

# Introductory Functional Analysis with Applications- Kreyszig

## Solutions to Exercises

### 1 Metric Spaces

#### 1.1 Metric Space

1. Show that the real line is a metric space.

*Proof.* To show that the real line is a metric space, we wish to show that there is a valid metric on this space. Obviously, the choice of the metric is  $d(x, y) = |x - y|$ . We now wish to show the following properties:

1.  $d$  is non-negative.
2.  $d(x, y) = 0 \iff x = y$ .
3.  $d(x, y) = d(y, x)$ .
4.  $d(x, y) \leq d(x, z) + d(z, y)$

Properties (1), (2), and (3) are obvious. We now show property (4) for the choice of  $d(x, y) = |x - y|$ . Consider

$$\begin{aligned}(x - y)^2 &= x^2 - 2xy + y^2 \\ &= x^2 - 2xz + z^2 + z^2 - 2zy + y^2 + (2xz + 2zy - 2xy - 2z^2) \\ &= (x - z)^2 + (z - y)^2 + 2x(z - y) + 2z(y - z) \\ &= |x - z|^2 + |z - y|^2 + 2(x - z)(z - y) \\ &\leq (|x - z| + |y - z|)^2\end{aligned}$$

where the inequality follows since  $x - y \leq |x - y|$ . Taking a square root on both sides finishes the proof. ■

2. Does  $d(x, y) = (x - y)^2$  define a metric on the set of all real numbers?

*Proof.* To show that  $d(x, y)$  is a valid metric, we must ensure that it is non-negative, symmetric, satisfies the triangle inequality, and satisfies the condition that  $(x - y)^2 = 0 \iff x = y$ . We see that the first, second, and last conditions are always true. We only need to check for the triangle inequality. Notice, from the answer of 1, we have

$$(x - y)^2 = (x - z)^2 + (y - z)^2 + 2(x - z)(z - y)$$

WLOG, assume  $x \leq y$ . If  $z \in [x, y]$ , then we have  $(x - z) \leq 0$  and  $(z - y) \leq 0$ , and hence,  $(x - z)(z - y) \geq 0$ , which results in

$$d(x, y) \geq d(x, z) + d(z, y)$$

which is a contradiction of the triangle inequality. Hence,  $d(x, y) = (x - y)^2$  is not a metric. ■

3. Show that  $d(x, y) = \sqrt{|x - y|}$  defines a metric on the set of all real numbers.

*Proof.* Once again, we can check that  $d(x, y)$  satisfies non-negativity, symmetry, and the property that  $d(x, y) = 0 \iff x = y$ . We wish to show  $d$  satisfies the triangle inequality. Since  $|x - y|$  is a valid metric (refer 1), using the triangle inequality for this along with the fact that  $\sqrt{\cdot}$  is monotonically increasing, we get

$$\sqrt{|x - y|} \leq \sqrt{|x - z| + |z - y|} \leq \sqrt{|x - z|} + \sqrt{|z - y|}$$

where the last inequality follows from the fact that if  $a, b \geq 0$ , then

$$\sqrt{a + b} \leq \sqrt{a + 2\sqrt{ab} + b} = \sqrt{(\sqrt{a} + \sqrt{b})^2} = \sqrt{a} + \sqrt{b}$$

.

4. Find all metrics on a set  $X$  consisting of two points and only one point.

*Proof.* If  $X$  consists of only one point, then the only metric that can be defined keeping in mind all properties (note that the triangle inequality does not apply since we do not have a second point for this) is  $d(x, x) = 0$ .

If  $X$  consists of two points, once again, the triangle inequality does not apply, and hence, it is easy to check that

$$d(x, y) = \begin{cases} 0 & x = y \\ c & x \neq y \end{cases}.$$

■

**5.** Let  $d$  be a metric on  $X$ . Determine all constants  $k$  such that  $kd$  and  $d + k$  is a metric on  $X$ .

*Proof.* Note that for  $kd$  to remain a metric, all we need is  $k \geq 0$  to ensure it is non-negative. All other properties continue to hold as it is.

For  $d + k$  to be a metric, the simplest starting point is that  $(d + k)(x, y) = 0 \iff x = y$ . However, we know that  $d(x, y) = 0 \iff x = y$ , and since,  $k$  is a constant, this property is satisfied only if  $k(x, y) = 0$ . Hence, no constant  $k$  can be added such that  $d + k$  is also a metric. ■

**6.** Show that  $d$  in 1.1-6 satisfies the triangle inequality.

*Proof.* The metric in 1.1-6 is for a sequence space  $\ell^\infty$  such that for  $x = (\xi_i)$  and  $y = (\eta_i)$  such that  $\xi \leq c_x$  and  $\eta_i \leq c_y$ , we have

$$d(x, y) = \sup_{i \in \mathbb{N}} |\xi_i - \eta_i|$$

Note that this is always non-negative and symmetric, owing to the absolute difference of the terms (since absolute difference is always non-negative, the least upper bound is also non-negative). Also, if  $x = y$ , i.e.,  $\xi_i = \eta_i$  for all  $i$ , then the supremum is 0. Finally, we show the triangle inequality as follows. For some  $z = (\zeta_i)$ , we have

$$d(x, y) = \sup_{i \in \mathbb{N}} |\xi_i - \eta_i| \leq \sup_{i \in \mathbb{N}} (|\xi_i - \zeta_i| + |\zeta_i - \eta_i|) \leq \sup_{i \in \mathbb{N}} |\xi_i - \zeta_i| + \sup_{i \in \mathbb{N}} |\zeta_i - \eta_i| = d(x, z) + d(z, y)$$

where we use the triangle inequality for  $|x - y|$  (check 1) and the fact that  $\sup(f+g)(x) \leq \sup f(x) + \sup g(x)$ . ■

**7.** If  $A$  is the subspace of  $\ell^\infty$  consisting of all sequences of zeros and ones, then what is the induced metric on  $A$ ?

*Proof.* If  $A$  is the subspace consisting of all sequences only containing zeros and ones, then the coordinate-wise absolute difference of any two sequences can only take on two values; 0, when the coordinates are the same, and 1 when the coordinates are different. Thus, the supremum of this absolute difference is 1. In other words, the metric can be expressed as:

$$d(x, y) = \mathbb{1}\{x \neq y\}$$

which is exactly the discrete metric. ■

**8.** Show that another metric  $\tilde{d}$  on the space  $X$  in 1.1-7 is defined by

$$\tilde{d}(x, y) = \int_a^b |x(t) - y(t)| dt$$

*Proof.* We wish to show that the metric  $d(x, y)$  is defined on the space of continuous functions  $C[a, b]$ , i.e.  $x$  and  $y$  are now continuous functions. Note that  $d(x, y)$  is clearly non-negative and symmetric owing to the point-wise absolute difference of the functions. Further, if  $x(c) = y(c) \forall c \in [a, b]$ , then,  $d(x, y) = \int_a^b 0 dt = 0$ . Thus, we wish to show the triangle inequality now. Using the fact that the absolute difference is a metric (check 1), using the triangle inequality, we have

$$\int_a^b |x(t) - y(t)| dt \leq \int_a^b (|x(t) - z(t)| + |z(t) - y(t)|) dt = \int_a^b |x(t) - z(t)| dt + \int_a^b |z(t) - y(t)| dt$$

which proves that  $d(x, y)$  satisfies the triangle inequality. ■

**9.** Show that  $d$  in 1.1-8 is a metric.

*Proof.* We wish to show that the discrete metric, defined as

$$d(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$$

is a metric. Clearly, this is non-negative and symmetric. Further, we have that  $d(x, y) = 0 \iff x = y$ . Finally, we also have

1. If  $x \neq y$ , then  $d(x, y) = 1$ . In such a case, if  $x = z$  or  $y = z$ , then  $d(x, z) + d(y, z) = 1$ . However, if  $x \neq z \neq y$ , then  $d(x, z) + d(y, z) = 2$ . Thus,  $d(x, y) \leq d(x, z) + d(z, y)$ .
2. If  $x = y$ , then  $d(x, y) = 0$ . Now, if  $x = y \neq z$ , then  $d(x, z) + d(z, y) = 2$ , and if  $x = y = z$ , then  $d(x, z) + d(z, y) = 0$ , and in both cases,  $d(x, z) + d(z, y) \geq d(x, y)$ .

Thus, the triangle inequality holds for  $d$ , and hence, it is a valid metric. ■

**10.** Let  $X$  be the set of all ordered triples of zeros and ones. Show that  $X$  consists of eight elements and a metric  $d$  on  $X$  is defined by  $d(x, y) = \text{number of places where } x \text{ and } y \text{ have different entries}$ .

*Proof.* Clearly,  $X$  has  $2^3 = 8$  elements. We wish to show that

$$d(x, y) = \sum_{i=1}^3 \mathbb{1}\{x_i \neq y_i\}$$

is a metric on  $X$  where  $x_i$  represents the  $i^{\text{th}}$  coordinate of  $x$ . Clearly, this metric is non-negative, symmetric, and satisfies the property that  $d(x, y) = 0 \iff x = y$ . Also, we have

$$\mathbb{1}\{x_i \neq y_i\} \leq \mathbb{1}\{x_i \neq z_i\} + \mathbb{1}\{z_i \neq y_i\}$$

This is easy to see using a case-by-case analysis.

1. If  $x_i = y_i$ , then  $\mathbb{1}\{x_i \neq y_i\} = 0$ . If  $z_i = x_i = y_i$ , we have  $\mathbb{1}\{x_i \neq z_i\} + \mathbb{1}\{z_i \neq y_i\} = 0$ , while if  $z_i \neq x_i = y_i$ , then  $\mathbb{1}\{x_i \neq z_i\} + \mathbb{1}\{z_i \neq y_i\} = 2$ , and hence,  $\mathbb{1}\{x_i \neq y_i\} \leq \mathbb{1}\{x_i \neq z_i\} + \mathbb{1}\{z_i \neq y_i\}$ .
2. If  $x_i \neq y_i$ , then  $\mathbb{1}\{x_i \neq y_i\} = 1$ . If  $z_i = x_i$  or  $z_i = y_i$ , we have  $\mathbb{1}\{x_i \neq z_i\} + \mathbb{1}\{z_i \neq y_i\} = 1$ , while if  $z_i \neq x_i \neq y_i$ , then  $\mathbb{1}\{x_i \neq z_i\} + \mathbb{1}\{z_i \neq y_i\} = 2$ , and hence,  $\mathbb{1}\{x_i \neq y_i\} \leq \mathbb{1}\{x_i \neq z_i\} + \mathbb{1}\{z_i \neq y_i\}$ .

Thus, the Hamming distance satisfies all the properties of a metric. ■

**11.** Prove (1).

*Proof.* We wish to prove the statement:

$$d(x_1, x_n) \leq \sum_{i=2}^n d(x_{i-1}, x_i)$$

We can do this by induction on  $n$ . The base case  $n = 3$  holds due to the triangle inequality. Assuming it holds for  $n = k$ , we wish to show that it holds for  $n = k + 1$ . This is easy to show as follows:

$$d(x_1, x_{k+1}) \leq d(x_1, x_k) + d(x_k, x_{k+1}) \leq \sum_{i=2}^k d(x_{i-1}, x_i) + d(x_k, x_{k+1}) = \sum_{i=2}^{k+1} d(x_{i-1}, x_i)$$

where the first inequality follows by the triangle inequality and the second follows from the induction hypothesis. ■

**12.** Using (1), show that

$$|d(x, y) - d(z, w)| \leq d(x, z) + d(y, w)$$

*Proof.* By the repeated application of triangle inequality, we have

$$d(x, y) \leq d(x, z) + d(z, y) \leq d(x, z) + d(z, w) + d(w, y)$$

which gives us

$$d(x, y) - d(z, w) \leq d(x, z) + d(y, w)$$

Similarly, we have

$$d(z, w) \leq d(z, x) + d(x, z) \leq d(z, x) + d(x, y) + d(y, w)$$

which gives us

$$d(z, w) - d(x, y) \leq d(x, z) + d(y, w)$$

Combining both these facts and using the symmetry property of metrics gives us the required result. ■

**13.** Using the triangle inequality, show that

$$|d(x, z) - d(y, z)| \leq d(x, y)$$

*Proof.* Note that  $d(x, z) \leq d(x, y) + d(y, z)$  and hence,  $d(x, z) - d(y, z) \leq d(x, y)$ . Similarly,  $d(y, z) \leq d(x, y) + d(x, z)$  and hence,  $d(y, z) - d(x, z) \leq d(x, y)$ . Combining both results finishes the proof. ■

**14.** Show that (M3) and (M4) could be obtained from (M2) and

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

*Proof.* We wish to show that the axioms  $d(x, y) = d(y, x)$  and  $d(x, y) \leq d(x, z) + d(z, y)$  can be obtained from the axioms  $d(x, y) = 0 \iff x = y$  and  $d(x, y) \leq d(z, x) + d(z, y)$ .

To prove symmetry, we have the following two equations:

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

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

Subtracting both equations results in

$$d(x, y) \leq d(y, x) \text{ and } d(y, x) \leq d(x, y) \implies d(x, y) = d(y, x).$$

Now, simply using symmetry, we have

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

since  $d(z, x) = d(x, z)$ . ■

**15.** Show that non-negativity of a metric follows from (M2) to (M4).

*Proof.* We wish to show that the non-negativity of a metric follows from the following axioms:

1.  $d(x, y) = d(y, x)$ .
2.  $d(x, y) = 0 \iff x = y$ .
3.  $d(x, y) \leq d(x, z) + d(z, y)$ .

Substituting  $y = x$  in the triangle inequality, we have

$$d(x, x) \leq d(x, z) + d(z, x)$$

Using the first and second axiom in tandem gives us  $d(x, x) \geq 0$ , which completes the proof. ■

## 1.2 Further Examples of Metric Spaces

**1.** Show that in 1.2-1, we can obtain another metric by replacing  $1/2^i$  with  $\mu_i$  such that  $\sum \mu_i$  converges.

*Proof.* We wish to show that for some  $\mu_j > 0$  such that  $\sum \mu_j$  converges, and two sequences  $x = (\xi_i)$  and  $y = (\eta_i)$ , we have

$$d(x, y) = \sum_{j=1}^{\infty} \mu_j \frac{|\xi_j - \eta_j|}{1 + |\xi_j - \eta_j|}$$

is a valid metric defined on the sequence space.

The proof follows on the same lines as the one given for 1.2-1. It is clear that  $d(x, y)$  is non-negative, symmetric, and satisfies  $d(x, y) = 0 \iff x = y$ . However, we wish to show this is bounded, i.e

$$d(x, y) = \sum_{j=1}^{\infty} \mu_j \frac{|\xi_j - \eta_j|}{1 + |\xi_j - \eta_j|} \leq \sup_{j \in \mathbb{N}} \frac{|\xi_j - \eta_j|}{1 + |\xi_j - \eta_j|} \sum_{j=1}^{\infty} \mu_j$$

Note that the RHS is bounded if and only if  $\sum \mu_j$  is convergent (since the supremum term can be upper bounded by 1). Further, the proof for the triangle inequality follows in the same fashion, define  $f(t) = \frac{t}{1+t}$ . Then,  $f(t)$  is monotonically increasing, and hence, using the triangle inequality for numbers, we get

$$f(|a + b|) \leq f(|a| + |b|)$$

and hence, substituting  $a = \xi_j - \zeta_j$  and  $b = \zeta_j - \eta_j$ , we get

$$\frac{|\xi_j - \eta_j|}{1 + |\xi_j - \eta_j|} \leq \frac{|\xi_j - \zeta_j| + |\zeta_j - \eta_j|}{1 + |\xi_j - \zeta_j| + |\zeta_j - \eta_j|} \leq \frac{|\xi_j - \zeta_j|}{1 + |\xi_j - \zeta_j|} + \frac{|\zeta_j - \eta_j|}{1 + |\zeta_j - \eta_j|}$$

Multiplying by  $\mu_j$  and summing over all  $j \in \mathbb{N}$  finishes the proof. ■

**2.** Using (6) show that the geometric mean of two positive numbers does not exceed the arithmetic mean.

*Proof.* We wish to show that the geometric mean does not exceed the arithmetic mean using the equation:

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

where  $p^{-1} + q^{-1} = 1$ .

Substituting  $x = \sqrt{a}, y = \sqrt{b}, p = q = 2$ , we get

$$\sqrt{ab} \leq \frac{a+b}{2}.$$

**3.** Show that the Cauchy Schwarz inequality implies

$$\left( \sum_{i=1}^n |\xi_i| \right)^2 \leq n \sum_{i=1}^n |\xi_i|^2$$

*Proof.* The Cauchy-Schwarz inequality states that

$$\left( \sum_{i=1}^n |\xi_i| |\eta_i| \right)^2 \leq \left( \sum_{i=1}^n |\eta_i|^2 \right) \left( \sum_{i=1}^n |\xi_i|^2 \right)$$

Putting  $\eta_i = 1$  for all  $i \in [n]$  completes the proof. ■

**4.** Find a sequence which converges to 0, but is not in any space  $\ell^p$ ,  $1 \leq p < \infty$ .

*Proof.* Recall that the  $\ell^p$  space consists of sequences  $x = (\xi_i)$  such that  $\sum |\xi_i|^p$  converges. The sequence  $\xi_i = \frac{1}{\log(i+1)}$  converges to 0 but is not part of any  $\ell^p$  space. To see this, note that  $\log(i+1)$  can be bounded above by  $i^{1/p}$  and hence,

$$\frac{1}{\log(i+1)} \geq i^{-1/p}.$$

Thus, we have

$$\sum_i \left( \frac{1}{\log(i+1)} \right)^p \geq \sum_i \left( \frac{1}{i^{1/p}} \right)^p \rightarrow \infty$$

**5.** Find a sequence which is in  $\ell^p$  with  $p > 1$ , but  $x \notin \ell^1$ .

*Proof.* We wish to find a sequence  $x = (\xi_i)$  such that  $\sum |\xi_i|^p$  converges for all  $p > 1$  but does not converge for  $p = 1$ . The simplest example is  $\xi_i = i^{-1}$ , the sum of which diverges. ■

**6.** The diameter  $\delta(A)$  of a nonempty set  $A$  in a metric space  $(X, d)$  is defined to be

$$\delta(A) = \sup_{x, y \in A} d(x, y)$$

$A$  is said to be bounded if  $\delta(A) < \infty$ . Show that  $A \subset B$  implies  $\delta(A) \leq \delta(B)$ .

*Proof.* First, by the definition of supremum, we have

$$\sup X = s \implies \forall x \in X, x \leq s$$

Thus, over the set  $B$ , if  $\delta(B)$  is the diameter of  $B$ , we have that

$$\forall (x, y) \in B \times B, d(x, y) \leq \delta(B).$$

Since,  $A \subset B$ , we have that all points in  $A$  also belong to  $B$  and hence, satisfy the property above, i.e

$$\forall(x, y) \in A \times A, d(x, y) \leq \delta(B)$$

and hence, taking the supremum on both sides, we get

$$\delta(A) = \sup_{x, y \in A} d(x, y) \leq \sup_{x, y \in A} \delta(B) = \delta(B).$$

**7.** Show that  $\delta(A) = 0$  if and only if  $A$  consists of a single point.

*Proof.* We wish to show

$$\delta(A) = 0 \iff |A| = 1.$$

We first show RHS implies LHS. Suppose  $\delta(A) = 0$ , which means

$$\sup_{x, y \in A} d(x, y) = 0 \implies d(x, y) \leq 0 \forall x, y \in A.$$

However, since  $d(x, y)$  is always non-negative, we have that

$$d(x, y) = 0 \forall x, y \in A \implies x = y \forall x, y \in A.$$

This tells us that there is only one unique point in  $A$ .

We now show LHS implies RHS. Suppose  $A$  has only one point  $x$ , then,

$$\delta(A) = \sup_A d(x, y) = \sup\{d(x, x)\} = \sup\{0\} = 0$$

This finishes the proof. ■

**8.** The distance  $D(A, B)$  between two non-empty sets of a metric space  $(X, d)$  is defined to be

$$D(A, B) = \inf_{\substack{a \in A \\ b \in B}} d(a, b)$$

Show that  $D$  does not define a metric on the power set of  $X$ .

*Proof.* An easy way to see that this does not define a metric on the power set of  $X$  (denoted by  $\mathcal{P}(X)$ ) is as follows: clearly, the metric is non-negative and symmetric. However, let  $A, B \in \mathcal{P}(X)$ . If  $A \cap B \neq \emptyset$ , then  $\exists x \in X$  such that  $x \in A$  and  $x \in B$ . In such a case,

$$D(A, B) = \inf_{\substack{a \in A \\ b \in B}} d(a, b) = d(x, x) = 0.$$

However,  $A \cap B \neq \emptyset \nRightarrow A = B$ , and hence, the condition

$$D(A, B) = 0 \iff A = B$$

fails. ■

**9.** If  $A \cap B \neq \emptyset$ , show that  $D(A, B) = 0$ . What about the converse?

*Proof.* Let  $x \in A$  and  $x \in B$ . Then,

$$\{d(a, b)\}_{\substack{a \in A \\ b \in B}} = \{d(x, x)\} \cup \{d(a, b)\}_{(a, b) \in A \times B \setminus \{(x, x)\}} = \{0\} \cup \{d(a, b)\}_{(a, b) \in A \times B \setminus \{(x, x)\}}$$

and hence,

$$\inf_{\substack{a \in A \\ b \in B}} d(a, b) = 0 = D(A, B).$$

The converse may not be true. The idea comes from accumulation points (defined in the next section).  $x$  is said to be an accumulation point of a set  $A$  if every neighborhood of  $x$  contains atleast one point from  $A$  distinct from  $x$ . Let  $x \in B$  and  $x \notin A$  be an accumulation point of  $A$ . Then, every  $\epsilon$ -neighborhood of  $x$  will contain  $a(\epsilon) \in A$  and hence, for every  $\epsilon > 0$ , you can find  $a(\epsilon)$  such that  $d(a(\epsilon), x) = \epsilon$ . Since  $\epsilon$  can be arbitrarily small, the infimum becomes zero. However,  $A \cap B = \emptyset$ . Thus, the converse may not be true. ■

**10.** The distance  $D(x)$  from a point  $x$  to a non-empty subset  $B$  of  $(X, d)$  is defined as

$$D(x, B) = \inf_{b \in B} d(x, b)$$

Show that for any  $x, y \in X$ ,

$$|D(x, B) - D(y, B)| \leq d(x, y).$$

*Proof.* First, consider the case when  $x \in B$  and  $y \in B$ . Clearly,  $D(x, B) = D(y, B) = 0$  and hence,

$$|D(x, B) - D(y, B)| = 0 \leq d(x, y).$$

by the properties of  $d$ . Now, consider the case when  $x \notin B$  and  $y \in B$ . Clearly,  $D(y, B) = 0$  and hence,

$$|D(x, B)| = \inf_{b \in B} d(x, b) \leq d(x, y)$$

since  $y \in B$ . Finally, consider the case when  $x \notin B$  and  $y \notin B$ . Then, we have

$$D(x, B) = \inf_{b \in B} d(x, b) \leq \inf_{b \in B} (d(x, y) + d(y, b)) = d(x, y) + \inf_{b \in B} d(y, b) = d(x, y) + D(y, B).$$

where the first inequality follows from the triangle inequality applied on  $d$ . Rearranging the terms gives us

$$D(x, B) - D(y, B) \leq d(x, y).$$

By symmetry, interchanging  $x$  and  $y$  results in

$$D(y, B) - D(x, B) \leq d(x, y).$$

This finishes the proof. ■

**11.** If  $(X, d)$  is any metric space, show that another metric on  $X$  is defined by

$$\tilde{d}(x, y) = \frac{d(x, y)}{1 + d(x, y)}$$

and  $X$  is bounded in the metric  $\tilde{d}$ .

*Proof.* First, the metric  $\tilde{d}$  is non-negative, symmetric and satisfies  $\tilde{d}(x, y) = 0 \iff x = y$  from the properties of  $d$ . Also

$$\lim_{d(x, y) \rightarrow \infty} \tilde{d}(x, y) = 1$$

and hence, it does remain bounded. Also, notice that

$$f(t) = \frac{t}{1 + t}$$

is always increasing and hence, by the non-negativity of the metric  $d$ , we have

$$\begin{aligned} f(d(x, y)) &\leq f(d(x, z) + d(z, y)) = \frac{d(x, z) + d(z, y)}{1 + d(x, z) + d(z, y)} = \frac{d(x, z)}{1 + d(x, z) + d(z, y)} + \frac{d(z, y)}{1 + d(x, z) + d(z, y)} \\ &\leq \frac{d(x, z)}{1 + d(x, z)} + \frac{d(z, y)}{1 + d(x, z)} \end{aligned}$$

which finishes the proof. ■

**12.** Show that the union of two bounded sets is also bounded.

*Proof.* Let  $A$  and  $B$  be two bounded sets such that  $\delta(A) < \infty$  and  $\delta(B) < \infty$ . Then, using the definition of  $\delta(A)$  from 6, we have

$$\forall a_1, a_2 \in A, d(a_1, a_2) \leq \delta(A) < \infty \text{ and } \forall b_1, b_2 \in B, d(b_1, b_2) \leq \delta(B) < \infty$$

Thus, for any  $a_1 \in A, b_1 \in B$ , we have

$$d(a_1, b_1) \leq d(a_1, a_2) + d(a_2, b_2) + d(b_1, b_2) \leq \delta(A) + \delta(B) + d(a_2, b_2)$$

and since  $d$  is a metric, it is also bounded, resulting in  $d(a_1, b_1)$  being bounded. ■

**13.** The cartesian product  $X = X_1 \times X_2$  of two metric spaces  $(X_1, d_1)$  and  $(X_2, d_2)$  can be made into a metric space  $(X, d)$  in many ways. For example, show that a metric  $d$  is defined by

$$d(x, y) = d_1(x_1, y_1) + d_2(x_2, y_2)$$

where  $x = (x_1, x_2)$  and  $y = (y_1, y_2)$ .

*Proof.* First, note that  $d$  is bounded, non-negative, and symmetric owing to the metric properties of  $d_1$  and  $d_2$ . Also, due to the metric properties of  $d_1$  and  $d_2$ , we have

$$d(x, y) = 0 \iff d_1(x_1, y_1) = 0 \text{ and } d_2(x_2, y_2) = 0 \iff x = (x_1, x_2) = (y_1, y_2) = y.$$

Finally, we show the triangle inequality for some  $z = (z_1, z_2)$  as follows:

$$d(x, y) = d_1(x_1, y_1) + d_2(x_2, y_2) \leq d_1(x_1, z_1) + d_1(z_1, y_1) + d_2(x_2, z_2) + d_2(z_2, y_2) = d(x, z) + d(z, y).$$

This finishes the proof. ■

**14.** Show that another metric on  $X$  is defined by

$$\tilde{d}(x, y) = \sqrt{d_1(x_1, y_1)^2 + d_2(x_2, y_2)^2}$$

*Proof.* Once again, note that  $\tilde{d}$  is bounded, non-negative, symmetric and satisfies  $\tilde{d}(x, y) = 0 \iff x = y$  owing to the metric properties of  $d_1$  and  $d_2$ . Finally, we show the triangle inequality for some  $z = (z_1, z_2)$  as follows:

$$\begin{aligned} \tilde{d}(x, y) &= \sqrt{d_1(x_1, y_1)^2 + d_2(x_2, y_2)^2} \\ &\leq \sqrt{(d_1(x_1, z_1) + d_1(z_1, y_1))^2 + (d_2(x_2, z_2) + d_2(z_2, y_2))^2} \\ &\leq \sqrt{d_1(x_1, z_1)^2 + d_2(x_2, z_2)^2 + d_1(z_1, y_1)^2 + d_2(z_2, y_2)^2 + 2d_1(x_1, z_1)d_1(z_1, y_1) + 2d_2(x_2, z_2)d_2(z_2, y_2)} \end{aligned}$$

Using Hölder's inequality, which says that for  $a_i, b_i \geq 0$  and  $p, q$  such that  $p^{-1} + q^{-1} = 1$ ,

$$\sum_{i=1}^N a_i b_i \leq \left( \sum_{i=1}^N a_i^p \right)^{\frac{1}{p}} + \left( \sum_{i=1}^N b_i^q \right)^{\frac{1}{q}}$$

we get

$$d_1(x_1, z_1)d_1(z_1, y_1) + d_2(x_2, z_2)d_2(z_2, y_2) \leq \sqrt{d_1(x_1, z_1)^2 + d_2(x_2, z_2)^2} \sqrt{d_1(z_1, y_1)^2 + d_2(z_2, y_2)^2}$$

and hence, we get

$$\tilde{d}(x, y) \leq \sqrt{d_1(x_1, z_1)^2 + d_2(x_2, z_2)^2} \sqrt{d_1(z_1, y_1)^2 + d_2(z_2, y_2)^2} = \tilde{d}(x, z) + \tilde{d}(z, y).$$

This finishes the proof. ■

**15.** Show that another metric on  $X$  is defined by

$$\tilde{\tilde{d}}(x, y) = \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}$$

*Proof.* Once again, note that  $\tilde{\tilde{d}}$  is bounded, non-negative, symmetric and satisfies  $\tilde{\tilde{d}}(x, y) = 0 \iff x = y$  owing to the metric properties of  $d_1$  and  $d_2$ . Finally, we show the triangle inequality for some  $z = (z_1, z_2)$  as follows:

$$\begin{aligned} \tilde{\tilde{d}}(x, y) &= \max\{d_1(x_1, y_1), d_2(x_2, y_2)\} \\ &\leq \max\{d_1(x_1, z_1) + d_1(z_1, y_1), d_2(x_2, z_2) + d_2(z_2, y_2)\} \\ &= d_i(x_i, z_i) + d_i(z_i, y_i) \quad i \in \{1, 2\} \\ &\leq \max\{d_1(x_1, z_1), d_2(x_2, z_2)\} + \max\{d_1(z_1, y_1), d_2(z_2, y_2)\} \\ &= \tilde{\tilde{d}}(x, z) + \tilde{\tilde{d}}(z, y) \end{aligned}$$

This finishes the proof. ■



### 1.3 Open set, Closed set, Neighborhood

1. Justify the terms open ball and closed ball by showing that an open ball is an open set and a closed ball is a closed set.

*Proof.* Let  $B(x_0, r)$  be a ball of radius  $r$  around  $x_0$ . If we wish to show that  $B(x_0, r)$  is open, it suffices to show that we can draw a ball around any point of  $B(x_0, r)$ . In other words, choose any point  $y_0 \in B(x_0, r)$ . For some radius  $r'$ ,  $B(y_0, r') \subseteq B(x_0, r)$ . Let  $r' \leq r - d(x_0, y_0)$ . Then, for some  $x \in B(y_0, r')$ , we have

$$d(x, x_0) \leq d(x, y_0) + d(y_0, x_0) \leq r$$

and hence,  $x \in B(x_0, r)$ . This shows that an open ball is also an open set.

Showing that a closed ball is a closed set is equivalent to showing the complement of a closed ball is open. Let  $X$  be a set and  $x_0 \in X$ . Let  $\tilde{B}(x_0, r)$  denote a closed ball of radius  $r$  in  $X$  and let  $\tilde{B}(x_0, r)^C$  denote the complement of  $\tilde{B}(x_0, r)$ . Choose some point  $y_0 \in \tilde{B}(x_0, r)^C$  and  $r' \leq d(x_0, y_0) - r$ . Clearly,  $d(x_0, y_0) \geq r$ . Then,  $B(y_0, r')$  is a ball about  $y_0$  with radius  $r'$ . Let  $x \in B(y_0, r')$ . Then,

$$d(x_0, y_0) \leq d(x_0, x) + d(y_0, x) \implies d(x_0, x) \geq d(x_0, y_0) - d(y_0, x) \geq r - d(y_0, x) \geq r.$$

and hence,  $x \in \tilde{B}(x_0, r)^C$ . Thus, the complement of the closed ball is open. This finishes the proof. ■

2. What is an open ball  $B(x_0, 1)$  in  $\mathbb{R}$  and  $\mathbb{C}$ ? What is an open ball in  $C[a, b]$ ?

*Proof.* In  $\mathbb{R}$ , the open ball is the set of points  $B(x_0, 1) = \{x_0 \pm \epsilon \mid |\epsilon| < 1\}$ . In  $\mathbb{C}$ , let  $x_0 = a + ib$ , then the open ball is defined as  $B(x_0, 1) = \{(a + \epsilon_1) + i(b + \epsilon_2) \mid \sqrt{\epsilon_1^2 + \epsilon_2^2} < 1\}$ . Finally, let  $x_0 \in C[a, b]$  be a function of  $t$ . Then, under the metric

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

we have the open ball to be defined as

$$B(x_0, 1) = \left\{ x_0(t) \pm \epsilon(t) \mid \max_{t \in [a, b]} \epsilon(t) < 1 \right\}.$$

3. Consider  $C[0, 2\pi]$  and determine the smallest  $r$  such that  $y \in \tilde{B}(x, r)$  where  $x(t) = \sin t$  and  $y(t) = \cos t$ .

*Proof.* We wish to find the smallest value of  $r$  such that  $y(t) \in \tilde{B}(x(t), r)$ . Let the metric on  $C[0, 2\pi]$  be defined as

$$d(x, y) = \max_{t \in [0, 2\pi]} |x(t) - y(t)|.$$

Then, we wish to find the smallest value of  $r$  such that  $d(x, y) \leq r$ . Substituting the values of  $x(t)$  and  $y(t)$ , we get

$$\max_{t \in [0, 2\pi]} |\sin t - \cos t| \leq r$$

Differentiating  $\sin t - \cos t$  w.r.t  $t$  and setting to 0, we find that the maximum value of the expression is obtained at  $t = 3\pi/4$  and the maximum value is  $\sqrt{2}$ . Thus, setting  $r = \sqrt{2}$  ensures  $y \in \tilde{B}(x, r)$ . ■

4. Show that any nonempty set  $A \subset (X, d)$  is open if and only if it is a union of open balls.

*Proof.* We first assume that a nonempty set  $A \subset (X, d)$  is open and wish to show it is the union of open balls. Choose any point  $x \in A$ . Then, there exists a ball  $B$  around  $x$  such that  $B \subset A$ . Since  $x$  is arbitrary, at each point in  $A$ , there exists a ball comprised entirely in  $A$  and hence,  $A$  is the union of all such open balls.

Now we assume  $A$  is a union of open balls. For some open ball  $B$ ,  $\exists x \in B$  such that we can draw another ball  $B_0$  around  $x$  and  $B_0 \subset B$ . Since  $A$  is the union of open balls,  $x \in A$  and  $B_0 \subset B \subset A$ . Since  $x$  is arbitrary,  $A$  is an open set. ■

5. It is important to realize that some sets may be open and closed at the same time. Show that this is always the case for  $X$  and  $\emptyset$ . Show that in a discrete metric space  $X$ , every subset is open and closed.

*Proof.* Note that  $\emptyset$  is always open since there are no points in it. This also ensures that the set  $X$  is closed (since its complement  $\emptyset$  is open). On the other hand,  $X$  is open because it consists of the entire space and all limit points, and hence, its complement  $\emptyset$  is closed.

In a discrete metric space, let  $A \subset X$ .  $A$  is open if we can draw a ball around each point in  $A$ . Let  $r < 1$ . Then, for each  $a \in A$ ,  $B(a, r) = \{a\} \subset A$ , and hence, the subsets are open. In a similar fashion, let  $A^C$  be the complement of  $A$ , and let  $a \in A^C$  be some point. Then, choosing  $r < 1$ , we have a ball  $B(a, r) = \{a\} \subset A^C$ , and hence,  $A^C$  is also open, resulting in  $A$  being closed. ■

6. If  $x_0$  is an accumulation point of  $A \subset (X, d)$ , show that any neighborhood of  $x_0$  contains infinitely many points of  $A$ .

*Proof.* We shall prove the claim by contradiction. By the definition of an accumulating point, the neighborhood of  $x_0$  contains at least one point distinct from  $x_0$  from  $A$ . Let  $B(x_0, \epsilon)$  be an  $\epsilon$ -neighborhood of  $x_0$ , which consists of finitely many points ( $\neq x_0$ )  $\{a_1, \dots, a_k\}$ . Let  $r \leq \min_{j \in [k]} d(x_0, a_j)$ . Then, the  $r$ -neighborhood of  $x_0$  consists of no points other than  $x_0$ , which is a contradiction since  $x_0$  is an accumulation point. Hence, there should be another point in the  $r$ -neighborhood, which contradicts our assumption of finitely many points. ■

7. Describe the closure of each of the sets: the integers on  $\mathbb{R}$ , the rational numbers on  $\mathbb{R}$ , the complex numbers with rational real and imaginary parts in  $\mathbb{C}$ , and the disk  $\{z \mid |z| < 1\} \subset \mathbb{C}$ .

*Proof.* The closure of a set includes all points of the set and all its limit points. The limit points of a set are points such every neighborhood of the point consists of at least one element of the set distinct from the point. Clearly, for the set of integers on  $\mathbb{R}$ , the closure is the set of integers itself. In other words, there are no limit points for the set of integers. Let  $z$  be an integer, then drawing an  $\epsilon$ -neighborhood around  $z$  with  $\epsilon < 1$  ensures that there is no integer apart from  $z$  in this neighborhood. Same is the case for any rational number  $q$ , any  $\epsilon$ -neighborhood around  $q$  may not consist of an integer.

For the set of rationals on  $\mathbb{R}$ , the closure is  $\mathbb{R}$ . This is because choosing any point in  $\mathbb{R}$  will have at least one rational number in its neighborhood. Same is the case for the set of complex numbers with rational real and imaginary parts. You can always find a new complex number with rational real and imaginary parts in a neighborhood.

Finally, for the open disk of radius 1 in the complex numbers, defined as

$$\{z \mid |z| < 1\} \subset \mathbb{C}$$

the accumulation points will be

$$\{z \mid |z| \leq 1\}.$$

Since any neighborhood of a "boundary point" will intersect with the disk, they also are accumulation points, and hence, the closure is simply the closed disk of radius 1. ■

8. Show that the closure  $\overline{B(x_0, r)}$  of an open ball  $B(x_0, r)$  in a metric space can differ from the closed ball  $\bar{B}(x_0, r)$ .

*Proof.* **TODO.** ■

9. Show that  $A \subset \bar{A}$ ,  $\overline{A \cup B} = \bar{A} \cup \bar{B}$ , and  $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$ .

