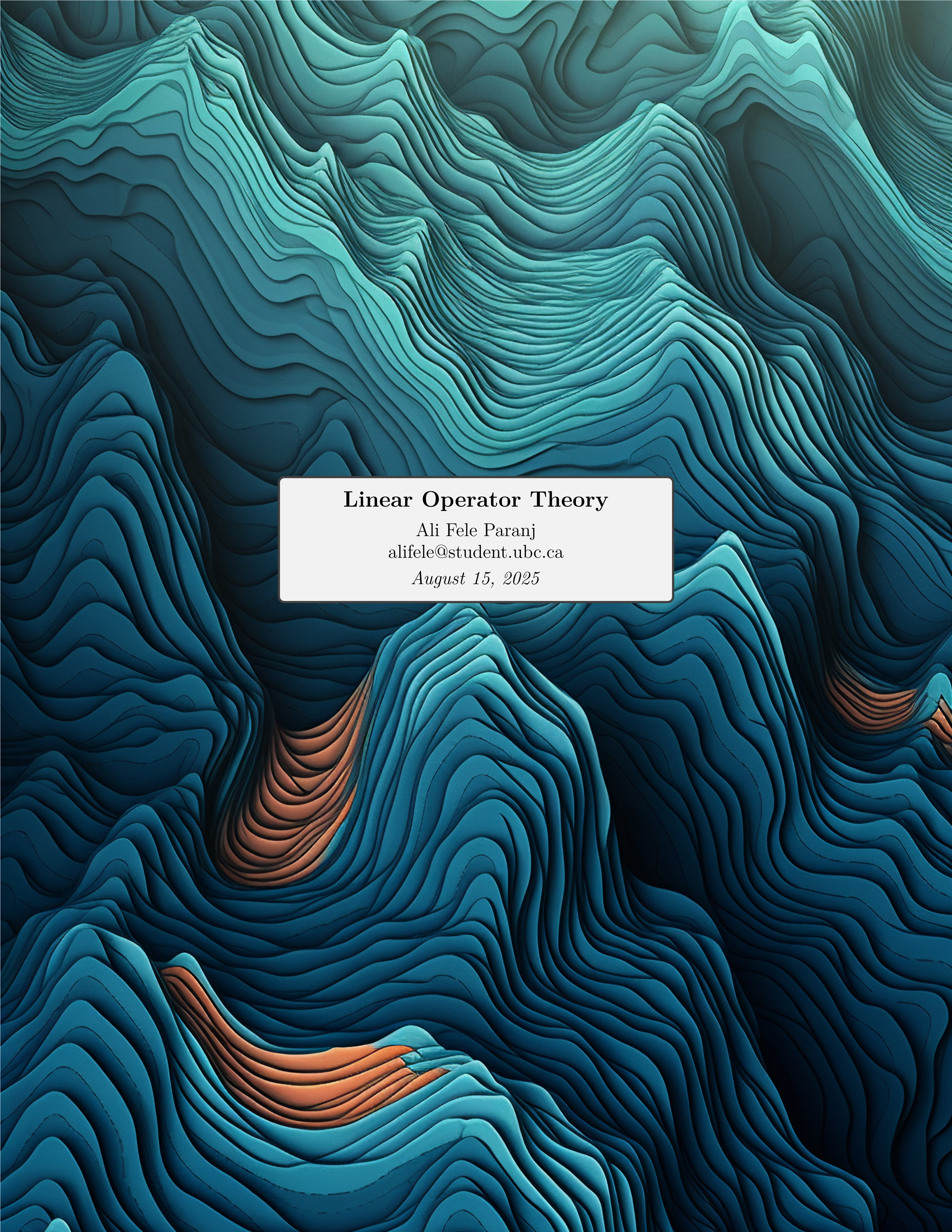


Linear Operator Theory

Ali Fele Paranj
alifele@student.ubc.ca

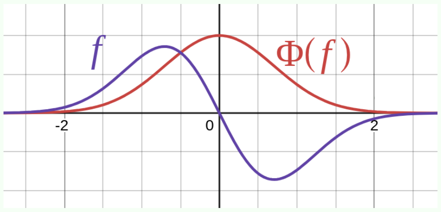
August 15, 2025

0. Contents

1. Topological Spaces

1.1 Random Notes

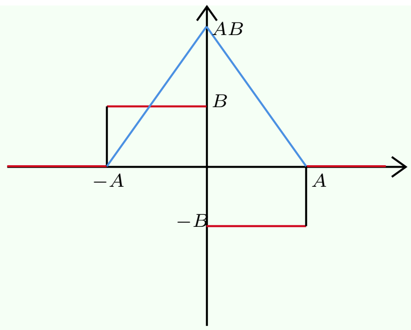
Observation 1.1.1 — Some notes on Example 2 page 63. In this example we study $X \subset L^2(-\infty, +\infty)$ the functions in $L^2(-\infty, +\infty)$ that their primitive functions also belongs to $L^2(-\infty, +\infty)$. I.e. they roughly look like the following function



Then we define the map $\Phi : X \to X$ given by

$$\Phi(x)(t) = \int_{-\infty}^{t} x(\tau)d\tau.$$

We want to show that this map is not continouse at any point of its domain. Since $L^2(-\infty, +\infty)$ is a vector space, it is enough to show this for the origin. So we want to show that this map is not continuous at the origin. Let x be the function whose graph is given as below (the red curve). It is easy to see that the blue curve is its primitive function.



The L^2 norm is roughly the area under the curve, so the L^2 norm of x is AB while the L^2 norm of $\Phi(x)$ is AB^2. Now to prove that the mapping Φ is not continuous at the origin, we need to change A, B such that the L^2 norm x goes to zero but the norm of $\Phi(x)$ gets larger and larger. Suppose we want to make the norm of x to be half and double the norm of $\Phi(x)$. So we need to solve

$$\frac{A'B'}{AB} = \frac{1}{2}, \quad \frac{A'B'^2}{AB^2} = 1.$$

It is easy to see that we need to have

$$\frac{A'}{A} = 4, \qquad \frac{B'}{B} = \frac{1}{2}.$$

