) + bf^{real}_{ext}(ix) - i[af^{real}_{ext}(ix) - bf^{real}_{ext}(x)] \\ &= (a+ib)[f^{real}_{ext}(x) - if^{real}_{ext}(ix)] \\ &= (a+ib)f^{real}_{ext}(x) \end{aligned}$$

Now we prove Task 2. Let $f_{ext}(\underline{0}) = 0$ which holds because $p(x) \geq 0 \ \forall x \in V$. Let $x \neq \underline{0}$ be such that $f_{ext}(\underline{0}) \neq 0$. Using the polar notation, $f_{ext}(x) = |f_{ext}(x)| \ exp(i\theta) \Rightarrow |f_{ext}(x)| = \exp(-i\theta)f_{ext}(x)$. Since $|f_{ext}(x)|$ is real, the absolute homogeneity property of the sublinear functional p yields

$$\begin{aligned} |f_{ext}(x)| &= f^{real}_{ext}\big(exp(-i\theta)x\big) \leq p\big(exp(-i\theta)x\big) \\ &= |exp(-i\theta(x)|p(x) = p(x) \end{aligned}$$

The proof is thus complete. ■

Further details are available in *Real and Complex Analysis* by Rudin (see Chapter 5, p. 105).

III. Application of Hahn-Banach Theorem to Bounded Linear Functionals

Theorem 3.1: (Hahn-Banach Theorem: Normed Spaces)

Let f be a bounded linear functional on a subspace U of a vector space V, defined on the real field $\mathbb{R}$ or the complex field $\mathbb{C}$. Then, there exists a bounded linear functional f_{ext} on V, which is an extension of f to V having the same norm,

$$||f_{ext}|| = ||f|| \tag{21}$$

Proof: If $U = \{0\}$, then $f = 0$ and consequently $f_{ext} = 0$. Let $f \neq 0$. Since we will use Theorem 2.2 to prove this theorem, we must first find an appropriate sublinear functional p. We have

$$|f(x)| \leq ||f||_U \ ||x|| \quad \forall x \in U$$

where we select $p(x) = ||f||_U \ ||x||$ (see Remark 2.1). Using Theorem 2.2, it follows that there exists a linear functional f_{ext}, which is an extension of f, satisfies the condition:

$$|f_{ext}(x)| \leq p(x) = ||f||_U \ ||x|| \quad \forall x \in V$$

Taking supremum over all unity norm $x \in V$, we obtain the inequality:

$$||f_{ext}||_V = sup_{||x||=1} \ |f_{ext}(x)| \leq ||f||_U \tag{22}$$

Since a norm cannot decrease under extension, we claim that

$$||f_{ext}||_V \geq ||f||_U \tag{23}$$

A combination of Eqs. (22) and (23) proves the theorem. ■

Corollary 3.1: Let V be a normed space and let $x^0 \neq 0$ be an arbitrary vector in V. Then, there exists a bounded linear functional g on V such that $||g|| = 1$ and $g(x^0) = ||x^0||_V$.

Proof: Let U be the subspace spanned by the vector x^0. Let us define a linear functional f on U as $f(\alpha x^0) = \alpha f(x^0) = \alpha||x^0||$, where α is a scalar. Then, f is bounded and $||f|| = 1$ because if $x = \alpha x^0$, then

$$|f(x)| = |f(\alpha x^0)| = |\alpha|||x^0|| = ||\alpha x^0|| = ||x||$$

Then, Theorem 3.1 implies that f has a linear extension from U to V of norm $||f_{ext}|| = ||f|| = 1$ because $f_{ext}(x^0) = f(x^0) = ||x^0||$. ■

Corollary 3.2: Let V be a normed vector space and $f \in V^*$. Then, every $x \in V$ has the following property:

$$||x||_V = sup_{||f||=1}|f(x)| \tag{24}$$

and if $x^0 \in V$ is such that $f(x^0) = 0 \;\forall\; f \in V^*$ for all $f \in V^*$, then $x^0 = 0$.

Proof: By replacing x^0 by x in Corollary 3.1, it follows that

$$sup_{x\in V^*-\{0\}} \frac{|f(x)|}{||f||} \geq \frac{|f_{ext}(x)|}{||f_{ext}||} = ||x||$$

and the proof follows from the fact that $|f(x| \leq ||f||||x||$. ■

Lemma 3.1: Let U be a proper closed subspace of a normed vector space V. Let $x^0 \in V - U$ be arbitrary and the distance from x^0 to U is defined as:

$$\delta = inf_{y\in U}||y - x^0|| > 0 \tag{25}$$

Then, there exists $f_{ext} \in V^*$ such that

$$||f_{ext}|| = 1; \;\; f_{ext}(y) = 0 \;\; \forall\, y \in U; \;\; \text{and} \;\; f_{ext}(x^0) = \delta \tag{26}$$

Proof: Let the subspace W be spanned by U and x^0. Let a bounded linear functional f be defined on W as:

$$f(z) = f(y + \alpha x^0) = \alpha\delta \tag{27}$$

We will first show that f satisfies Eq. (26) and then extend f to f_{ext} on V by Theorem 3.1.