*Proof.*  $A \subset \bar{A}$  follows from the definition of a closure. Now, we show  $\overline{A \cup B} = \bar{A} \cup \bar{B}$ . To show this, we show that  $\overline{A \cup B} \subset \bar{A} \cup \bar{B}$  and  $\bar{A} \cup \bar{B} \subset \overline{A \cup B}$ . First, let  $x \in \bar{A} \cup \bar{B}$ . Then,

1.  $x \in A \cup B$ . This means that  $x \in A$  or  $x \in B$ . WLOG, assume  $x \in A$ , then by the definition of closure,  $x \in \bar{A}$ , and hence,  $x \in \bar{A} \cup \bar{B}$ .
2.  $x \notin A \cup B$  and  $x$  is an accumulation point of  $A \cup B$ . This means that  $x$  is also an accumulation point of either  $A$  or  $B$ . This is clear if every neighborhood of  $x$  contains some point of  $A$  or if every neighborhood contains a point of  $B$ . Now consider the case where some  $\epsilon_A$ -neighborhood of  $x$  contains only points of  $A$  and  $\epsilon_B$ -neighborhood contains points of only  $B$ . WLOG, let  $\epsilon_A \geq \epsilon_B$ . Then,  $\epsilon_A$ -neighborhood also contains points from  $B$ . Thus,  $x$  would be a accumulation point for either  $A$  or  $B$ , which implies that  $x \in \bar{A}$  or  $x \in \bar{B}$ , resulting in  $x \in \bar{A} \cup \bar{B}$ .

This shows that  $x \in \bar{A} \cup \bar{B}$ .

Now we show that  $\bar{A} \cup \bar{B} \subset \overline{A \cup B}$ . Let  $x \in \bar{A} \cup \bar{B}$ . WLOG, assume  $x \in \bar{A}$ . Then,

1.  $x \in A$ , which means that  $x \in A \cup B$  and hence,  $x \in \overline{A \cup B}$ .
2.  $x \notin A$  and  $x$  is an accumulation point for  $A$ . This means that every neighborhood of  $x$  contains some point from  $A$ . That means that every neighborhood of  $x$  also contains some point from  $A \cup B$ , and hence,  $x$  is an accumulation point for  $A \cup B$ . Thus,  $x \in \overline{A \cup B}$ .

Thus,  $\bar{A} \cup \bar{B} \subset \overline{A \cup B}$ . This finishes the proof.

Finally, we show that  $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$ . Let  $x \in \overline{A \cap B}$ . Then,

1.  $x \in A \cap B$ . This means that  $x \in A$  and  $x \in B$ , and hence,  $x \in \bar{A}$  and  $x \in \bar{B}$ , resulting in  $x \in \bar{A} \cap \bar{B}$ .

2.  $x \notin A \cap B$  but  $x$  is an accumulating point of  $A \cap B$ . That means that every neighborhood of  $x$  has a point from  $A \cap B$ , which means the point in the neighborhood belongs to both  $A$  and  $B$ . Thus, every neighborhood contains a point from  $A$  and  $B$ , making  $x$  an accumulation point for both  $A$  and  $B$ . Hence,  $x \in \bar{A}$  and  $x \in \bar{B}$ , resulting in  $x \in \bar{A} \cap \bar{B}$ .

This finishes the proof. ■

**10.** A point  $x$  not belonging to a closed set  $M \subset (X, d)$  always has a nonzero distance from  $M$ . To prove this, show that  $x \in \bar{A}$  if and only if  $D(x, A) = 0$ , where  $A$  is any nonempty subset.

*Proof.* First, we define

$$D(x, A) = \inf_{a \in A} d(x, a).$$

We show that  $x \in \bar{A} \implies D(x, A) = 0$ . If  $x \in A$ , then by definition of  $D$ , we have  $D(x, A) = 0$ . However, let  $x \notin A$ , then for every  $\epsilon$ -neighborhood of  $x$ ,  $B(x, \epsilon) \cap A \neq \emptyset$ , which means that for any  $\epsilon > 0$ , we can find a point distinct from  $x$  belonging to  $A$ . As  $\epsilon$  can be arbitrarily small, the infimum would be 0.

Next we show that  $D(x, A) = 0 \implies x \in \bar{A}$ . If  $x \in A$ , then it is trivial. However, if  $x \notin A$ , then we have that for all  $\epsilon > 0$ , there exists some point  $a(\epsilon) \in A$  such that  $d(a(\epsilon), x) = \epsilon$ . Since  $\epsilon$  can be made arbitrarily small, the infimum results in 0. However, this also means that for every  $\epsilon$ -neighborhood of  $x$ , there exists some point  $a(\epsilon) \in A$  and hence,  $x$  is an accumulation point of  $A$ . Thus,  $x \in \bar{A}$ . This finishes the proof. ■

**11.** A boundary point of  $A \subset (X, d)$  is a point of  $X$  (which may or may not be in  $A$ ) such that every neighborhood of  $x$  contains points of  $A$  as well as points not belonging to  $A$ . The boundary of  $A$  is defined to be the set of all boundary points of  $A$ . Find the boundaries of  $(-1, 1)$ ,  $[-1, 1)$ ,  $[-1, 1]$  on  $\mathbb{R}$ , the set of rationals on  $\mathbb{R}$ , and the disks  $\{z \mid |z| < 1\}$  and  $\{z \mid |z| \leq 1\} \subset \mathbb{C}$ .

*Proof.* The boundary points of  $(-1, 1)$  are  $\{-1, 1\}$ . This is because any  $\epsilon$ -neighborhood of  $-1$  will contain  $-1 + \epsilon \in (-1, 1)$  and  $-1 - \epsilon \notin (-1, 1)$ . Note that  $x \in (-1, 1)$  is not a boundary point because every neighborhood will not contain any point which does not belong to  $(-1, 1)$ . The boundary points for  $(-1, 1]$  and  $[-1, 1]$  are also  $\{-1, 1\}$ .

The boundary points for the set of rationals on  $\mathbb{R}$  is  $\mathbb{R}$ . This is because for any  $q \in \mathbb{R}$ , the neighborhood of  $q$  will contain rationals as well as irrationals.

Finally, the boundary points for the disks  $\{z \mid |z| < 1\}$  and  $\{z \mid |z| \leq 1\} \subset \mathbb{C}$  is simply the set  $\{z \mid |z| = 1\}$ . ■

**12.** Show that  $B[a, b]$  is not separable.

*Proof.* We shall use the same methodology as the one used to show  $\ell^p$  is not separable. We will first construct an uncountable set of functions  $x_n$  such that  $d(x_n, x_m) > \epsilon$ . Now, if we take balls of radius  $\frac{\epsilon}{3}$  around each  $x_i$ , we will have uncountable balls that do not overlap. Suppose  $M$  is a dense subset, then by the definition of dense, each of the balls around  $x_i$  should have an element of  $M$ . This means that  $M$  has to intersect with every ball and since the number of balls are uncountable and disjoint,  $M$  is uncountable. Since any arbitrary dense set is uncountable, the space is not separable.

Now, to construct this uncountable set of functions, note that  $[a, b]$  is an uncountable set, and hence, having functions labeled with  $c \in [a, b]$  would result in uncountable functions. Define

$$x_c(d) = \mathbb{1}\{c = d\} \forall d \in [a, b].$$

This function is bounded, and moreover there are uncountable number of functions. Further,

$$d(x_m, x_n) = \sup_{t \in [a, b]} |x_m(t) - x_n(t)| = 1$$

and hence, substituting  $\epsilon = 1$  will finish the proof. ■

**13.** Show that a metric space  $X$  is separable if and only if  $X$  has a countable subset  $Y$  with the following property: For every  $\epsilon > 0$  and every  $x \in X$ , there is a  $y \in Y$  such that  $d(x, y) < \epsilon$ .

*Proof.* We first show that if a metric space  $X$  is separable, then it contains a countable subset with the property mentioned above. First, if  $X$  is separable it would contain separable dense subsets  $Y$ . Since  $Y$  is dense, we have that  $\bar{Y} = X$ , which means that every point in  $x \in X$  either belongs to  $Y$  or is an accumulation point of  $Y$ . In such a case, if  $x \in Y$ , then it is trivial. However, if  $x \notin Y$  and is an accumulation point, then by the definition of an accumulation point, every  $\epsilon$ -neighborhood of  $x$  would contain some  $y \in Y$ , and hence,  $d(x, y) < \epsilon$ . This finishes the proof.

We now prove the converse, i.e if there is a countable subset  $Y$  that satisfies the above mentioned property, then  $X$  is separable. In such a case, the property mentions that for every  $\epsilon > 0$  and for every  $x \in X$ , there

exists  $y \in Y$  such that  $d(x, y) < 0$ . If  $x \in Y$ , then this is trivially true. However, if  $x \notin Y$ , then, by the above mentioned property,  $x$  is an accumulation point for  $Y$ , which means that all points in  $X$  are either in  $Y$  or are accumulation points for  $Y$ . Thus,  $\overline{Y} = X$  and hence,  $Y$  is dense in  $X$ . Since  $Y$  is countable, we have that  $X$  is a separable set. This finishes the proof. ■

**14.** Show that a mapping  $T : X \mapsto Y$  is continuous if and only if the inverse image of any closed set  $M \subset Y$  is a closed set in  $X$ .

*Proof.* We instead show that if  $T$  is a continuous mapping if and only if the inverse image of any open set in  $Y$  is open in  $X$ .

Suppose  $T$  is a continuous mapping. Then for every  $\epsilon > 0, \exists \delta > 0$  such that

$$d(x_0, x) \leq \delta \implies d(Tx_0, Tx) \leq \epsilon.$$

Thus, consider the point  $Tx_0 \in Y$  and the set  $B(Tx_0, \epsilon)$  which is open. Clearly, due to the definition of  $T$ ,  $B(Tx_0, \epsilon)$  maps back to  $B(x_0, \delta)$  which is again open, and hence, if  $T$  is continuous, then the inverse image of an open set is an open set.

Now, we show the converse. Suppose we take the set  $B(Tx_0, \epsilon) \subset Y$ , which is open. Then, we have that the inverse image of  $B(Tx_0, \epsilon)$  is also open, say some set  $N \in X$  such that  $x_0 \in N$ . Also, since  $N$  is an open set,  $B(x_0, \delta) \subset N$ , and hence, the definition of continuity is fulfilled.

Now, we shall show the statement. Suppose  $T$  is continuous and  $M \subset Y$  is closed. This means that  $M^C$  is open and hence, the inverse image of  $M$ , say  $A \subset X$  is open. Thus,  $A^C$  is closed. Similarly, let  $M \subset Y$  is closed and it's inverse image  $A \subset X$  is also closed. This means that  $M^C$  and  $A^C$  is open and hence,  $T$  is continuous. ■

**15.** Show that the image of an open set under a continuous mapping need not be open.

*Proof.* Consider a function  $f$  on the set  $(a, b)$  that maps to  $[c, d]$  such that for some  $a < a_1, b_1 < b$ , we have  $f(a_1) = c$  and  $f(b_1) = d$ . Any such function will satisfy the above claim. ■

## 1.4 Convergence, Cauchy Sequences, Completeness

**1.** If a sequence  $(x_n)$  in a metric space  $X$  is convergent and has limit  $x$  show that every subsequence  $(x_{n_k})$  of  $(x_n)$  is also convergent and has the same limit  $x$ .

*Proof.* Since  $(x_n)$  is convergent, we have that for every  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for all  $n \geq N(\epsilon)$ , we have that  $d(x, x_n) < \epsilon$ . This also means that for all  $n_k \geq N(\epsilon)$ ,  $d(x_{n_k}, x) \leq \epsilon$ . Hence, the subsequence  $(x_{n_k})$  also follows the definition of convergence. ■

**2.** If  $(x_n)$  is Cauchy and has a convergent subsequence, say  $x_{n_k} \rightarrow x$ , show that  $(x_n)$  is convergent with the limit  $x$ .

*Proof.* Since  $(x_n)$  is Cauchy, we have that for every  $\epsilon > 0$ , there exists  $N_1(\epsilon)$  such that for all  $m, n \geq N_1(\epsilon)$ , we have  $d(x_m, x_n) \leq \epsilon$ . Now, since the subsequence  $(x_{n_k})$  is converging to  $x$ , we have that for every  $\epsilon > 0$ , there exists  $N_2(\epsilon)$  such that for all  $n_k \geq N_2(\epsilon)$ ,  $d(x_{n_k}, x) \leq \epsilon$ . Thus, fixing the value of  $\epsilon$ , and choosing the value of  $N(\epsilon) = \max\{N_1(\epsilon), N_2(\epsilon)\}$ , we have that for some  $n, n_k \geq N(\epsilon)$

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) \leq 2\epsilon.$$

This finishes the proof. ■

**3.** Show that  $x_n \rightarrow x$  if and only if for every neighborhood  $V$  of  $x$  there is an integer  $n_0$  such that  $x_n \in V$  for all  $n > n_0$ .

*Proof.* Suppose  $x_n \rightarrow x$ . Then, for every  $\epsilon > 0$ , there exists some  $N(\epsilon)$  such that for all  $n \geq N(\epsilon)$ ,  $d(x_n, x) \leq \epsilon$ . This means that for every  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for  $n \geq N(\epsilon)$ ,  $x_n \in B(x, \epsilon)$ . Thus, for every neighborhood  $V$ , there exists  $\epsilon > 0$  such that  $B(x, \epsilon) \subset V$ , and hence, for  $n \geq N(\epsilon) = n_0$ ,  $x_n \in V$ .

We now show the converse. Suppose, for every neighborhood  $V$ , there is an integer  $n_0$  such that  $x_n \in V$  for  $n > n_0$ . Assume that  $B(x, \epsilon) \subset V$ . Thus, setting  $n_0 = N(\epsilon)$ , for all  $n > N(\epsilon)$ ,  $x_n \in B(x, \epsilon)$ , and hence, the definition of convergence is fulfilled. ■

**4.** Show that a Cauchy sequence is bounded.

*Proof.* A Cauchy sequence  $(x_n)$  is defined as follows: for every  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for all  $m, n \geq N(\epsilon)$ , we have that  $d(x_m, x_n) \leq \epsilon$ . Fix an  $\epsilon$  and  $x_N = x_{N(\epsilon)}$  as our reference point. Now we know that for  $n \geq N$ ,  $d(x_n, x_N) \leq \epsilon$ . Thus, choose

$$r = \max_{i \in [N]} d(x_i, x_N)$$

and  $r_0 = \max\{r, \epsilon\}$ . Then, for all  $n > 0$ , we have  $d(x_n, x_N) \leq r_0$  which shows that the entire sequence is bounded in a ball of radius  $r_0$  centered at  $x_N$ . ■

**5.** *Is boundedness of a sequence in a metric space sufficient for the sequence to be Cauchy or convergent?*

*Proof.* No, the boundedness property is not sufficient. An example is  $x_n = \sin \frac{n\pi}{2}$ , which is always bounded by 1 but is neither Cauchy nor convergent. ■

**6.** *If  $(x_n)$  and  $(y_n)$  are Cauchy sequences in a metric space  $(X, d)$ , show that  $(a_n)$ , where  $a_n = d(x_n, y_n)$  is convergent.*

*Proof. Note: this question seems incorrect, and should only be true for complete spaces.*

Since  $(x_n)$  is Cauchy, for every  $\epsilon > 0$ , there exists  $N_x(\epsilon)$  such that for every  $m, n > N_x(\epsilon)$ , we have that  $d(x_m, x_n) < \epsilon$ . Similarly, since  $(y_n)$  is Cauchy, for every  $\epsilon > 0$ , there exists  $N_y(\epsilon)$  such that for every  $m, n > N_y(\epsilon)$ , we have that  $d(y_m, y_n) < \epsilon$ . Now, fix  $\epsilon$  and let  $N(\epsilon) = \max\{N_x(\epsilon), N_y(\epsilon)\}$ . Then for every  $m, n > N(\epsilon)$ , we have

$$d(x_n, y_n) \leq d(x_n, x_m) + d(y_n, y_m) + d(x_m, y_m) < 2\epsilon + d(x_m, y_m)$$

and hence,

$$d(x_n, y_n) - d(x_m, y_m) \leq 2\epsilon$$

Similarly, by symmetry

$$d(x_m, y_m) - d(x_n, y_n) \leq 2\epsilon$$

and hence,

$$|d(x_n, y_n) - d(x_m, y_m)| \leq 2\epsilon.$$

Thus, the sequence  $(a_n)$  is Cauchy, and since the space should be complete, the sequence is also convergent. ■

**7.** *Give an indirect proof of Lemma 1.4-2(b).*

*Proof.* What an indirect proof is seems unclear. We wish to prove that if a sequence  $(x_n) \rightarrow x$  and  $(y_n) \rightarrow y$ , then  $d(x_n, y_n) \rightarrow d(x, y)$ . By the properties of convergence, we have

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0 \text{ and } \lim_{n \rightarrow \infty} d(y_n, y) = 0.$$

Thus, using the triangle inequality and non-negativity of the metric, we have

$$0 \leq d(x_n, y_n) \leq d(x_n, x) + d(y_n, y) + d(x, y).$$

As  $n$  grows, we have  $0 \leq \lim_{n \rightarrow \infty} d(x_n, y_n) \leq d(x, y)$ . Similarly, we can also obtain  $0 \leq d(x, y) \leq \lim_{n \rightarrow \infty} d(x_n, y_n)$ , which gives us the result that

$$\lim_{n \rightarrow \infty} d(x_n, y_n) = d(x, y).$$

■

**8.** *If  $d_1$  and  $d_2$  are metrics on the same set  $X$  and there are positive numbers  $a$  and  $b$  such that for all  $x, y \in X$*

$$ad_1(x, y) \leq d_2(x, y) \leq bd_1(x, y),$$

*show that the Cauchy sequences in  $(X, d_1)$  and  $(X, d_2)$  are the same.*

*Proof.* Let  $(x_n)$  be a Cauchy sequence on  $(X, d_1)$ . Then, for every  $\epsilon > 0$ , we have some  $N(\epsilon)$  such that for all  $n, m > N(\epsilon)$ ,  $d_1(x_m, x_n) < \epsilon$ . Using the relation between  $d_1$  and  $d_2$ , we have that for every  $m, n > N(\frac{1}{b}(\epsilon))$ ,  $d_2(x_m, x_n) < b\epsilon$  and hence, every sequence  $(x_n)$  that is Cauchy in  $(X, d_1)$  is also Cauchy in  $(X, d_2)$ . Similarly, if a sequence  $(y_n)$  is Cauchy in  $(X, d_2)$ , for some  $\epsilon > 0$ , we have some  $N(\epsilon)$  such that for all  $m, n > N(\epsilon)$ ,  $d_2(y_m, y_n) < \epsilon$ . Using the relation once again, we have that for  $m, n > N(a(\frac{\epsilon}{b}))$ ,  $d_1(y_m, y_n) < \frac{\epsilon}{a}$  and hence,  $(y_n)$  is Cauchy in  $(X, d_1)$ . This finishes the proof. ■

**9.** *Using 8, show that the metric spaces in 13, 14, 15 from section 1.2 have the same Cauchy sequences.*

*Proof.* Define

$$\begin{aligned} d(x, y) &= d_1(x_1, y_1) + d_2(x_2, y_2) \\ \tilde{d}(x, y) &= \sqrt{d_1(x_1, y_1)^2 + d_2(x_2, y_2)^2} \\ \tilde{\tilde{d}}(x, y) &= \max\{d_1(x_1, y_1), d_2(x_2, y_2)\} \end{aligned}$$

on the metric space  $X = X_1 \times X_2$ . It is easy to see the following:

1.  $\tilde{d} \leq \sqrt{2}\tilde{\tilde{d}}$ .
2.  $d \leq 2\tilde{d}$ .
3.  $d \leq \tilde{d}$ .
4.  $\tilde{\tilde{d}} \leq d$ .

Combining these gives us

$$\tilde{\tilde{d}} \leq d \leq \tilde{d} \leq 2\tilde{\tilde{d}}.$$

This finishes the proof. ■

**10.** Using the completeness of  $\mathbb{R}$ , prove the completeness of  $\mathbb{C}$ .

*Proof.* Since  $\mathbb{R}$  is complete, every Cauchy sequence converges to some point in  $\mathbb{R}$ . Now, every complex number has a real part and an imaginary part. Let  $(c_n)$  be some Cauchy sequence of complex numbers such that  $c_n = a_n + \iota b_n$ . Clearly, the sequence  $(a_n)$  is Cauchy and convergent since  $\mathbb{R}$  is complete. Also, multiplying every element of a sequence does not change the convergence properties and hence,  $(\iota b_n)$  is also Cauchy and convergent. Thus,  $(a_n + \iota b_n) = (c_n)$  is also Cauchy and convergent. ■

## 1.5 Examples: Completeness Proofs

**1.** Let  $a, b \in \mathbb{R}$  and  $a < b$ . Show that the open interval  $(a, b)$  is an incomplete subspace of  $\mathbb{R}$  while the closed interval  $[a, b]$  is complete.

*Proof.* First, due to Theorem 1.4-7, we have that a closed subspace is complete and hence,  $[a, b]$  is complete. We shall prove this theorem again which states that *A subspace  $M$  of a metric space  $X$  is complete if and only if the set  $M$  is closed.* First, let  $M$  be complete, this means that for any Cauchy sequence  $(x_n) \in M$ ,  $x_n \rightarrow x \in M$ . This means that all the limit points (accumulation points) of  $M$  belong to  $M$  and hence,  $\bar{M} = M$ . Thus,  $M$  is closed. Now, we show the converse. Let  $M$  be closed. This means that all the limit points (accumulation points) of  $M$  belong to  $M$ , since  $\bar{M} = M$ , and hence, any Cauchy sequence  $(x_n) \in M$  will converge to some  $x \in M$ , and hence, the space is complete.

Further, we can construct a counterexample to show that  $(a, b)$  is incomplete. Consider the sequence  $x_n = a + \frac{1}{n}$ . Clearly, as  $n \rightarrow \infty$ ,  $x_n \rightarrow a$ , however,  $a \notin (a, b)$ . Thus, the set  $(a, b)$  is incomplete. We can also see that  $a$  is an accumulation point of  $(a, b)$  and hence, the closure of  $(a, b)$  is  $[a, b]$ . Since the closure is not equal to the set, it is incomplete. ■

**2.** Let  $X$  be the space of all ordered  $n$ -tuples  $x = (\xi_1, \dots, \xi_n)$  of real numbers and

$$d(x, y) = \max_i |\xi_i - \eta_i|$$

where  $y = (\eta_j)$ . Show that  $(X, d)$  is complete.

*Proof.* First, assume that  $(x_n) = (\xi_j^{(n)})$  is a Cauchy sequence. This means that for every  $\epsilon > 0$  there exists  $N(\epsilon)$  such that for all  $m, n > N(\epsilon)$ , we have

$$d(x_n, x_m) < \epsilon.$$

This means that

$$\max_i |\xi_i^{(n)} - \xi_i^{(m)}| < \epsilon \implies \forall i \in [n] \quad |\xi_i^{(n)} - \xi_i^{(m)}| < \epsilon$$

By the completeness of real numbers, we have that the sequences  $(\xi_i^{(1)}, \xi_i^{(2)}, \dots)$  is Cauchy, and hence, converges to some value, say  $\xi_i$ . Hence, define  $x = (\xi_i)$ . Clearly,  $x \in X$ . We now wish to show that the sequence  $(x_n)$  converges. Recall that for  $\epsilon > 0$  and  $m, n > N(\epsilon)$ , we have that  $d(x_m, x_n) < \epsilon$ . As  $m \rightarrow \infty$ , we have

$$d(x_n, x) < \epsilon$$

and hence,  $x$  is a limit point for  $x_n$ . Thus, the space is complete. ■

**3.** Let  $M \subset \ell^\infty$  be the subspace of all sequences  $x = (\xi_j)$  with at most finitely many nonzero terms. Find a Cauchy sequence in  $M$  that does not converge in  $M$  so that  $M$  is not complete.

*Proof.* The easiest way to show that the space is not complete is to find a Cauchy sequence that converges to an all-zero sequence. Consider the sequence

$$x_n = \left(1, \frac{1}{2}, \dots, \frac{1}{n}, 0, \dots, 0\right)$$

Then, choose  $\epsilon > 0$  and  $N = \frac{1}{\epsilon}$  and for some  $n > m > N$ ,

$$d(x_n, x_m) = \sup \left\{ \frac{1}{m+1}, \dots, \frac{1}{n} \right\} = \frac{1}{m+1} \leq \frac{1}{m} \leq \frac{1}{N} = \epsilon.$$

Thus, it is a Cauchy sequence. However, note that the sequence will converge to  $x = (1, \frac{1}{2}, \dots, \frac{1}{n}, \dots) \notin M$ . ■

**4.** Show that the subspace  $M$  in the previous question 3 is not complete using Theorem 1.4-7.

*Proof.* Theorem 1.4-7 says that a subspace is complete if and only if the subspace is closed. We wish to show that the subspace is not closed, i.e.  $M \neq \overline{M}$ , or in other words, there exists some  $x$  which is an accumulation point and  $x \notin M$ . This is exactly what we showed in the previous question. ■

**5.** Show that the set  $X$  of all integers with metric  $d(x, y) = |x - y|$  is a complete metric space.

*Proof.* Recall that the set of integers is closed in  $\mathbb{R}$  (Section 1.3 7) and hence, it is a complete metric space. We shall also attempt to show it using Cauchy sequences. Consider some sequence  $(x_n)$ . Then for every  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for every  $n, m > N(\epsilon)$

$$d(x_m, x_n) < \epsilon.$$

However, if we choose  $\epsilon < 1$ , then at some point, all these points would become equal, i.e.

$$d(x_m, x_n) < 1 \implies x_m = x_n$$

and hence, the sequence converges to some integer  $x$ , which shows that the space is complete. ■

**6.** Show that the set of all real numbers is incomplete if we choose

$$d(x, y) = |\arctan x - \arctan y|$$

*Proof.* Consider the sequence  $x_n = \tan\left(\frac{\pi}{2} - \frac{1}{n}\right)$ . Clearly,

$$\lim_{n \rightarrow \infty} x_n = \infty$$

and hence, the sequence is diverging. However, for  $m > n$ ,

$$d(x_n, x_m) = \left| \frac{1}{m} - \frac{1}{n} \right| = \frac{1}{n} - \frac{1}{m}.$$

Clearly, for some  $\epsilon > 0$ ,  $N(\epsilon) = \frac{1}{\epsilon}$ , and  $m > n > N(\epsilon)$ , we have that  $d(x_n, x_m) < \epsilon$  and hence, the sequence is Cauchy. Thus, the space is not complete. ■

**7.** Let  $X$  be the set of all positive integers and  $d(m, n) = |m^{-1} - n^{-1}|$ . Show that this is not complete.

*Proof.* Consider the sequence  $(x_n) = n$ , then for some  $\epsilon > 0$  and  $N(\epsilon) = \frac{1}{\epsilon}$ , choosing  $n > m > N$ , we have

$$d(x_m, x_n) = \frac{1}{m} - \frac{1}{n} < \frac{1}{N} = \epsilon$$

and hence, the sequence is Cauchy. However, clearly, it is not convergent. ■

**8.** Show that the subspace  $Y \subset C[a, b]$  consisting of all  $x \in C[a, b]$  such that  $x(a) = x(b)$  is complete.

*Proof.* Suppose there is some Cauchy sequence  $x_n$  such that for some  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for all  $m, n > N(\epsilon)$ , we have  $d(x_n, x_m) < \epsilon$ , where  $d(x, y)$  is defined as

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

Suppose  $x_n \rightarrow x$ . To show that  $Y$  is complete, we wish to show  $x \in Y$ . Since  $x_n \rightarrow x$ , we have that for some  $\epsilon > 0$  and  $n > N(\epsilon)$ ,

$$d(x_n, x) < \epsilon \implies \max_{t \in [a, b]} |x_n(t) - x(t)| < \epsilon \implies |x_n(t) - x(t)| < \epsilon \forall t \in [a, b].$$

Thus,

$$d(x(a), x(b)) < d(x(a), x_n(a)) + d(x_n(a), x_n(b)) + d(x_n(b), x(b)) < 2\epsilon$$

since  $x_n(a) = x_n(b)$ . Thus, we have

$$0 \leq d(x(a), x(b)) < 2\epsilon$$

for all  $\epsilon > 0$ . As  $\epsilon$  becomes arbitrarily small, we have  $d(x(a), x(b)) = 0 \implies x(a) = x(b)$  and hence,  $x \in Y$ . ■

**9.** Prove that if a sequence  $(x_m)$  of continuous functions on  $[a, b]$  converges on  $[a, b]$  and the convergence is uniform, then the limit function  $x$  is also continuous on  $[a, b]$ .

*Proof.* First, since  $x_m$  is continuous, we have for a fixed  $t_0 \in [a, b]$  and for some  $\epsilon > 0$ , there exists some  $\delta > 0$  such that

$$d(t, t_0) \leq \delta \implies d(x_m(t), x_m(t_0)) \leq \epsilon \forall m.$$

Also, since the sequence  $(x_m)$  converges pointwise to  $x$ , we have for some  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for  $n > N(\epsilon)$ ,

$$d(x_n(t), x(t)) \leq \epsilon \forall t \in [a, b].$$

Thus, for a fixed  $t_0 \in [a, b]$  and  $\epsilon > 0$ , corresponding  $\delta > 0$ , and  $n > N(\epsilon)$ , we have that if  $d(t, t_0) < \delta$

$$d(x(t), x(t_0)) \leq d(x(t), x_n(t)) + d(x_n(t), x_n(t_0)) + d(x_n(t_0), x(t_0)) \leq 3\epsilon.$$

This shows that  $x$  is continuous at  $t_0$ . Since  $t_0$  was an arbitrary point, it shows that  $x$  is continuous. ■

**10.** Show that the discrete metric space is complete.

*Proof.* Recall the discrete metric space is defined as  $d(x, x) = 0$  and  $d(x, y) = 1$  if  $x \neq y$ . Recall that this space has no accumulation points, and hence, the space is closed. Thus, the only Cauchy sequences are constant sequences, and these definitely converge. This also makes sense since setting  $\epsilon < 1$  ensures a constant sequence, while  $\epsilon \geq 1$  brings the whole space into play, and thus, no sequence will ever converge. Thus, convergence only occurs with  $\epsilon < 1$ , which results in a constant sequence. Also, the only Cauchy sequences are constant. Thus, the space is complete. ■

**11.** Show that in the space  $s$ , we have  $x_n \rightarrow x$  if and only if  $\xi_j^{(n)} \rightarrow \xi_j$  for all  $j$ .

*Proof.* Let  $x_n = \xi_j^{(n)}$  and the metric is defined as

$$d(x, y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|\xi_j - \eta_j|}{1 + |\xi_j - \eta_j|}$$

Suppose  $x_n \rightarrow x$ . This means

$$\lim_{n \rightarrow \infty} d(x, x_n) = 0 \implies \lim_{n \rightarrow \infty} \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|\xi_j^{(n)} - \xi_j|}{1 + |\xi_j^{(n)} - \xi_j|} = 0$$

However, since the sum comprises of non-negative terms, the sum is zero only if each term is zero, or

$$\lim_{n \rightarrow \infty} |\xi_j^{(n)} - \xi_j| = 0 \implies \lim_{n \rightarrow \infty} d(\xi_j^{(n)}, \xi_j) = 0 \implies \xi_j^{(n)} \rightarrow \xi_j$$

On the other hand, if  $\xi_j^{(n)} \rightarrow \xi_j$ , then

$$\lim_{n \rightarrow \infty} |\xi_j^{(n)} - \xi_j| = 0.$$

Thus,

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0.$$

This finishes the proof. ■



**12.** Using 11 show that the sequence space  $s$  is complete.

*Proof.* Suppose  $(x_n)$  is a Cauchy sequence where  $x_n = (\xi_1^{(n)}, \dots)$ . Then, for some  $\epsilon > 0$ , there exists some  $N(\epsilon)$  and some  $m, n > N(\epsilon)$  such that

$$d(x_m, x_n) < \epsilon \implies \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|\xi_j^{(n)} - \xi_j^{(m)}|}{1 + |\xi_j^{(n)} - \xi_j^{(m)}|} < \epsilon \implies |\xi_j^{(n)} - \xi_j^{(m)}| < \frac{\epsilon}{2}$$

or, in other words,  $(\xi_j^{(1)}, \xi_j^{(2)}, \dots)$  is Cauchy, and hence, converges. Since the space of complex numbers is complete, we also have  $\xi^{(k)_j} \rightarrow \xi_j$  (say), and hence, from 11, we can say  $x_n \rightarrow x$ . This shows that  $x$  is complete. ■

**13.** Show that another Cauchy sequence is given by

$$x_n(t) = n \text{ if } 0 \leq t \leq \frac{1}{n^2} \text{ and } x_n(t) = t^{-\frac{1}{2}} \text{ if } \frac{1}{n^2} \leq t \leq 1$$

under the metric

$$d(x, y) = \int_0^1 |x(t) - y(t)| dt.$$

*Proof.* Clearly, for some  $m > n$ ,

$$\begin{aligned} d(x_m, x_n) &= \int_0^{\frac{1}{m^2}} |m - n| dt + \int_{\frac{1}{m^2}}^{\frac{1}{n^2}} |t^{-\frac{1}{2}} - n| dt \\ &= \frac{m - n}{m^2} + \int_{\frac{1}{m^2}}^{\frac{1}{n^2}} t^{-\frac{1}{2}} - n dt \\ &= \frac{m - n}{m^2} + 2 \left( \frac{1}{n} - \frac{1}{m} \right) - n \left( \frac{1}{n^2} - \frac{1}{m^2} \right) \\ &= \frac{1}{n} - \frac{1}{m} \\ &\leq \epsilon \end{aligned}$$

for some  $\epsilon > 0$  and choosing  $N(\epsilon) = \frac{1}{\epsilon}$  and  $m > n > N(\epsilon)$ . Thus, the sequence is Cauchy. ■

**14.** Show that the Cauchy sequence in 13 does not converge.

*Proof.* Note, from 13, we have for some  $m > n$ ,

$$d(x_m, x_n) = \int_0^{\frac{1}{m^2}} |m - n| dt + \int_{\frac{1}{m^2}}^{\frac{1}{n^2}} |t^{-\frac{1}{2}} - n| dt$$

and hence,

$$\lim_{m \rightarrow \infty} d(x_m, x_n) = d(x, x_n) = \int_0^{\frac{1}{n^2}} |t^{-\frac{1}{2}} - n| dt$$

Clearly,  $x$  would be unbounded at  $t = 0$  while the function would then be defined as

$$x(t) = t^{-\frac{1}{2}} \forall t \in (0, 1]$$

Hence, there is no convergence at  $t = 0$ . ■

**15.** Let  $X$  be the metric space of all real sequences that has only finitely many nonzero terms, and  $d(x, y) = \sum |\xi_j - \eta_j|$ , where  $y = (\eta_j)$ . Show that with  $x_n = (\eta_j^{(n)})$ ,

$$\eta_j^{(n)} = j^{-2} \text{ for } n \in [j] \text{ and } \eta_j^{(n)} = 0 \text{ for } j > n$$

is Cauchy but does not converge.

*Proof.* Let  $\epsilon > 0$  be fixed. Then, there exists  $N(\epsilon)$  such that for  $m > n > N(\epsilon)$ , we have

$$d(x_m, x_n) = \sum_{j=n}^m \frac{1}{j^2} \leq \frac{m-n}{n^2} < \epsilon.$$

However, say we fix some  $N$ , then for  $n > N$ , we have

$$d(x_n, x) = |1 - \xi_1| + \left| \frac{1}{4} - \xi_2 \right| + \dots + \frac{1}{(N+1)^2} + \dots + \frac{1}{n^2} \geq \frac{1}{(N+1)^2}$$

which will never be arbitrarily small since  $N$  is fixed. This equation uses the fact that if  $x \in X$ , then, after some  $N$ , all its terms would be zero. ■

## 1.6 Completion of Metric Spaces

1. Show that if a subspace  $Y$  of a metric space consists of finitely many points, then  $Y$  is complete.

*Proof.* First, suppose there is a Cauchy sequence  $x_n$ . This means that for every  $\epsilon > 0$ , there exists  $N(\epsilon)$  and some  $m, n > N(\epsilon)$  such that  $d(x_m, x_n) < \epsilon$ . However, due to a finite number of points available in the set, the sequence has finitely many terms, and thus, we can express  $\epsilon_0 = \min_{a, b \in Y} d(a, b)$ . Thus, if we choose  $\epsilon = \epsilon_0$ , the sequence will have to become constant, which always converges. Thus, all Cauchy sequences become constant eventually, which converge, and hence, the space is complete. ■

2. What is the completion of  $(X, d)$  where  $X$  is the set of all rational numbers and  $d(x, y) = |x - y|$ ?

*Proof.* The completion would be the set of all reals,  $\mathbb{R}$ . ■

3. What is the completion of a discrete metric space?

*Proof.* Recall that a discrete metric space is always complete. Hence, the completion is the entire space  $X$  itself. ■

4. If  $X_1$  and  $X_2$  are isometric and  $X_1$  is complete, show that  $X_2$  is complete.

*Proof.*  $X_1$  and  $X_2$  are isometric spaces if there exists some bijective isometry  $T : X_1 \mapsto X_2$ . By an isometry, we mean that for some  $x, y \in X_1$

$$d_1(x, y) = d_2(Tx, Ty).$$

If  $X_1$  is complete, then  $X_1$  is also closed, i.e.,  $\overline{X_1} = X_1$ . Suppose  $x \in X_1$  is an accumulation point of  $X_1$ . This means that for every  $\epsilon > 0$ , there exists some  $x_\epsilon \in X_1$  such that

$$d_1(x, x_\epsilon) < \epsilon.$$

Now, since  $X_1$  and  $X_2$  are isometric spaces

$$d_2(Tx, Tx_\epsilon) < \epsilon$$

which means that  $Tx \in X_2$  is also an accumulation point for  $X_2$  (since  $x_\epsilon$  is arbitrary). Thus, all accumulation points of  $X_2$  lie in  $X_2$  and hence,  $X_2$  is also complete. ■

5. A homeomorphism is a continuous bijective mapping  $T : X \mapsto Y$  whose inverse is continuous. The metric spaces  $X$  and  $Y$  are said to be homeomorphic. Show that if  $X$  and  $Y$  are isometric, then they are also homeomorphic.