Observation 1.1.2 — Some notes on Example 4, page 64. Putting this example in contrast to the example above, teaches us a lot! In the example above we observed that the "integration" of $L^(-\infty, +\infty)$ functions is nowhere continuous. However, in this example, we will see that the integration operator in $L^2[0,T]$ is a uniformly continuous operator! This is a huge contrast and the surprising fact is that in both cases we are dealing with the "same" operator and the only difference is their domain. To see this let $X = L^2[0,T]$ and let $\Phi : L^2[0,T] \to L^2[0,T]$ defined as

$$[\Phi(x)](t) = \int_0^t x(\tau) d\tau.$$

■ **Remark 1.1** We need to check to see if $\Phi(x)$ really belongs to $L^2[0,T]$. However the argument will be very similar to what we will see below, so we have skipped this part.

Since $L^2[0,T]$ is a vector space and Φ is a linear map, it is enough to show the continuity at the origin. Let $x \in L^2[0,T]$. Then

$$\begin{aligned}
d(\Phi(x),0)^2 &= \int_0^T \left| \int_0^t [\Phi(x)](\tau) d\tau \right|^2 dt \le \int_0^T \left(\int_0^t \left|[\Phi(x)](\tau)\right| d\tau \right)^2 dt \\
&\le \int_0^T \left(\int_0^T \left|[\Phi(x)](\tau)\right| d\tau \right)^2 dt \\
&\le \int_0^T \left(\int_0^T 1 d\tau \right) \left(\int_0^T \left|[\Phi(x)](t)\right|^2 d\tau \right) dt \\
&= T^2 \int_0^T \left|[\Phi(x)](t)\right|^2 d\tau \\
&= T^2 d(x,0).
\end{aligned}$$

So we have

$$d(\Phi(x),0) \le T d(x,0).$$

Since the expansion factor T does not depend on x, we can conclude that Φ is actually uniformly continuous.