Linearity of f is readily seen. Since U is closed and $\delta > 0$, it follows that $f \neq 0$. It follows from Eq. (26) that $f(y) = 0$ and $f(x^0) = \delta$ by setting α to 0 and 1, respectively.

For $\alpha = 0$, $f(z) = 0$. For $\alpha \neq 0$, it follows from Eq. (25) that

$$\begin{aligned} |f(z)| = |\alpha|\delta = |\alpha| \; inf_{y\in U} \; ||y - x^0|| \\ \leq |\alpha| \; || - \alpha^{-1}y - x^0|| \\ = ||y + \alpha x^0|| \end{aligned}$$

Therefore, $|f(z)| \leq ||z|| \;\; \forall\, z \in W$. Hence, f is bounded and $||f|| \leq 1$.

Next we show that $||f|| \geq 1$. By definition, U contains a sequence $\{y^k\}$ such that $||y^k - x^0|| \to \delta$ as $k \to \infty$. Let $z^k \triangleq y^k - x^0$. Then, $f(z^k) = -\delta$ by setting $\alpha = -1$ in Eq. (27). Furthermore,

$$||f|| = sup_{z\in W-\{0\}} \frac{|f(z)|}{||z||} \geq \frac{|f(z^k)|}{||z^k||} = \frac{\delta}{||z^k||} \to 1 \quad k \to \infty$$

Hence, $||f|| \geq 1$ which implies that $||f|| = 1$. By Theorem 3.1, f is extended to V without increasing the norm. ■

A. Dual Spaces and Separability

Theorem 3.2: (Separability) For a normed vector space V, if the dual space V^* is separable, then V itself is separable.

Proof: Given that V^* be separable, the unit ball $U^* \triangleq \{f \in V^* : ||f|| = 1\} \subset V^*$ contains a countable dense subset, say, $\{f^k : k \in \mathbb{N}\}$, where $||f^k|| = sup_{||x||=1}|f^k(x)| = 1$. Therefore, there exist unit vectors $x^k \in V$ such that $f^k(x^k) \in [0, 1]$. Let $f^k(x^k) \geq 0.5$.

Let W be the closure of the space spanned by $\{x^k\}$. Then, W is separable because W has a countable dense subset, namely, the set of all linear combinations of the vectors x^k with rational coefficients.

To show that $W = V$ by contradiction, let us assume that $W \neq V$. Since W is closed, it follows from Lemma 3.1 that there exists $f_{ext} \in V^*$ with $||f_{ext}|| = 1$ and $f_{ext}(y) = 0 \;\forall\; y \in W$. Since $x^k \in W$, we have $f_{ext}(x^k) = 0 \;\forall\; k$, which implies

$$\begin{aligned} 0.5 \leq |f^k(x^k)| = |f^k(x^k) - f_{ext}(x^k)| = |(f^k - f_{ext})(x^k)| \\ \leq ||(f^k - f_{ext})||||(x^k)|| = ||(f^k - f_{ext})|| \end{aligned}$$

The assertion $||(f^k - f_{ext})|| \geq 0.5$ is a contradiction because $\{f^k\}$ is dense in U^*; in fact, $||(f_{ext})|| = 1$. ■

Corollary 3.3: ℓ_∞^* is not isometrically isomorphic to ℓ_1.

Proof: Let us assume that ℓ_∞^* is isometrically isomorphic to ℓ_1. Since ℓ_1 is separable, so is ℓ_∞^*. By Theorem 3.2, ℓ_∞ must be separable. This is a contradiction. ■

Remark 3.1: L_∞^* is not isometrically isomorphic to L_1 by the same argument as in Corollary 3.3.

Remark 3.2: It follows from Corollary 3.3 that the converse of Theorem 3.2 is false.

Remark 3.3: The space c_0 of all sequences of scalars converging to zero is separable and c_0^* is isometrically isomorphic to ℓ_1.