*Proof.* First, an isometry implies the existence of a bijective map, so we already have a bijection. We need to ensure both the map and its inverse is continuous. The definition of continuity is as follows: for every  $\epsilon > 0$ , there exists  $\delta > 0$ , such that  $d_1(x, x_0) < \delta \implies d_2(Tx, Tx_0) < \epsilon$  ensures that the map  $T$  is continuous as  $x_0$ . Now, from the condition of isometric spaces, we have that

$$d_1(x, x_0) = d_2(Tx, Tx_0)$$

and hence, if we set  $\delta = \epsilon$ , we see that the definition of continuity is always fulfilled. Thus, the mapping from  $X$  to  $Y$  is always continuous. Making a similar argument for  $T^{-1}$  results in the conclusion that  $T^{-1}$  is continuous and hence,  $X$  and  $Y$  are also homeomorphic. ■

6. Show that  $C[0, 1]$  and  $C[a, b]$  are isomorphic.

*Proof.* To show that  $C[0, 1]$  and  $C[a, b]$  is isometric, we wish to find some mapping  $T : C[0, 1] \mapsto C[a, b]$  such that for some  $x, y \in C[0, 1]$ ,

$$d_1(x, y) = d_2(Tx, Ty),$$

where  $d_1$  and  $d_2$  are metrics on  $C[0, 1]$  and  $C[a, b]$  respectively.

Let  $x, y \in C[0, 1]$ . Let  $T : C[0, 1] \mapsto C[a, b]$ . Then, we can define  $Tx \in C[a, b]$  as

$$(Tx)(t) = x\left(\frac{t-a}{b-a}\right).$$

In such a case,

$$d(Tx, Ty) = \max_{t \in [a, b]} |(Tx)(t) - (Ty)(t)| = \max_{t \in [a, b]} \left| x\left(\frac{t-a}{b-a}\right) - y\left(\frac{t-a}{b-a}\right) \right| = \max_{t \in [0, 1]} |x(t) - y(t)| = d(x, y)$$

and hence, the map is distance preserving. We can similarly define for  $x, y \in C[a, b]$ ,  $T^{-1} : C[a, b] \mapsto C[0, 1]$  such that  $T^{-1}x \in C[0, 1]$

$$(T^{-1}x)(t) = x((b-a)t + a)$$

and hence, we have

$$d(T^{-1}x, T^{-1}y) = \max_{t \in [0, 1]} |(T^{-1}x)(t) - (T^{-1}y)(t)| = \max_{t \in [a, b]} |x(t) - y(t)| = d(x, y)$$

and hence,  $T^{-1}$  is also distance preserving. Further, because of the linear mapping, the mapping is bijective. Hence, the spaces are isometric. ■

**7.** If  $(X, d)$  is complete, show that  $(X, \tilde{d})$  where  $\tilde{d} = \frac{d}{1+d}$  is complete.

*Proof.* Let  $x_m$  be some Cauchy sequence in the space  $(X, d)$ , such that for every  $\epsilon > 0$ , there exists some  $N(\epsilon)$ , such that for all  $m, n > N(\epsilon)$ ,  $d(x_m, x_n) < \epsilon$ . Now, consider the same sequence under the metric  $\tilde{d}$ . Thus, we have

$$\tilde{d}(x_m, x_n) = \frac{d(x_m, x_n)}{1 + d(x_m, x_n)} < \epsilon$$

and hence,  $x_m$  is a Cauchy sequence in  $(X, \tilde{d})$  too. Since  $(X, d)$  is complete, we have that every Cauchy sequence converges to some point in  $X$ . Let  $x_m \rightarrow x$ , then we have that for some  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for  $n > N(\epsilon)$ ,  $d(x_n, x) < \epsilon$ . Similarly, we have for the same sequence  $x_n$ ,

$$\tilde{d}(x, x_n) = \frac{d(x, x_n)}{1 + d(x, x_n)} < \epsilon$$

and hence, the sequence converges in  $(X, \tilde{d})$ . Thus, this space is also complete. ■

**8.** Show that in 7, completeness of  $(X, \tilde{d})$  implies completeness of  $(X, d)$ .

*Proof.* Since  $(X, \tilde{d})$  is complete, we have that all accumulation points of  $X$  belong to  $X$  itself. Let  $x \in X$  be an accumulation point. Then, for every  $\epsilon > 0$ , there exists  $x_\epsilon$  such that

$$\tilde{d}(x, x_\epsilon) < \epsilon \implies \frac{d(x, x_\epsilon)}{1 + d(x, x_\epsilon)} < \epsilon \implies d(x, x_\epsilon) < \frac{\epsilon}{1 - \epsilon}$$

Thus, if  $\epsilon < 1$ ,  $x_\epsilon$  is also an accumulation point for  $X$  under the metric  $d$ . On the other hand, if  $\epsilon > 1$ , then this is trivially true. ■

**9.** If  $(x_n)$  and  $(x'_n)$  on  $(X, d)$  are such that (1) holds and  $x_n \rightarrow l$ , show that  $(x'_n)$  converges and has the limit  $l$ .

*Proof.* (1) is the following:

$$\lim_{n \rightarrow \infty} d(x_n, x'_n) = 0$$

Since  $x_n$  converges to  $l$ , we have that

$$\lim_{n \rightarrow \infty} d(x_n, l) = 0$$

Thus, we have

$$0 \leq \lim_{n \rightarrow \infty} d(x'_n, l) \leq \lim_{n \rightarrow \infty} d(x'_n, x_n) + \lim_{n \rightarrow \infty} d(x_n, l) = 0$$

and hence,  $x'_n$  converges to  $l$ . ■

**10.** If  $(x_n)$  and  $(x'_n)$  are convergent sequences in a metric space  $(X, d)$  and have the same limit  $l$ , show that they satisfy (1).

*Proof.* (1) is the following:

$$\lim_{n \rightarrow \infty} d(x_n, x'_n) = 0$$

Since, both  $(x_n)$  and  $(x'_n)$  converge to the same limit  $l$ , we have that

$$\lim_{n \rightarrow \infty} d(x_n, l) = 0$$

$$\lim_{n \rightarrow \infty} d(x'_n, l) = 0$$

Thus, we have

$$0 \leq \lim_{n \rightarrow \infty} d(x_n, x'_n) \leq \lim_{n \rightarrow \infty} d(x_n, l) + \lim_{n \rightarrow \infty} d(l, x'_n) = 0$$

This finishes the proof. ■

**11.** Show that (1) defines an equivalence relation on the set of all Cauchy sequences of elements of  $X$ .

*Proof.* (1) claims that two Cauchy sequences  $(x_n)$  and  $(x'_n)$  are equivalent (represented by  $x_n \sim x'_n$ ) if

$$\lim_{n \rightarrow \infty} d(x_n, x'_n) = 0$$

An equivalence relation  $R \subset X \times X$  satisfies the following three claims:

1.  $R(x, x) \forall x \in X$ .
2.  $R(x, y) = R(y, x) \forall x, y \in X$ .
3.  $R(x, y)$  and  $R(y, z)$  implies  $R(x, z) \forall x, y, z \in X$ .

Clearly,  $\lim_{n \rightarrow \infty} d(x_n, x_n) = 0$  and also  $\lim_{n \rightarrow \infty} d(x_n, x'_n) = \lim_{n \rightarrow \infty} d(x'_n, x_n) = 0$ . Suppose,  $\lim_{n \rightarrow \infty} d(x_n, z_n) = 0$  and  $\lim_{n \rightarrow \infty} d(y_n, z_n) = 0$ , then

$$0 \leq \lim_{n \rightarrow \infty} d(x_n, y_n) \leq \lim_{n \rightarrow \infty} d(x_n, z_n) + \lim_{n \rightarrow \infty} d(z_n, y_n) = 0 \implies \lim_{n \rightarrow \infty} d(x_n, y_n) = 0$$

Thus, the relation satisfies all the conditions of being an equivalence relation. ■

**12.** If  $(x_n)$  is Cauchy in  $(X, d)$  and  $(x'_n)$  in  $X$  satisfies (1), show that  $(x'_n)$  is Cauchy in  $X$ .

*Proof.* Since  $x_n$  is Cauchy in  $X$ , we have that for every  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for all  $m, n > N(\epsilon)$ ,  $d(x_m, x_n) < \epsilon$ . Also, (1) claims that

$$\lim_{n \rightarrow \infty} d(x_n, x'_n) = 0.$$

Thus, for some  $\epsilon > 0$  and  $m, n > N(\epsilon)$ , we have

$$d(x'_n, x'_m) \leq d(x'_n, x_n) + d(x_n, x_m) + d(x'_m, x_m) \leq 3\epsilon.$$

Thus,  $(x'_n)$  is also Cauchy. ■

**13.** A finite pseudometric on a set  $X$  is a function  $d : X \times X \mapsto \mathbb{R}$  satisfying (M1), (M3), (M4) and

$$(M2^*)d(x, x) = 0.$$

What is the difference between a metric and pseudometric. Show that  $d(x, y) = |\xi_1 - \eta_1|$  defines a pseudometric on the set of all ordered pairs of real numbers where  $x = (\xi_1, \xi_2)$  and  $y = (\eta_1, \eta_2)$ .

*Proof.* Let us reinterpret the original condition M2 for a metric

$$d(x, y) = 0 \iff x = y$$

This is a two sided condition, i.e if  $d(x, y) = 0$ , then  $x = y$ . On the other hand, if  $x = y$ , then  $d(x, y) = 0$ . However the condition for a pseudometric seems to be one sided, in other words the distance between the same point is 0. However, if the distance between two points is 0, it does not necessarily mean it is the same point. In other words,  $x = y \implies d(x, y) = 0$ . However,  $d(x, y) = 0 \not\Rightarrow x = y$ . This is clear from the example. Suppose  $x = y = (\xi_1, \xi_2)$ . Then,  $d(x, y) = |\xi_1 - \xi_1| = 0$ . However,  $d(x, y) = |\xi_1 - \eta_1| = 0 \implies \xi_1 = \eta_1$ . However, this does not mean  $\xi_2 = \eta_2$  and hence,  $x$  and  $y$  may not be equal points. ■

14. Does

$$d(x, y) = \int_a^b |x(t) - y(t)| \, dt$$

define a metric or pseudometric on  $X$  if  $X$  is the set of all real-valued continuous functions on  $[a, b]$  and if  $X$  is the set of all real-valued Riemann integrable functions on  $[a, b]$ ?

*Proof.* If  $X$  is the set of all real-valued continuous functions, then it is evident that if  $x = y$ , then  $d(x, y) = 0$ . However, if  $d(x, y) = 0$ , since the metric is the integral of positive quantities, the only way the sum is zero is if each of the quantities is zero, i.e.  $x(t) = y(t)$  for all  $t \in [a, b]$ .

On the other hand, if  $X$  is the set of all real-valued Riemann integrable functions, then it is once again clear that if  $x = y$ ,  $d(x, y) = 0$ . However,  $d(x, y) = 0$  does not imply  $x = y$ . For example, consider the function  $x(t) = c_1$  and  $y(t) = c_2$  if  $t \in C$  where  $C$  is a finite set, and  $c_1$  otherwise. Then, the pointwise difference does not matter, and the integral shall remain 0. ■

15. If  $(X, d)$  is a pseudometric space, we call a set

$$B(x_0, r) = \{x \in X \mid d(x, x_0) < r\}$$

an open ball in  $X$  with center  $x_0$  and radius  $r$ . What are open balls of radius 1 in 13?

*Proof.* In problem 13, we have  $X$  to be the set of ordered real value pairs such that  $x = (\xi_1, \xi_2)$  and  $y = (\eta_1, \eta_2)$ , then,  $d(x, y) = |\xi_1 - \eta_1|$ . Suppose,  $x_0 = (a, b)$ . Then, any point within distance 1 of  $x_0$  has first coordinate  $a \pm 1$ , while the second coordinate does not matter. Thus, the open ball would be of the form

$$B((a, b), 1) = \{(a \pm 1, r) \mid r \in \mathbb{R}\}$$

which results in vertical strips of width 2 centered at  $a$ . ■

## 2 Normed Spaces. Banach Spaces.

### 2.1 Vector Space

1. Show that the set of all real numbers with the usual addition and multiplication constitutes a one-dimensional real vector space and the set of all complex numbers constitutes a one-dimensional complex vector space.

*Proof.* In the set of all real numbers, it is easy to see that that addition satisfies associativity, commutativity, and consists of an identity (0), and an inverse  $(-x) \in \mathbb{R}$ . Also, multiplication satisfies associativity, has an identity (1), and satisfies distributivity. Thus, it is a vector space.

For the set of complex numbers, addition satisfies associativity, commutativity, and consists of an identity  $(0 + i0)$ , and an inverse  $(-x) \in \mathbb{C}$ . Also, multiplication satisfies associativity, has an identity  $(1 + i0)$ , and satisfies distributivity. Thus, it is a vector space. ■

2. Prove (1) and (2).

*Proof.* Denote the zero vector as  $\theta$ . Then, we wish to prove:

$$0x = \theta$$

$$\alpha\theta = \theta$$

$$(-1)x = -x$$

For the first statement, notice that for some  $\alpha \in K$  (where  $K$  is the scalar field), we have

$$\theta = \alpha x + (-\alpha)x = (\alpha + (-\alpha))x = 0x$$

where the first equality follows from the properties of addition on the vector space, the second equality follows from the distributive property for vector spaces, and the third follows from the properties of addition on  $K$  (there exists an additive inverse).

For the second statement, substituting  $x = \theta$  in the first statement gives us the result.

For the third statement, we have

$$(1 + (-1))x = \theta \implies (-1)x = -x$$

where the implication follows from distributive property. ■

3. Describe the span of  $M = \{(1, 1, 1), (0, 0, 2)\} \in \mathbb{R}^3$ .

*Proof.* The span of  $M$  is described as all possible linear combinations of the vectors in  $M$ . Let the coefficients be  $\alpha \in \mathbb{R}$  and  $\beta \in \mathbb{R}$  for  $(1, 1, 1)$  and  $(0, 0, 2)$  respectively. Then, we have

$$\text{span } M = \{(\alpha, \alpha, \alpha + 2\beta)\}$$

which represents a plane passing through the line  $x = y$ . ■

4. Which of the following subsets of  $\mathbb{R}^3$  constitute a subspace of  $\mathbb{R}^3$ ? Here,  $x = (\xi_1, \xi_2, \xi_3)$ .

1. All  $x$  with  $\xi_1 = \xi_2$  and  $\xi_3 = 0$ ?
2. All  $x$  with  $\xi_1 = \xi_2 + 1$ .
3. All  $x$  with positive  $\xi_1, \xi_2, \xi_3$ .
4. All  $x$  with  $\xi_1 - \xi_2 + \xi_3 = k$ .

*Proof.* Denote the subsets by  $X$ .

1. Consider two points  $x_1 = (a_1, a_1, 0) \in X$  and  $x_2 = (a_2, a_2, 0) \in X$ . Then, we have  $x_1 + x_2 = (a_1 + a_2, a_1 + a_2, 0) \in X$ . Also, let  $\alpha$  be some scalar, then,  $\alpha x_1 = (\alpha a_1, \alpha a_1, 0) \in X$ . Hence,  $X$  is a subspace.
2. Consider two points  $x_1 = (1 + a_1, a_1, b_1) \in X$  and  $x_2 = (1 + a_2, a_2, b_2) \in X$ . Then,  $x_1 + x_2 = (2 + a_1 + a_2, a_1 + a_2, b_1 + b_2) \notin X$  since  $2 + a_1 + a_2 \neq 1 + (a_1 + a_2)$ . Thus,  $X$  is not a subspace.
3. Let  $x_1 = (a_1, b_1, c_1) \in X$ , i.e.,  $a_1, b_1, c_1 > 0$ . Consider some  $\alpha < 0$ , then,  $\alpha x_1 \notin X$ , since the coordinates now become negative. Hence,  $X$  is not a subspace.

4. Consider  $x_1 = (a_1, b_1, c_1) \in X$  and let  $\alpha$  be some scalar. Then,  $\alpha x_1 = (\alpha a_1, \alpha b_1, \alpha c_1)$ . However,  $\alpha(a_1 - b_1 + c_1) \neq k$  and hence,  $X$  is not a subspace. ■

5. Show that  $\{x_1, \dots, x_n\}$  where  $x_j(t) = t^j$  is a linearly independent set in the space  $C[a, b]$ .

*Proof.* We wish to show that

$$\sum_{j=1}^n \alpha_j x_j = 0 \implies \alpha_j = 0 \forall j \in [n].$$

Note that this convergence has to be pointwise. Consider the function  $\sum_{j=1}^n \alpha_j x_j$ . Then,

$$\left( \sum_{j=1}^n \alpha_j x_j \right) (t) = \sum_{j=1}^n \alpha_j t^j = 0$$

Clearly, if  $t > 0$ , then this only holds if all the coefficients  $\alpha_j = 0$ . ■

6. Show that in a  $n$ -dimensional vector space  $X$ , the representation of any  $x$  as a linear combination of given basis vector is unique.

*Proof.* Suppose not. Then, let the two sets of coefficients be  $\{\alpha_i\}$  and  $\{\beta_i\}$ . We have that

$$x = \sum_{i=1}^n \alpha_i e_i = \sum_{i=1}^n \beta_i e_i.$$

Thus, we also have that

$$\sum_{i=1}^n (\alpha_i - \beta_i) e_i = 0.$$

But since the set of vectors  $e_i$  are linearly independent, this means that for all  $i$ ,  $\alpha_i - \beta_i = 0 \implies \alpha_i = \beta_i$ . This finishes the proof. ■

7. Let  $\{e_1, \dots, e_n\}$  be a basis for a complex vector space  $X$ . Find a basis for  $X$  regarded as a real vector space. What is the dimension of  $X$  in either case?

*Proof.* Suppose  $X$  were a real space. Then, we would need a separate basis for the real parts and the imaginary parts. This leads to the basis of  $\{e_1, \dots, e_n, \iota e_1, \dots, \iota e_n\}$ , resulting in a dimension of  $2n$ . ■

8. If  $M$  is linearly dependent set in a complex vector space  $X$ , is  $M$  linearly dependent in  $X$ , regarded as a real vector space?

*Proof.* Suppose  $M = \{a_i + \iota b_i\}_{i=1}^n$  and let

$$a_i + \iota b_r = \sum_{i=1, i \neq r}^n \frac{\alpha_i}{\alpha_r} (a_i + \iota b_i)$$

Equating the real and imaginary parts results in

$$\alpha_r = \sum_{i=1, i \neq r}^n \alpha_r a_i \text{ and } b_r = \sum_{i=1, i \neq r}^n \alpha_r b_i.$$

Thus, regarding  $X$  as real vector space, each of the coefficients can still be expressed as a linear combination of the other vectors. Thus, it is still linearly dependent. ■

9. On a fixed interval  $[a, b] \subset \mathbb{R}$ , consider the set  $X$  consisting of all polynomials with real coefficients and of degree not exceeding a given  $n$  and the polynomial  $x = 0$ . Show that  $X$  with the usual addition and multiplication by real numbers is a real vector space of dimension  $n+1$ . Find a basis for  $X$ . Show that we can obtain a complex vector space  $\overline{X}$  in a similar fashion if we let those coefficients be complex. Is  $X$  a subspace of  $\overline{X}$ ?

*Proof.* First, note that for addition, associativity and commutativity holds. Also, the zero polynomial is a part of  $X$ , and hence, there exists an additive identity. Further, the additive inverse is simply given by negating all coefficients, and hence, the additive inverse also belongs to the set  $X$ . Similarly, the associative property for multiplication also holds, and a multiplicative identity is simply the polynomial  $x = 1$ . Thus, it is a valid vector space. If every polynomial has degree at most  $n$ , then we require the monomials  $x^n, x^{n-1}, \dots, x, 1$  to build each polynomial, and hence, the degree is  $n + 1$ . The basis is the set of monomials  $\{x^j\}_{j=0}^n$ . If each of the coefficients are complex, then the overall polynomial can be split into a real polynomial and an imaginary polynomial, resulting in the space of complex polynomials. However,  $X$  is not a subspace of  $\overline{X}$  because allowing complex coefficients with polynomials of  $X$  would result in a complex polynomial, which will not belong to  $X$ . ■

**10.** Show that if  $Y$  and  $Z$  are subspaces of a vector space  $X$ , show that  $Y \cap Z$  is a subspace of  $X$  but  $Y \cup Z$  need not be one.

*Proof.* Let the set of vectors in  $Y \cap Z$  be denoted by  $u_i$ . Then, by the definition of  $u_i \in Y$  and  $u_i \in Z$ . Then, any linear combination  $\sum \alpha_j u_j$  will belong to both  $Y$  and  $Z$  and hence, to  $Y \cap Z$ . Thus,  $Y \cap Z$  is a subspace. On the other hand, the union of subspaces may not be a subspace. Let  $Y$  be the subspace of polynomials of the form  $\alpha + \beta x$ , while let  $Z$  be the subspace of polynomials of the form  $\alpha + \beta x^2$ . Then, the union is simply all polynomials of the form  $\alpha + \beta x$  or  $\alpha + \beta x^2$ . However, a linear combination of the vectors in  $Y \cup Z$  may result in polynomials of the form  $\alpha + \beta x + \gamma x^2$ , which do not belong to  $Y \cup Z$ . ■

**11.** If  $M \neq \emptyset$  is any subset of a vector space  $X$ , show that  $\text{span } M$  is a subspace of  $X$ .

*Proof.* Let  $M = \{x_1, \dots, x_n\}$ . By the definition of span, we know that

$$\text{span } M = \left\{ \sum_{j=1}^n \alpha_j x_j \mid (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n \right\}.$$

Suppose we take two vectors  $\sum_{j=1}^n \alpha_j x_j \in \text{span } M$  and  $\sum_{j=1}^n \beta_j x_j \in \text{span } M$ . Representing  $\gamma_j = \alpha_j + \beta_j$ , we get the sum of the vectors to be  $\sum_{j=1}^n \gamma_j x_j$ , which belongs to  $M$ . Similarly, let  $\alpha_j \beta = \gamma_j$ , then multiplying the vector  $\sum_{j=1}^n \alpha_j x_j$  by  $\beta$  results in the vector  $\sum_{j=1}^n \gamma_j x_j \in \text{span } M$ . Thus,  $M$  is a subspace. ■

**12.** Show that the set of all real two rowed square matrices forms a vector space  $X$ . What is the zero vector in  $X$ ? Determine  $\dim X$ . Find a basis for  $X$ . Give examples of subspaces for  $X$ . Do the symmetric matrices  $x \in X$  form a subspace? The singular matrices?

*Proof.* Clearly, addition for the two-rowed square matrices is associative and commutative, since the addition is element-wise. Further, the additive identity can be represented as

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Also, the additive inverse can be defined as follows:

$$x = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies -x = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$$

Also, multiplication is associative (and not commutative!), and we can define the multiplicative identity as

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Thus, this space is a vector space. Clearly,  $\dim X = 4$  and the basis is the following set:

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

The set of symmetric matrices is a subspace because the addition of two symmetric matrices is also symmetric, while multiplying by a constant maintains symmetry. However, the set of all singular matrices is not a subspace. Denote the set of singular matrices by  $S$ . Then,

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in S \text{ and } \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in S \text{ but } \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \notin S.$$

■



**13.** Show that the Cartesian product  $X = X_1 \times X_2$  of two vector spaces over the same field becomes a vector space if we define the two algebraic operations by

$$(x_1, x_2) + (y_1, y_2) = (x_1 + x_2, y_1 + y_2),$$

$$\alpha(x_1, x_2) = (\alpha x_1, \alpha x_2)$$

.

*Proof.* We first have

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2) = (y_1, y_2) + (x_1, x_2)$$

$$((x_1, x_2) + (y_1, y_2)) + (z_1, z_2) = (x_1 + y_1 + z_1, x_2 + y_2 + z_2) = (x_1, x_2) + ((y_1, y_2) + (z_1, z_2))$$

$$(x_1 + x_2) + (-x_1, -x_2) = (0, 0)$$

and hence, all properties of addition are satisfied. Similarly, we can verify for scalar multiplication:

$$\alpha(\beta(x_1, x_2)) = \alpha(\beta x_1, \beta x_2) = (\alpha \beta x_1, \alpha \beta x_2) = \beta(\alpha x_1, \alpha x_2) = \beta(\alpha x)$$

$$1(x_1, x_2) = (x_1, x_2)$$

and can verify the distributive laws. Thus, this space is a vector space. ■

**14.** Let  $Y$  be a subspace of a vector space  $X$ . The coset of an element  $x \in X$  with respect to  $Y$  is denoted by  $x + Y$  and is defined to be the set

$$x + Y = \{v \mid v = x + y, y \in Y\}.$$

Show that the distinct cosets form a partition of  $X$ . Show that under algebraic operations defined by

$$(w + Y) + (x + Y) = (w + x) + Y$$

$$\alpha(x + Y) = \alpha x + Y$$

these cosets constitute the elements of a vector space. This space is called the quotient space of  $X$  by  $Y$  and is denoted by  $X/Y$ . Its dimension is called the co-dimension of  $Y$  and is denoted by  $\text{codim } Y = \dim(X/Y)$ .

*Proof.* We show that no two cosets can have the same element, and if they do, then they are essentially the same coset. Let  $a \in X$  be some element which belongs to two different cosets:  $x + Y$  and  $w + Y$ . Thus, there exists  $y_x \in Y$  and  $y_w \in Y$  such that

$$a = x + y_x = w + y_w \implies x = a - y_x \text{ and } w = a - y_w.$$

Thus, we can write the cosets as

$$x + Y = \{a - y_x + y \mid y \in Y\}$$

$$w + Y = \{a + y_w - y \mid y \in Y\}.$$

Since  $Y$  is a subspace,  $-y_x + y$  for all  $y \in Y$  will always result in some vector belonging to  $Y$ , and similarly,  $-y_w + y$  will always result in some vector belonging to  $Y$ . Thus, as we circle through all elements of  $Y$ , the sets  $\{y - x + y \mid y \in Y\}$  and  $\{-y_w + y \mid y \in Y\}$  would be equal sets. Hence, the cosets  $x + Y$  and  $w + Y$  would be equal. Also, since  $X$  is a vector space, for any  $x \in X$  and  $y \in Y$ ,  $x - y$  has to belong to  $X$  and hence, to one of the cosets. This shows that the distinct cosets form partitions of  $X$ .

Consider the sets :

$$x + Y = \{x + y \mid y \in Y\},$$

$$w + Y = \{w + y \mid y \in Y\}.$$

Then, we have

$$(w + Y) + (x + Y) = \{w + y + x + y \mid y \in Y\},$$

but since  $2y \in Y$  (since  $Y$  is a subspace), denoting  $2y = y'$ , we have

$$(w + Y) + (x + Y) = \{w + x + y' \mid y' \in Y\} = (w + x) + Y.$$

Similarly,

$$\alpha(x + Y) = \{\alpha(x + y) \mid y \in Y\}.$$

However, since  $\alpha y \in Y$  (since  $Y$  is a subspace), denoting  $\alpha y = y'$ , we have

$$\alpha(x + Y) = \{\alpha x + y' \mid y' \in Y\} = \alpha x + Y.$$

This finishes the proof. ■

**15.** Let  $X = \mathbb{R}^3$  and  $Y = \{(\xi_1, 0, 0) \mid \xi_1 \in \mathbb{R}\}$ . Find  $X/Y$ ,  $X/X$ , and  $X/\{0\}$ .

*Proof.* Note that  $X/Y = \{(\xi_1, \xi_2, \xi_3) \mid \xi_2, \xi_3 \in \mathbb{R} \mid \xi_1 \in \mathbb{R}\}$ . This means that each coset is obtained by fixing the  $\xi_1$  coordinate and varying  $\xi_2$  and  $\xi_3$ . Thus, each of the cosets is a plane parallel to  $\xi_1$  axis. Now,  $X/X$  is simply  $\{0\}$  because for any  $x \in X$ , the coset  $x + X = \{x + x_0 \mid x_0 \in X\}$  is simply  $X$  and hence, there is no space of distinct cosets, i.e this quotient space is empty. Finally,  $X/\{0\}$  is simply  $X$  since each point in itself would be a coset, and the space of these cosets is  $X$  itself. ■

## 2.2 Normed Space. Banach Space.

**1.** Show that the norm  $\|x\|$  of  $x$  is the distance of  $x$  from 0.

*Proof.* By the definition of a norm, we have

$$d(x, y) = \|x - y\|.$$

Substituting  $y = 0$ , we get

$$\|x\| = d(x, 0).$$

**2.** Verify that the usual length of a vector in the plane or in three-dimensional space has the properties (N1) to (N4).

*Proof.* The properties referred to are:

1.  $\|x\| \geq 0$ .
2.  $\|x\| = 0 \iff x = 0$
3.  $\|\alpha x\| = |\alpha| \|x\|$ .
4.  $\|x + y\| \leq \|x\| + \|y\|$ .

Now, the length of vector  $x = (a, b, c)$  in three dimensions is given by

$$\|x\| = \sqrt{a^2 + b^2 + c^2}.$$

Clearly, this value is always non-negative. Also,

$$\sqrt{a^2 + b^2 + c^2} = 0 \implies a = b = c = 0.$$

Similarly,  $a = b = c$  gives us  $\|x\| = 0$ . Finally,

$$\|\alpha x\| = \sqrt{\alpha^2 a^2 + \alpha^2 b^2 + \alpha^2 c^2} = \sqrt{\alpha^2} \sqrt{a^2 + b^2 + c^2} = |\alpha| \|x\|.$$

We now show that the triangle inequality also holds. Let  $x = (a_1, b_1, c_1)$  and  $y = (a_2, b_2, c_2)$ . Then,

$$\|x + y\| = \sqrt{(a_1 + a_2)^2 + (b_1 + b_2)^2 + (c_1 + c_2)^2}.$$

Recall Hölder's inequality for  $a_i, b_i \geq 0$ :

$$\sum_{i=1}^n a_i b_i \leq \left( \sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \left( \sum_{i=1}^n b_i^q \right)^{\frac{1}{q}},$$

where  $p^{-1} + q^{-1} = 1$ . Using this, we get

$$a_1 a_2 + b_1 b_2 + c_1 c_2 \leq \sqrt{a_1^2 + b_1^2 + c_1^2} \sqrt{a_2^2 + b_2^2 + c_2^2}.$$

and hence,

$$\|x + y\| \leq \sqrt{a_1^2 + a_2^2 + b_1^2 + b_2^2 + c_1^2 + c_2^2 + 2\sqrt{a_1^2 + b_1^2 + c_1^2} \sqrt{a_2^2 + b_2^2 + c_2^2}} = \sqrt{a_1^2 + b_1^2 + c_1^2} + \sqrt{a_2^2 + b_2^2 + c_2^2},$$

and thus,

$$\|x + y\| \leq \|x\| + \|y\|.$$

**3. Prove (2).**

*Proof.* We wish to show

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Using the triangle inequality, we have

$$\|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\| \implies \|x\| - \|y\| \leq \|x - y\|.$$

Similarly,

$$\|y\| - \|x\| \leq \|y - x\| = \|x - y\|$$

, and hence, combining both inequalities results in

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

■

**4. Show that we may replace (N2) by**

$$\|x\| = 0 \implies x = 0$$

*without altering the concept of a norm. Show that non-negativity of a norm follows from (N3) and (N4).*

*Proof.* First, we wish to show that we can replace

$$\|x\| = 0 \iff x = 0$$

by

$$\|x\| = 0 \implies x = 0.$$

This means, we wish to show that the condition

$$x = 0 \implies \|x\| = 0$$

is redundant. Since we have already shown that  $\|x\| = d(x, 0)$ , substituting  $x = 0$  gives us the claim. Hence, this condition follows from the definition of a metric induced by a norm, and is hence, redundant. Another way to see this is: we know that  $\|\alpha x\| = |\alpha| \|x\|$ . Substituting  $\alpha = 0$  gives us

$$\|0\| = 0.$$

We now wish to show that we can prove the non-negativity of a norm using the following conditions:

$$\|\alpha x\| = |\alpha| \|x\|$$

$$\|x + y\| \leq \|x\| + \|y\|.$$

We have that

$$0 = \|x - x\| \leq \|x\| + |-1| \|x\| = 2 \|x\|$$

and hence,  $\|x\| \geq 0$ .

■

**5. Show that (3) defines a norm.**

*Proof.* We wish to show that

$$\|x\| = \left( \sum_{j=1}^n |\xi_j|^2 \right)^{1/2} = \sqrt{|\xi_1|^2 + \dots + |\xi_n|^2}$$

is a norm. We have already shown this in 2 for  $n = 3$ , and hence, the proof follows in exactly the same manner. We demonstrate the proof for triangle inequality once again. Let  $x = (\xi_1, \dots, \xi_n)$  and  $y = (\eta_1, \dots, \eta_n)$ . Then,

$$\begin{aligned}\|x + y\| &= \sqrt{\sum_{i=1}^n (\xi_i + \eta_i)^2} \\ &= \sqrt{\sum_{i=1}^n \xi_i^2 + \sum_{i=1}^n \eta_i^2 + 2 \sum_{i=1}^n \xi_i \eta_i} \\ &\leq \sqrt{\sum_{i=1}^n \xi_i^2 + \sum_{i=1}^n \eta_i^2 + 2 \sqrt{\sum_{i=1}^n \xi_i^2} \sqrt{\sum_{i=1}^n \eta_i^2}} \\ &= \sqrt{\sum_{i=1}^n \xi_i^2} + \sqrt{\sum_{i=1}^n \eta_i^2} \\ &= \|x\| + \|y\|\end{aligned}$$

where the inequality makes use of Hölder's inequality. ■

**6.** Let  $X$  be the vector space of all ordered pairs  $x = (\xi_1, \xi_2)$ ,  $y = (\eta_1, \eta_2)$  of real numbers. Show that norms on  $X$  are defined by

$$\begin{aligned}\|x\|_1 &= |\xi_1| + |\xi_2| \\ \|x\|_2 &= \sqrt{|\xi_1|^2 + |\xi_2|^2} \\ \|x\|_\infty &= \max\{|\xi_1|, |\xi_2|\}.\end{aligned}$$

*Proof.* We first show  $\|x\|_1$  is a valid norm. Clearly, it is always non-negative, and since it is the sum of non-negative terms, it is zero iff each of the terms is zero, resulting in the condition  $\|x\|_1 = 0 \iff x = 0$ . Now, we have

$$\|\alpha x\|_1 = |\alpha \xi_1| + |\alpha \xi_2| = |\alpha| |\xi_1| + |\alpha| |\xi_2| = |\alpha| \|x\|_1.$$

Also, using the triangle inequality for the absolute values on real numbers gives:

$$\|x + y\|_1 = |\xi_1 + \eta_1| + |\xi_2 + \eta_2| \leq |\xi_1| + |\xi_2| + |\eta_1| + |\eta_2| = \|x\|_1 + \|y\|_1.$$

Thus,  $\|x\|_1$  is a valid norm.

We have shown  $\|x\|_2$  is a valid norm for three-dimensional points in 3 and for  $n$ -dimensional points in 5.

We now show  $\|x\|_\infty$  is also a valid norm. First, by definition it is always non-negative, and the maximum of two non-negative terms can be zero iff each of the terms is zero, resulting in the condition  $\|x\|_\infty = 0 \iff x = 0$ . Further,

$$\|\alpha x\|_\infty = \max\{|\alpha \xi_1|, |\alpha \xi_2|\} = |\alpha| \max\{|\xi_1|, |\xi_2|\} = |\alpha| \|x\|_\infty.$$

Finally, the triangle inequality follows since:

$$\|x + y\|_\infty = \max\{|\xi_1 + \eta_1|, |\xi_2 + \eta_2|\} = |\xi_i + \eta_i| \leq |\xi_i| + |\eta_i| \leq \max\{|\xi_1|, |\xi_2|\} + \max\{|\eta_1|, |\eta_2|\} = \|x\|_\infty + \|y\|_\infty$$

Thus,  $\|x\|_\infty$  is a valid norm. ■

**7.** Verify that (4) satisfies (N1) to (N4).

*Proof.* We wish to show that the  $\ell_p$  norm, defined as

$$\|x\|_p = \left( \sum_{j=1}^n |\xi_j|^p \right)^{1/p}$$

is a valid norm. First, it is easy to see that it is the sum of non-negative terms, and hence, is always non-negative. Also, it would only be zero if each of the terms is zero, resulting in the condition that  $\|x\|_p = 0 \iff x = 0$ . Further,

$$\|\alpha x\|_p = \left( \sum_{j=1}^n |\alpha \xi_j|^p \right)^{1/p} = \left( |\alpha|^p \sum_{j=1}^n |\xi_j|^p \right)^{1/p} = |\alpha| \left( \sum_{j=1}^n |\xi_j|^p \right)^{1/p} = |\alpha| \|x\|_p.$$

Finally, the triangle inequality follows from Minkowski's inequality:

$$\|x + y\|_p = \left( \sum_{j=1}^n |\xi_j + \eta_j|^p \right)^{1/p} \leq \left( \sum_{j=1}^n |\xi_j|^p \right)^{1/p} + \left( \sum_{j=1}^n |\eta_j|^p \right)^{1/p} = \|x\|_p + \|y\|_p$$

8. Show that these norms on the vector space of ordered  $n$ -tuples of numbers are valid norms.

$$\begin{aligned} \|x\|_1 &= |\xi_1| + \dots + |\xi_n| \\ \|x\|_2 &= \sqrt{|\xi_1|^2 + \dots + |\xi_n|^2} \\ \|x\|_\infty &= \max\{|\xi_1|, \dots, |\xi_n|\} \end{aligned}$$

*Proof.* We first begin with  $\|x\|_1$ , which is clearly non-negative and can only be zero iff  $x = 0$ . Also, using the fact that  $|\alpha x| = |\alpha| |x|$ , we have

$$\|\alpha x\|_1 = \sum_{i=1}^n |\alpha \xi_i| = \sum_{i=1}^n |\alpha| |\xi_i| = |\alpha| \|x\|_1.$$

Finally, the triangle inequality follows from the triangle inequality for real numbers:

$$\|x + y\|_1 = \sum_{i=1}^n |\xi_i + \eta_i| \leq \sum_{i=1}^n |\xi_i| + \sum_{i=1}^n |\eta_i| = \|x\|_1 + \|y\|_1.$$

Thus,  $\|x\|_1$  is a valid norm.