Observation 1.1.3 — Discrete version of the example above. We can construct a discrete version of the example above, where the underlying space is $\mathbb{R}^n$, and the integration operator Φ is a linear map $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$\Phi : \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_1 + x_2 + x_3 \\ \vdots \\ x_1 + \cdots + x_n \end{bmatrix}.$$

This can be represented by the matrix

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

However, from the Example 1 from page 62 we know that a linear map between vector spaces is continuous with the operator norm at most $\sum_{ij} \left|m_{i,j}\right|$. So the operator norm of this matrix will be

$$\|\Phi\| \le \left(\frac{n(n+1)}{2} \right)^{1/2} \approx n.$$

Summary ✒ 1.1 — Some notes on Example 5 on page 65. The example 5 demonstrates the observation above in a very useful way. It explicitly shows that the example above is just a very special matrix and we can have more general linear maps with arbitrary elements in their matrix representation. So the matrix operators between finite dimensional vector spaces correspond to the integration kernels. consider $L_2[0,T]$ and let $k \in L^2([0,T] \times [0,T])$, i.e.

$$\int_0^T \int_0^T \left|k(t,\tau)\right|^2 d\tau dt = M^2 < \infty.$$

And we define the linear map $\Phi_k : L_2[0,T] \mapsto L_2[0,T]$

$$[\Phi_k(x)](t) = \int_0^T k(t,\tau)x(\tau)d\tau.$$

The Example 5 shows that this is a continouse map (bounded linear map) with operator norm

$$\|\Phi_k\| \leq M.$$

Compare this with Example 1.

■ **Remark 1.2** One special kind of kernel integrals are convolutions where the kernel is of the form $k(t,\tau) = h(t-\tau)$ for some function h of appropriate type (I think it is enough to have $L_2[0,T]$, with the definition that $h(t-\tau) = 0$ if $t - \tau \notin [0,T]$). We often write convolution as

$$(f * h)(t) = \int_0^t h(t-\tau)f(\tau)d\tau.$$

Summary ✒ 1.2 — Checking Continuity. By definition, a map between two topological spaces $F : U \to V$ is continuous if and only if the pre-image of open sets are open. However, instead of checking the pre-image of every open sets in $\mathcal{T}_V$, it is enough to check it on the basis of the topology. Since in a metric space, the basis of the topology is the set of all open balls, then to check to see if a map is continuous, it is enough to check to see if the pre-image of every open ball in V is open in U, i.e. contains an open ball of U.

Summary ✒ 1.3 — Isometric embedding of metric space in a Banach space. In Problem 1.2.5.3 we showed a very important fact! We showed that for any metric space (X,d) one can come up with in isometric embedding of X onto $BC(X,\mathbb{R})$, the space of all continuous and bounded functions defined on X. For $x_0 \in X$ the embedding $\Phi : X \to BC(X,\mathbb{R})$ is given as

$$[\Phi(y)](x) = d(x,y) - d(x,x_0).$$

This is known as Kuratowski embedding. Also, note that $BC(X,\mathbb{R})$ is a Banach space, because $\mathbb{R}$ is complete.

1.2 Solved Problems

1.2.1 Metric Spaces

Summary 1.4 — A metric for bounded functions defined on any set S. Let S be a nonempty set and let $X = B(S)$ denote the collection of all bounded real-valued functions defined on S. Then

$$d(f,g) = \sup\{\left|f(s) - g(s)\right| : s \in S\},$$

is a metric on X.

Summary 1.5 — A metric for differential operators. Let $X(n)$ denote all the differential operators of the form

$$P(D) = p_0 D_0 + p_1 D_1 + \cdots + p_n D_n.$$

Then d is a metric on $X(n)$: $d(P(D), Q(D)) = \sum_{i=0}^{n} |p_i - q_i|$.