B. Bounded Linear Functionals on $C[a,b]$

This section presents a general representation formula for bounded linear functionals on $C[a,b]$, where $C[a,b]$ is the space of continuous functions on a fixed compact interval $[a,b]$ with the metric defined as:

$$d(x,y) = max_{t\in[a,b]}\ |x(t) - y(t)| \tag{28}$$

Definition 3.1: A function w on $[a,b]$ is defined to be of *bounded variation* on $[a,b]$ if its total variation is finite, i.e.,

$$Var(w) = sup\ \sum_{j=1}^{n} |w(t_j) - w(t_{j-1})| < \infty \tag{29}$$

where the supremum is taken over all partitions

$$\mathcal{P}_n \triangleq \{a = t_0 < t_1 < \cdots < t_n = b\} \text{ for some } n \in \mathbb{N} \tag{30}$$

Theorem 3.3: (Reisz Theorem on Functionals)

Every bounded linear functional f on $C[a,b]$ can be represented by a Riemann-Stieltjes integral

$$f(x) = \int_a^b x(t)dw(t) \tag{31}$$

where w is of bounded variation on $[a,b]$ and has the total variation

$$Var(w) = ||f|| \tag{32}$$

Proof: It follows from Theorem 3.1 that f has an extension f_{ext} from $C[a,b]$ to the space of all bounded functions on $[a,b]$ with the norm defined as:

$$||x|| = sup_{t\in[a,b]}\ |x(t)| \tag{33}$$

and the bounded functional f_{ext} has the same norm as f, i.e., $||f_{ext}|| = ||f||$. If the function w in Eq. (31) is real-valued, then it is defined as:

$$w(t) = f_{ext}(\chi_{[a,t]})$$

where $t \in [a,b]$ and the characteristic function $\chi_{[a,t]} = 1$ on the support $[a,t]$. In general, for a complex-valued w, we use the polar notation to express $w(t) = |w(t)|\ exp(i\ \theta)$, where $\theta = arg(w(t))$.

For any partition (see Eq. (30)), we have:

$$\sum_{j=1}^{n}\ |w(t_j) - w(t_{j-1})|$$
$$= |f_{ext}(x_1) + \sum_{j=2}^{n}\ |f_{ext}(x_j) - f_{ext}(x_{j-1})|$$
$$= \varepsilon_1 f_{ext}(x_1) + \sum_{j=2}^{n}\ \varepsilon_j\left(f_{ext}(x_j) - f_{ext}(x_{j-1})\right)$$
$$= f_{ext}\big(\varepsilon_1 x_1 + \sum_{j=2}^{n}\ \varepsilon_j\left(f_{ext}(x_j) - x_{j-1}\right)\big)$$
$$\leq ||f_{ext}||\ ||\big(\varepsilon_1 x_1 + \sum_{j=2}^{n}\ \varepsilon_j(x_j - x_{j-1})\big)||$$

On the right hand side, $||f_{ext}|| = ||f||$ and the other factor $||\cdots||$ equals to 1 because $|\varepsilon_j| = 1$ and only one of the terms $x_1,\ x_2 - x_1,\ \cdots$ is nonzero and its norm is equal to 1. This is true because of the choice of the structure of x_t as the characteristic function $\chi_{[a,t]}$. On the left we take supremum over all partitions of $[a,b]$. Then, it follows that

$$Var(w) \leq ||f|| \tag{34}$$

Hence, w is of bounded variation on $[a,b]$.

Next we prove Eq. (31). For every partition P_n of the form similar to Eq. (30) on $C[a,b]$, we define a function as:

$$z_{P_n} = x_0 x_1 + \sum_{j=2}^{n} x_{j-1}(x_j - x_{j-1})$$

implying that f_{ext} is bounded on $[a,b]$. By definition of w,

$$f_{ext}(z_{P_n})$$
$$= x_0 f_{ext}(x_1) + \sum_{j=2}^{n} x_{j-1}\left(f_{ext}(x_j) - f_{ext}(x_{j-1})\right)$$
$$= x_0 w(t_1) + \sum_{j=2}^{n} t_{j-1}\left(w(t_j) - w(t_{j-1})\right)$$
$$= \sum_{j=1}^{n} t_{j-1}\left(w(t_j) - w(t_{j-1})\right) \tag{35}$$

where the last equality follows from $w(t_0) = w(a) = 0$. By making the partition P_n finer and taking $n \to \infty$, the sum on the right hand side of Eq. 35 approaches the integral in Eq. (31). Since $f_{ext}(z_{P_n}) \to f_{ext}(x)$, the integral in Eq. (31) becomes equal to $f(x)$ because $x \in C[a,b]$.

Since $\left|\int_a^b x(t)\ \ dw(t)\right| \leq max_{t\in[a,b]}\ |x(t)|Var(w)$, it follows from the supremum over all $x \in C[a,b]$ that

$$Var(w) \leq ||f|| \tag{36}$$

The combination of Eqs. (34) and (36) yields the equality in Eq. (32). The proof is complete. ∎