We have already shown  $\|x\|_2$  is a valid norm in 5.

We now show that  $\|x\|_\infty$  is a valid norm. Clearly, it is non-negative, and the maximum of non-negative terms can be zero iff each of the terms is zero, thus resulting in the condition  $\|x\| = 0 \iff x = 0$ . Further,

$$\|\alpha x\|_\infty = \max_{i \in [n]} |\alpha \xi_i| = |\alpha| \max_{i \in [n]} |\xi_i| = |\alpha| \|x\|_\infty.$$

The triangle inequality follows since:

$$\|x + y\|_\infty = \max_{i \in [n]} |\xi_i + \eta_i| = |\xi_j + \eta_j| \leq |\xi_j| + |\eta_j| \leq \max_{i \in [n]} |\xi_i| + \max_{i \in [n]} |\eta_i| = \|x\|_\infty + \|y\|_\infty.$$

Thus,  $\|x\|_\infty$  is a valid norm.

9. Verify that (5) defines a norm.

*Proof.* We wish to show that

$$\|x\| = \max_{t \in J} |x(t)|$$

defines a norm. By definition, this is non-negative, and the only way  $\|x\| = 0$  is if for each  $t \in J$ ,  $x(t) = 0$ , or in other words,  $x = 0$ . Further,

$$\|\alpha x\| = \max_{t \in J} |\alpha x(t)| = |\alpha| \max_{t \in J} |x(t)| = |\alpha| \|x\|.$$

The triangle inequality follows since:

$$\|x + y\| = \max_{t \in J} |x(t) + y(t)| \leq |x(t_0) + y(t_0)| \leq |x(t_0)| + |y(t_0)| \leq \max_{t \in J} |x(t)| + \max_{t \in J} |y(t)| = \|x\| + \|y\|.$$

Thus, it is a valid norm.

10. The sphere

$$S(0; 1) = \{x \in X \mid \|x\| = 1\}$$

in a normed space  $X$  is called the unit sphere. Show that for the norms in 6 and for the norm defined by

$$\|x\|_4 = (\xi_1^4 + \xi_2^4)^{1/4}$$

The unit spheres look as shown in Fig.16.

*Proof.* It is not clear on how to show this. We will attempt to show it in the following way:  $\|x\|_1$ ,  $\|x\|_2$ , and  $\|x\|_\infty$  in  $\mathbb{R}^2$  depict a diamond, circle, and square, respectively. Then, we wish to show that if  $x_1$  satisfies  $\|x_1\|_2 = 1$ ,  $x_2$  satisfies  $\|x_2\|_4 = 1$ , and  $x_3$  satisfies  $\|x_3\|_\infty = 1$ , then ,

$$\|x_1\|_2 \leq \|x_2\|_2 \leq \|x_3\|_\infty$$

, or in other words, the distance from the origin of these points increases.

First,

$$\|x\|_1 = 1 \implies |\xi_1| + |\xi_2| = 1.$$

This results in the following 4 lines:  $\pm\xi_1 \pm \xi_2 = 1$ , which results in the diamond. Next,

$$\|x\|_2 = 1 \implies \xi_1^2 + \xi_2^2 = 1$$

results in the equation of a circle centered around the origin with radius 1. Finally,

$$\|x\|_\infty = 1 \implies \max\{|\xi_1|, |\xi_2|\} = 1$$

results in a square. Now, for  $\|x_3\|_\infty = 1$ , any point of the form  $(\pm 1, t)$  or  $(t, \pm 1)$  works, where  $t \in [-1, 1]$  and thus,  $\|x_3\|_2 \leq \sqrt{1+1} = \sqrt{2}$  and  $\|x_3\|_2 \geq \sqrt{1} = 1$ . Finally, we see that  $x_2 = (\sqrt{\cos\theta}, \sqrt{\sin\theta})$  satisfies  $\|x_2\|_4 = 1$ , and hence,

$$\|x_2\|_2 = \sqrt{\cos\theta + \sin\theta} = \sqrt{\sqrt{2} \sin\left(\frac{\pi}{4} + \theta\right)} \leq 2^{1/4}$$

and hence. we have

$$\|x_1\|_2 \leq \|x_2\|_2 \leq \|x_3\|_\infty$$

, which shows that the distance of points which satisfy  $\|\cdot\|_4 = 1$  is more than that of points which satisfy  $\|\cdot\|_2 = 1$  but less than that of  $\|\cdot\|_\infty = 1$ . ■

**11.** A subset  $A$  of vector space  $X$  is said to be convex if  $x, y \in A$  implies

$$M = \{z \in X \mid z = \alpha x + (1 - \alpha)y, 0 \leq \alpha \leq 1\} \subset A.$$

$M$  is called a closed segment with boundary points  $x$  and  $y$ , any other  $z \in M$  is called an interior point of  $M$ . Show that the closed unit ball

$$\tilde{B}(0, 1) = \{x \in X \mid \|x\| \leq 1\}$$

in a normed space is convex.

*Proof.* Choose two points  $x, y \in \tilde{B}(0, 1)$  such that  $\|x\| \leq 1$  and  $\|y\| \leq 1$ . Then, the norm of any convex combination of  $x$  and  $y$  is given by

$$\|\alpha x + (1 - \alpha)y\| \leq |\alpha| \|x\| + |1 - \alpha| \|y\| \leq 1$$

which shows that any convex combination of the points lies inside the closed ball. Here, the first inequality follows by using the triangle inequality and scalar multiplication property of the norm in tandem. ■

**12.** Using 11, show that

$$\varphi(x) = \left( \sqrt{|\xi_1|} + \sqrt{|\xi_2|} \right)^2$$

does not define a norm on the vector space of all ordered pairs of real numbers.

*Proof.* In 11 we have shown that the closed ball of radius 1 is a convex set with respect to any norm. Suppose,  $\varphi(x)$  defines a norm such that  $\varphi(x) \leq 1$  and  $\varphi(y) \leq 1$ . Then, we have

$$\begin{aligned} \varphi(\alpha x + (1 - \alpha)y) &= \left( \sqrt{|\alpha\xi_1 + (1 - \alpha)\eta_1|} + \sqrt{|\alpha\xi_2 + (1 - \alpha)\eta_2|} \right)^2 \\ &\leq \left( \sqrt{\alpha|\xi_1| + (1 - \alpha)|\eta_1|} + \sqrt{\alpha|\xi_2| + (1 - \alpha)|\eta_2|} \right)^2 \\ &= \alpha(|\xi_1| + |\xi_2|) + (1 - \alpha)(|\eta_1| + |\eta_2|) + 2\sqrt{(\alpha|\xi_1| + (1 - \alpha)|\eta_1|)(\alpha|\xi_2| + (1 - \alpha)|\eta_2|)} \end{aligned}$$

Now, since  $\phi(x) = (|\xi_1| + |\xi_2|) + s\sqrt{|\xi_1\xi_2|} \leq 1$ , we have that  $(|\xi_1| + |\xi_2|) \leq 1$ . Similarly,  $(|\eta_1| + |\eta_2|) \leq 1$  and hence, we get

$$\varphi(\alpha x + (1 - \alpha)y) \leq 1 + 2\sqrt{(\alpha|\xi_1| + (1 - \alpha)|\eta_1|)(\alpha|\xi_2| + (1 - \alpha)|\eta_2|)}$$

Since each of the terms inside the square root are non-negative, we have that the the additional quantity is non-negative, and hence,  $\varphi(\alpha x + (1 - \alpha)y)$  may be greater than 1, leading to the fact that the subset is no longer convex. ■

**13.** Show that the discrete metric on a vector space  $X \neq \{0\}$  cannot be obtained from a norm.

*Proof.* Let  $d$  be a metric induced by a norm. Then,  $d$  satisfies the following:

$$d(x + a, y + a) = \|x - y\| = d(x, y),$$

$$d(\alpha x, \alpha y) = \|\alpha x - \alpha y\| = |\alpha| \|x - y\| = |\alpha| d(x, y).$$

Suppose  $x \neq y$ . Then,  $d(x, y) = 1$ . However, for  $\alpha \neq \pm 1$ ,  $d(x, y) \neq d(\alpha x, \alpha y)$ , and hence, the metric is not induced from a norm. ■

**14.** If  $d$  is a metric on a vector space  $X \neq \{0\}$  obtained from a norm, and  $\tilde{d}$  is defined by

$$\tilde{d}(x, x) = 0, \quad \tilde{d}(x, y) = d(x, y) + 1 \text{ for } x \neq y.$$

Show that  $\tilde{d}$  cannot be obtained from a norm.

*Proof.* Suppose  $\tilde{d}$  was induced from a norm. Then,

$$\tilde{d}(\alpha x, \alpha y) = d(\alpha x, \alpha y) + 1 = |\alpha| d(x, y) + 1.$$

However, we also have

$$\tilde{d}(\alpha x, \alpha y) = |\alpha| \tilde{d}(x, y) = |\alpha| d(x, y) + |\alpha|.$$

Thus, if  $\alpha \neq \pm 1$ , then,  $\tilde{d}$  is not a metric induced by a norm. Since, this should hold for all values of  $\alpha$ ,  $\tilde{d}$  is not a metric induced by a norm. ■

**15.** Show that a subset  $M$  in a normed space  $X$  is bounded if and only if there is a positive number  $c$  such that  $\|x\| \leq c$  for every  $x \in M$ .

*Proof.* Recall that a subset  $M$  is bounded if

$$\delta(M) = \sup_{x, y \in M} d(x, y) < \infty.$$

Suppose the set is bounded. Then,  $\sup_{M \times M} d(x, y) < \infty$ . Let  $d(x, y) < c$  for all pairs  $(x, y) \in M \times M$ . Then,

$$\|x\| \leq \|x - y\| + \|y\| \leq c + \|y\|$$

which is bounded above by some other constant  $c'$ .

On the other hand, we have

$$d(x, y) = \|x - y\| \leq \|x\| + \|y\| \leq 2c$$

and hence,

$$\sup_{(x, y) \in M \times M} d(x, y) = 2c.$$

This finishes the proof. ■

## 2.3 Further properties of normed spaces

**1.** Show that  $c \subset \ell^\infty$  is a vector subspace of  $\ell^\infty$  and so is  $c_0$ , the space of all sequences of scalars converging to zero.

*Proof.* Recall that  $c \subset \ell^\infty$  is the space of all convergent sequences  $(x_n)$  with the metric induced from the  $\ell^\infty$  space. To show that  $c$  is a vector space, we need to ensure that all properties of addition and multiplication hold. First, for addition we wish to ensure that the sum of two convergent sequences is convergent, i.e if  $x_n \rightarrow x$  and  $y_n \rightarrow y$ , then

$$0 \leq \lim_{n \rightarrow \infty} \|(x_n + y_n) - (x + y)\| \leq \lim_{n \rightarrow \infty} \|x_n - x\| + \lim_{n \rightarrow \infty} \|y_n - y\| = 0$$

and hence,  $x_n + y_n \rightarrow x + y$ . Clearly, the commutative and associative properties of addition hold. Also, the additive identity is simply the constant zero sequence, which is convergent.

For multiplication, if  $x_n \rightarrow x$ , then

$$\lim_{n \rightarrow \infty} \|\alpha x_n - \alpha x\| = |\alpha| \lim_{n \rightarrow \infty} \|x_n - x\| = 0$$

and hence,  $\alpha x_n \rightarrow \alpha x$ . The associativity for multiplication holds, and the multiplicative identity is simply the scalar 1.

In a similar fashion, if  $x_n \rightarrow 0$  and  $y_n \rightarrow 0$ , then  $x_n + y_n \rightarrow 0$ , and the addition property satisfies associativity and commutativity. Further, the zero vector is the additive identity, and since the zero vector converges to 0, it belongs to  $c_0$ . The multiplication of a sequence  $x_n \rightarrow 0$  does not change any convergence properties, and the multiplicative identity continues to remain the scalar 1. ■

2. Show that  $c_0$  is a closed subspace of  $\ell^\infty$ , so that  $c_0$  is complete.

*Proof.* Let  $(x_1, x_2, \dots)$  be a sequence converging to  $x = (\xi_1, \xi_2, \dots)$ . Each term  $x_i = (\xi_1^{(i)}, \xi_2^{(i)}, \dots)$  is itself a sequence that converges to zero. Thus, we have that for some  $\epsilon > 0$ ,  $\exists N_1(\epsilon)$  such that for  $n > N_1(\epsilon)$

$$d(x_n, 0) < \epsilon.$$

Now, since  $(x_1, x_2, \dots)$  also converges to  $x$  under the metric induced by the  $\ell^\infty$  space, so we have that for some  $\epsilon > 0$ ,  $\exists N_2(\epsilon)$  such that for  $n > N_2(\epsilon)$ , we have

$$d(x_n, x) = \sup_i |\xi_i^{(n)} - \xi_i| < \epsilon \implies |\xi_i^{(n)} - \xi_i| < \epsilon \forall i.$$

We wish to show that the limit point  $x = (\xi_1, \xi_2, \dots)$  also converges to zero, so that  $x \in c_0$ . Thus, we have that for some  $\epsilon > 0$  and  $N(\epsilon) = \max\{N_1(\epsilon), N_2(\epsilon)\}$ , for all  $n > N(\epsilon)$ , we have

$$d(\xi_i, 0) \leq d(\xi_i, \xi_i^{(n)}) + d(\xi_i^{(n)}, 0) \leq 2\epsilon.$$

Thus,  $x \in c_0$ , and hence the space is closed, which in turn also means, the space is complete. ■

3. In  $\ell^\infty$ , let  $Y$  be the subset of all sequences with only finitely many non-zero terms. Show that  $Y$  is a subspace of  $\ell^\infty$  by not a closed subspace.

*Proof.* Consider the sequence of sequences  $(x_1, x_2, \dots)$  where

$$x_i = \left(1, \frac{1}{2}, \dots, \frac{1}{i}, 0, 0, \dots\right)$$

Then, this sequence converges to

$$\left(1, \frac{1}{2}, \dots, \frac{1}{n}, \frac{1}{n+1}, \dots\right)$$

which does not belong to  $Y$ . ■

4. Show that in a normed space  $X$ , vector addition and multiplication by scalars are continuous operations with respect to the norm, i.e., the mappings defined by  $(x, y) \mapsto x + y$  and  $(\alpha, x) \mapsto \alpha x$  are continuous.

*Proof.* Recall the definition of continuity of a map  $T : X \mapsto Y$  at a point  $x_0$ : for some  $\epsilon > 0$ ,  $\exists \delta > 0$  such that  $\|x - x_0\| < \delta \implies \|Tx - Tx_0\| < \epsilon$ .

First, for addition, consider the map  $T : X \times X \mapsto X$  such that  $T(x, y) = x + y$ . Then, we show that

$$\|(x, y) - (x_0, y_0)\| < \delta \implies \|T(x, y) - T(x_0, y_0)\| < \epsilon.$$

First, we have

$$\|(x, y) - (x_0, y_0)\| = \|(x - x_0, y - y_0)\| < \|x - x_0\| + \|y - y_0\| < \delta.$$

Then,

$$\|T(x, y) - T(x_0, y_0)\| = \|x + y - x_0 - y_0\| \leq \|x - x_0\| + \|y - y_0\| < \delta.$$

Thus, choosing  $\delta = \epsilon$  shows that addition is a continuous map.

Now, consider the map  $T : K \times X \mapsto X$  such that  $T(\alpha, x) = \alpha x$ . Now, we show that

$$\|(\alpha, x) - (\alpha_0, x_0)\| < \delta \implies \|T(\alpha, x) - T(\alpha_0, x_0)\| < \epsilon.$$

First, we have that

$$\|(\alpha, x) - (\alpha_0, x_0)\| = \|(\alpha - \alpha_0, x - x_0)\| = \|\alpha - \alpha_0\| + \|x - x_0\| < \delta$$

Now,

$$\|T(\alpha, x) - T(\alpha_0, x_0)\| = \|\alpha x - \alpha_0 x_0\| = \|(\alpha - \alpha_0)x + \alpha_0(x - x_0)\| \leq \|\alpha - \alpha_0\| \|x\| + \|\alpha_0\| \|x - x_0\|.$$

Now, setting  $\|\alpha - \alpha_0\| < \delta$  and  $\|x - x_0\| < \delta$ , we get

$$\|T(\alpha, x) - T(\alpha_0, x_0)\| < \delta(\|x_0\| + \delta) + \|\alpha_0\| \delta < \epsilon.$$

Thus, assuming  $\delta < 1$ , we get

$$\delta^2 + \delta(\|x_0\| + \|\alpha_0\|) < \delta(\|x_0\| + \|\alpha_0\| + 1) < \epsilon,$$

and solving for  $\delta$  gives us

$$\delta = \min \left\{ \frac{\epsilon}{2(\|x_0\| + \|\alpha_0\| + 1)}, 1 \right\}.$$

Thus, the multiplication map is also continuous. ■



5. Show that  $x_n \rightarrow x$  and  $y_n \rightarrow y$  implies  $x_n + y_n \rightarrow x + y$ . Show that  $\alpha_n \rightarrow \alpha$  and  $x_n \rightarrow x$  implies  $\alpha_n x_n \rightarrow \alpha x$ .

*Proof.* Since  $x_n \rightarrow x$  and  $y_n \rightarrow y$ , we have that

$$\lim_{n \rightarrow \infty} \|x_n - x\| = 0 \text{ and } \lim_{n \rightarrow \infty} \|y_n - y\| = 0,$$

and hence,

$$0 \leq \lim_{n \rightarrow \infty} \|(x_n + y_n) - (x + y)\| \leq \lim_{n \rightarrow \infty} \|x_n - x\| + \lim_{n \rightarrow \infty} \|y_n - y\| = 0$$

and hence,  $x_n + y_n \rightarrow x + y$ . Similarly, since  $\alpha_n \rightarrow \alpha$  and  $x_n \rightarrow x$ , we have

$$\lim_{n \rightarrow \infty} |\alpha_n - \alpha| = 0 \text{ and } \lim_{n \rightarrow \infty} \|x_n - x\| = 0,$$

since  $\alpha$  is a part of the scalar field. Thus,

$$\lim_{n \rightarrow \infty} \|\alpha_n x_n - \alpha x\| = \lim_{n \rightarrow \infty} \|\alpha_n(x_n - x)\| + \lim_{n \rightarrow \infty} \|(\alpha_n - \alpha)x\| = \lim_{n \rightarrow \infty} |\alpha_n| \|x_n - x\| + \lim_{n \rightarrow \infty} |\alpha_n - \alpha| \|x\| = 0.$$

6. Show that the closure  $\bar{Y}$  of a subspace  $Y$  of a normed space  $X$  is again a vector space.

*Proof.* First, since  $Y$  is a subspace, for any  $x, y \in Y$ , we have

$$\alpha x + \beta y \in Y$$

. Now, consider  $\bar{Y}$ , which consists of  $Y$  and the accumulation points of  $Y$ . If the accumulation points of  $Y$  belong to  $Y$  itself, we are done. However, say  $x$  and  $y$  are two accumulation points which do not belong to  $Y$ . Then, there exist sequences  $(x_n)$  and  $(y_n)$  in  $Y$  that converge to  $x$  and  $y$  respectively. From 5, we have that  $\alpha x_n + \beta y_n \rightarrow \alpha x + \beta y$  and hence,  $\alpha x + \beta y$  is also an accumulation point. Thus,  $\alpha x + \beta y$  also belongs to the set of accumulation points. One question that should arise is what if one point  $y \in Y$  and  $x \notin Y$  but  $x$  is an accumulation point. In such a case, the constant sequence  $(y, y, \dots) \rightarrow y$  and another sequence  $x_n \rightarrow x$ , and hence,  $\alpha x + \beta y$  would also be an accumulation point. Thus, the set  $\bar{Y}$  is also a subspace. ■

7. Show that the convergence of  $\|y_1\| + \|y_2\| + \|y_3\| \dots$  may not imply convergence of  $y_1 + y_2 + y_3 + \dots$ .

*Proof.* We shall make use of the hint given in the question. Consider the space in 3, a space of sequences here each sequence consists of finitely many non-zero terms. Consider the sequence  $(y_1, y_2, \dots)$ , where  $y_n = (\eta_j^{(n)})$  and

$$\eta_j^{(n)} = \begin{cases} \frac{1}{n^2} & j = n \\ 0 & j \neq n \end{cases}.$$

In such a case, we know that

$$\sum_{i=1}^{\infty} \|y_i\| = \sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}.$$

However, consider the series  $y_1 + y_2 + y_3 + \dots$ . Consider the partial sum  $s_n$ :

$$s_n = \left(1, \frac{1}{4}, \dots, \frac{1}{n^2}, 0, 0, \dots\right).$$

This sequence of partial sums will converge to a term which will not have finitely many non-zero terms and hence, does not converge in the space. ■

8. If in a normed space  $X$ , if absolute convergence of any series always implies convergence of that series, show that  $X$  is complete.

*Proof.* Let  $(x_n)$  be a sequence that converges absolutely, and hence, also converges. Since it converges absolutely, we know that the partial sums of the norms converge, and hence, the sequence of the partial sums of the norms is Cauchy. Let  $s_n$  denote the  $n^{th}$  sequence of partial sums. Then, for some  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for  $m, n > N(\epsilon)$ , we have

$$\|s_m - s_n\| = \left\| \sum_{i=n+1}^m \|x_i\| \right\| \leq \sum_{i=n+1}^m \|x_i\| \leq \epsilon.$$

Note that the final expression is exactly the partial sums of the sequence  $(x_1, x_2, \dots)$ ,

$$\|s_m - s_n\| = \left\| \sum_{i=1}^n x_i - \sum_{i=1}^m x_i \right\| = \left\| \sum_{i=n+1}^m x_i \right\| < \sum_{i=n+1}^m \|x_i\| < \epsilon$$

and hence, the sequence of partial sums is Cauchy. Since absolute convergence implies convergence, the Cauchy sequence will converge if and only if the space is complete. ■

**9.** Show that in a Banach space, an absolutely convergent series is convergent.

*Proof.* The answer to this is similar to the last one. A Banach space is a complete normed space. Suppose  $(x_n)$  is an absolutely convergent series. Then, the sequence of partial sums of the norm is convergent, and hence, Cauchy. Thus, representing the partial sum as  $s_n$ , we have

$$\|s_m - s_n\| = \left\| \sum_{i=n+1}^m \|x_i\| \right\| \leq \sum_{i=n+1}^m \|x_i\| \leq \epsilon.$$

This is exactly the difference of the partial sums of the series  $(x_n)$  and hence, shows that the sequence of partial sums of  $(x_n)$  is Cauchy. Since it is a Banach space, Cauchy sequences converge, and hence, the sequence of partial sums converges, and hence, the series  $(x_n)$  ■

**10.** Show that if a normed space has a Schauder basis, it is separable.

*Proof.* Recall the definition of a Schauder basis: a sequence  $(e_n)$  is called a Schauder basis if for every  $x \in X$ , there exists a unique sequence of scalars  $\alpha_i$  such that

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \alpha_i e_i = x.$$

To show that  $X$  is separable, we need to show it has a countable dense subset. Let  $Y$  be a subset. Let  $x \in Y$  be such that

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \alpha_i e_i = x.$$

Now, since  $\alpha_i$  belong to the underlying field  $K$ , and hence, there exists a sequence  $\beta_i$  such that

$$|\alpha_i - \beta_i| < \epsilon.$$

Let us denote some point  $y$  such that

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \beta_i e_i = y.$$

Then, we have that for some  $\epsilon > 0$ ,  $\exists N(\epsilon)$  such that for  $n > N(\epsilon)$ , we have

$$\begin{aligned} \|x - y\| &\leq \left\| x - \sum_{i=1}^n \alpha_i e_i \right\| + \left\| y - \sum_{i=1}^n \beta_i e_i \right\| + \left\| \sum_{i=1}^n (\alpha_i - \beta_i) e_i \right\| \\ &\leq \epsilon + \epsilon + n\epsilon \|e_i\| \\ &\leq \epsilon' \end{aligned}$$

and hence, for every  $x$  there exists some point  $y$  such that  $x \in B(y, \epsilon)$ . Thus, the subset  $Y$  is dense. ■

**11.** Show that  $(e_n)$  where  $e_n = (\delta_{nj})$ , is a Schauder basis for  $\ell^p$ , where  $1 \leq p < \infty$ .

*Proof.* Let  $x = (\xi_1, \xi_2, \dots) \in \ell^p$  such that

$$\|x\|_p = \left( \sum_{i=1}^{\infty} |\xi_i|^p \right)^{1/p} < \infty.$$

In other words,  $(\sum_{i=1}^{\infty} |\xi_i|^p)^{1/p}$  converges, and hence, for some  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for all  $n > N(\epsilon)$ , we have

$$\left( \sum_{i=n+1}^{\infty} |\xi_i|^p \right)^{1/p} < \epsilon.$$

Now, consider the sequence  $(\delta_{nj})$  with the coefficients  $\xi_1, \xi_2 \dots$ . Represent

$$x_n = \sum_{i=1}^n \xi_i \delta_{ni}.$$

Then, we have

$$\|x - x_n\| = \left\| \sum_{i=n+1}^{\infty} \xi_i \delta_{ni} \right\| \leq \sum_{i=n+1}^{\infty} |\xi_i| \|\delta_{ni}\| \leq \epsilon.$$

and hence,  $x_n \rightarrow x$ . Thus, it is a Schauder basis. ■

**12.** A seminorm on a vector space  $X$  is a mapping  $p : X \mapsto \mathbb{R}$  satisfying (N1), (N3), (N4). Show that  $p(0) = 0$  and  $|p(y) - p(x)| < p(y - x)$ . This also means that if  $p(x) = 0 \implies x = 0$ , then it is a norm.

*Proof.* We are given that  $p$  satisfies the following properties:

1.  $p \geq 0$
2.  $p(\alpha x) = |\alpha| p(x)$ .
3.  $p(x + y) \leq p(x) + p(y)$ .

Now, substituting  $\alpha = 0$  gives us  $p(0) = 0$ . Also, we have

$$p(x) \leq p(x - y) + p(y) \implies p(x) - p(y) \leq p(x - y).$$

Similarly, we have

$$p(y) \leq p(y - x) + p(x) = |-1| p(x - y) + p(x) \implies p(y) - p(x) \leq p(x - y).$$

Combining both these results, we get

$$|p(x) - p(y)| \leq p(x - y).$$
■

**13.** Show that in 12, the elements  $x \in X$  such that  $p(x) = 0$  form a subspace  $N$  of  $X$  and a norm on  $X/N$  is defined by  $\|\hat{x}\|_0 = p(x)$  where  $x \in \hat{x}$  and  $\hat{x} \in X/N$ .

*Proof.* Let  $x, y \in N$  such that  $p(x) = p(y) = 0$ . We wish to show that  $\alpha x + \beta y \in N$ . This will only happen if  $p(\alpha x + \beta y) = 0$ .

$$p(\alpha x + \beta y) \leq p(\alpha x) + p(\beta y) = |\alpha| p(x) + |\beta| p(y) = 0$$

and thus,  $\alpha x + \beta y \in N$ . Hence,  $N$  is a subset. Recall the definitions of a coset and a quotient space as follows: for some element  $x$ , we define the coset w.r.t  $Y$  as

$$x + Y = \{x + y \mid y \in Y\}.$$

These distinct cosets partition  $X$ , and the space of these cosets is called the quotient space and is represented by  $X/Y$ .

Now, to show that  $\|\hat{x}\|_0 = p(x)$  is a norm on  $X/N$ , we have to show that  $p(x)$  satisfies all properties of a norm on  $X/N$ . From 13, we have that  $p(x)$  satisfies non-negativity, scalar multiplication, and the triangle inequality. We have also shown that  $p(0) = 0$ . Thus, to show that  $p(x)$  is a norm, we wish to show that  $p(x) = 0 \implies x = 0$ . Let  $\hat{x}$  define a coset and  $x \in \hat{x}$ . Then,

$$\hat{x} = \{x + n \mid n \in N\}.$$

We first show that  $p(x) = p(x + n)$  for any  $n \in N$ . We have that

$$p(x) = p((x + n) - n) \leq p(x + n) + p(n) \leq p(x) + p(n) = p(x)$$

Thus, all the inequalities have to be equalities and we get

$$p(x) = p(x + n) + p(n) = p(x + n).$$

Thus,  $p(x) = 0 \implies x \in N$ . But,  $x \in \hat{x}$ , which means that  $\hat{x} = N$ , and hence,  $\hat{x}$  is the zero element of  $X/N$ . This is easy to see. Assume  $\hat{y} = y + N$  and  $\hat{x} = x + N$ , where  $x \in N$ . Then,  $\hat{x} + \hat{y} = (x + y) + N = y + N$  since  $x + N = N$ . ■

14. Let  $Y$  be a closed subspace of a normed space  $(X, \|\cdot\|)$ . Show that a norm  $\|\cdot\|$  on  $X/Y$  is defined by

$$\|\hat{x}\|_0 = \inf_{x \in \hat{x}} \|x\|$$

where  $\hat{x} \in X/Y$ , i.e.,  $\hat{x}$  is any coset of  $Y$ .

*Proof.* Clearly,  $\|\hat{x}\|_0$  is non-negative. Also, if  $\hat{x} = Y$ , which is the zero element of  $X/Y$ , then

$$\|\hat{x}\|_0 = \|Y\|_0 = \inf_{y \in Y} \|y\| = 0$$

since  $Y$  is a closed subspace,  $0 \in Y$ . Also, if  $\|\hat{x}\|_0 = 0$ , then

$$\|\hat{x}\|_0 = 0 = \inf_{x \in \hat{x}} \|x\| \implies \exists x \in \hat{x} \text{ such that } \|x\| = 0 \implies x = 0 \implies \hat{x} = Y.$$

where the second to last implication follows since  $\|\cdot\|$  is a valid norm. Now,

$$\|\alpha \hat{x}\|_0 = \inf_{x \in \alpha \hat{x}} \|x\| = \inf_{x \in \hat{x}} \|\alpha x\| = |\alpha| \inf_{x \in \hat{x}} \|x\| = |\alpha| \|\hat{x}\|_0.$$

Finally, suppose  $\hat{x} = x + Y$  and  $\hat{y} = y + Y$ , then

$$\hat{x} + \hat{y} = (x + y) + Y.$$

Thus,

$$\|\hat{x} + \hat{y}\|_0 = \inf_{a \in \hat{x} + \hat{y}} \|a\| = \inf_{x \in \hat{x}, y \in \hat{y}} \|x + y\| \leq \inf_{x \in \hat{x}, y \in \hat{y}} \|x\| + \inf_{x \in \hat{x}, y \in \hat{y}} \|y\| = \inf_{x \in \hat{x}} \|x\| + \inf_{y \in \hat{y}} \|y\| = \|\hat{x}\|_0 + \|\hat{y}\|_0.$$

Thus, it is a valid norm. ■

15. If  $(X_1, \|\cdot\|_1)$  and  $(X_2, \|\cdot\|_2)$  are normed spaces, show that the product vector space  $X = X_1 \times X_2$  becomes a normed space if

$$\|x\| = \max\{\|x_1\|_1, \|x_2\|_2\}.$$

*Proof.* Clearly, for  $x = (x_1, x_2)$ ,  $\|x\| \geq 0$ . Also, if  $x = (0, 0)$ , then  $\|x\| = 0$ . Also, if  $\|x\| = 0$ , then both  $\|x_1\|_1 = 0$  and  $\|x_2\|_2 = 0$ , resulting in  $x = 0$ . Now,

$$\|\alpha x\| = \max\{\|\alpha x_1\|_1, \|\alpha x_2\|_2\} = |\alpha| \max\{\|x_1\|_1, \|x_2\|_2\} = |\alpha| \|x\|.$$

Finally,

$$\|x + y\| = \max\{|x_1 + y_1|, |x_2 + y_2|\} = |x_j + y_j| \leq |x_j| + |y_j| \leq \max\{|x_1|, |x_2|\} + \max\{|y_1|, |y_2|\} = \|x\| + \|y\|$$

Thus, it is a valid norm. ■

## 2.4 Finite Dimensional Normed Spaces

1. Give examples of subspaces of  $\ell^\infty$  and  $\ell^2$  which are not closed.

*Proof.* Consider the subspaces with finitely many non-zero terms. For example, take the sequence

$$x_n = \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, 0, \dots\right)$$

Then, this sequence belongs to both  $\ell^2$  and  $\ell^\infty$  spaces, however, the limit, does not have finitely many non-zero terms, and hence, does not belong to the subspaces. Thus, the subspaces are not closed. ■

2. What is the largest possible  $c$  if  $X = \mathbb{R}^2$  and  $x_1 = (1, 0), x_2 = (0, 1)$ ? If  $X = \mathbb{R}^3$  and  $x_1 = (1, 0, 0), x_2 = (0, 1, 0)$ , and  $x_3 = (0, 0, 1)$ ?

*Proof.* We know that

$$\left\| \sum_{i=1}^n \alpha_i x_i \right\| \geq c \sum_{i=1}^n |\alpha_i|.$$

In the first case, consider

$$\|(\alpha_1, \alpha_2)\| \geq c(|\alpha_1| + |\alpha_2|).$$

If we use,  $\|(x, y)\| = |x| + |y|$ , since the  $\ell^1$  norm is maximum, then we have

$$|\alpha_1| + |\alpha_2| \geq c(|\alpha_1| + |\alpha_2|)$$

which shows  $c \leq 1$ . The same holds for  $\mathbb{R}^3$ . ■

3. Show that in Def 2.4-4 the axioms of an equivalence relation hold.

*Proof.* Def 2.4-4 says that two norms  $\|\cdot\|$  and  $\|\cdot\|_0$  are equivalent if

$$a \|x\|_0 \leq \|x\| \leq b \|x\|_0.$$

We now check if the conditions for equivalent relations hold:

1.  $\|\cdot\|$  and  $\|\cdot\|$  should be equivalent, which is true by setting  $a = b = 1$ .
2. If  $\|\cdot\| \sim \|\cdot\|_0$ , then  $\|\cdot\|_0 \sim \|\cdot\|$  which is true since

$$\frac{1}{b} \|x\| \leq \|x\|_0 \leq \frac{1}{a} \|x\|.$$

3. Suppose  $\|\cdot\| \sim \|\cdot\|_0$  and  $\|\cdot\|_0 \sim \|\cdot\|_1$ , then  $\|\cdot\| \sim \|\cdot\|_1$ . First notice that

$$a \|x\|_0 \leq \|x\| \leq b \|x\|_0$$

and

$$c \|x\|_1 \leq \|x\|_0 \leq d \|x\|_1.$$

Then,

$$\|x\|_1 \leq \|x\| \leq \|x\|_1$$

Then, we can write the following:

$$\begin{aligned} c \|x\|_1 \leq \|x\|_0 \leq \frac{1}{a} \|x\|_1 &\implies ac \|x\|_1 \leq \|x\|_0 \\ \frac{1}{b} \|x\| \leq \|x\|_0 \leq d \|x\|_1 &\implies \|x\|_0 \leq bd \|x\|_1. \end{aligned}$$

This finishes the proof. ■

4. Show that equivalent norms on a vector space  $X$  induce the same topology on  $X$ .

*Proof.* We wish to show that the set of open sets induced by equivalent norms is the same. Let  $\|\cdot\|_0$  and  $\|\cdot\|_1$  be two equivalent norms such that

$$c_1 \|x\|_1 \leq \|x\|_0 \leq c_2 \|x\|_1.$$

Let  $A$  be an open set under the norm  $\|\cdot\|_0$ . That means that for all  $x \in A$ , there exists a ball of radius  $\epsilon$  contained in  $A$ . In other words,

$$B(x, \epsilon) \in A \implies \forall a \in A, \|x - a\|_0 < \epsilon \implies \|x - a\|_1 < \frac{\epsilon}{c_1}.$$

This means that  $A$  is also an open set under the norm  $\|\cdot\|_1$ . Thus, all the open sets induced by the metric  $\|\cdot\|_0$  is contained in the set of open sets induced by  $\|\cdot\|_1$ .

In a similar way, let  $A$  be an open set under the norm  $\|\cdot\|_1$ . That means that for all  $x \in A$ , there exists a ball of radius  $\epsilon$  contained in  $A$ . In other words,

$$B(x, \epsilon) \in A \implies \forall a \in A, \|x - a\|_1 < \epsilon \implies \|x - a\|_0 < c_2 \epsilon.$$

This means that  $A$  is also an open set under the norm  $\|\cdot\|_0$ . Thus, all the open sets induced by the metric  $\|\cdot\|_1$  is contained in the set of open sets induced by  $\|\cdot\|_0$  and hence, the set of all open sets induced, or the topologies are equal. ■

5. If  $\|\cdot\|$  and  $\|\cdot\|_0$  are equivalent norms on  $X$ , show that the Cauchy sequences in  $(X, \|\cdot\|)$  and  $(X, \|\cdot\|_0)$  are the same.

*Proof.* Let

$$a \|x\|_0 \leq \|x\| \leq b \|x\|_0.$$

Let  $(x_n)$  be a Cauchy sequence in the space  $(X, \|\cdot\|)$ , then, for some  $\epsilon > 0$ , we have  $N(\epsilon)$  and for all  $m, n > N(\epsilon)$ , we have

$$\|x_m - x_n\| < \epsilon \implies \|x_m - x_n\|_0 < \frac{\epsilon}{a}.$$

This means that  $(x_n)$  is also a Cauchy sequence in the space  $(X, \|\cdot\|_0)$ .

In a similar fashion, let  $(x_n)$  be a Cauchy sequence in the space  $(X, \|\cdot\|_0)$ . Then, for some  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for all  $m, n > N(\epsilon)$ , we have

$$\|x_m - x_n\|_0 < \epsilon \implies \|x_m - x_n\| < b\epsilon$$

which means that  $(x_n)$  is also a Cauchy sequence in the space  $(X, \|\cdot\|)$ . Thus, the Cauchy sequences are equal. ■

**6.** Show that  $\|\cdot\|_2$  and  $\|\cdot\|_\infty$  are equivalent.

*Proof.* First, we know that

$$\|x\|_2 = \left( \sum_{i=1}^n |\xi_i|^2 \right)^{1/2}.$$

and

$$\|x\|_\infty = \max_{i \in [n]} |\xi_i|.$$

Thus, we have

$$\|x\|_2 \leq \left( \sum_{i=1}^n \left( \max_{i \in [n]} |\xi_i| \right)^2 \right)^{1/2} = \max_{i \in [n]} |\xi_i| \sqrt{n} = \sqrt{n} \|x\|_\infty$$

On the other hand, we have

$$\|x\|_\infty^2 = \left( \max_{i \in [n]} |\xi_i| \right)^2 \leq \sum_{i=1}^n |\xi_i|^2 = \|x\|_2^2$$

and thus,

$$\|x\|_\infty \leq \|x\|_2.$$

■

**7.** Let  $\|x\|_2$  be defined on the set of all  $n$ -tuples on  $\mathbb{R}^2$  and let  $\|\cdot\|$  be any norm on that vector space, say  $X$ . Show directly, using that there exists  $b > 0$  such that  $\|x\| \leq b \|x\|_2$  for all  $x$ .