Summary 1.6 Let X denote the collection of all bounded closed intervals $[a,b]$ from the real line. Let $[a,b]$ and $[c,d]$ be two sets. Then the symmetric difference $[a,b]\delta[c,d]$ is the union of at most two bounded intervals. $\rho([a,b],[c,d])$ to be the sum of the lengths of these intervals. Then ρ is a pseudometric. In particular, the distance between any two single points is zero.

■ **Remark 1.3** TODO: I am still thinking why the triangle inequality of the metric defined above holds.

■ **Problem 1.2.1.1** Let d be a metric on X and let $d_\alpha = \alpha d$, where $0 < \alpha < 1$. Show that $d_\alpha \not\equiv d$ if X has more than two points.

Solution Let X have more than one point. Then $\exists x, y \in X$ such that $d(x,y) \neq 0$. Then $d(x,y) \neq \alpha d(x,y) = d_\alpha(x,y)$. So $d \not\equiv d_\alpha$.

■ **Problem 1.2.1.2** Show that $d(x,y) = |x - y|$ is a metric on the real number $\mathbb{R}$. Show that it is also a metric on the set of complex numbers $\mathbb{C}$.

Solution First, we show $(\mathbb{R}, d)$ is a metric space. We need to show:

(i) First, we show that d is always positive. Since $|x| = \max\{x, -x\}$, we have $\pm x \leq |x|$. Adding to two inequalities we will get $0 \leq |x|$. Now we want to show that $|x-y| = 0$ then $x = y$. $|x-y| = 0$ implies $x - y = 0$, and then we will have $x = y$.

(ii) Since $|x-y| = \max x - y, y - x$, we will have $|y - x| = \max y - x, x - y = |x - y|$.

(iii) Triangle inequality: From the definition of absolute value we have $|x| = \max\{x, -x\}$. So $x \leq |x|$ and $-|x| \leq x$. Let $a, b \in \mathbb{R}$. Then

$$a \leq |a|\,,\ b \leq |b| \quad \Longrightarrow \quad a + b \leq |a| + |b|\,.$$

Furthermore,

$$-|a| \leq a,\ -|b| \leq b \quad \Longrightarrow \quad -|a| - |b| \leq a + b.$$

Combining these two inequalities, we will get

$$-|a| - |b| \leq a + b \leq |a| + |b|.$$

So

$$|a+b| \leq |a| + |b|.$$

Now we show that $(\mathbb{C}, d)$ is a metric space. For this purpose we can utilize the properties of the complex conjugation and work with the definition $|z_1 - z_2| = \sqrt{(z_1 - z_2)\overline{(z_1 - z_2)}}$. But instead, we want to show that there is a bijection from $(\mathbb{C}, d)$ (we don't know yet if it is a metric space, otherwise we could call this map an isometric bijection) to $(\mathbb{R}^2, d_{\mathbb{R}^2})$. Recall

$$\mathbb{C} = \{a + ib : \ a, b \in \mathbb{R}\}.$$

Let $\phi : \mathbb{C} \to \mathbb{R}^2$ given by $a + ib \mapsto (a, b)$. This is a bijection, since for any $a + ib \in \mathbb{C}$ we have a unique $(a, b) \in \mathbb{R}^2$ (uniquely determined by a, b), and for any $(a, b) \in \mathbb{R}^2$ we have a uniquely determined $a + ib \in \mathbb{C}$. Also we claim that

$$d(z, w) = d_{\mathbb{R}^2}(\phi(z), \phi(w)).$$

This is true because

$$d(z,w) = d(z_1+iz_2, w_1+iw_2) = \sqrt{(z_1 - w_1)^2 + (z_2 - w_2)^2} = d_{\mathbb{R}^2}((z_1, z_2), (w_1, w_2)) = d_{\mathbb{R}^2}(\phi(z), \phi(w)).$$

We can use this map to transfer all of the properties of $d_{\mathbb{R}^2}$ in $\mathbb{R}^2$ to d in $\mathbb{C}$. So $(\mathbb{C}, d)$ is indeed a metric space. So now we can call ϕ an isometric bijection.