*Proof.* Let  $x = \sum_{i=1}^n \alpha_i e_i$ . Then, we have that

$$\|x\| = \left\| \sum_{i=1}^n \alpha_i e_i \right\| \leq \sum_{i=1}^n |\alpha_i| \|e_i\|$$

and similarly,

$$\|x\|_2^2 \leq \sum_{i=1}^n |\alpha_i|^2.$$

Set  $\|e_i\| = \sqrt{\sum_{i=1}^n |\alpha_i|^2}$ , then we get

$$\|x\| \leq \left( \sum_{i=1}^n |\alpha_i| \right) \|x\|_2$$

■

**8.** Show that norms  $\|\cdot\|_1$  and  $\|\cdot\|_2$  in Section 2.2, 8 satisfy

$$\frac{1}{\sqrt{n}} \|x\|_1 \leq \|x\|_2 \leq \|x\|_1.$$

*Proof.* We have that

$$\|x\|_1 = \sum_{i=1}^n |\xi_i| \text{ and } \|x\|_2 = \left( \sum_{i=1}^n |\xi_i|^2 \right)^{1/2}.$$

By the application of Cauchy-Schwarz, we get

$$\|x\|_1 = \sum_{i=1}^n |\xi_i| \leq \sqrt{\sum_{i=1}^n 1} \sqrt{\sum_{i=1}^n |\xi_i|^2} = \sqrt{n} \|x\|_2.$$

Also,

$$\|x\|_2 = \sqrt{\sum_{i=1}^n |\xi_i|^2} \leq \sqrt{\sum_{i=1}^n |\xi_i|^2 + 2 \sum_{i=1}^n \frac{\prod_{j=1}^n |\xi_j|}{|\xi_i|}} \leq \left| \sum_{i=1}^n |\xi_i| \right| = \|x\|_1.$$

■

**9.** If two norms  $\|\cdot\|$  and  $\|\cdot\|_0$  on a vector space  $X$  are equivalent, show that  $\|x_n - x\| \rightarrow 0$  implies  $\|x_n - x\|_0 \rightarrow 0$  and vice-versa.

*Proof.* Let

$$a \|x\|_0 \leq \|x\| \leq b \|x\|_0.$$

Then, if  $\|x_n - x\| \rightarrow 0$ , we have that for every  $\epsilon > 0$ , there exists  $N(\epsilon)$  such that for every  $n > N(\epsilon)$

$$\|(x_n - x) - 0\| < \epsilon \implies \|(x_n - x) - 0\|_0 < \frac{\epsilon}{a}$$

which shows that  $\|x_n - x\|_0 \rightarrow 0$ . In a similar fashion, we can show that

$$\|(x_n - x) - 0\|_0 \leq \epsilon \implies \|(x_n - x) - 0\| \leq b\epsilon$$

and hence,  $\|x_n - x\| \rightarrow 0$ . ■

**10.** Show that all complex  $m \times n$  matrices  $A = (a_{jk})$  with fixed  $m$  and  $n$  constitute a  $mn$ -dimensional vector space  $Z$ . Show that all norms on  $Z$  are equivalent.

*Proof.* Note that every  $m \times n$  matrix can be written in the basis  $\{E_{ij}\}_{i \in [m], j \in [n]}$  where  $E_{ij}$  is the matrix with 1 in the  $(i, j)^{th}$  position and zero otherwise. Hence, the dimension is  $mn$ . Since, it is a finite dimension space, all norms are equivalent. This is easy to see since let  $X = \sum_{i,j} \alpha_{i,j} E_{i,j}$ . Then, we have

$$\|X\| \geq c \sum_{i,j} |\alpha_{i,j}|.$$

Also,

$$\|X\|_0 = \left\| \sum_{i,j} \alpha_{i,j} E_{i,j} \right\|_0 \leq \sum_{i,j} |\alpha_{i,j}| \|E_{i,j}\|_0 \leq \frac{\max_{i,j} \|E_{i,j}\|_0}{c} \|X\|.$$

and similarly, we can exchange the norms to obtain the other side of the inequality. ■

## 2.5 Compactness and Finite Dimension

**1.** Show that  $\mathbb{R}^n$  and  $\mathbb{C}^n$  are not compact.

*Proof.* First, note that  $\mathbb{R}^n$  and  $\mathbb{C}^n$  are both  $n$ -dimensional, i.e, finite dimensional. We know that a finite dimensional subspace is compact if and only if it is closed and bounded. However,  $\mathbb{R}^n$  and  $\mathbb{C}^n$  are not bounded, and hence, not compact. ■

**2.** Show that a discrete metric space  $X$  consisting of infinitely many points is not compact.

*Proof.* Consider any sequence  $(x_n)$  with all elements distinct. Then, we have that for all  $m, n$

$$d(x_m, x_n) = 1.$$

Thus, this sequence can never converge, and neither can any subsequence. What adds to this is the fact that  $X$  has infinite points, which means that none of the elements or subsequences would repeat. Thus, the space is not compact. ■

**3.** Give examples of compact and non-compact curves in the plane  $\mathbb{R}^2$ .

*Proof.* Any closed and bounded curve would be compact, such as the closed ball of radius  $r$ . The open ball of radius  $r$  is open and hence, not compact. ■

**4.** Show that for an infinite subset  $M$  in the space  $s$  to be compact, it is necessary that there are numbers  $\gamma_1, \gamma_2, \dots$  such that for all  $x = (\xi_k(x)) \in M$  we have  $|\xi_k(x)| \leq \gamma_k$ .

*Proof.* Recall that  $s$  is the space of all bounded or unbounded sequences. The metric defined on this space is

$$d(x, y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|\xi_j - \eta_j|}{1 + |\xi_j + \eta_j|}.$$

We can show this space is complete as follows: Consider  $(x_n)_n = ((\xi_j^{(n)})_j)_n$  to be a sequence that is Cauchy, i.e

$$d(x_m, x_n) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|\xi_j - \eta_j|}{1 + |\xi_j + \eta_j|} < \epsilon \implies \max_j \frac{|\xi_j^{(n)} - \xi_j^{(m)}|}{1 + |\xi_j^{(n)} - \xi_j^{(m)}|} < 2\epsilon \implies \max_j |\xi_j^{(n)} - \xi_j^{(m)}| < \frac{2\epsilon}{1 - 2\epsilon}.$$

Thus, for a fixed  $j$ , the sequence  $(\xi_j^{(1)}, \xi_j^{(2)}, \dots)$  is Cauchy, and hence, converges to  $\xi_j$ . Thus, considering  $x = (\xi_1, \xi_2, \dots) \in M$ , we have that  $x_n \rightarrow x$ . Hence, this space is complete. For  $s$  to be compact, we require it to be bounded. This means that bounding the absolute value of each of the terms for all sequences is sufficient. ■

5. A metric space  $X$  is said to be locally compact if every point of  $X$  has a neighborhood that is compact. Show that  $\mathbb{R}^n$  and  $\mathbb{C}^n$  is locally compact.

*Proof.* Take any point  $x \in \mathbb{R}^n$ . We know that any neighborhood of  $x$  contains an  $\epsilon$ -ball around  $x$ . Let the neighborhood be the closed ball of radius  $\epsilon$ . Since,

$$B(x, \epsilon) \subset \tilde{B}(x, \epsilon),$$

this closed ball is a valid neighborhood. Now, this neighborhood is bounded (since we can bound  $\tilde{B}(x, \epsilon)$  with another open ball of radius  $\epsilon + \delta, \delta > 0$ ) and is closed, and hence, is compact. Thus,  $\mathbb{R}^n$  is locally compact. ■

6. Show that a compact metric space  $X$  is locally compact.

*Proof.* Consider some point  $x \in X$ . Since  $X$  belongs to set of all open sets in  $X$ , there exists a ball around  $x$  completely contained in  $X$  and hence,  $X$  is a neighborhood itself. Since  $X$  is compact, it follows that  $X$  is locally compact. ■

7. If  $\dim Y < \infty$  in Reisz Lemma, show that one can choose  $\theta = 1$ .

*Proof.* **TODO.** ■

8. In 7, section 2.4, show that there is an  $a > 0$  such that  $a \|x\|_2 < \|x\|$ .

*Proof.* **TODO** ■

9. If  $X$  is a compact metric space and  $M \subset X$  is closed, show that  $M$  is compact.

*Proof.* Let  $(x_n)$  be any sequence in  $M$  that converges to  $x$ . Since  $M$  is closed,  $x \in M$ . Also, since  $X$  is compact, we have that atleast one subsequence of  $(x_n)$  converges. Since  $x_n \rightarrow x$ , the subsequence also converges to  $x$ , and  $x \in M$ . Thus, we have a sequence in  $M$  that has a subsequence that converges to some point in  $M$ . Hence,  $M$  is compact. ■

10. Let  $X$  and  $Y$  be metric spaces,  $X$  is compact, and  $T : X \mapsto Y$  is bijective and continuous. Show that  $T$  is a homomorphism.

*Proof.* Recall the definition of a homomorphism: a continuous and bijective mapping whose inverse is also continuous.

Now, we wish to show that  $T^{-1} : Y \mapsto X$  is continuous given that  $X$  and  $Y$  are compact. Let  $(y_n)$  be some sequence in  $M \subset Y$  that has a convergent subsequence  $(y_{n_k})$  that converges to  $y$ . Now, consider the sequence  $(x_n) = (T^{-1}y_n)$ . Since  $X$  is compact, there exists some subsequence  $(x_{n_j})$  that converges to  $x_0$ . Also, assume  $Tx = y$ . Now, since  $Tx_n = y_n$ , using the continuity of  $T$ , we have

$$y_n = Tx_n \rightarrow Tx_0$$

since a sequence converges to the same limit as that of its subsequences. Now, since the limits are unique, we have that  $Tx_0 = y$ . But since  $T$  is bijective, and hence, one-one,  $x_0 = x$ . Thus, we have that under the mapping  $T^{-1}$ , a sequence in  $Y$  and its corresponding image in  $X$  are convergent, and hence,  $T^{-1}$  is continuous. ■

## 2.6 Linear Operators

1. Show that the operators in 2.6-2, 2.6-3, and 2.6-4 are linear.

*Proof.* We are required to show that

1. The identity operator:  $I : X \mapsto X$  defined as  $Ix = x$  is linear. Note that  $I(\alpha x + \beta y) = \alpha x + \beta y = \alpha Ix + \beta Iy$ . Hence, the operator is linear.

2. The zero operator, defined as  $0 : X \mapsto X$  such that  $0x = 0$  is linear. Indeed,

$$0(\alpha x + \beta y) = 0 = \alpha 0x + \beta 0y$$

and hence, the operator is linear.

3. The differentiation operator, defined on  $X$ , the space of all polynomials on  $[a, b]$  as

$$Tx(t) = x'(t)$$

is linear. Notice that

$$T(\alpha x + \beta y)(t) = (\alpha x + \beta y)'(t) = \alpha x'(t) + \beta y'(t) = (\alpha Tx + \beta Ty)(t)$$



2. Show that the operators  $T_1, \dots, T_4$  from  $\mathbb{R}^2$  to  $\mathbb{R}^2$  defined by

1.  $(\xi_1, \xi_2) \mapsto (\xi_1, 0)$
2.  $(\xi_1, \xi_2) \mapsto (0, \xi_2)$
3.  $(\xi_1, \xi_2) \mapsto (\xi_2, \xi_1)$
4.  $(\xi_1, \xi_2) \mapsto (\gamma\xi_1, \gamma\xi_2)$

respectively are linear, and interpret these operations geometrically.

*Proof.* 1. We have,

$$\begin{aligned} T_1(\alpha(\xi_1, \xi_2) + \beta(\eta_1, \eta_2)) &= T_1((\alpha\xi_1 + \beta\eta_1, \alpha\xi_2 + \beta\eta_2)) = (\alpha\xi_1 + \beta\eta_1, 0) = \alpha(\xi_1, 0) + \beta(\eta_1, 0) \\ &= \alpha T_1((\xi_1, \xi_2)) + \beta T_1((\eta_1, \eta_2)) \end{aligned}$$

and hence, the operator is linear. Ideally, this is a projection onto the  $\xi_1$ -axis.

2. We have,

$$\begin{aligned} T_2(\alpha(\xi_1, \xi_2) + \beta(\eta_1, \eta_2)) &= T_2((\alpha\xi_1 + \beta\eta_1, \alpha\xi_2 + \beta\eta_2)) = (0, \alpha\xi_2 + \beta\eta_2) = \alpha(0, \xi_2) + \beta(0, \eta_2) \\ &= \alpha T_2((\xi_1, \xi_2)) + \beta T_2((\eta_1, \eta_2)) \end{aligned}$$

and hence, the operator is linear. Ideally, this is a projection onto the  $\xi_2$ -axis.

3. We have,

$$\begin{aligned} T_3(\alpha(\xi_1, \xi_2) + \beta(\eta_1, \eta_2)) &= T_3((\alpha\xi_1 + \beta\eta_1, \alpha\xi_2 + \beta\eta_2)) = (\alpha\xi_2 + \beta\eta_2, \alpha\xi_1 + \beta\eta_1) = \alpha(\xi_2, \xi_1) + \beta(\eta_2, \eta_1) \\ &= \alpha T_3((\xi_1, \xi_2)) + \beta T_3((\eta_1, \eta_2)) \end{aligned}$$

and hence, the operator is linear. Ideally, this is a reflection about the line  $\xi_1 = \xi_2$ .

4. We have,

$$\begin{aligned} T_4(\alpha(\xi_1, \xi_2) + \beta(\eta_1, \eta_2)) &= T_4((\alpha\xi_1 + \beta\eta_1, \alpha\xi_2 + \beta\eta_2)) = (\alpha\gamma\xi_1 + \beta\gamma\eta_1, \alpha\gamma\xi_2 + \beta\gamma\eta_2) \\ &= \alpha(\gamma\xi_1, \gamma\xi_2) + \beta(\gamma\eta_1, \gamma\eta_2) \\ &= \alpha T_4((\xi_1, \xi_2)) + \beta T_4((\eta_1, \eta_2)) \end{aligned}$$

and hence, the operator is linear. Ideally, this is a scaling operation by the factor  $\gamma$ .

3. What are the domain, range, and null space of  $T_1, T_2, T_3$  in 2?

*Proof.* We have,

1.  $\mathcal{D}(T_1) = \mathbb{R}^2, \mathcal{R}(T_1) = \xi_1\text{-axis}, \mathcal{N}(T_1) = \xi_2\text{-axis}.$
2.  $\mathcal{D}(T_2) = \mathbb{R}^2, \mathcal{R}(T_2) = \xi_2\text{-axis}, \mathcal{N}(T_2) = \xi_1\text{-axis}.$
3.  $\mathcal{D}(T_3) = \mathbb{R}^2, \mathcal{R}(T_3) = \mathbb{R}^2, \mathcal{N}(T_3) = \text{origin}.$

4. What is the null space of  $T_4$  in 2? Of  $T_1$  and  $T_2$  in 2.6-7? of  $T$  in 2.6-4?

*Proof.* The null space of  $T_4$  in 2 is the origin if  $\gamma \neq 0$  and  $\mathbb{R}^2$  if  $\gamma = 0$ . If  $\gamma = 0$ , then  $T_4$  is the zero operator.

In 2.6-7,  $T_1$  is the cross-product defined on  $\mathbb{R}^3$ . Keeping one input fixed, the null space is simply all the scaled versions of the first input.  $T_2$  is the dot product defined on  $\mathbb{R}^3$ . Keeping one input fixed, the null space is the set of all vectors which are orthogonal to the first.

$T$  in 2.6-4 is the differentiation operator on the space of all polynomials defined in  $[a, b]$ . In such a case, we wish to find the set of polynomials such that the differentiation results in a zero polynomial, or a polynomial that is zero at each point in  $[a, b]$ . Thus, all the constant polynomials constitute the null space.

5. Let  $T : X \mapsto Y$  be a linear operator. Show that the image of a subspace  $V$  of  $X$  is a vector space, and so is the inverse image of a subspace  $W$  of  $Y$ .

*Proof.* Let  $V$  be a subspace of  $X$ . We wish to show that  $T(V)$  is also a subspace. In other words, let  $v_1, v_2 \in V$  and  $Tv_1, Tv_2 \in T(V)$ . We wish to show that  $\alpha Tv_1 + \beta Tv_2 \in T(V)$ , i.e.,  $\exists v_3 \in V$  such that  $Tv_3 = \alpha Tv_1 + \beta Tv_2$ . Since,  $T$  is a linear operator, we have

$$\alpha Tv_1 + \beta Tv_2 = T(\alpha v_1) + T(\beta v_2) = T(\alpha v_1 + \beta v_2)$$

and since  $V$  is a subspace,  $\alpha v_1 + \beta v_2 \in V$ . Thus, setting  $v_3 = \alpha v_1 + \beta v_2$  completes the proof.

Let  $T^{-1}$  denote the inverse mapping of  $T$ . Let  $w_1, w_2 \in W$  and  $T^{-1}w_1, T^{-1}w_2 \in T^{-1}(W)$ . We wish to show that  $\exists w_3 \in W$  such that  $T^{-1}w_3 = \alpha T^{-1}w_1 + \beta T^{-1}w_2$ . This implies that  $\alpha T^{-1}w_1 + \beta T^{-1}w_2 \in T^{-1}(W)$  and hence,  $T^{-1}(W)$  is a subspace. Since  $T$  is a linear mapping we have

$$T(\alpha T^{-1}w_1 + \beta T^{-1}w_2) = \alpha TT^{-1}w_1 + \beta TT^{-1}w_2 = \alpha w_1 + \beta w_2$$

where we use the fact that  $TT^{-1} = I$ . Thus, we have

$$\alpha T^{-1}w_1 + \beta T^{-1}w_2 = T^{-1}(\alpha w_1 + \beta w_2)$$

and hence, setting  $w_3 = \alpha w_1 + \beta w_2$  completes the proof. ■

**6.** If the product of two linear operator exists, show that it is linear.

*Proof.* Let  $T_1 : X \mapsto Z$  and  $T_2 : Z \mapsto Y$  be two linear functions such that their product is defined as  $T_2T_1 : X \mapsto Y$ . Then, we have for  $x_1, x_2 \in X$

$$T_2T_1(\alpha x_1 + \beta x_2) = T_2(\alpha T_1x_1 + \beta T_1x_2) = \alpha T_2(T_1x_1) + \beta T_2(T_1x_2) = \alpha T_2T_1x_1 + \beta T_2T_1x_2$$

where the first equality follows since  $T_1$  is linear, and the second follows from the fact that  $T_2$  is linear. Hence, their composite is also linear. ■

**7.** Let  $X$  be any vector space and  $S : X \mapsto X$  and  $T : X \mapsto X$  be any operators.  $S$  and  $T$  are said to commute if  $ST = TS$ . Do  $T_1$  and  $T_3$  in 2 commute?

*Proof.* Recall

$$T_1((\xi_1, \xi_2)) = (\xi_1, 0)$$

$$T_3((\xi_1, \xi_2)) = (\xi_2, \xi_1)$$

Then, we have

$$T_1T_3((\xi_1, \xi_2)) = T_1((\xi_2, \xi_1)) = (\xi_2, 0)$$

while

$$T_3T_1((\xi_1, \xi_2)) = T_3((\xi_1, 0)) = (0, \xi_1)$$

Hence, they do not commute. ■

**8.** Write the operators in 2 using matrices.

*Proof.* We have

$$T_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, T_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, T_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, T_4 = \begin{pmatrix} \gamma & 0 \\ 0 & \gamma \end{pmatrix}$$

**9.** In 2.6-8, write  $Ax = y$  in terms of components, show that  $T$  is linear.

*Proof.* We have that for some  $A \in \mathbb{R}^{m \times n}$ ,  $y \in \mathbb{R}^m$ ,  $x \in \mathbb{R}^n$ ,

$$\eta_i = \sum_{j=1}^n A_{ij} \xi_j.$$

This is linear because for  $x_1, x_2 \in \mathbb{R}^n$

$$y_i = A(\alpha x_1 + \beta x_2)_i = \sum_{j=1}^n A_{ij}(\alpha \xi_j^{(1)} + \beta \xi_j^{(2)}) = \alpha \sum_{j=1}^n A_{ij} \xi_j^{(1)} + \beta \sum_{j=1}^n A_{ij} \xi_j^{(2)} = \alpha (Ax_1)_i + \beta (Ax_2)_i$$

**10.** Formulate the condition in 2.6-10(a) in terms of the null space of  $T$ .

*Proof.* The condition in 2.6-10 reads that the inverse  $T^{-1} : \mathcal{R}(T) \mapsto \mathcal{D}(T)$  exists if and only if

$$T(x) = 0 \implies x = 0.$$

This can be reformulated as  $\mathcal{N}(T) = \{0\}$ . ■

**11.** Let  $X$  be the vector space of all complex  $2 \times 2$  matrices and define  $T : X \mapsto X$  by  $Tx = bx$  where  $b \in X$  is fixed and  $bx$  denotes the usual product of matrices. Show that  $T$  is linear. Under what condition does  $T^{-1}$  exist?

*Proof.* Matrix multiplication is linear and hence,  $T$  is linear. For  $T^{-1}$  to exist, the null space should only contain the zero matrix. In other words, the eigenvalues should be non-zero and hence, the matrix  $b$  should be non-singular. ■

**12.** Does the inverse of  $A$   $T$  in 2.6-4 exist?

*Proof.* 2.6-4 is the differentiation operator. In 4, we have shown that the null space for this operator is the set of all constant polynomials. Since the null space consists of more elements than just the zero polynomial, the inverse does not exist. ■

**13.** Let  $T : \mathcal{D}(T) \mapsto Y$  be a linear operator whose inverse exists. If  $\{x_1, \dots, x_n\}$  is a linearly independent set in  $\mathcal{D}(T)$ , show that the set  $\{Tx_1, \dots, Tx_n\}$  is linear independent.

*Proof.* Since  $\{x_1, \dots, x_n\}$  is linearly independent, we have

$$\sum_{i=1}^n \alpha_i x_i = 0 \implies \alpha_i = 0 \forall i \in [n].$$

Since  $T$  is a linear operator, we have

$$T\left(\sum_{i=1}^n \alpha_i x_i\right) = T0 = 0$$

where the last inequality follows since  $T^{-1}$  exists. Thus, we have

$$\sum_{i=1}^n \alpha_i Tx_i = 0 \implies \alpha_i = 0$$

and hence,  $\{Tx_1, \dots, Tx_n\}$  is a linearly independent set. ■

**14.** Let  $T : X \mapsto Y$  be a linear operator and  $\dim X = \dim Y = n < \infty$ . Show that  $\mathcal{R}(T) = Y$  if and only if  $T^{-1}$  exists.

*Proof.* First, assume  $\mathcal{R}(T) = Y$ . Let  $\{e_1, \dots, e_n\}$  be some basis of  $X$ . To show the existence of  $T^{-1}$ , we wish to show that

$$Ta = Tb \implies a = b.$$

Let  $a = \sum_{i=1}^n \alpha_i e_i$  and  $b = \sum_{i=1}^n \beta_i e_i$ . Then,

$$Ta - Tb = T\left(\sum_{i=1}^n (\alpha_i - \beta_i) e_i\right) = \sum_{i=1}^n (\alpha_i - \beta_i) Te_i.$$

If  $Ta = Tb$ , then  $\alpha_i = \beta_i \forall i \in [n]$ , since  $\{e_1, \dots, e_n\}$  are linearly independent. Thus,  $a = b$ .

On the other hand, let  $T^{-1}$  exist. Then, we wish to show that  $\mathcal{R}(T) = Y$ . We already know that  $\mathcal{R}(T) \subset Y$ . Thus, showing  $Y \subset \mathcal{R}(T)$  suffices. Thus, we wish to show that for any  $y \in Y$ , there exists some  $x \in X$  such that  $Tx = y$ . From 13, we have that if  $\{e_1, \dots, e_n\}$  is a basis for  $X$ , then  $\{Te_1, \dots, Te_n\}$  is a basis for  $Y$ . Hence, for any  $y = \sum_{i=1}^n \alpha_i Te_i$ , we have

$$Tx = T\left(\sum_{i=1}^n \alpha_i e_i\right) = \sum_{i=1}^n \alpha_i Te_i$$

or in other words, any  $y = \sum_{i=1}^n \alpha_i Te_i$  can be written as  $Tx$  where  $x = \sum_{i=1}^n \alpha_i e_i$ . This finishes the proof. ■

**15.** Consider the vector space  $X$  of all real-valued functions which are defined on  $\mathbb{R}$  and have derivatives of all orders everywhere on  $\mathbb{R}$ . Define  $T : X \mapsto X$  by  $y(t) = Tx(t) = x'(t)$ . Show that  $\mathcal{R}(T)$  is all of  $X$  but  $T^{-1}$  does not exist.

*Proof.* Suppose  $x \in X$  is some function that has derivatives of all orders defined. This means that there exists some function  $x_0 \in X$  such that  $x'_0 = x$  and at the same time  $x_1 \in X$  such that  $x_1 = x'$ . Thus, for every function, we can find one derivative higher and one derivative lower, and similarly, every function is either the derivative of, or the integral of another function. Hence, the range is  $X$ .  $T^{-1}$  does not exist because the null space is all constant polynomials. ■

## 2.7 Bounded and Continuous Linear Operators

1. *Prove (7).*

*Proof.* We wish to show that  $\|T_1 T_2\| \leq \|T_1\| \|T_2\|$  and  $\|T^n\| \leq \|T\|^n$ . The second follows from repeatedly applying the first. We show the first one now. Let  $T_1 : Z \mapsto Y$  and  $T_2 : X \mapsto Z$ , then  $T_1 T_2 : X \mapsto Y$ . We have

$$\|T_1 T_2\| = \sup_{\|x\|=1} \|T_1 T_2 x\| \leq \sup_{\|x\|=1} \|T_1\| \|T_2 x\| = \|T_1\| \sup_{\|x\|=1} \|T_2 x\| = \|T_1\| \|T_2\|.$$

2. *Let  $X$  and  $Y$  be normed spaces. Show that a linear operator  $T : X \mapsto Y$  is bounded if and only if  $T$  maps bounded sets in  $X$  to bounded sets in  $Y$ .*

*Proof.* Suppose  $T$  is bounded. Then,

$$\|Tx\| \leq c \|x\|.$$

Suppose,  $M \subset X$  is a bounded set such that for all  $x_1, x_2 \in M$ ,

$$\|x_1 - x_2\| \leq \delta(M).$$

Then,

$$\|Tx_1 - Tx_2\| = \|T(x_1 - x_2)\| \leq c \|x_1 - x_2\| = c\delta(M).$$

In other words, the set that  $M$  is mapped into is also bounded.

Now, consider bounded sets in  $X$  are mapped into bounded sets in  $Y$ . Let  $M \subset X$  be a bounded set that maps into  $N \subset Y$  and for some  $x_1, x_2, x \in M$  and  $\epsilon > 0$ , define

$$x_1 = x_2 + (\delta(M) - \epsilon) \frac{x}{\|x\|} \implies \|x_1 - x_2\| = |\delta(M) - \epsilon| \leq \delta(M).$$

Then, we have

$$\|Tx_1 - Tx_2\| = \left\| T(\delta(M) - \epsilon) \frac{x}{\|x\|} \right\| = |\delta(M) - \epsilon| \frac{\|Tx\|}{\|x\|} \leq \delta(N) \implies \|Tx\| \leq \frac{\delta(N)}{|\delta(M) - \epsilon|} \|x\|$$

and hence,  $T$  is bounded. ■

3. *If  $T \neq 0$  is a bounded linear operator, show that for any  $x \in \mathcal{D}(T)$  such that  $\|x\| < 1$ , we have the strict inequality,  $\|Tx\| < \|T\|$ .*

*Proof.* We have,

$$\|Tx\| \leq \|T\| \|x\| < 1.$$

4. *Give a direct proof of 2.7-9(b), without using 2.7-9(a).*

*Proof.* We have to show that if a bounded linear operator  $T$  is continuous at one point, say  $x_0$ , then, it is continuous throughout its domain. Since  $T$  is continuous at  $x_0$ , we have

$$\|x_0 - x\| < \delta \implies \|Tx_0 - Tx\| < \epsilon.$$

Also, since  $T$  is bounded, we have

$$\|Tx\| \leq c \|x\|.$$

Now, suppose we have to test continuity at some point  $x_1$  such that  $\|x_1 - x_0\| < \delta'$ . Then, we have

$$\|x_1 - x\| = \|x_1 - x_0 + x_0 - x\| \leq \|x_1 - x_0\| + \|x_0 - x\| < \delta' + \delta.$$

Also,

$$\|T(x_1 - x)\| = \|T(x_1 - x_0 + x_0 - x)\| \leq \|T(x_1 - x_0)\| + \|T(x_0 - x)\| < c \|x_1 - x_0\| + \epsilon = c\delta' + \epsilon.$$

Thus, continuity at any other point is implied. ■

**5.** Show that a operator  $T : \ell^\infty \mapsto \ell^\infty$  defined by  $y = (\eta_j) = Tx, \eta_j = \xi_j/j, x = (\xi_j)$  is linear and bounded.

*Proof.* First, we have

$$(T(\alpha x_1 + \beta x_2))_j = \frac{\alpha \xi_j^{(1)} + \beta \xi_j^{(2)}}{j} = \alpha \frac{\xi_j^{(1)}}{j} + \beta \frac{\xi_j^{(2)}}{j} = \alpha (Tx_1)_j + \beta (Tx_2)_j.$$

Thus,  $T$  is linear. Now, the norm in the  $\ell^\infty$  space is defined as

$$\|T\| = \sup_{\|x\|=1} \|Tx\|_\infty = \sup_{\|x\|=1} \sup_j |(Tx)_j|.$$

Let  $x$  be any sequence such that  $\|x\|_\infty = 1$ . In other words,

$$\sup_j |\xi_j| = 1 \implies |\xi_j| \leq 1.$$

Then,

$$\sup_j |(Tx)_j| = \sup_j \left| \frac{\xi_j}{j} \right| \leq \sup_j |\xi_j| = 1.$$

Thus,  $T$  is also bounded. ■

**6.** Show that the range  $\mathcal{R}(T)$  of a bounded linear operator  $T : X \mapsto Y$  need not be closed in  $Y$ .

*Proof.* Consider the example from 5:

$$y = Tx = (\eta_j), \eta_j = \xi_j/j, x = (\xi_j).$$

where  $x, y \in \ell^\infty$ . Let

$$x_n = \left(1, \frac{1}{2}, \dots, \frac{1}{n}, 0, \dots, 0\right) \implies Tx_n = \left(1, \frac{1}{4}, \dots, \frac{1}{n^2}, 0, \dots, 0\right).$$

In such a case,

$$Tx_n \rightarrow \left(1, \frac{1}{4}, \dots, \frac{1}{n^2}, \frac{1}{(n+1)^2}, \dots\right),$$

however, the corresponding element from  $\ell^\infty$  to produce the limit point would be

$$\left(1, \frac{1}{2}, \dots, \frac{1}{n}, \frac{1}{n+1}, \dots\right) \notin \ell^\infty.$$

Thus, the limit point does not belong to the range and hence, the range is not closed. ■

**7.** Let  $T$  be a bounded linear operator from a normed space  $X$  onto a normed space  $Y$ . If there is a positive  $b$  such that

$$\|Tx\| \geq b \|x\| \quad \forall x \in X$$

then show that  $T^{-1} : Y \mapsto X$  exists and is bounded.

*Proof.* To show that  $T^{-1}$  exists, we show that the null space only consists of the zero element. Suppose not. Then, let  $Tx = 0, x \neq 0$ . Then,

$$\|Tx\| = 0 \geq b \|x\| \geq 0 \implies \|x\| = 0 \implies x = 0$$

which shows that  $Tx = 0 \implies x = 0$ . Thus,  $T^{-1}$  exists.

Now, to show  $T^{-1}$  is bounded, note that

$$b \|x\| \leq \|Tx\| \leq c \|x\|.$$

Let  $y = Tx$ , then

$$\|T^{-1}y\| = \|x\| \leq \frac{1}{b} \|y\|$$

and hence,  $T^{-1}$  is bounded. ■

**8.** Show that the inverse  $T^{-1} : \mathcal{R}(T) \mapsto X$  of a bounded linear operator  $T : X \mapsto Y$  need not be bounded.

*Proof.* Consider the example in 5,

$$y = (\eta_j) = Tx, \eta_j = \xi_j/j, x = (\xi_j).$$

Clearly, in this case,  $T^{-1}y = (j\eta_j)$ . Suppose  $y \in \ell^\infty$  such that  $\|y\|_\infty = 1$ . Then,

$$\|y\|_\infty = 1 \implies \sup_j |\eta_j| = 1 \implies |\eta_j| \leq 1.$$

Now, consider

$$\|T^{-1}y\| = \sup_{\|y\|=1} \sup_j \|j\eta_j\| \leq \sup_j |\eta_j| j \leq \sup_j j$$

which is clearly unbounded. Thus, the inverse mapping need not be bounded. ■

**9.** Let  $T : C[0, 1] \mapsto C[0, 1]$  such that

$$y(t) = \int_0^t x(\tau) d\tau.$$

Find  $\mathcal{R}(T)$  and  $T^{-1} : \mathcal{R}(T) \mapsto C[0, 1]$ . Is  $T^{-1}$  linear and bounded?

*Proof.* The range of  $T$ ,  $\mathcal{R}(T)$  is the subspace of all continuous functions on  $[0, 1]$  such that  $y(0) = 0$ . Clearly,  $T^{-1}y(t) = y'(t)$ .  $T^{-1}$  is linear since differentiation is a linear operation. However, it is not bounded. Consider  $y(t) = t^n$ . Then,

$$\|y\| = \max_{t \in [0, 1]} |y(t)| = 1.$$

Now,

$$\|T^{-1}y\| = \max_{t \in [0, 1]} |nt^{n-1}| = n = n\|y\|$$

and hence,  $\|T^{-1}\| = n$  which is unbounded. ■

**10.** On  $C[0, 1]$ , define  $S$  and  $T$  by

$$y(s) = s \int_0^1 x(t) dt, \quad y(s) = sx(s).$$

Do  $S$  and  $T$  commute? Find  $\|S\|, \|T\|, \|ST\|, \|TS\|$ .

*Proof.* Consider

$$STx(t) = S(tx(t)) = t \int_0^1 tx(t) dt = t^2 \int_0^1 x(t) dt.$$

On the other hand,

$$TSx(t) = Tt \int_0^1 x(t) dt = t^2 \int_0^1 1x(t) dt.$$

Thus,  $S$  and  $T$  commute. Now, let  $\|x\| = 1$ . Thus,

$$\|x\| = \max_{t \in [0, 1]} |x(t)| = 1 \implies |x(t)| \leq 1.$$

Now,

$$\begin{aligned} \|S\| &= \sup_{\|x\|=1} \|Sx\| = \sup_{\|x\|=1} \max_{t \in [0, 1]} t \int_0^1 x(t) dt \leq 1 \cdot \int_0^1 1 dt = 1 \\ \|T\| &= \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\|=1} \max_{t \in [0, 1]} tx(t) \leq 1 \cdot 1 = 1 \\ \|ST\| &= \|TS\| = \sup_{\|x\|=1} \|STx\| = \sup_{\|x\|=1} \max_{t \in [0, 1]} t^2 \int_0^1 x(t) dt \leq 1 \cdot \int_0^1 1 dt = 1 \end{aligned}$$
■

**11.** Let  $X$  be the normed space of all bounded real-valued functions on  $\mathbb{R}$  with norm defined by

$$\|x\| = \sup_{t \in \mathbb{R}} |x(t)|$$

and let  $T : X \mapsto X$  be defined by

$$y(t) = Tx(t) = x(t - \Delta).$$

Here,  $\Delta > 0$  is a constant. Is  $T$  linear and bounded?

*Proof.* First, to check if  $T$  is linear, we have

$$T(\alpha x_1 + \beta x_2)(t) = (\alpha x_1 + \beta x_2)(t - \Delta) = \alpha x_1(t - \Delta) + \beta x_2(t - \Delta) = \alpha T x_1(t) + \beta T x_2(t).$$

where the last inequality follows since  $x_1$  and  $x_2$  are bounded real-valued functions. Thus,  $T$  is linear. Now, let  $\|x\| = 1$ .

$$\|x\| = \max_{t \in \mathbb{R}} |x(t)| = 1 \implies |x(t)| \leq 1 \quad \forall t \in \mathbb{R}.$$

Thus,

$$\|T\| = \sup_{\|x\|=1} \|Tx\| = \max_{t \in \mathbb{R}} |x(t - \Delta)| \leq 1.$$

since  $t \in \mathbb{R} \implies t - \Delta \in \mathbb{R}$  since  $\Delta > 0$  is fixed. Thus,  $T$  is bounded. ■

**12.** From 2.7-7, we know that an  $r \times n$  matrix  $A = (a_{jk})$  defines a linear operator from the vector space  $X$  of all ordered  $n$ -tuples of numbers into the vector space  $Y$  of all ordered  $r$ -tuples of numbers. Suppose that any norm  $\|\cdot\|_1$  is given on  $X$  and any norm  $\|\cdot\|_2$  is given on  $Y$ . there could be various norms on the space  $Z$  of all matrices. A norm  $\|\cdot\|$  on  $Z$  is said to be compatible with  $\|\cdot\|_1$  and  $\|\cdot\|_2$  if

$$\|Ax\|_2 \leq \|A\| \|x\|_1.$$

Show that the norm defined by

$$\|A\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_1}$$

is compatible with  $\|\cdot\|_1$  and  $\|\cdot\|_2$ . This norm is often called the natural norm defined by  $\|\cdot\|_1$  and  $\|\cdot\|_2$ . If we choose  $\|x\|_1 = \max_j |\xi_j|$  and  $\|y\|_2 = \max_i |\eta_i|$ , show that the natural norm is

$$\|A\| = \max_j \sum_{k=1}^n |a_{jk}|.$$

*Proof.* First, the norm

$$\|A\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_1}$$

is always non-negative. Let  $\|A\| = c$ , then for all  $x \in X$  such that  $x \neq 0$ , we have

$$\frac{\|Ax\|_2}{\|x\|_1} \leq c \implies \|Ax\|_2 \leq c \|x\|_1 \implies \|Ax\|_2 \leq \|A\| \|x\|_1.$$

Thus, the given norm is a natural norm for  $\|\cdot\|_1$  and  $\|\cdot\|_2$ .

Suppose, we now consider the norms

$$\|x\|_1 = \max_j |\xi_j|, \quad \|y\|_2 = \max_i |\eta_i|.$$

Then, we have

$$\|y\|_2 = \max_{j \in [r]} \left| \sum_{k=1}^n a_{jk} \xi_k \right| \leq \max_{j \in [r]} \sum_{k=1}^n |a_{jk}| |\xi_k| \leq \max_{j \in [r]} \sum_{k=1}^n |a_{jk}| \left( \max_{i \in [n]} |\xi_i| \right) = \|x\|_1 \max_{j \in [r]} \sum_{k=1}^n |a_{jk}|$$

and hence,  $\|A\| = \max_{j \in [r]} \sum_{k=1}^n |a_{jk}|$ . ■

**13.** Show that if  $A \in \mathbb{R}^{n \times n}$ , then a compatible norm is defined by

$$\|A\| = \left( \sum_{j=1}^n \sum_{k=1}^n a_{jk}^2 \right)^{1/2}$$

but for  $n > 1$  this is not the natural norm defined by the Euclidian norm on  $\mathbb{R}^n$ .

*Proof.* Let  $y = Ax$ . Then,

$$\begin{aligned}
\|y\| &= \left( \sum_{i=1}^n \left( \sum_{j=1}^n a_{ij} x_j \right)^2 \right)^{1/2} \\
&\leq \left( \sum_{i=1}^n \left( \sqrt{\sum_{j=1}^n a_{ij}^2} \sqrt{\sum_{j=1}^n x_j^2} \right)^2 \right)^{1/2} \\
&= \left( \|x\|^2 \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \right)^{1/2} \\
&= \|x\| \left( \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \right)^{1/2}
\end{aligned}$$

where the first inequality follows from Hölder's inequality, which gives us

$$\sum |x_i y_i| \leq \left( \sum |x_i|^p \right)^{1/p} \left( \sum |y_i|^q \right)^{1/q}, \quad p^{-1} + q^{-1} = 1.$$

Thus, this norm, also called the Frobenius norm is compatible with the Euclidian norm. Now, we shall find the natural norm. Recall, the natural norm is defined as

$$\|A\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Ax\|}{\|x\|}.$$

Let  $A = I$ . Then, the Frobenius norm results in  $\|A\| = \sqrt{n}$ , while the natural norm results in  $\|A\| = 1$ . Thus, these are not the same norms, unless  $n = 1$ . ■

**14.** *If in 12, we choose*

$$\|x\|_1 = \sum_{k=1}^n |\xi_k|, \quad \|y\|_2 = \sum_{j=1}^r |a_{jk}|,$$

*then show that a compatible norm is defined by*

$$\|A\| = \max_k \sum_{j=1}^r |a_{jk}|.$$

*Proof.* We have,

$$\begin{aligned}
\|y\| &= \sum_{j=1}^r |\eta_j| = \sum_{j=1}^r \left| \sum_{k=1}^n a_{jk} \xi_k \right| \\
&\leq \sum_{j=1}^r \sum_{k=1}^n |a_{jk}| |\xi_k| \\
&\leq \sum_{j=1}^r \sum_{k=1}^n \left( \max_{k \in n} |a_{jk}| \right) |\xi_k| \\
&= \sum_{j=1}^r \left( \max_{k \in n} |a_{jk}| \right) \|x\|_1
\end{aligned}$$

and hence, we get a compatible norm as

$$\|A\| = \sum_{j=1}^r \left( \max_{k \in n} |a_{jk}| \right).$$

■



**15.** Show that for  $r = n$ , the norm in 14 is the natural norm corresponding to  $\|\cdot\|_1$  and  $\|\cdot\|_2$  as defined in the problem.

*Proof.* The norm is defined as

$$\|A\| = \max_{k \in [n]} \sum_{j=1}^n |a_{jk}|$$

From 14, we had

$$\|Ax\|_2 \leq \|A\| \|x\| \implies \|A\| \geq \frac{\|Ax\|_2}{\|x\|_1} \quad \forall x \neq 0 \implies \|A\| = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_1}$$

which is exactly the definition of the natural norm. ■

## 2.8 Linear Functionals

**1.** Show that the functionals in 2.8-7 and 2.8-8 are linear.

*Proof.* In 2.8-7, we have the functional

$$f_1(x) = x(t_0)$$

over the space  $C[a, b]$  and for a fixed  $t_0 \in [a, b]$ . Note that

$$f(\alpha x + \beta y) = (\alpha x + \beta y)(t_0) = \alpha x(t_0) + \beta y(t_0) = \alpha f(x) + \beta f(y)$$

and hence,  $f$  is linear.

In 2.8-8, we have the functional

$$f(x) = \sum_{j=1}^{\infty} \xi_j a_j$$

for  $x = (\xi_j) \in \ell^2$  and  $a = (a_j) \in \ell^2$ . Note that

$$f(\alpha x + \beta y) = \sum_{j=1}^{\infty} a_j(\alpha \xi_j + \beta \eta_j) = \alpha \sum_{j=1}^{\infty} a_j \xi_j + \beta \sum_{j=1}^{\infty} a_j \eta_j = \alpha f(x) + \beta f(y)$$

and hence,  $f$  is linear. ■

**2.** Show that the functionals defined on  $C[a, b]$  by

$$f_1(x) = \int_a^b x(t) y_0(t) dt \quad y_0 \in C[a, b]$$

$$f_2(x) = \alpha x(a) + \beta x(b)$$

are linear and bounded.

*Proof.* First, we have

$$f_1(\alpha x + \beta y) = \int_a^b (\alpha x + \beta y)(t) y_0(t) dt = \alpha \int_a^b x(t) y_0(t) dt + \beta \int_a^b y(t) y_0(t) dt = \alpha f_1(x) + \beta f_1(y)$$

and hence,  $f_1$  is linear. We now show  $f_1$  is bounded as follows:

$$|f_1(x)| = \left| \int_a^b x(t) y_0(t) dt \right| \leq \int_a^b |x(t)| |y_0(t)| dt \leq \int_a^b \left( \max_{t \in [a, b]} |x(t)| \right) |y_0(t)| dt = \|x\| \int_a^b |y_0(t)| dt$$

and hence, we have

$$\|f\| \leq \int_a^b |y_0(t)| dt$$

Now, we have

$$f_2(\gamma x + \delta y) = \alpha(\gamma x + \delta y)(a) + \beta(\gamma x + \delta y)(b) = \gamma(\alpha x(a) + \beta x(b)) + \delta(\alpha y(a) + \beta y(b)) = \gamma f_2(x) + \delta f_2(y),$$

and hence,  $f_2$  is linear. To show  $f_2$  is bounded, we have

$$|f_2(x)| = |\alpha x(a) + \beta x(b)| \leq |\alpha| \max_{t \in [a, b]} |x(t)| + |\beta| \max_{t \in [a, b]} |x(t)| = (|\alpha| + |\beta|) \|x\|$$

and hence,

$$\|f_2\| \leq (|\alpha| + |\beta|). \quad \text{■}$$

3. Find the norm of the linear functional  $f$  defined on  $C[-1, 1]$  by

$$f(x) = \int_{-1}^0 x(t) dt - \int_0^1 x(t) dt$$

*Proof.* Recall the definition of the norm:

$$\|f\| = \sup_{\|x\|=1} |f(x)|.$$

Consider some  $x$  with norm 1,

$$\|x\| = 1 \implies \max_{t \in [-1, 1]} |x(t)| = 1 \implies |x(t)| \leq 1 \forall t \in [-1, 1].$$

Now,

$$\|f\| = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\|=1} \left| \int_{-1}^0 x(t) dt - \int_0^1 x(t) dt \right| \leq \sup_{\|x\|=1} \int_{-1}^0 |x(t)| dt + \sup_{\|x\|=1} \int_0^1 |x(t)| dt \leq 2.$$

On the other hand, using  $|f(x)| \leq \|f\| \|x\|$ , for  $x(t) = 1 \forall t \in [-1, 0]$  and  $x(t) = -1 \forall t \in [0, 1]$  and hence,  $\|x\| = 1$ , we get

$$\|f\| \geq \frac{|f(x)|}{\|x\|} = |f(x)| = \left| \int_{-1}^0 1 dt - \int_0^1 -1 dt \right| = 2$$

and hence,  $\|f\| = 2$ . ■

4. Show that

$$f_1(x) = \max_{t \in [a, b]} x(t)$$

$$f_2(x) = \min_{t \in [a, b]} x(t)$$

define functionals on  $C[a, b]$ . Are they linear and bounded?

*Proof.* Note that both  $f_1$  and  $f_2$  are functionals since they take as input an element of the space, in this case,  $C[a, b]$ , and return the pointwise maximum (or minimum), which is an element of the underlying scalar field.

Now,

$$f_1(\alpha x + \beta y) = \max_{t \in [a, b]} (\alpha x + \beta y)(t) \leq \alpha \max_{t \in [a, b]} x(t) + \beta \max_{t \in [a, b]} y(t) = \alpha f_1(x) + \beta f_2(x)$$

and hence, the functional is not linear. Now, we have

$$|f(x)| = \left| \max_{t \in [a, b]} x(t) \right| \leq \max_{t \in [a, b]} |x(t)| = 1 \cdot \|x\|,$$

and hence  $\|f_1\| \leq 1$ . Now, choose  $x(t) = 1 \forall t \in [a, b]$ , then  $\|x\| = 1$  and

$$\|f_1\| \geq \frac{|f_1(x)|}{\|x\|} = \frac{|\max_{t \in [a, b]} x(t)|}{\|x\|} = 1,$$

and hence,  $\|f_1\| = 1$ . So,  $f_1$  is bounded.

Similarly, we have

$$f_2(\alpha x + \beta y) \geq \alpha f_2(x) + \beta f_2(y)$$

and hence,  $f_2$  is not linear. Now, we have

$$|f_2(x)| = \left| \min_{t \in [a, b]} x(t) \right| \leq \left| \max_{t \in [a, b]} x(t) \right| \leq \max_{t \in [a, b]} |x(t)| \leq 1 \cdot \|x\|$$

and hence,  $f_2$  is bounded. ■

5. Show that on any sequence space  $X$  we can define a linear functional  $f$  by setting  $f(x) = \xi_n$  ( $n$  fixed), where  $(\xi_j)$ . Is  $f$  bounded if  $X = \ell^\infty$ ?

*Proof.* First,  $f(x) = \xi_n$  is a functional since it takes as input a sequence  $x$  and outputs an element from the sequence, which belongs to the underlying scalar field. Also,

$$f(\alpha x + \beta y) = (\alpha x + \beta y)_n = \alpha \xi_n + \beta \eta_n = \alpha f(x) + \beta f(y).$$

Hence, it is linear. Also, if  $x \in \ell^\infty$ , we have

$$|f(x)| = |\xi_n| \leq \max_j |\xi_j| \leq 1 \cdot \|x\|_\infty$$

and hence,  $f$  is bounded. ■

**6.** The space  $C^1[a, b]$  or  $C'[a, b]$  is the normed space of all continuously differentiable functions on  $J = [a, b]$  with norm defined by

$$\|x\| = \max_{t \in [a, b]} |x(t)| + \max_{t \in [a, b]} |x'(t)|.$$

Show that the axioms of a norm are satisfied. Show that  $f(x) = x'(c)$ ,  $c = (a + b)/2$ , defines a bounded linear functional on  $C'[a, b]$ . Show that  $f$  is not bounded, considered as a functional on the subspace of  $C[a, b]$  which consists of all continuously differentiable functions.

*Proof.* First,  $\|x\|$  is clearly non-negative. Also, if  $x = 0$ , then  $\|x\| = 0$ . However, if  $\|x\| = 0$ , then

$$\max_{t \in [a, b]} |x(t)| = 0 \implies x(t) = 0 \forall t \in [a, b]$$

$$\max_{t \in [a, b]} |x'(t)| = 0 \implies x'(t) = 0 \forall t \in [a, b] \implies x(t) = k \forall t \in [a, b]$$

Hence,  $\|x\| = 0 \implies x = 0$ . Now, we have

$$\|\alpha x\| = \max_{t \in [a, b]} |\alpha x(t)| + \max_{t \in [a, b]} |\alpha x'(t)| = |\alpha| \|x\|.$$

And finally, we have

$$\|x + y\| = \max_{t \in [a, b]} |x(t) + y(t)| + \max_{t \in [a, b]} |x'(t) + y'(t)| \leq \max_{t \in [a, b]} |x(t)| + \max_{t \in [a, b]} |x'(t)| + \max_{t \in [a, b]} |y(t)| + \max_{t \in [a, b]} |y'(t)| = \|x\| + \|y\|.$$

Thus,  $\|x\|$  defines a valid norm. Now, define  $f(x) = x'(c)$ ,  $c = (a + b)/2$ . Then, we have

$$f(\alpha x + \beta y) = (\alpha x + \beta y)'(c) = \alpha x'(c) + \beta y'(c) = \alpha f(x) + \beta f(y)$$

and hence,  $f$  is linear. Also,

$$|f(x)| = |x'(c)| \leq \max_{t \in [a, b]} |x'(t)| \leq \max_{t \in [a, b]} |x(t)| + \max_{t \in [a, b]} |x'(c)| = \|x\|.$$

Hence,  $f$  is also bounded.

Now, consider  $f$  on the subspace of  $C[a, b]$  with all continuously differentiable functions. Here, the norm is defined as

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

Suppose  $\|x\| = 1 \implies |x(t)| \leq 1 \forall t \in [a, b]$ . We wish to find some  $x$  such that

$$x'((a + b)/2) \geq 1.$$

**TODO.** ■

**7.** If  $f$  is a bounded linear functional on a complex normed space, is  $\bar{f}$  linear and bounded? Here defined complex conjugate.

*Proof.* Consider  $\bar{f}(\alpha x) = \overline{f(\alpha x)} = \overline{\alpha f(x)} = \bar{\alpha} \bar{f}(x)$ . Hence, it is not linear. However,  $|\bar{f}(x)| = \left| \overline{f(x)} \right| = |f(x)| \leq \|f\| \|x\|$ . Since  $f$  is bounded,  $\bar{f}$  is also bounded. ■

**8.** The null space  $N(M^*)$  of a set  $M^* \subset X^*$  is defined to be the set of all  $x \in X$  such that  $f(x) = 0$  for all  $f \in M^*$ . Show that  $N(M^*)$  is a vector space.

*Proof.* Suppose  $x_1, x_2 \in N(M^*)$ . Then,  $f(x_1) = 0$  and  $f(x_2) = 0$  for all  $f \in M^*$ . Then, we have that for any  $f \in M^*$

$$f(\alpha x_1 + \beta x_2) = \alpha f(x_1) + \beta f(x_2) = 0$$

and hence,  $\alpha x_1 + \beta x_2 \in N(M^*)$ . The equality follows since by definition,  $X^*$  contains the linear functionals. Hence,  $N(M^*)$  is a vector space. ■

**9.** Let  $f \neq 0$  be any linear functional on a vector space  $X$  and  $x_0$  any fixed element of  $X - \mathcal{N}(f)$  where  $\mathcal{N}(f)$  is the null space of  $f$ . Show that any  $x \in X$  has a unique representation  $x = \alpha x_0 + y$ , where  $y \in \mathcal{N}(f)$ .

*Proof.* We have that  $y = x - \alpha x_0$ . Thus,

$$f(y) = f(x) - f(\alpha x_0) = f(x) - \alpha f(x_0)$$

Thus, setting  $\alpha = f(x)/f(x_0)$  results in  $f(y) = 0 \implies y \in \mathcal{N}(f)$ . To prove that it is unique, let  $x = \alpha_1 x_0 + y_1 = \alpha_2 x_0 + y_2$ . Then,

$$f(x) = \alpha_1 f(x_0) = \alpha_2 f(x_0)$$

since  $f(y_1) = f(y_2) = 0$ . Now, since  $x_0 \notin \mathcal{N}(f)$ , we have  $\alpha_1 = \alpha_2$  which results in  $y_1 = y_2$ . ■

**10.** Show that in 9, two elements  $x_1, x_2 \in X$  belong to the same element of the quotient space  $X/\mathcal{N}(f)$  if and only if  $f(x_1) = f(x_2)$ . Show that  $\text{codim } \mathcal{N}(f) = 1$ .

*Proof.* Suppose  $x_1, x_2 \in X$  belong to the same element of the quotient space  $X/\mathcal{N}(f)$ . Then, there exists some  $x \in X$  such that for some  $y_1, y_2 \in \mathcal{N}(f)$

$$x_1 = x + y_1, \quad x_2 = x + y_2.$$

Then,

$$f(x_1) = f(x + y_1) = f(x) + f(y_1) = f(x).$$

$$f(x_2) = f(x + y_2) = f(x) + f(y_2) = f(x).$$

Thus,  $f(x_1) = f(x_2)$ .

Now, suppose  $f(x_1) = f(x_2)$ . Then,

$$f(x_1) - f(x_2) = 0 = f(x_1 - x_2) \implies x_1 - x_2 \in \mathcal{N}(f)$$

Thus, it is easy to see that if  $x_1 - x_2 \in \mathcal{N}(f)$ , then  $x_1$  and  $x_2$  differ by an additive factor of some  $y$ ,  $y \in \mathcal{N}(f)$ , and hence, they belong to the same coset.

Since we can write any  $x = \alpha x_0 + y = \alpha(x_0 + \frac{1}{\alpha}y)$ ,  $y \in \mathcal{N}(f)$ , the codimension of the cosets is 1. ■

**11.** Show that two linear functionals  $f_1 \neq 0$  and  $f_2 \neq 0$  which are defined on the same vector space and have the same null space are proportional.

*Proof.* First, note that any element in the kernel space, say  $x \in \mathcal{N}(f_1)$  can be written as  $x_1 + \alpha x_2$  where  $\alpha = -\frac{f_1(x_1)}{f_1(x_2)}$ . Since  $\mathcal{N}(f_1) = \mathcal{N}(f_2)$ ,  $x \in \mathcal{N}(f_2)$ , and hence, we can also write  $x = x_1 + \alpha x_2$  where  $\alpha = -\frac{f_2(x_1)}{f_2(x_2)}$ . Thus, we get, for some constant  $k$

$$-\frac{f_1(x_1)}{f_1(x_2)} = -\frac{f_2(x_1)}{f_2(x_2)} = k \implies f_1(x) = k f_2(x) \quad \forall x \in X$$

Hence, the functions are proportional. ■

**12.** If  $Y$  is a subspace of a vector space  $X$  and  $\text{codim } Y = 1$ , then every element of  $X/Y$  is called a hyperplane parallel to  $Y$ . Show that for any linear functional  $f \neq 0$  on  $X$ , the set  $H_1 = \{x \in X \mid f(x) = 1\}$  is a hyperplane parallel to the null space  $\mathcal{N}(f)$  of  $f$ .

*Proof.* We wish to show that  $H_1 = \{x \in X \mid f(x) = 1\}$  is an element of  $X/\mathcal{N}(f)$  and that  $\text{codim } \mathcal{N}(f) = 1$ . We have shown that the codimension of the null space is 1 in 10. Now, we show that  $H_1$  is an element of  $X/\mathcal{N}(f)$ . Let  $x = \alpha x_0 + y$  for some fixed  $x_0$ ,  $y \in \mathcal{N}(f)$ , and  $f(x) = 1$ . Then,  $\alpha = \frac{1}{f(x_0)}$ . Note that  $f(\alpha x_0) = 1$ . Hence, we have expressed  $H_1 = \{\alpha x_0 + y \mid y \in \mathcal{N}(f)\}$  and hence, it is a coset parallel to the null space (shifted by  $\alpha x_0$ ). ■

**13.** If  $Y$  is a subspace of a vector space  $X$  and  $f$  is a linear functional on  $X$  such that  $f(Y)$  is not the whole scalar field of  $X$ , show that  $f(y) = 0 \quad \forall y \in Y$ .

*Proof.* Since  $Y$  is a subspace, the zero element belongs to  $Y$ . Hence, 0 has to be one of the values in the range of  $f$ . Suppose  $f(y_1) = k \neq 0$ . Then, for any  $k_0 \in K$ , we have that

$$f\left(\frac{k_0}{k}y\right) = k_0$$

or, in other words,  $k_0$  will also belong to the range of  $f$ . Since  $k_0$  was arbitrary, the entire scalar field will be the range of  $f$ . Hence,  $f(y) = 0 \forall y \in Y$  so that the complete scalar field is not the range. ■

**14.** Show that the norm  $\|f\|$  of a bounded linear functional  $f \neq 0$  on a normed space  $X$  can be interpreted geometrically as the reciprocal of the distance  $\tilde{d} = \inf\{\|x\| \mid f(x) = 1\}$  of the hyperplane  $H_1 = \{x \in X \mid f(x) = 1\}$  from the origin.

*Proof.* We know that

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}.$$

Suppose we condition on the hyperplane  $\{x \mid f(x) = 1\}$ . Then,

$$\|f\| = \sup_{x: f(x)=1, x \neq 0} \frac{1}{\|x\|} = \inf_{x: f(x)=1} \|x\|$$

where the last inequality follows from the fact that finding the least upper bound of  $1/\|x\|$  is the same as finding the lowest upper bound of  $\|x\|$ . ■

**15.** Let  $f \neq 0$  be a bounded linear functional on a real normed space  $X$ . Then, for any scalar  $c$ , we have a hyperplane  $H_c = \{x \in X \mid f(x) = c\}$  and  $H_c$  determines the two half spaces:

$$X_{c1} = \{x \mid f(x) \leq c\}, \quad X_{c2} = \{x \in X \mid f(x) \geq c\}.$$

Show that the closed unit ball lies in  $X_{c1}$  where  $c = \|f\|$ , but for no  $\epsilon > 0$ , the half space  $X_{c1}$  with  $c = \|f\| - \epsilon$  contains that ball.

*Proof.* The closed unit ball is the set of all points  $\tilde{B}(0, 1) = \{x \mid \|x\| \leq 1\}$ . Thus, for any  $x \in \tilde{B}(0, 1)$ , we have

$$f(x) \leq \|f\| \|x\| \leq \|f\|$$

and hence,  $\tilde{B}(0, 1) \subset X_{c1}$  where  $c = \|f\|$ .

Now, we also know that

$$\|f\| = \sup_{\|x\|=1} |f(x)|.$$

Now, from the definition of a supremum, we know that for any  $\epsilon > 0$ , there exists some  $x \in \tilde{B}(0, 1)$  such that  $f(x) > \|f\| - \epsilon$ . Hence,  $\tilde{B}(0, 1)$  is no longer a subset of  $X_{c1}$  for  $c = \|f\| - \epsilon$ . ■

## 2.9 Linear Operators and Functionals on Finite Dimensional Spaces

**1.** Determine the null space of the operator  $T : \mathbb{R}^3 \mapsto \mathbb{R}^2$  represented by

$$\begin{bmatrix} -1 & 3 & 2 \\ -2 & 1 & 0 \end{bmatrix}$$

*Proof.* Let the null space consist of all points  $[x, y, z]$ . Then, we have

$$\begin{bmatrix} 1 & 3 & 2 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{bmatrix} x + 3y + 2z \\ -2x + y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solving these equations results in  $y = 2x, z = -3.5x$ . Thus, any point of the form

$$\alpha \begin{bmatrix} 1 \\ 2 \\ -3.5 \end{bmatrix}$$

belongs to the null space. ■

2. Let  $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$  be defined by  $(\xi_1, \xi_2, \xi_3) \mapsto (\xi_1, \xi_2, -\xi_1 - \xi_2)$ . Find  $\mathcal{R}(T), \mathcal{N}(T)$ , and the matrix which represents  $T$ .

*Proof.* First, the matrix that represents  $T$  is given by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix}$$

Clearly, the null space  $\mathcal{N}(T)$  consists of  $\xi_3$ -axis, i.e any point of the form  $(0, 0, \xi_3)$ . Thus, the range  $\mathcal{R}(T) \subset \mathbb{R}^3$ . We can denote the range as

$$\mathcal{R}(T) = \{(x, y, z) \mid x + y + z = 0\}$$

and hence, the range is simply the plane where  $x + y + z = 0$ . ■

3. Find the dual basis of the basis  $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  for  $\mathbb{R}^3$ .

*Proof.* We know, the dual basis  $f_1, f_2, f_3$  is simply given by  $\delta_{jk}$  and hence, the basis is  $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ . ■

4. Let  $\{f_1, f_2, f_3\}$  be the dual basis of  $\{e_1, e_2, e_3\}$  for  $\mathbb{R}^3$ , where  $e_1 = (1, 1, 1)$ ,  $e_2 = (1, 1, -1)$ ,  $e_3 = (1, -1, -1)$ . Find  $f_1(x), f_2(x), f_3(x)$  for  $x = (1, 0, 0)$ .

*Proof.* First, let  $x = \sum_{i=1}^3 \xi_i e_i$ . Then,

$$f(x) = f\left(\sum_{i=1}^3 \xi_i e_i\right) = \sum_{i=1}^3 \xi_i f(e_i).$$

Denote,  $f(e_i) = f_i$ . Then, we call  $f_i$  the dual basis. Let  $f_i(e_j) = \mathbb{1}\{e_i = e_j\}$  simply be the Kronecker Delta. In other words,  $f_i(e_j) = 0$  and  $f_i(e_i) = 1$ .

Now,  $x = 0.5e_1 + 0.5e_3$ . Thus,  $f_1(x) = 0.5$  and  $f_3(x) = 0.5$ . However,  $f_2(x) = 0$ . ■

5. If  $f$  is a linear functional on a  $n$ -dimensional vector space  $X$ , what dimension can the null space  $\mathcal{N}(f)$  have?

*Proof.* If the functional  $f = 0$ , then the null space is the entire space, and hence, has dimension  $n$ . However, if  $f \neq 0$ , then, say  $x = \sum_{i=1}^n \alpha_i e_i \in \mathcal{N}(f)$  and  $x \neq 0$ . Thus,  $f(x) = 0$  implies that atleast one of the coefficients  $\alpha_i$  is non-zero. Thus, we can renormalize and set  $n-1$  coefficients, and fix  $n^{th}$  coefficient would be pre-determined. In other words, the null space is

$$\{(x_1, \dots, x_n) \mid x_1 + \dots + x_n = 0\}$$

which is an  $(n-1)$ -dimensional space. ■

6. Find a basis for the null space of the functional  $f$  defined on  $\mathbb{R}^3$  by  $f(x) = \xi_1 + \xi_2 - \xi_3$ .

*Proof.* First, clearly, the null space is given by

$$\{(\xi_1, \xi_2, \xi_3) \mid \xi_1 + \xi_2 = \xi_3\}.$$

Thus, if we fix any two coordinates, the third coordinate gets fixed automatically. Hence, the dimension of the null space is 2 and the basis is  $\{(1, 0, 1), (0, 1, 1)\}$  such that any vector written in terms of these basis

$$\alpha[1, 0, 1] + \beta[0, 1, 1] = [\alpha, \beta, \alpha + \beta]$$

satisfies the condition to be a part of the null space. ■

7. Find a basis for the null space of the functional  $f$  defined on  $\mathbb{R}^3$  by  $f(x) = \alpha_1 \xi_1 + \alpha_2 \xi_2 + \alpha_3 \xi_3$ , where  $\alpha_1 \neq 0$ .

*Proof.* Suppose  $\alpha_2 = \alpha_3 = 0$ . Then, the null space is given by the  $\xi_2 \xi_3$ -plane, i.e any points of the form  $(0, \xi_2, \xi_3)$ . Similarly, if  $\alpha_2$  or  $\alpha_3$  is zero, then the null space is given by  $\{(x, y, z) \mid z = \frac{-\alpha_3}{\alpha_1} x, y \in R\}$  and  $\{(x, y, z) \mid y = \frac{-\alpha_2}{\alpha_1} x, z \in R\}$  respectively. Instead of setting a condition on  $\alpha_2$  or  $\alpha_3$ , we can also express it as  $\{(x, 0, z) \mid \frac{-\alpha_3}{\alpha_1} x\}$  and  $\{(x, y, 0) \mid y = \frac{-\alpha_2}{\alpha_1} x\}$ . The union of these two sets results in the null space. ■

8. If  $Z$  is an  $(n-1)$ -dimensional subspace of an  $n$ -dimensional vector space  $X$ , show that  $Z$  is the null space of a suitable linear functional  $f$  on  $X$ , which is uniquely determined to within a scalar multiple.

*Proof.* Let the basis for  $Z$  be denoted by  $\{z_1, \dots, z_{n-1}\}$ , that can be extended to the basis of  $X$  as  $\{z_1, \dots, z_{n-1}, z_n\}$ . Define the dual basis to be  $\{f_1, \dots, f_n\}$  where  $f_i(z_j) = \mathbb{1}\{z_i = z_j\}$ . We define  $f$  as

$$f = \sum_{i=1}^n f(z_i) f_i.$$

Then, for some  $z \in Z = \sum_{i=1}^{n-1} \alpha_j z_j$ , we have

$$f(z) = \sum_{i=1}^n f(z_i) f_i(z) = \sum_{i=1}^{n-1} f(z_i) \alpha_i.$$

Thus, to show the existence of a linear functional that contains  $Z$  as a null space, it suffices to set  $f(z_i) = 0 \forall i \in [n-1]$ . Set  $f(z_n)$  to be some constant  $c$ , then, we have for any  $x = \sum_{i=1}^n \beta_i z_i$ ,

$$f(x) = \sum_{i=1}^n f(z_i) f_i(x) = \sum_{i=1}^n f(z_i) \beta_i = f(z_n) \beta_n = c \beta_n.$$

Thus, for a fixed  $x$ , all such functionals differ by a scalar multiple of the last coordinate. ■

**9.** Let  $X$  be the vector space of all real polynomials of a real variable and of degree less than a given  $n$ , together with the polynomial  $x = 0$ . Let  $f(x) = x^{(k)}(a)$ , the value of the  $k^{\text{th}}$  derivative of  $x \in X$  at a fixed  $a \in \mathbb{R}$ . Show that  $f$  is a linear functional in  $X$ .

*Proof.* If  $k > n$ , then  $f(x) = 0$  which is always linear. We deal with the more interesting case, where  $k \leq n$ . Suppose  $x, y \in X$ . Assume both  $x$  and  $y$  are polynomials of degree  $< k$ . Then,  $\alpha x + \beta y$  is also a polynomial of degree  $< k$ . Hence,  $f(\alpha x + \beta y) = 0$ , which is again linear since  $f(x) = f(y) = 0$ . So, we deal with the case where  $x$  and  $y$  have degrees at least  $k$ . In such a case,  $x^{(k)}$  and  $y^{(k)}$  will be polynomials of degrees  $\deg(x) - k$  and  $\deg(y) - k \leq (n - k)$ . Thus, we can define the functional  $f$  as an operator  $T$  that maps the input polynomials to a space with polynomials of degree at most  $(n - k)$  and then a functional evaluator  $f'$  that evaluates the polynomial at a fixed point  $a \in \mathbb{R}$ . Thus, we wish to show  $f = f'T$  is linear. Now, we first show  $T$  is linear:

$$T(\alpha x + \beta y) = (\alpha x + \beta y)^{(k)} = ((\alpha x + \beta y)^{(k-1)})' = \dots = (((\dots((\alpha x + \beta y)')') \dots)')'.$$

Now, repeatedly expanding  $((\dots((\alpha x + \beta y)')') \dots)')$  from the inside would result in  $\alpha x^{(k)} + \beta y^{(k)}$ . Thus  $T$  is linear. Also,

$$f'(\alpha x + \beta y) = (\alpha x + \beta y)(a) = \alpha x(a) + \beta y(a) = \alpha f'(x) + \beta f'(y)$$

and hence,  $f'$  is linear. Thus, the composition of  $T$  and  $f'$  is also linear. ■

**10.** let  $Z$  be a proper subspace of an  $n$ -dimensional vector space  $X$  and let  $x_0 \in X - Z$ . Show that there is a linear functional  $f$  on  $X$  such that  $f(x_0) = 1$  and  $f(x) = 0$  for all  $x \in Z$ .

*Proof.* First, let the basis of  $Z$  be  $\{z_1, \dots, z_m\}$ . Then,  $\{z_1, \dots, z_m, x_0\}$  is linearly independent since  $x_0 \in Z$ . We can extend this to a basis for  $X$  as  $\{z_1, \dots, z_m, x_0, x_{m+2}, \dots, x_n\}$ . Define a functional  $f$  such that  $f(z_i) = 0$  and  $f(x_0) = 1$ . Thus, we can define  $f$  as

$$f = \sum_{i=1}^m f(z_i) \mathbb{1}\{\cdot = z_i\} + f(x_0) \mathbb{1}\{\cdot = x_0\} + \sum_{i=m+2}^n f(x_i) \mathbb{1}\{\cdot = x_i\} = f(x_0) \mathbb{1}\{\cdot = x_0\} + \sum_{i=m+2}^n f(x_i) \mathbb{1}\{\cdot = x_i\}$$

Then, for any  $z \in Z = \sum_{i=1}^m \beta_i z_i$ , we have  $f(z) = 0$ . Hence, we have defined a function by specifying the values on the basis vectors. ■

**11.** If  $x \neq y \in X$  and  $X$  is finite dimensional, show that there is a linear functional  $f$  such that  $f(x) \neq f(y)$ .

*Proof.* Suppose not. Then,  $f(x) = f(y) \implies f(x) - f(y) = 0 = f(x - y)$  for all possible functionals  $f \in X^*$ . Thus,  $x - y = 0$ , which is a contradiction. ■

**12.** If  $f_1, \dots, f_p$  are linear functionals on an  $n$ -dimensional vector space  $X$ , where  $p < n$ , show that there is a vector  $x \neq 0$  in  $X$  such that  $f_1(x) = 0, \dots, f_p(x) = 0$ .

*Proof.* Let  $x = \sum_{i=1}^n \alpha_i e_i$ . Then, using the facts that  $f_1(x) = 0, \dots, f_p(x) = 0$ , we get

$$\begin{bmatrix} f_1(e_1) & f_1(e_2) & \dots & f_1(e_n) \\ f_2(e_1) & f_2(e_2) & \dots & f_2(e_n) \\ \vdots & \vdots & \dots & \vdots \\ f_p(e_1) & f_p(e_2) & \dots & f_p(e_n) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

which results in  $p$  equations with  $n$  unknowns. Thus, there has to be a non-trivial solution. Another way to look at it is as follows: suppose  $x = 0$  is the only such vector. Then, define a mapping  $T : X \mapsto \mathbb{R}^p$  such that  $T(x) = (f_1(x), \dots, f_p(x))$ . Then,  $T(0) = 0$ , but there exists no other vector  $x$  such that  $T(x) = 0$ . Hence,  $T$  is injective. Thus,  $\dim X \leq p \implies n \leq p$ . This is a contradiction. ■

**13.** Let  $Z$  be a proper subspace of an  $n$ -dimensional vector space  $X$ , and let  $f$  be a linear functional on  $Z$ . Show that  $f$  can be extended linearly to  $X$ , that is, there is a linear functional  $\tilde{f}$  on  $X$  such that  $\tilde{f}|_Z = f$ .

*Proof.* Let  $\{z_1, \dots, z_m\}$  be the basis for  $Z$  that can be extended to the basis of  $X$ , as in  $\{z_1, \dots, z_m, z_{m+1}, \dots, z_n\}$ . Define  $\{f_1, \dots, f_n\}$  to be the dual basis as  $f_i(z_j) = \mathbb{1}\{z_i = z_j\}$ . Thus, we can define  $\tilde{f}$  as

$$\tilde{f} = \sum_{i=1}^n f(z_i) f_i.$$

Now, set  $\tilde{f}(z_i) = f(z_i) \forall i \in [m]$ . Then, for any  $z = \sum_{i=1}^m \alpha_i z_i \in Z$ , we have

$$\tilde{f}(z) = \sum_{i=1}^n \tilde{f}(z_i) f_i \left( \sum_{i=1}^m \alpha_i z_i \right) = \sum_{i=1}^m f(z_i) \alpha_i$$

where we use the fact that  $\tilde{f}(z_i) = f(z_i) \forall i \in [m]$ . Thus,  $\tilde{f}$  restricted to  $Z$  results in  $f$ , while  $\tilde{f}$  is linear. ■

**14.** Let the functional  $f$  on  $\mathbb{R}^2$  be defined by  $f(x) = 4\xi_1 - 3\xi_2$ . Regard  $\mathbb{R}^2$  as a subspace of  $\mathbb{R}^3$  given by  $\xi_3 = 0$ . Determine all linear extensions  $\tilde{f}$  of  $f$  from  $\mathbb{R}^2$  to  $\mathbb{R}^3$ .

*Proof.* Let  $\{f_1, f_2, f_3\}$  represent the dual basis of the extension  $\tilde{f}$ . Thus,

$$\tilde{f} = \tilde{f}(e_1) f_1 + \tilde{f}(e_2) f_2 + \tilde{f}(e_3) f_3.$$

Now, for any  $x \in \mathbb{R}^2$  of the form  $(\xi_1, \xi_2, 0)$ , we have

$$\tilde{f}(x) = \tilde{f}(e_1) \xi_1 + \tilde{f}(e_2) \xi_2$$

Since the restriction of  $\tilde{f}$  on  $\mathbb{R}^2$  should be the same as  $f$ , we get  $\tilde{f}(e_1) = 4$  and  $\tilde{f}(e_2) = -3$ . Now, for any point  $x \in \mathbb{R}^3$  of the form  $(\xi_1, \xi_2, \xi_3)$ , we have

$$\tilde{f}(x) = 4\xi_1 - 3\xi_2 + \tilde{f}(e_3) \xi_3.$$

Thus, we have ensured that the restriction of  $\tilde{f}$  is  $f$  on  $\mathbb{R}^2$ , and setting  $\tilde{f}(e_3)$  to various values would result in various different linear functionals. ■

**15.** Let  $Z \subset \mathbb{R}^3$  be the subspace represented by  $\xi_2 = 0$  and let  $f$  on  $Z$  be defined by  $f(x) = (\xi_1 - \xi_3)/2$ . Find a linear extension  $\tilde{f}$  of  $f$  to  $\mathbb{R}^3$  such that  $\tilde{f}(x_0) = k$ , where  $x_0 = (1, 1, 1)$ . Is  $\tilde{f}$  unique?

*Proof.* We will use the ideas from 14. Define  $\{f_1, f_2, f_3\}$  to be the dual basis and  $\tilde{f}$  to be the extension of  $f$  from  $Z$  to  $\mathbb{R}^3$ . Then, we have

$$\tilde{f} = \tilde{f}(e_1) f_1 + \tilde{f}(e_2) f_2 + \tilde{f}(e_3) f_3.$$

Now, evaluating  $\tilde{f}$  on  $x \in Z$  should result in the functional  $f$ , or in other words, for  $x = (\xi_1, 0, \xi_3)$ , we have

$$\tilde{f}(x) = \tilde{f}(e_1) \xi_1 + \tilde{f}(e_3) \xi_3 = f(x) \implies \tilde{f}(e_1) = 1/2, \tilde{f}(e_3) = -1/2.$$

Now, to obtain  $\tilde{f}(x_0)$ , where  $x_0 = (1, 1, 1)$ , we get

$$\tilde{f}(x_0) = \tilde{f}(e_1) + \tilde{f}(e_2) + \tilde{f}(e_3) = \tilde{f}(e_2) = k \implies \tilde{f}(e_2) = k$$

Thus, the extension  $\tilde{f}$  is given by  $(\xi_1 + 2\xi_2 - \xi_3)/2$ . Yes, this extension is unique. ■



## 2.10 Normed Spaces of Operators. Dual Space

**1.** What is the zero element of the vector space  $B(X, Y)$ ? The inverse of a  $T \in B(X, Y)$  in the sense of Def 2.1-1?

*Proof.* The zero element is simply the zero operator:  $T : X \mapsto Y, Tx = 0$ . The definition 2.1-1 defines a vector space which should constitute addition and multiplication operations. The additive inverse then is  $-T$ . ■

**2.** The operators and functionals considered in the text are defined on the entire space  $X$ . Show that without that assumption, in the case of functionals we still have the following theorem. If  $f$  and  $g$  are bounded linear functionals with domains in a normed space  $X$ , then for any nonzero scalars  $\alpha$  and  $\beta$  the linear combination  $h = \alpha f + \beta g$  is a bounded linear functional with domain  $\mathcal{D}(h) = \mathcal{D}(f) \cap \mathcal{D}(g)$ .

*Proof.* If  $\alpha \neq 0$  and  $\beta \neq 0$ , then  $x \in \mathcal{D}(h) \implies x \in \mathcal{D}(f)$  and  $x \in \mathcal{D}(g)$ . Thus,  $\mathcal{D}(h) \subset \mathcal{D}(f) \cap \mathcal{D}(g)$ . Similarly, for some  $x \in \mathcal{D}(f) \cap \mathcal{D}(g)$ ,  $f(x)$  and  $g(x)$  are defined, and hence,  $h(x)$  is defined. Thus,  $\mathcal{D}(f) \cap \mathcal{D}(g) \subset \mathcal{D}(h)$ . This gives us  $\mathcal{D}(h) = \mathcal{D}(f) \cap \mathcal{D}(g)$ . Now, for some  $x, y \in \mathcal{D}(h)$ , we have

$$h(\gamma x + \delta y) = (\alpha f + \beta g)(\gamma x + \delta y) = \alpha f(\gamma x + \delta y) + \beta g(\gamma x + \delta y) = \alpha \gamma f(x) + \beta \gamma g(x) + \alpha \delta f(y) + \beta \delta g(y) = \gamma h(x) + \delta h(y)$$

Hence,  $h$  is linear. Further, we have

$$|h(x)| = |(\alpha f + \beta g)(x)| \leq \|\alpha f + \beta g\| \|x\| \leq (|\alpha| \|f\| + |\beta| \|g\|) \|x\|$$

Since,  $f$  and  $g$  is bounded, we have  $\|h\| = |\alpha| \|f\| + |\beta| \|g\|$  and hence,  $h$  is also bounded. ■

**3.** Extend the theorem in 2 to bounded linear operators  $T_1$  and  $T_2$ .

*Proof.* Let  $T = \alpha T_1 + \beta T_2$ . Once again, let  $x \in \mathcal{D}(T)$ , then  $x \in \mathcal{D}(T_1)$  and  $x \in \mathcal{D}(T_2)$  and hence,  $\mathcal{D}(T) \subset \mathcal{D}(T_1) \cap \mathcal{D}(T_2)$ . Let  $x \in \mathcal{D}(T_1) \cap \mathcal{D}(T_2)$ , then both  $T_1 x$  and  $T_2 x$  is defined, and hence,  $Tx$  is defined. Hence,  $x \in \mathcal{D}(T)$  resulting in  $\mathcal{D}(T_1) \cap \mathcal{D}(T_2) \subset \mathcal{D}(T)$ . Thus, we get  $\mathcal{D}(T) = \mathcal{D}(T_1) \cap \mathcal{D}(T_2)$ . Now, for some  $x, y \in \mathcal{D}(T)$ , we have

$$T(\gamma x + \delta y) = (\alpha T_1 + \beta T_2)(\gamma x + \delta y) = \alpha T_1(\gamma x + \delta y) + \beta T_2(\gamma x + \delta y) = \alpha \gamma T_1 x + \beta \gamma T_2 x + \alpha \delta T_1 y + \beta \delta T_2 y = \gamma T x + \delta T y$$

and hence,  $T$  is linear. Also,

$$\|Tx\| = \|(\alpha T_1 + \beta T_2)x\| \leq \|\alpha T_1 + \beta T_2\| \|x\| \leq (|\alpha| \|T_1\| + |\beta| \|T_2\|) \|x\|$$

and hence,  $\|T\| = |\alpha| \|T_1\| + |\beta| \|T_2\|$ . Since  $T_1$  and  $T_2$  is bounded,  $T$  is bounded. ■

**4.** Let  $X$  and  $Y$  be normed spaces and  $T_n : X \mapsto Y$  be bounded linear operators. Show that convergence  $T_n \rightarrow T$  implies that for every  $\epsilon > 0$ , there is a  $N$  such that for all  $n > N$  all  $x$  in a given closed ball  $\|T_n x - Tx\| < \epsilon$ .

*Proof.* Let  $x \in X$ . Now, we have

$$\|T_n x - Tx\| \leq \|T_n - T\| \|x\|.$$

Since  $T_n \rightarrow T$ , we have for every  $\epsilon' > 0$ ,  $\exists N$  such that for all  $n > N$ ,  $\|T_n - T\| < \epsilon'$ . Substituting this back, we get that for every  $n > N$ ,

$$\|T_n x - Tx\| \leq \epsilon' \|x\| = \epsilon.$$

Thus, putting  $\epsilon' = \epsilon \|x\|$  prove the claim. ■

**5.** Show that 2.8-5 is in agreement with 2.10-5.

*Proof.* 2.8-5 says the dot product functional  $f : \mathbb{R}^3 \mapsto \mathbb{R}^3$  with one factor keep fixed can be defined as

$$f(x) = x \cdot a = \sum_{i=1}^3 \xi_i \alpha_i$$

where  $a = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$ . In such a case,  $\|f\| = \|a\|$ . Also, 2.10-5 says that the dual space of  $\mathbb{R}^n$  is  $\mathbb{R}^n$ .

Let  $n = 3$ . Denote the dot product functional  $f \in \mathbb{R}^{3'}$  as

$$f(x) = \sum_{i=1}^3 \xi_i f(e_i)$$

and from the definition of the dot product, we have

$$f(x) = \sum_{i=1}^3 \xi_i \alpha_i.$$

Comparing values gives us  $f(e_i) = \alpha_i \forall i \in [3]$ . Denote  $a = (\alpha_1, \alpha_2, \alpha_3)$ . Now, from 2.10-5, there exists an isomorphic mapping  $T$  such that  $Tf = a$  and hence,  $\|f\| = \|a\|$ . This is in direct accordance with 2.8-5. ■

6. If  $X$  is the space of all ordered  $n$ -tuples of real numbers and  $\|x\| = \max_i |\xi_i|$ , where  $x = (\xi_1, \dots, \xi_n)$ , what is the corresponding norm on the dual space  $X'$ .

*Proof.* Let  $f \in X'$ . Denote the basis of  $X$  as  $\{e_1, \dots, e_n\}$ . We can write  $f$  as

$$f(x) = \sum_{i=1}^n \xi_i f(e_i).$$

Then, we have

$$|f(x)| \leq \sum_{i=1}^n |\xi_i f(e_i)| \leq \left( \max_{i \in [n]} |\xi_i| \right) \sum_{i=1}^n |f(e_i)| = \left( \sum_{i=1}^n |f(e_i)| \right) \|x\|.$$

Thus, we have

$$\|f\| = \sup_{\|x\|=1} |f(x)| \leq \sum_{i=1}^n |f(e_i)|.$$

At the same time, for  $x = \sum_{i=1}^n \frac{|f(e_i)|}{f(e_i)} e_i$ , we have,  $\|x\| = 1$ , and hence,

$$|f(x)| = \left| \sum_{i=1}^n |f(e_i)| \right| = \sum_{i=1}^n |f(e_i)| \leq \|f\| \|x\| \implies \|f\| \geq \sum_{i=1}^n |f(e_i)|$$

and hence, we get

$$\|f\| = \sum_{i=1}^n |f(e_i)|$$

■

7. What conclusion can we draw from 2.10-6 with respect to the space  $X$  of all ordered  $n$ -tuples of real numbers.

*Proof.* If the norm on  $X$  is  $l_\infty$  norm, then the norm on the dual space is  $l_1$  norm. Moreover, we can find a mapping  $T : X' \mapsto X$  that is norm-preserving. ■

8. Show that the dual space of the space  $c_0$  is  $\ell^1$ .

*Proof.* Recall that the space  $c_0$  is a subspace of  $\ell^\infty$  consisting of sequences of scalars converging to zero. Let  $(x_n)$  be some sequence in  $c_0$ , such that each  $x_m = (\xi_1^{(m)}, \xi_2^{(m)}, \dots)$  and  $\xi_n^{(m)} \rightarrow 0$ .

Now, for the  $\ell^\infty$  space, we can have a Schauder basis  $(e_1, e_2, \dots)$  such that  $x_m = \sum_{i=1}^\infty \xi_i^{(m)} e_i$ , and we have a sequence  $(x_1, x_2, \dots)$ . Define a functional  $f$  such that

$$f(x_n) = \sum_{i=1}^\infty f(e_i) \xi_i^{(n)}.$$

Then, we have

$$|f(x_n)| \leq \sum_{i=1}^n |\xi_i^{(n)} f(e_i)| \leq \left( \max_{i \in [n]} |\xi_i^{(n)}| \right) \sum_{i=1}^n |f(e_i)| = \left( \sum_{i=1}^n |f(e_i)| \right) \|x\|.$$

Thus,

$$\sup_{\|x\|=1} |f(x)| \leq \sum_{i=1}^n |f(e_i)|.$$

Now, at the same time, define  $\xi_i^{(n)} = \frac{|f(e_i)|}{f(e_i)}$  if  $i \leq k$  and  $f(e_i) \neq 0$ , else  $\xi_i^{(n)} = 0$ . This gives us  $\|x\| = 1$ , and hence, we have

$$|f(x_n)| = \sum_{i=1}^k |f(e_i)| \leq \|f\| \|x\| \implies \|f\| \geq \sum_{i=1}^k |f(e_i)|.$$

Since  $k$  is arbitrary, taking  $k \rightarrow \infty$ , we get that

$$\|f\| \geq \sum_{i=1}^\infty |f(e_i)|.$$

Consider the point  $c = (f(e_1), f(e_2), \dots)$ . Then, we have

$$\|f\| = \|c\|_1.$$

Hence, we can define a mapping  $T : c'_0 \mapsto \ell^1$  such that it is norm preserving. ■

**9.** Show that a linear functional  $f$  on a vector space  $X$  is uniquely determined by its values on a Hamel basis for  $X$ .

*Proof.* Let the Hamel basis be defined as  $\{e_1, e_2\}$ , and choose values  $\gamma_i$  corresponding to  $e_i$ . Define two functionals  $f$  and  $g$  with the same values on the Hamel basis. Then, for some fixed  $x = \sum \xi_i e_i$ , we have

$$f(x) = \sum \gamma_i \xi_i, g(x) = \sum \gamma_i \xi_i \implies (f - g)(x) = 0 \forall x \in X \implies f = g$$

Thus, it is a unique functional. ■

**10.** Let  $X$  and  $Y \neq \{0\}$  be normed spaces, where  $\dim X = \infty$ . Show that there is at least one unbounded linear operator  $T : X \mapsto Y$ .

*Proof.* Consider a map  $T : X \mapsto Y$  such that  $Tx_n = ny$  for some fixed  $y$  such that  $\|y\| = 1$ . Also, WLOG, assume  $\|x_i\| = 1 \forall i \in [n]$ . Then,

$$\|T\| = \sup_{\|x\|=1} \|Tx\| = n \|y\|$$

which is unbounded for each  $n$ . ■

**11.** If  $X$  is a normed space and  $\dim X = \infty$ , show that the dual space  $X'$  is not identical with the algebraic dual space  $X^*$ .

*Proof.* First, set  $Y = \mathbb{R}$  or  $\mathbb{C}$ , so that we can talk about operators. From 10, we see that when  $\dim X = \infty$ , there exists an unbounded linear operator, and hence  $X^* \neq X'$ . ■

**12.** Examples in the text can be used to prove the completeness of certain spaces. How? And for what spaces?

*Proof.* We can use the concept of isometric spaces. Since isometries are norm preserving, achieving completeness in one space is equivalent to achieving completeness in the other. Let  $X$  and  $\tilde{X}$  be isometric spaces such that  $T : X \mapsto \tilde{X}$  is an isometry and is bounded. For some Cauchy sequence  $(x_n) \in X$ , we have that for every  $\epsilon > 0$ ,  $\exists N(\epsilon)$  such that for all  $m, n > N(\epsilon)$ , we have  $\|x_m - x_n\| < \epsilon$ . Thus, we also have

$$\|Tx_m - Tx_n\| < \|T\| \epsilon.$$

Hence,  $(Tx_n)$  is also Cauchy. Further, since  $X$  is complete, say  $(x_n) \rightarrow x$ , and hence,  $\|Tx_n - Tx\| < \|T\| \|x_n - x\| = \|T\| \epsilon$ . Thus, the sequence  $(Tx_n)$  also converges, and hence,  $\tilde{X}$  is complete. ■

**13.** Let  $M \neq \emptyset$  be any subset of a normed space  $X$ . The annihilator  $M^a$  of  $M$  is defined to be the set of all bounded linear functionals on  $X$  which are zero everywhere on  $M$ . Thus,  $M^a$  is a subset of the dual space  $X'$  of  $X$ . Show that  $M^a$  is a vector subspace of  $X'$  and is closed. What are  $X^a$  and  $\{0\}^a$ ?

*Proof.* Since  $M^a$  consists of bounded linear functionals, it is easy to see that addition and multiplication operations hold. Hence, it is a vector subspace. Now, let  $f \in \overline{M^a}$ . Then, by the an accumulation point, there exists some sequence  $f_n \in M^a$  such that  $f_n \rightarrow f$ . We wish to show that  $f \in M^a$ .  $f_n \rightarrow f$ , we have that for all  $\epsilon > 0$ ,  $\exists N(\epsilon)$  such that for all  $n > N(\epsilon)$ ,

$$\|f_n - f\| < \epsilon.$$

Thus, we have  $f = (f - f_n) + f_n \implies \|f\| \leq \epsilon + \|f_n\|$ , and hence,  $f$  is bounded. Further, for some  $m \in M$ , we have

$$\|f_n m - f m\| \leq \|f_n - f\| \|m\| \implies \lim_{n \rightarrow \infty} \|f_n m - f m\| = 0$$

and hence,  $f m = 0$ , thus,  $f \in M^a$ .

Now,  $X^a$  will be the bounded linear functionals that are zero everywhere on  $X$ , and hence,  $X^a = \{0\}$ . Similarly,  $\{0\}^a$  is the subset of bounded linear functionals that are zero on  $\{0\}$ , and hence,  $\{0\}^a = X'$ . ■

**14.** If  $M$  is a  $m$ -dimensional subspace of an  $n$ -dimensional normed space  $X$ , show that  $M^a$  is  $(n-m)$ -dimensional subspace of  $X'$ . Formulate this as a theorem about solutions of a system of linear equations.

*Proof.* Let the basis for  $M$  be  $\{e_1, \dots, e_m\}$  that is extended to the basis of  $X$  as  $\{e_1, \dots, e_m, e_{m+1}, \dots, e_n\}$ . Define a functional  $f$  such that it has the dual basis  $\{f_1, \dots, f_n\}$ . We have

$$f = \sum_{i=1}^n f(e_i) f_i.$$

If  $f \in M^a$ , then for any  $m \in M$ ,  $f(m) = 0$ . Since this is true for all  $m \in M$ , setting  $f(e_i) = 0 \forall i \in [m]$ , ensures that  $f(m) = 0$  for any  $m = \sum_{i=1}^m \alpha_i e_i$ . Thus,  $f$  is now  $(n-m)$ -dimensional. ■

**15.** Let  $M = \{(1, 0, -1), (1, -1, 0), (0, 1, -1)\} \subset \mathbb{R}^3$ . Find a basis for  $M^a$ .

*Proof.* For any  $x = (\xi_1, \xi_2, \xi_3)$ , define

$$f(x) = \sum_{i=1}^3 f(e_i)\xi_i$$

If  $f \in M^a$ , we have  $f(x) = 0 \forall x \in M$ . This gives us  $f(e_1) = f(e_2) = f(e_3)$ . We know that  $\mathbb{R}^{3'}$  is isomorphic to  $\mathbb{R}^3$  and hence, we can define a mapping  $T : \mathbb{R}^{3'} \mapsto \mathbb{R}^3$ ,  $Tf \mapsto (f(e_1), f(e_2), f(e_3))$ . Since,  $f \in M^a \implies f(e_1) = f(e_2) = f(e_3)$ , we have that the basis is simply  $\{(1, 1, 1)\} \in \mathbb{R}^3$ . Thus, any vector of the form  $\alpha(1, 1, 1) \in \mathbb{R}^3$  uniquely defines a functional in  $M^a$ . ■

### 3 Inner Product Spaces. Hilbert Spaces.

#### 3.1 Inner Product Space. Hilbert Space.

1. Prove (4)

*Proof.* We wish to prove

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

We know that  $\|x\| = \langle x, x \rangle^{1/2}$ , and hence,

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle + \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle \\ &= 2(\langle x, x \rangle + \langle y, y \rangle) \\ &= 2(\|x\|^2 + \|y\|^2) \end{aligned}$$

2. If  $x \perp y$  in an inner product space  $X$ , show that

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

*Extend the formula to  $m$  mutually orthogonal vectors.*

*Proof.* Using the definitions of norms ( $\|x\| = \langle x, x \rangle^{1/2}$ ) and orthogonality ( $x \perp y \iff \langle x, y \rangle = 0$ ), we get

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \langle x, x \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2.$$

where we use the fact that  $\langle x, y \rangle = 0 \implies \overline{\langle x, y \rangle} = 0 = \langle y, x \rangle$ .

For  $m$  mutually orthogonal vectors, this gets extended as

$$\left\| \sum_{i=1}^m x_i \right\|^2 = \sum_{i=1}^m \|x_i\|^2$$

3. If an inner product space  $X$  is real, show that conversely, the given relation in 2 implies  $x \perp y$ . Show that this may not hold if  $X$  is complex.

*Proof.* If  $X$  is real, then, we have that  $\langle x, y \rangle = \langle y, x \rangle$ , and hence,

$$0 = \langle x + y, x + y \rangle - \langle x, x \rangle - \langle y, y \rangle = \langle x, y \rangle + \langle y, x \rangle \implies \langle x, y \rangle = 0 \implies x \perp y.$$

However, if  $X$  is complex, then

$$0 = \langle x + y, x + y \rangle - \langle x, x \rangle - \langle y, y \rangle = \langle x, y \rangle + \langle y, x \rangle = \langle x, y \rangle + \overline{\langle x, y \rangle} \implies \Re \langle x, y \rangle = 0 \not\Rightarrow x \perp y.$$

4. If an inner product space  $X$  is real, show that the condition  $\|x\| = \|y\|$  implies  $\langle x + y, x - y \rangle = 0$ . What does this mean geometrically if  $X = \mathbb{R}^2$ ? What does the condition imply if  $X$  is complex?

*Proof.* If  $X$  is real, we have

$$\langle x + y, x - y \rangle = \langle x, x \rangle + \langle y, x \rangle - \langle x, y \rangle - \langle y, y \rangle = 0$$

since  $\langle x, x \rangle = \langle y, y \rangle$  and since  $X$  is real,  $\langle x, y \rangle = \langle y, x \rangle$ . Geometrically, this says that  $x + y$  and  $x - y$  are perpendicular. If  $X$  is complex, we have

$$\langle x + y, x - y \rangle = \langle x, x \rangle + \langle y, x \rangle - \langle x, y \rangle - \langle y, y \rangle = 0 \implies \langle x, y \rangle = \langle y, x \rangle = \overline{\langle x, y \rangle} \implies \Im \langle x, y \rangle = 0.$$

5. Verify that for any elements in an inner product space

$$\|z - x\|^2 + \|z - y\|^2 = \frac{1}{2} \|x - y\|^2 + 2 \left\| z - \frac{1}{2}(x + y) \right\|^2.$$

Show that this identity can also be obtained from the parallelogram equality.

*Proof.* We have that

$$\begin{aligned}
\|z - x\|^2 + \|z - y\|^2 &= 2\langle z, z \rangle + \langle x, x \rangle + \langle y, y \rangle - \langle x, z \rangle - \langle z, x \rangle - \langle y, z \rangle - \langle z, y \rangle \\
&= 2\langle z, z \rangle - \langle x + y, z \rangle - \langle z, x + y \rangle + \langle x, x \rangle + \langle y, y \rangle \\
&= 2\left\langle z - \frac{1}{2}(x + y), z - \frac{1}{2}(x + y) \right\rangle - \frac{1}{2}\langle x + y, x + y \rangle + \langle x, x \rangle + \langle y, y \rangle \\
&= 2\left\|z - \frac{1}{2}(x + y)\right\|^2 + \frac{1}{2}(\langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle) \\
&= 2\left\|z - \frac{1}{2}(x + y)\right\|^2 + \frac{1}{2}\|x - y\|^2
\end{aligned}$$

Now, recall the parallelogram law:

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Substituting  $x = z - \frac{1}{2}(x + y)$  and  $y = \frac{1}{2}(x + y)$  results in the same claim. ■

**6.** Let  $x \neq 0$  and  $y \neq 0$ . (a) If  $x \perp y$ , show that  $\{x, y\}$  is a linearly independent set. (b) Extend the result to mutually orthogonal non-zero vectors  $\{x_1, \dots, x_m\}$ .

*Proof.* For some  $\alpha$  and  $\beta$  consider the sum

$$\alpha x + \beta y = 0.$$

Taking an inner product with  $x$  on both sides results in

$$\langle \alpha x + \beta y, x \rangle = \alpha \langle x, x \rangle = 0 \implies x = 0 \text{ or } \alpha = 0.$$

Similarly, taking an inner product with  $y$  on both sides results in  $y = 0$  or  $\beta = 0$ . Since  $x \neq 0$  and  $y \neq 0$ , we have  $\alpha = \beta = 0$ . Thus, the vectors are linearly independent.

A similar result can be shown for  $m$  non-zero mutually orthogonal vectors: consider the sum

$$\sum_{i=1}^m \alpha_i x_i = 0.$$

Thus, an inner product with  $x_j$  results in

$$\left\langle \sum_{i=1}^m \alpha_i x_i, x_j \right\rangle = \alpha_j x_j = 0 \implies \alpha_j = 0$$

and in this way, we can show that  $\{x_1, \dots, x_m\}$  form a linearly independent set. ■

**7.** If in an inner product space,  $\langle x, u \rangle = \langle x, v \rangle$  for all  $x$ , show that  $u = v$ .

*Proof.* First, note that

$$\langle x, y - z \rangle = \overline{\langle y - z, x \rangle} = \overline{\langle y, x \rangle - \langle z, x \rangle} = \overline{\langle y, x \rangle} - \overline{\langle z, x \rangle} = \langle x, y \rangle - \langle x, z \rangle.$$

Thus, we are given that  $\langle x, u - v \rangle = 0$  for all  $x$ . Then, this holds for  $x = u - v$  as well. We get

$$\langle u - v, u - v \rangle = 0 = \|u - v\|^2 \implies \|u - v\| = 0 \implies u - v = 0 \implies u = v. \quad \text{■}$$

**8.** Prove (9).

*Proof.* We wish to show that for a real inner product

$$\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$

Now, we know that

$$\|a + b\|^2 = \|a\|^2 + \|b\|^2 + \langle a, b \rangle + \langle b, a \rangle.$$

Thus, expanding the RHS gives

$$\frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) = \frac{1}{4}(\|x\|^2 + \|y\|^2 + \langle x, y \rangle + \langle y, x \rangle - \|x\|^2 - \|y\|^2 + \langle x, y \rangle + \langle y, x \rangle).$$

Finally, since  $\langle x, y \rangle = \langle y, x \rangle$  for real inner products, we get the resulting claim. ■

**9. Prove (10).**

*Proof.* We wish to show that for a complex inner product, we have

$$\Re \langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

$$\Im \langle x, y \rangle = \frac{1}{4}(\|x + \iota y\|^2 - \|x - \iota y\|^2).$$

From 8, we get

$$\frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) = \frac{1}{4}(\langle x, y \rangle + \langle y, x \rangle + \langle x, y \rangle + \langle y, x \rangle) = \frac{1}{4} = \frac{1}{2}(\langle x, y \rangle + \overline{\langle x, y \rangle}) = \Re \langle x, y \rangle.$$

Similarly, using 8 and substituting  $y = \iota y$ , we get

$$\frac{1}{4}(\|x + \iota y\|^2 - \|x - \iota y\|^2) = \frac{1}{2}(\langle x, \iota y \rangle + \langle \iota y, x \rangle) = \frac{1}{2}(\iota \langle x, y \rangle + \iota \overline{\langle x, y \rangle}) = \frac{i}{2}(\langle x, y \rangle - \overline{\langle x, y \rangle}) = \Im \langle x, y \rangle \iota.$$

■

**10.** Let  $z_1$  and  $z_2$  denote complex numbers. Show that  $\langle z_1, z_2 \rangle = z_1 \overline{z_2}$  defines an inner product, which yields the usual metric on the complex plane. Under what conditions do we have orthogonality?

*Proof.* First, we show that the conditions for an inner product are fulfilled:

1.  $\langle z_1 + z_3, z_2 \rangle = (z_1 + z_3) \overline{z_2} = z_1 \overline{z_2} + z_3 \overline{z_2} = \langle z_1, z_2 \rangle + \langle z_3, z_2 \rangle.$
2.  $\langle z_1, \alpha z_2 \rangle = z_1 \overline{\alpha z_2} = \overline{\alpha} z_1 \overline{z_2} = \overline{\alpha} \langle z_1, z_2 \rangle.$
3.  $\langle z_1, z_2 \rangle = z_1 \overline{z_2} = \overline{z_2 \overline{z_1}} = \overline{\langle z_2, z_1 \rangle}.$
4.  $\langle z, z \rangle = z \overline{z} = |z|^2 \geq 0.$
5.  $\langle z, z \rangle = 0 \implies |z|^2 = 0 \implies z = 0.$

Thus, it is a valid inner product that defines the norm:

$$\|z\|^2 = \langle z, z \rangle = |z|^2.$$

$z_1$  and  $z_2$  are orthogonal to each other if  $\langle z_1, z_2 \rangle = z_1 \overline{z_2} = 0.$

■

**11.** Let  $X$  be the vector space of all ordered pairs of complex numbers. Can we obtain the norm defined on  $X$  by

$$\|x\| = |\xi_1| + |\xi_2|$$

from an inner product?

*Proof.* We know that a norm is obtainable from an inner product if the parallelogram law holds, i.e

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Consider  $x = (1, \iota)$  and  $y = (1, -\iota)$ . Then,

$$\|x + y\|^2 = |2|^2 = 4, \|x - y\|^2 = |2\iota|^2 = 4, 2\|x\|^2 = 2(|1| + |\iota|)^2 = 8, 2\|y\|^2 = 2(|1| + |-\iota|)^2 = 8.$$

Hence, we get  $\|x + y\|^2 + \|x - y\|^2 \neq 2\|x\|^2 + 2\|y\|^2$ . Thus, it cannot be derived from an inner product.

■

**12.** What is  $\|x\|$  in 3.1-6 if  $x = \{\xi_1, \xi_2, \dots\}$ , where (a)  $\xi_n = 2^{-n/2}$ , (b)  $\xi_n = 1/n$ ?

*Proof.* In 3.1-6, we consider the Hilbert sequence space  $\ell^2$ , where the inner product is defined by

$$\langle x, y \rangle = \sum_{i=1}^{\infty} \xi_i \overline{\eta_i}.$$

Now,

$$(a) \|x\| = \langle x, x \rangle^{1/2} = \sqrt{\sum_{n=1}^{\infty} 2^{-n/2} \cdot 2^{-n/2}} = \sqrt{\sum_{i=1}^{\infty} 2^{-n}} = 1.$$

$$(b) \|x\| = \langle x, x \rangle^{1/2} = \sqrt{\sum_{n=1}^{\infty} 1/n^2} = \frac{\pi}{\sqrt{6}}.$$

13. Verify that for continuous functions, the inner product in 3.1-5 satisfies (IP1) to (IP4). ■

*Proof.* We wish to show that the inner product defined via

$$\langle x, y \rangle = \int_a^b x(t) \overline{y(t)} dt$$

satisfies all the properties of an inner product.

1.  $\langle x + z, y \rangle = \int_a^b (x + z)(t) \overline{y(t)} dt = \int_a^b x(t) \overline{y(t)} dt + \int_a^b z(t) \overline{y(t)} dt = \langle x, y \rangle + \langle z, y \rangle.$
2.  $\langle x, \alpha y \rangle = \int_a^b x(t) \overline{\alpha y(t)} dt = \overline{\alpha} \int_a^b x(t) \overline{y(t)} dt = \overline{\alpha} \langle x, y \rangle.$
3.  $\langle x, y \rangle = \int_a^b x(t) \overline{y(t)} dt = \int_a^b \overline{y(t) \overline{x(t)}} dt = \overline{\int_a^b y(t) \overline{x(t)} dt} = \overline{\langle y, x \rangle}.$
4.  $\langle x, x \rangle = \int_a^b x(t) \overline{x(t)} dt = \int_a^b |x(t)|^2 dt \geq 0.$  Also,  $x = 0 \implies \langle x, x \rangle = 0.$  Also,  $\langle x, x \rangle = \int_a^b |x(t)|^2 dt = 0 \implies x = 0.$

Thus, it satisfies all properties of an inner product. ■

14. Show that the norm on  $C[a, b]$  is invariant under a linear transformation  $t = \alpha\tau + \beta$ . Use this to prove the statement in 3.1-8 by mapping  $[a, b]$  onto  $[0, 1]$  and then considering the functions defined by  $\tilde{x}(\tau) = 1$ ,  $\tilde{y}(\tau) = \tau$ , where  $\tau \in [0, 1]$ .

*Proof.* Recall that the norm on  $C[a, b]$  is defined as follows:

$$\|x\| = \max_{t \in [a, b]} |x(t)|.$$

Define  $t' = \frac{t-a}{b-a}$ . Then,  $t' \in [0, 1]$ . We have

$$\|x\| = \max_{t' \in [0, 1]} |x((b-a)t' + a)|.$$

Now  $\tilde{x}(\tau) = 1 \forall \tau \in [0, 1] \implies \|\tilde{x}\| = 1$ . Also,  $\tilde{y}(\tau) = \tau \forall \tau \in [0, 1] \implies \|\tilde{y}\| = 1$ . Thus, we have

$$\|x + y\|^2 = \max_{\tau \in [0, 1]} |1 + \tau|^2 = 4, \quad \|x - y\|^2 = \max_{\tau \in [0, 1]} |1 - \tau|^2 = 1, \quad 2\|x\|^2 = 2, \quad 2\|y\|^2 = 2$$

which gives us  $\|x + y\|^2 + \|x - y\|^2 \neq 2\|x\|^2 + 2\|y\|^2$ . ■

15. If  $X$  is a finite dimensional vector space and  $(e_j)$  is a basis for  $X$ , show that an inner product on  $X$  is completely determined by its values  $\gamma_{jk} = \langle e_j, e_k \rangle$ . Can we choose such scalars  $\gamma_{jk}$  in a completely arbitrary fashion?

*Proof.* Let the basis for  $X$  be  $\{e_1, \dots, e_n\}$ . Then, for some  $x = \sum_{i=1}^n \alpha_i e_i$  and  $y = \sum_{i=1}^n \beta_i e_i$ , we have

$$\langle x, y \rangle = \left\langle \sum_{i=1}^n \alpha_i e_i, \sum_{i=1}^n \beta_i e_i \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\beta_j} \langle e_i, e_j \rangle.$$

Thus, simply knowing  $\langle e_i, e_j \rangle$  determines the inner product between any two points. No, we cannot choose these values arbitrarily because we need to satisfy the properties of an inner product such as:

$$\overline{\langle y, x \rangle} = \overline{\sum_{i=1}^n \sum_{j=1}^n \beta_j \overline{\alpha_i} \langle e_i, e_j \rangle} = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\beta_j} \overline{\langle e_i, e_j \rangle} \implies \langle e_i, e_j \rangle = \overline{\langle e_j, e_i \rangle}.$$

$$\langle x, x \rangle = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\alpha_j} \langle e_i, e_j \rangle \geq 0 \implies \langle e_i, e_j \rangle \geq 0$$



### 3.2 Further Properties of Inner Product spaces

1. What is the Schwarz inequality in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ ? Give another proof of it in these cases.

*Proof.* Recall the Schwarz inequality is given by

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Consider  $\mathbb{R}^2$ . Then, the Schwarz inequality is given by

$$\sum_{i=1}^2 |x_i y_i| \leq \sqrt{\sum_{i=1}^2 |x_i|^2} \sqrt{\sum_{i=1}^2 |y_i|^2}$$

which is precisely Hölder's inequality with  $p = q = 1/2$ . The same can be said for  $\mathbb{R}^3$ . ■

2. Give examples of subspaces of  $\ell^2$ .

*Proof.*  $\ell^2$  consists of sequences such that their  $\ell_2$ -norm is bounded. One subspace could simply be the space of all sequences that converge to 0. This holds since the sum of two sequences converging to zero would also converge to zero, and multiplying by a scalar would not change convergence properties. ■

3. Let  $X$  be the inner product space consisting of the polynomial  $x = 0$  and all real polynomials in  $t$  of degree not exceeding 2, considered for real  $t \in [a, b]$  with inner product defined by Section 3.1, (7). Show that  $X$  is complete. Let  $Y$  consist of all  $x \in X$  such that  $x(a) = 0$ . Is  $Y$  a subspace of  $X$ . Do all  $x \in X$  of degree 2 form a subspace of  $X$ ?

*Proof.*  $X$  is the inner product space consisting of the polynomial  $x = 0$  and all real polynomials in  $t$  of degree not exceeding 2, considered for real  $t \in [a, b]$  with inner product defined by

$$\langle x, y \rangle = \int_a^b x(t)y(t) dt.$$

Now,  $X$  is finite dimensional, since the basis is  $\{1, x, x^2\}$ . Hence, it is closed, and hence, it is also complete.

Now, consider  $y_1, y_2 \in Y$  such that  $y_1(a) = y_2(a) = 0$ . Consider the polynomial  $\alpha y_1 + \beta y_2$ . It is easy to see that  $(\alpha y_1 + \beta y_2)(a) = 0$  and hence, it forms a subspace.

However, all polynomials of degree 2 do not form a subspace. Consider the polynomials  $x = t^2$  and  $x = -t^2$ . Their addition is not a polynomial of degree 2. ■

4. Show that  $y \perp x_n$  and  $x_n \rightarrow x$  together implies  $x \perp y$ .

*Proof.* Since  $y \perp x_n$ , we have that  $\langle x_n, y \rangle = 0$ . Now, since  $x_n \rightarrow x$ , we have that using the continuity of inner products

$$\lim_{n \rightarrow \infty} x_n = x \implies \langle x, y \rangle = \left\langle \lim_{n \rightarrow \infty} x_n, y \right\rangle = \lim_{n \rightarrow \infty} \langle x_n, y \rangle = 0$$

and hence,  $x \perp y$ . ■

5. Show that for a sequence  $(x_n)$  in an inner product space the conditions  $\|x_n\| \rightarrow \|x\|$  and  $\langle x_n, x \rangle \rightarrow \langle x, x \rangle$  imply convergence  $x_n \rightarrow x$ .

*Proof.* We wish to show that  $\langle x_n - x, x_n - x \rangle \rightarrow 0$ . Expanding this inner product, we get

$$\langle x_n - x, x_n - x \rangle = \langle x_n, x_n \rangle - \langle x, x_n \rangle - \langle x_n, x \rangle + \langle x, x \rangle \rightarrow \langle x, x \rangle - \langle x, x \rangle + \langle x, x \rangle + \langle x, x \rangle = 0$$

since  $\langle x_n, x \rangle \rightarrow \langle x, x \rangle \in \mathbb{R} \implies \langle x, x_n \rangle \rightarrow \langle x, x \rangle$ . ■

6. Prove 5 for complex plane.

*Proof.* In the complex plane, let  $\langle x_n, x \rangle = c_n + id_n$ . Then, since  $\langle x_n, x \rangle \rightarrow \langle x, x \rangle$ , we have  $d_n \rightarrow 0$ . Hence,  $\langle x, x_n \rangle = c_n - id_n \rightarrow \langle x, x \rangle$ , and hence, the same proof as 5 follows. ■

7. Show that in an inner product space,  $x \perp y$  if and only if we have  $\|x + \alpha y\| = \|x - \alpha y\|$  for all scalars  $\alpha$ .

*Proof.* Suppose  $x \perp y$ . Then, we have that  $\langle x, y \rangle = 0 = \langle y, x \rangle$ . Thus, for any scalar  $\alpha$ ,

$$\begin{aligned} 0 &= 2\langle x, \alpha y \rangle + 2\langle \alpha y, x \rangle \\ &= \langle x, x \rangle + \langle x, \alpha y \rangle + \langle \alpha y, x \rangle + \langle \alpha y, \alpha y \rangle - \langle x, x \rangle + \langle x, \alpha y \rangle + \langle \alpha y, x \rangle - \langle \alpha y, \alpha y \rangle \\ &= \langle x + \alpha y, x + \alpha y \rangle - \langle x - \alpha y, x - \alpha y \rangle \\ &= \|x + \alpha y\|^2 - \|x - \alpha y\|^2 \end{aligned}$$

Now, conversely, let  $\|x + \alpha y\| = \|x - \alpha y\|$ . Squaring and expanding, we get

$$\langle x, \alpha y \rangle + \langle \alpha y, x \rangle = 0 \implies \Re \langle \alpha y, x \rangle = 0$$

holds for  $\alpha$ . If the space is real, then we immediately get  $\langle x, y \rangle = 0 \implies x \perp y$ . However, if the space is complex, let  $\langle y, x \rangle = c + \iota d$ . Then, choosing  $\alpha = \iota$  gives us  $d = 0$ , while choosing  $\alpha = 1$  gives us  $c = 0$ , resulting in  $\langle x, y \rangle = 0 \implies x \perp y$ . ■

**8.** Show that in an inner product space,  $x \perp y$  if and only if  $\|x + \alpha y\| \geq \|x\|$  for all scalars  $\alpha$ .

*Proof.* Suppose  $x \perp y$ . Then,  $\langle x, y \rangle = 0 = \langle y, x \rangle$ . Then, we have for all scalars  $\alpha$ ,

$$\begin{aligned} 0 &= \langle x, \alpha y \rangle + \langle \alpha y, x \rangle \\ &= \langle x, x \rangle + \langle x, \alpha y \rangle + \langle \alpha y, x \rangle - \langle x, x \rangle \\ &\leq \langle x, x \rangle + \langle x, \alpha y \rangle + \langle \alpha y, x \rangle - \langle x, x \rangle + \langle \alpha y, \alpha y \rangle \\ &= \|x + \alpha y\|^2 - \|x\|^2 \end{aligned}$$

which results in  $\|x + \alpha y\|^2 \geq \|x\|^2$ . The inequality follows since  $\langle \alpha y, \alpha y \rangle \geq 0$ .

Now, conversely, assume  $\|x + \alpha y\| \geq \|x\|$  for all scalars  $\alpha$ . Squaring, we get

$$\langle x, \alpha y \rangle + \langle \alpha y, x \rangle + \langle \alpha y, \alpha y \rangle \geq 0 \implies \alpha \langle y, x \rangle + \bar{\alpha} \langle x, y \rangle + \alpha \langle y, y \rangle \geq 0.$$

Since this holds for all  $\alpha$ , it also holds for  $\alpha = -\frac{\langle x, y \rangle}{\langle y, y \rangle}$ . Substituting, we get

$$-\frac{\langle x, y \rangle^2}{\langle y, y \rangle} \geq 0 \implies \langle x, y \rangle^2 \leq 0 \implies \langle x, y \rangle = 0.$$

Thus,  $x \perp y$ . ■

**9.** Let  $V$  be the vector space of all continuous complex-valued functions on  $J = [a, b]$ . Let  $X_1 = (V, \|\cdot\|_\infty)$ , where  $\|x\|_\infty = \max_{t \in J} |x(t)|$  and let  $X_2 = (V, \|\cdot\|_2)$ , where

$$\|x\|_2 = \langle x, x \rangle^{1/2}, \quad \langle x, y \rangle = \int_a^b x(t) \overline{y(t)} dt$$

Show that the identity mapping  $x \mapsto x$  of  $X_1$  onto  $X_2$  is continuous.

*Proof.* We wish to show that for all  $\epsilon > 0$ ,  $\exists \delta > 0$ , such that  $\|x_0 - x\|_\infty < \delta \implies \|x_0 - x\|_2 < \epsilon$ . First, if  $\|x_0 - x\|_\infty < \delta$ , we have that

$$\|x_0 - x\|_\infty = \max_{t \in [a, b]} |x_0(t) - x(t)| < \delta \implies |x_0(t) - x(t)| < \delta \quad \forall t \in [a, b].$$

Now, consider  $\|x_0 - x\|_2 = \langle x_0 - x, x_0 - x \rangle^{1/2}$ . We get

$$\langle x_0 - x, x_0 - x \rangle = \int_a^b |x_0(t) - x(t)|^2 dt < (b - a)\delta^2.$$

Thus,  $\|x - x_0\|_2 < \sqrt{b - a}\delta < \epsilon$ . Hence, substituting  $\delta = \frac{\epsilon}{\sqrt{b - a}}$  shows that the map is continuous. ■

**10.** Let  $T : X \mapsto X$  be a bounded linear operator on a complex inner product space  $X$ . If  $\langle Tx, x \rangle = 0$  for all  $x \in X$ , show that  $T = 0$ . Show that this does not hold in case of a real inner product space.

*Proof.* First, since  $T$  is a bounded linear operator, we have

$$0 = \langle T(x + y), x + y \rangle = \langle Tx, x \rangle + \langle Ty, y \rangle + \langle Tx, y \rangle + \langle Ty, x \rangle = \langle Tx, y \rangle + \langle Ty, x \rangle.$$

We also have

$$0 = \langle T(x + \iota y), x + \iota y \rangle = \langle Tx, x \rangle + \langle T\iota y, \iota y \rangle + \langle Tx, \iota y \rangle + \langle T\iota y, x \rangle = -\iota \langle Tx, y \rangle + \iota \langle Ty, x \rangle \implies \langle Ty, x \rangle = \langle Tx, y \rangle$$

Solving both equations gives us  $\langle Tx, y \rangle = \langle Ty, x \rangle = 0$  for all  $x, y$ . This means that  $T = 0$ .

However, in the real case,  $T$  could simply be a rotation by  $\pi/2$  which results in  $\langle Tx, x \rangle = 0 \quad \forall x \in X$ . ■

### 3.3 Orthogonal Complements and Direct Sums

1. Let  $H$  be a Hilbert Space,  $M \subset H$  is convex,  $(x_n)$  is a sequence in  $M$  such that  $\|x_n\| \rightarrow d$ , where  $d = \inf_{x \in M} \|x\|$ . Show that  $(x_n)$  converges in  $H$ .

*Proof.* Since  $H$  is complete, it suffices to show  $(x_n)$  is Cauchy. First, we have that  $\forall x \in M, d \leq \|x\|$ . Also, assume  $\|x\|_n = d_n$  such that  $d_n \rightarrow d$ . Then, we have

$$\|x_n - x_m\|^2 = 2\|x_n\|^2 + 2\|x_m\|^2 - \|x_n + x_m\|^2 = 2d_n^2 + 2d_m^2 - 4\left\|\frac{x_n + x_m}{2}\right\|^2 \leq 2d_n^2 + 2d_m^2 - 4d^2 \rightarrow 0.$$

where the last inequality follows since  $\frac{x_n + x_m}{2} \in M$ . Thus,  $(x_n)$  is Cauchy. ■

2. Show that  $M = \{y = (\eta_j) \mid \sum \eta_j = 1\}$  of complex space  $\mathbb{C}^n$  is complete and convex. Find the vector of minimum norm in  $M$ .

*Proof.* First, we show  $M$  is convex. Consider any  $y_1, y_2 \in M$ . Then,  $\alpha y_1 + (1 - \alpha)y_2 = (\alpha\eta_j^{(1)} + (1 - \alpha)\eta_j^{(2)})$ . Clearly

$$\sum (\alpha\eta_j^{(1)} + (1 - \alpha)\eta_j^{(2)}) = \alpha \sum \eta_j^{(1)} + (1 - \alpha) \sum \eta_j^{(2)} = \alpha + (1 - \alpha) = 1.$$

since  $\sum \eta_j^{(1)} = \sum \eta_j^{(2)} = 1$ . Thus  $M$  is convex. Also, let  $(y_n)$  be some Cauchy sequence in  $M$ . Then, we have  $\forall \epsilon > 0$ , there exists  $N(\epsilon)$  such that  $\forall m, r > N(\epsilon)$ ,  $\|y_m - y_r\| < \epsilon$ . Using the  $\ell^2$  norm, we basically get  $|\eta_j^{(m)} - \eta_j^{(r)}|^2 < \epsilon^2$ , and hence, the sequence  $(\eta_j^{(r)})_r$  is convergent, to say  $\eta_j$ . Thus, defining  $y = (\eta_j)$ , we wish to show that  $y_r \rightarrow y$  and  $y \in M$ , i.e  $\sum \eta_j = 1$ . First, it is easy to see that

$$\lim_{r \rightarrow \infty} \|y_r - y_m\| = \lim_{r \rightarrow \infty} \|y_r - y\| = 0.$$

Finally, we have

$$\eta_j^{(m)} \rightarrow \eta_j \implies 1 = \sum \eta_j^{(m)} \rightarrow \sum \eta_j$$

Thus,  $y \in M$ , and hence, the subset  $M$  is closed. Suppose we wish to minimize the  $\ell^p$  norm. We have that

$$\sum \eta_j = 1 = \left| \sum \eta_j \right| \leq \|y\|_p \|(1, 1, \dots, 1)\|_q \implies \|y\|_p \geq \frac{1}{n^{1/q}} = n^{1/p-1}.$$

To find the corresponding  $y$  that achieves the minimum norm, we need an equality in Hölder's inequality, when  $y$  is a scalar multiple of  $(1, 1, \dots, 1)$ , resulting in the point  $(\frac{1}{n}, \dots, \frac{1}{n})$ . ■

3. (a) Show that the vector space  $X$  of all real valued continuous functions on  $[-1, 1]$  is the direct sum of the set of all even continuous functions and the set of all odd continuous functions on  $[-1, 1]$ . (b) Give examples of representations on  $\mathbb{R}^3$  as a direct sum of (i) of a subspace and its orthogonal complement and (ii) of any complementary pair of subspaces.

*Proof.* Recall an even continuous function is a function  $f$  such that  $f(-x) = f(x)$  and an odd continuous function is a function  $f$  such that  $f(-x) = -f(x)$ . Define  $g(x) = \frac{f(x) + f(-x)}{2}$  and  $h(x) = \frac{f(x) - f(-x)}{2}$ . It is easy to see,  $g(-x) = g(x)$  while  $h(-x) = -h(x)$ , and hence,  $g$  is even while  $h$  is odd. Also,  $f(x) = g(x) + h(x)$ . Thus, we can express the space of all real continuous functions as the direct sum of even and odd continuous functions.

(a) Consider the subspace  $x$ -axis, so that the orthogonal subspace is the  $yz$ -plane. Clearly,  $\mathbb{R}^3$  is the direct sum of both these spaces. (b) The example in (a) suffices. ■

4. (a) Show that the conclusion of Theorem 3.3-1 also holds if  $X$  is a Hilbert space and  $M \subset X$  is a closed subspace. (b) How could we use Appolonius' identity in the proof of Theorem 3.3-1?

*Proof.* Theorem 3.3- says that for an inner product space  $X$  and a nonempty convex subset  $M$  which is complete, for every  $x \in X$ , there exists a unique  $y \in M$  such that

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| = \|x - y\|.$$

Now, since  $X$  is Hilbert (complete) and  $M$  is closed,  $M$  is also complete. Moreover, since  $M$  is a subspace, it is convex, and hence, all conditions of the Theorem hold.

(b) **TODO**. ■

5. Let  $X = \mathbb{R}^2$ . Find  $M^\perp$  if  $M$  is (a)  $\{x\}$ , where  $x = (\xi_1, \xi_2) \neq 0$  (b) a linearly independent set  $\{x_1, x_2\} \subset X$ .

*Proof.* (a) If  $x = (\xi_1, \xi_2)$ , then the set of points of the form  $\{(\eta, -\frac{\eta\xi_1}{\xi_2}) \mid \eta \in \mathbb{R}\}$  is perpendicular to  $M$ .

(b) Let  $y \perp M$ . Then,  $y = \alpha_1 x_1 + \alpha_2 x_2$ , since  $\dim X = 2$ , any linearly independent set with two vectors will then be a basis. We now have,  $\langle y, x_1 \rangle = 0$  and  $\langle y, x_2 \rangle = 0$  which gives us  $\alpha_1 = 0$  and  $\alpha_2 = 0$ , giving us  $M^\perp = \{0\}$ . ■

6. Show that  $Y = \{x \mid x = (\xi_j) \in \ell^2, \xi_{2n} = 0, n \in \mathbb{N}\}$  is a closed subspace of  $\ell^2$  and find  $Y^\perp$ . What is  $Y^\perp$  if  $Y = \text{span}\{e_1, \dots, e_n\} \subset \ell^2$ , where  $e_j = (\delta_{jk})$ ?

*Proof.* Let  $x \in \bar{Y}$ . Then, there exists a sequence  $(x_n) = (\xi_j^{(n)}) \in Y$  such that  $x_n \rightarrow x$ . For a fixed  $j = 2k$ ,  $k \in \mathbb{N}$ , we have  $\xi_2^{(n)} k = 0 \forall n$ , and hence  $\xi_2^{(n)} k \rightarrow 0$ . Since a series converges to the limits of its subsequences,  $x_{2k} = 0 \forall k \in \mathbb{N}$ . Thus,  $x \in Y$ , and hence  $Y$  is closed.

Now, we claim that

$$Y^\perp = \{z \mid z_{2n-1} = 0 \forall n \in \mathbb{N}\}.$$

Denote  $Z = \{z \mid z_{2n-1} = 0 \forall n \in \mathbb{N}\}$ . Clearly, for any  $z \in Z$ ,  $\langle z, y \rangle = 0 \forall y \in Y$ , and hence,  $Z \subset Y^\perp$ . Now, let  $m \in Y^\perp$ . Then, define some new sequence  $l$  such that  $l_{2n} = m_{2n}$  and  $l_{2n+1} = 0$ . Clearly,  $l \in Y^\perp$  and  $l \in Z$ , and hence,  $Y^\perp \subset Z$ . Thus,  $Y^\perp = Z$ .

If  $Y = \text{span}\{e_1, \dots, e_n\} \subset \ell^2$ , then  $Y^\perp = \text{span}\{e_{n+1}, \dots\} \subset \ell^2$ . ■

7. Let  $A$  and  $A \subset B$  be nonempty subsets of an inner product space  $X$ . Show that (a)  $A \subset A^{\perp\perp}$ , (b)  $B^\perp \subset A^\perp$ , (c)  $A^{\perp\perp\perp} = A^\perp$ .

*Proof.* (a) Let  $x \in A$ . Then,  $x \perp y$  for all  $y \in A^\perp$ . Now, for some  $y \in A^\perp$ ,  $y \perp z$  for all  $z \in A^{\perp\perp}$ . But  $y \perp x$  and hence,  $x \in A^{\perp\perp}$ . Thus,  $A \subset A^{\perp\perp}$ .

(b) We know that  $A \subset B$ . Let  $x \in B^\perp$ . Then,  $x \perp b$  for all  $b \in B$ . Since  $A \subset B$ , all points  $a \in A$  also belong to  $B$ , and hence,  $x \perp a$  for all  $a \in A$ . Thus,  $x \in A^\perp$ , and hence,  $B^\perp \subset A^\perp$ .

(c) First, we show that  $A^{\perp\perp\perp} \subset A^\perp$ . Denote  $C = A^\perp$ . Then, this is equivalent to showing  $C^{\perp\perp} \subset C$  which follows from (a). Now, we show that  $A^\perp \subset A^{\perp\perp\perp}$ . Let  $x \in A^\perp$ . Then,  $x \perp A^{\perp\perp}$ . Thus,  $x \in A^{\perp\perp\perp}$ , and hence,  $A^\perp \subset A^{\perp\perp\perp}$ . ■

8. Show that the annihilator  $M^\perp$  of a set  $M \neq \emptyset$  in an inner product space  $X$  is a closed subspace of  $X$ .

*Proof.* We know that  $M^\perp$  is always a subspace. We wish to show it is closed. Consider  $x \in \overline{M^\perp}$ . Then, there exists a sequence  $(x_n)$  such that  $x_n \rightarrow x$ . Then, by the continuity of inner products, we have, for some  $m \in M$ ,

$$\langle x_n, m \rangle \rightarrow \langle x, m \rangle$$

or,  $x \in M^\perp$ . Thus,  $M^\perp$  is closed. ■

9. Show that a subspace  $Y$  of a Hilbert space  $H$  is closed in  $H$  iff  $Y = Y^{\perp\perp}$ .

*Proof.* First we show that if  $Y$  is closed then  $Y = Y^{\perp\perp}$ . We know that  $Y \subset Y^{\perp\perp}$ . Suppose  $x \in Y^{\perp\perp}$ . Then, since  $Y$  is closed  $x = y + z$ , where  $y \in Y \subset Y^{\perp\perp}$  and  $z \in Y^\perp$ . Then,  $z = x - y \in Y^{\perp\perp}$ . This means that  $z \in Y^\perp \cap Y^{\perp\perp} = \{0\}$ , and hence  $x \in Y$ . Thus,  $Y = Y^{\perp\perp}$ .

For the converse, assume  $Y = Y^{\perp\perp}$ . Let  $x \in \bar{Y}$ . Then, there exists some sequence  $y_n \in Y$  such that  $y_n \rightarrow x$ . We know that for all  $z \in Y^\perp$ ,

$$\langle y_n, z \rangle = 0 \implies \lim_{n \rightarrow \infty} \langle y_n, z \rangle = \langle x, z \rangle = 0.$$

and hence  $x \in Y$ . Thus,  $Y$  is closed. ■

10. If  $M \neq \emptyset$  is any subset of a Hilbert space  $H$ , show that  $M^{\perp\perp}$  is the smallest closed subspace of  $H$  which contains  $M$ , that is  $M^{\perp\perp}$  is contained in any closed subspace  $Y \subset H$  such that  $M \subset Y$ .

*Proof.* **TODO.** ■

### 3.4 Orthonormal Sets and Sequences

1. Show that an inner product space of finite dimension  $n$  has a basis  $b_1, \dots, b_n$  of orthonormal vectors.

*Proof.* Since the inner product space has finite dimension  $n$ , it has a basis  $\{e_1, e_2, \dots, e_n\}$  (say, all of them have norm 1 for convenience). Now, the existence of the set of orthonormal vectors follows from the Gram-Schmidt orthogonalization procedure.

Let  $b_1 = e_1$ . Define  $b_2 = e_2 - \langle e_2, b_1 \rangle e_1 \implies \langle b_1, b_2 \rangle = \langle e_2, b_1 \rangle - \langle e_2, b_1 \rangle \langle e_1, b_1 \rangle = 0$ . Thus,  $b_2 \perp b_1$ . Also,  $b_2 \neq 0$  since this would mean that  $e_2 = \alpha e_1$ , i.e.  $e_2$  and  $e_1$  are linearly dependent, which is not the case. In this way, we can continue constructing  $b_3, \dots$  ■

2. How can we interpret (12\*) geometrically in  $\mathbb{R}^r$ ,  $r \geq n$ .

*Proof.* (12\*) is Bessel's inequality, which says that for any orthonormal sequence  $(e_n)$ , we have

$$\sum_{k=1}^n |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

If we choose an orthonormal sequence which is not a basis, then we would be able to obtain the inequality, since there is some component which is not getting expressed by the orthonormal sequence. ■

3. Obtain the Schwarz inequality from (12\*).

*Proof.* (12\*) is Bessel's inequality, which says that for any orthonormal sequence  $(e_n)$ , we have

$$\sum_{k=1}^n |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

For  $n = 1$ , we get

$$|\langle x, e \rangle|^2 \leq \|x\|^2.$$

Setting  $e = y/\|y\|$ , finishes the proof. ■

4. Give an example of an  $x \in \ell^2$  such that we have strict inequality in (12\*).

*Proof.* Whenever we have an orthonormal basis, we would always obtain equality. Thus, finding an orthonormal sequence that is not a basis suffices. A very nice example (found on the internet) was  $(e_n) = \delta_{2n}$ , i.e., a sequence which has a 1 in  $2n$ th position, zero elsewhere, and cannot express the sequences which contain non-zero elements in the odd positions, and hence, is not a basis. ■

5. If  $(e_k)$  is an orthonormal sequence in an inner product space  $X$ , and  $x \in X$ , show that  $x = y$  with  $y$  given by

$$y = \sum_{k=1}^n \alpha_k e_k$$

where  $\alpha_k = \langle x, e_k \rangle$  is orthogonal to the subspace  $Y_n = \text{span}\{e_1, \dots, e_n\}$ .

*Proof.* Let  $z = \sum_{i=1}^n \langle z, e_i \rangle e_i \in \text{span}\{e_1, \dots, e_n\}$ . Then,  $\langle x - y, z \rangle = \langle x, z \rangle - \langle y, z \rangle$ . We have that

$$\langle x, z \rangle = \left\langle x, \sum_{i=1}^n \langle z, e_i \rangle e_i \right\rangle = \sum_{i=1}^n \langle z, e_i \rangle \langle x, e_i \rangle.$$

$$\langle y, z \rangle = \left\langle \sum_{i=1}^n \alpha_i e_i, \sum_{i=1}^n \langle z, e_i \rangle e_i \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \langle z, e_j \rangle \langle e_i, e_j \rangle = \sum_{i=1}^n \alpha_i \langle z, e_i \rangle.$$

Since  $\alpha_i = \langle x, e_i \rangle$ , we have  $\langle x, z \rangle - \langle y, z \rangle = 0 \implies x - y \perp z$ . ■

6. Let  $\{e_1, \dots, e_n\}$  be an orthonormal set in an inner product space  $X$ , where  $n$  is fixed. Let  $x \in X$  be any fixed element and  $y = \beta_1 e_1 + \dots + \beta_n e_n$ . Then,  $\|x - y\|$  depends on  $\beta_1, \dots, \beta_n$ . Show by direct calculation that  $\|x - y\|$  is minimum iff  $\beta_i = \langle x, e_i \rangle$ .

*Proof.* First, we have that

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \|x\|^2 + \langle y, x \rangle + \langle x, y \rangle + \langle y, y \rangle = \|x\|^2 - \sum_{i=1}^n \beta_i \langle e_i, x \rangle - \sum_{i=1}^n \overline{\beta_i} \langle x, e_i \rangle + \sum_{i=1}^n \beta_i \overline{\beta_i}$$

We can rewrite this as

$$\|x - y\|^2 = \|x\|^2 - \sum_{i=1}^n \beta_i \langle e_i, x \rangle - \sum_{i=1}^n \overline{\beta_i} (\langle x, e_i \rangle - \beta_i)$$

Clearly, setting  $\beta_i = \langle x, e_i \rangle$  sets the terms inside the parenthesis to 0 and gives  $\|x - y\|^2 = \|x\|^2 - \sum_{i=1}^n |\langle x, e_i \rangle|^2$ . Since the norm is non-negative, we get

$$\sum_{i=1}^n |\langle x, e_i \rangle|^2 \leq \|x\|^2.$$

Now, suppose  $\beta_i = \langle x, e_i \rangle$ . Then,  $\|x - y\|^2 = \|x\|^2 - \sum_{i=1}^n |\langle x, e_i \rangle|^2$ . Since the norm is non-negative, we get

$$\sum_{i=1}^n |\langle x, e_i \rangle|^2 \leq \|x\|^2.$$

**7.** Let  $(e_k)$  be any orthonormal sequence in an inner product space  $X$ . Show that for any  $x, y \in X$ ,

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle \langle y, e_k \rangle| \leq \|x\| \|y\|.$$

*Proof.* Using Cauchy Schwarz, we get

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle \langle y, e_k \rangle| \leq \sqrt{\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2} \sqrt{\sum_{k=1}^{\infty} |\langle y, e_k \rangle|^2} \leq \|x\| \|y\|$$

where the last inequality follows from Bessel Inequality. ■

**8.** Show that an element  $x$  of an inner product space  $X$  cannot have too many Fourier coefficients  $\langle x, e_k \rangle$  which are big. Show that for an orthonormal sequence  $(e_n)$ , the number  $n_m$  of  $\langle x, e_k \rangle$  such that  $|\langle x, e_k \rangle| > 1/m$  must satisfy  $n_m \leq m^2 \|x\|^2$ .

*Proof.* Using Bessel's inequality, we have

$$\sum_{i=1}^n |\langle x, e_i \rangle|^2 \leq \|x\|^2.$$

Let  $n_m$  coefficients be such that  $|\langle x, e_k \rangle| > 1/m$ . Then, we have

$$\|x\|^2 \geq \sum_{i=1}^n |\langle x, e_i \rangle|^2 \geq \sum_{|\langle x, e_k \rangle| > 1/m} |\langle x, e_k \rangle|^2 + \sum_{|\langle x, e_k \rangle| < 1/m} |\langle x, e_k \rangle|^2 \geq \frac{n_m}{m^2}.$$

Rearranging gives us the result. ■

**9.** Orthonormalize the first three terms of the sequence  $(x_0, x_1, x_2, \dots)$ , where  $x_j(t) = t^j$  on the interval  $[-1, 1]$ , and

$$\langle x, y \rangle = \int_{-1}^1 x(t)y(t) dt.$$

*Proof.* We are given the sequence  $(1, t^1, t^2, \dots)$  which we are required to orthonormalize. We first calculate the norms for each term:

$$\|1\| = \sqrt{\int_{-1}^1 1 \cdot 1 dt} = \sqrt{2}.$$

$$\|t\| = \sqrt{\int_{-1}^1 t^2 dt} = \sqrt{\frac{2}{3}}.$$

Now, using the Gram-Schmidt orthogonalization, we get

$$b_1 = \frac{1}{\|1\|} = \frac{1}{\sqrt{2}}.$$

$$b'_2 = t - \langle t, b_1 \rangle b_1 = t - \frac{1}{\sqrt{2}} \int_{-1}^1 \frac{t}{\sqrt{2}} dt = t \implies b_2 = \frac{t}{\|t\|} = \frac{\sqrt{3}t}{\sqrt{2}}.$$

$$b'_3 = t^2 - \langle t^2, b_2 \rangle b_2 - \langle t^2, b_1 \rangle b_1 = t^2 - \frac{\sqrt{3}t}{\sqrt{2}} \int_{-1}^1 \frac{\sqrt{3}t^3}{\sqrt{2}} dt - \frac{1}{\sqrt{2}} \int_{-1}^1 \frac{t^2}{\sqrt{2}} dt = t^2 - \frac{1}{3} \implies b_3 = \frac{3t^2 - 1}{\|3t^2 - 1\|}$$

We can calculate the norm as

$$\|3t^2 - 1\| = \sqrt{\int_{-1}^1 (3t^2 - 1)^2 dt} = \sqrt{\frac{18}{5} + 2 - 4} = \sqrt{\frac{8}{5}}.$$

Thus,

$$b_3 = \frac{\sqrt{5}(3t^2 - 1)}{\sqrt{8}}$$

**10.** Let  $x_1(t) = t^2$ ,  $x_2(t) = t$  and  $x_3(t) = 1$ . Orthonormalize  $x_1, x_2, x_3$  in this order on the interval  $[-1, 1]$  with respect to the inner product given in 9.

*Proof.* We are given the sequence  $(t^2, t, 1, \dots)$  which we are required to orthonormalize. Calculating the norm:

$$\|t^2\| = \sqrt{\int_{-1}^1 t^4 \cdot 1 dt} = \sqrt{\frac{2}{5}}.$$

Select the first vector as

$$b_1 = \frac{t^2}{\|t^2\|} = \frac{\sqrt{2}t^2}{\sqrt{5}}.$$

Now, the second vector is calculated as

$$b'_2 = t - \langle t, b_1 \rangle b_1 = t - \frac{\sqrt{2}t^2}{\sqrt{5}} \int_{-1}^1 \frac{\sqrt{2}t^3}{\sqrt{5}} dt = t$$

and the norm is

$$\|b'_2\| = \|t\| = \sqrt{\int_{-1}^1 t^2 dt} = \sqrt{\frac{2}{3}}$$

resulting in

$$b_2 = \frac{t}{\|t\|} = \frac{\sqrt{3}t}{\sqrt{2}}.$$

Finally, we get

$$b'_3 = 1 - \langle 1, b_2 \rangle b_2 - \langle 1, b_1 \rangle b_1 = 1 - \frac{\sqrt{3}t}{\sqrt{2}} \int_{-1}^1 \frac{\sqrt{3}t}{\sqrt{2}} dt - \frac{\sqrt{2}t^2}{\sqrt{5}} \int_{-1}^1 \frac{\sqrt{2}t^2}{\sqrt{5}} dt = \frac{15 - 4t^2}{15}$$

and the norm is given by

$$\|b'_3\| = \sqrt{\int_{-1}^1 (15 - 4t^2)^2 dt} = 550 + \frac{32}{5} - 80 = \sqrt{\frac{2382}{5}}$$

and hence,

$$b_3 = \frac{\sqrt{5}(15 - 4t^2)}{\sqrt{2382}}.$$

### 3.5 Series Related to Orthonormal Sequences

1. If (6) converges with sum  $x$ , show that (7) has the sum  $\|x\|^2$ .

*Proof.* We wish to show that if

$$x = \sum_{k=1}^{\infty} \alpha_k e_k$$

then,

$$\sum_{k=1}^{\infty} |\alpha_k|^2 = \|x\|^2$$

Since  $\sum_{k=1}^{\infty} \alpha_k e_k$  converges, the sequence of partial sums converges, say  $s_n$ .

$$\|s_n\|^2 = \left\langle \sum_{k=1}^n \alpha_k e_k, \sum_{k=1}^n \alpha_k e_k \right\rangle = \sum_{k=1}^n |\alpha_k|^2.$$

Now, since  $s_n \rightarrow x$ , we have

$$\langle s_n, s_n \rangle - \langle x, x \rangle = \langle s_n - x, s_n \rangle + \langle x, s_n - x \rangle \leq \|s_n - x\| \|s_n\| + \|x\| \|s_n - x\| \rightarrow 0.$$

and hence,  $\|s_n\|^2 = \|x\|^2$ . Now, note that the partial sums for  $\sum_{k=1}^{\infty} |\alpha_k|^2$  is precisely  $\|s_n\|^2$ , and hence, the quantity  $\sum_{k=1}^{\infty} |\alpha_k|^2$  converges to  $\|x\|^2$ . ■

2. Derive from (1) and (2) a Fourier series representation of a function  $\tilde{x}$  (as a function of  $\tau$ ) of arbitrary period  $p$ .

*Proof.* We have that  $x(t) = a_0 + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt)$ , where

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} x(t) dt \quad a_k = \frac{1}{\pi} \int_0^{2\pi} x(t) \cos kt dt \quad b_k = \frac{1}{\pi} \int_0^{2\pi} x(t) \sin kt dt$$

TODO ■

3. Illustrate with an example that a convergent series  $\sum \langle x, e_k \rangle e_k$  need not have the sum  $x$ .

*Proof.* A simple example could be one where we do not consider all components of the basis. For example, let  $\{e_1, \dots, e_n\}$  be a basis and  $x = \sum \langle x, e_k \rangle e_k$ . Consider only one term of the sum  $\langle x, e_1 \rangle e_1$ . This is convergent, but the sum does not equal to  $x$ . ■

4. If  $(x_j)$  is a sequence in an inner product space  $X$  such that the series  $\|x_1\| + \|x_2\| + \dots$  converges, show that  $(s_n)$  is a Cauchy sequence, where  $s_n = x_1 + \dots + x_n$ .

*Proof.* To show  $(s_n)$  is Cauchy, we wish to show that for all  $\epsilon > 0$ , we have some  $N(\epsilon)$  such that for all  $m, n > N(\epsilon)$ ,  $\|\sum_{i=m+1}^n x_i\| < \epsilon$ .

Now, since  $\|x_1\| + \|x_2\| + \dots$  converges, the sequence of partial sums converges, and hence, is Cauchy. This means that for all  $\epsilon > 0$ , we have some  $N(\epsilon)$  such that for all  $m, n > N(\epsilon)$  and

$$\left\| \sum_{i=m+1}^n \|x_i\| \right\| \leq \sum_{i=m+1}^n \|x_i\| < \epsilon$$

and hence, for the sequence of partial sums we have

$$\|s_m - s_n\| = \left\| \sum_{i=m+1}^n x_i \right\| \leq \sum_{i=m+1}^n \|x_i\| < \epsilon$$

5. Show that in a Hilbert space  $H$ , convergence of  $\sum \|x_j\|$  implies convergence of  $\sum x_j$ . ■



*Proof.* Suppose  $\sum \|x_j\|$  converges, then the sequence of partial sums is Cauchy, which means

$$\forall \epsilon > 0 \exists N(\epsilon); \forall m, n > N(\epsilon), \left\| \sum_{i=m}^n \|x_j\| \right\| \leq \sum_{i=m}^n \|x_j\| \leq \epsilon.$$

Thus, consider the partial sums of the sequence  $\sum x_j$ , we get

$$\left\| \sum_{i=m}^n x_j \right\| \leq \sum_{i=m}^n \|x_j\| \leq \epsilon.$$

Thus, the sequence of partial sums is Cauchy, and hence, due to the completeness of the Hilbert space, the sequence is converging. ■

**6.** Let  $(e_j)$  be an orthonormal sequence in a Hilbert space  $H$ . Show that if

$$x = \sum_{j=1}^{\infty} \alpha_j e_j \quad y = \sum_{j=1}^{\infty} \beta_j e_j \implies \langle x, y \rangle = \sum_{j=1}^{\infty} \alpha_j \bar{\beta}_j,$$

the series being absolutely convergent.

*Proof.* It is easy to see

$$\langle x, y \rangle = \left\langle \sum_{j=1}^{\infty} \alpha_j e_j, \sum_{j=1}^{\infty} \beta_j e_j \right\rangle = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \alpha_j \bar{\beta}_k \langle e_j, e_k \rangle = \sum_{j=1}^{\infty} \alpha_j \bar{\beta}_j.$$

**7.** Let  $(e_j)$  be an orthonormal sequence in a Hilbert space  $H$ . Show that for every  $x \in H$ ,

$$y = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j$$

exists in  $H$  and  $x - y$  is orthogonal to every  $e_k$ .

*Proof.* First, we know that  $y = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j$  converges if and only if  $\sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2$  converges, which is taken care of by Bessel's inequality. Now, we have

$$\langle x - y, e_k \rangle = \langle x, e_k \rangle - \langle y, e_k \rangle = \langle x, e_k \rangle - \sum_{j=1}^{\infty} \langle x, e_j \rangle \langle e_k, e_j \rangle = 0 \implies (x - y) \perp e_k.$$

**8.** Let  $(e_j)$  be an orthonormal sequence in a Hilbert space  $H$  and let  $M = \text{span}(e_k)$ . Show that for any  $x \in H$  we have  $x \in \overline{M}$  if and only if  $x$  can be represented by (6) with coefficients  $\alpha_k = \langle x, e_k \rangle$ .

*Proof.* Let  $x \in \overline{M}$ . Then, there exists some sequence  $x_n \in M$  such that  $x_n \rightarrow x$ . Let  $x_n = \sum \langle x_n, e_k \rangle e_k$ . Since,  $x_n \rightarrow x$ , we have that  $\|x_n - x\| \rightarrow 0$ . Then, we have

$$\langle x_n - x, e_k \rangle \leq \|x_n - x\| \|e_k\| \rightarrow 0 \implies \langle x_n, e_k \rangle \rightarrow \langle x, e_k \rangle.$$

and hence,  $x_n \rightarrow x = \sum \langle x, e_k \rangle e_k$ .

On the other hand, let  $x = \sum \langle x, e_k \rangle e_k$ . Then, each of the partial sums, say  $s_n$  converges to  $x$ , and  $s_n \in M$ . Thus,  $x \in \overline{M}$ . ■

**9.** Let  $(e_n)$  and  $(\tilde{e}_n)$  be orthonormal sequences in a Hilbert space  $H$ , and let  $M_1 = \text{span}(e_n)$  and  $M_2 = \text{span}(\tilde{e}_n)$ . Using 8, show that  $\overline{M_1} = \overline{M_2}$  if and only if

$$(a) \quad e_n = \sum_{m=1}^{\infty} \alpha_{nm} \tilde{e}_m \quad (b) \quad \tilde{e}_n = \sum_{m=1}^{\infty} \overline{\alpha_{nm}} e_m, \quad \alpha_{nm} = \langle e_n, \tilde{e}_m \rangle.$$

*Proof.* From 8, we have that if  $e_n$  and  $\tilde{e}_n$  have the particular representation, then  $e_n \in \overline{M_2}$  and  $\tilde{e}_n \in \overline{M_1}$ . This means that  $(e_n) \in \overline{M_2}$  and  $(\tilde{e}_n) \in \overline{M_1}$ . Thus,  $\overline{M_1} = \overline{M_2}$ , since  $e_n$  (which is in  $\overline{M_2}$ ) can be written in terms of  $(\tilde{e}_n)$ , which is in  $\overline{M_1}$ . ■

**10.** *Work out the details of the proof of Lemma 3.5-3.*

*Proof.* We wish to show that for any  $x$ , we can have countably many nonzero Fourier coefficients. From Bessel's inequality, we have

$$\sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2 \leq \|x\|^2.$$

Consider the coefficients  $\langle x, e_j \rangle > \frac{1}{m}$ . Let the number of such coefficients be  $n_m$ . Then, we have

$$\|x\|^2 \geq \sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2 \geq \sum_{|\langle x, e_j \rangle| > 1/m} |\langle x, e_j \rangle|^2 + \sum_{|\langle x, e_j \rangle| < 1/m} |\langle x, e_j \rangle|^2 \geq \frac{n_m}{m^2}.$$

Thus, we get  $n_m \leq \|x\|^2 m^2$  which is countable for each  $m$ , and hence, the total number of coefficients with non-zero values are also countable. ■