■ **Remark 1.4** In the question above, for $(\mathbb{R}, d)$ we used the definition $|x| = \max\{x, -x\}$, and for $(\mathbb{C}, d)$ we used the definition $|x| = \sqrt{x\overline{x}}$.

■ **Problem 1.2.1.3** Let $d(x, y)$ be a metric on X. Show that

$$d_1(x,y) = \frac{d(x,y)}{1 + d(x,y)}, \qquad d_2(x,y) = \min\{1, d(x,y)\},$$

are also metrics on X. Show that every set in the metric space (X, d_1) and (X, d_2) is bounded.

Solution (i) Showing that $d_1(x, y)$ is a metric: Positive definiteness: since the enumerator and denumenator are both positive, then $d_1(x, y)$ is also positive for all $x, y \in X$. Let $x, y \in X$ such that $d_1(x, y) = 0$. This implies $d(x, y) = 0$, hence $x = y$. Symmetry follows from d being symmetric. For the triangle inequality we use the fact that $f(x) = \frac{x}{1+x}$ is strictly increasing for $x \geq 0$ (because it has positive derivative), and also it is subadditive. To see the subadditivity, observe that

$$f(x+y) \leq \frac{x+y}{1+x+y} = \frac{x}{1+x+y} + \frac{y}{1+x+y} \leq \frac{x}{1+x} + \frac{y}{1+y} = f(x) + f(y).$$

Using the fact that for all $x, y, z \in X$ we have

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

using the monotinicity of f and then its sub additivity, we will get

$$d_2(x,z) \leq d_2(x,y) + d_2(y,z).$$

(See the summary box below for a detailed argument of a more general setting).

(ii) Showing that $d_2(x,y) = d(x,y) \wedge 1$ is a metric: Positive definiteness: Since $d(x,y) \geq 0$ then $d(x,y) \wedge 1 \geq 0$. Let $x, y \in X$ such that $d(x,y) \wedge 1 = 0$. So $d(x,y) = 0$. This implies $x = y$. So d_2 is positive definite. Symmetric: Let $x, y \in X$. Then $d_2(x,y) = d(x,y) \wedge 1 = d(y,x) \wedge 1 = d_2(y,x)$, where we have used the fact that $d(x,y) = d(y,x)$. So $d_2(x,y) = d_2(y,x)$. For the triangle inequality we use the following lemma:

Lemma 1.1 Let $a, b \geq 0$. Then

$$a < b \qquad \Longrightarrow \qquad a \wedge 1 \leq b \wedge 1.$$

Also

$$(a+b) \wedge 1 \leq a \wedge 1 + b \wedge 1.$$

Proof. For the first implication, when $a < b$, then there are three cases: $0 \leq a < b \leq 1$, $0 \leq a \leq 1 < b, 1 < a < b$. In the first case $a \wedge 1 = a, b \wedge 1 = b$ So $a \wedge 1 \leq b \wedge 1$ holds. This holds for the other cases as well.

For the second implication, again we will show in cases: When $a+b < 1$, $a+b = 1$, $a+b > 1$. For the first case we can only have $a, b < 1$. So $(a+b) \wedge 1 = a+b$, $a \wedge 1 = a$, $b \wedge 1 = b$. So the desired inequality holds. For the second case, again $a = a \wedge 1$, $b = b \wedge 1$, and $(a+b) \wedge 1 = a+b = 1$. So the desired inequality holds. For the third case, WLOG we can assume that $a \leq b$. Since $a+b \geq 1$, then at least one of a or b should be larger than 1. WLOG we can assume $a \geq 1$. So $(a+b) \wedge 1 = 1$ and $a \wedge 1 = 1$. So the desired inequality holds. □

So by the triangle inequality $d(x,z) \leq d(x,y) + d(y,z)$, and using the lemma above we can conclude that

$$d(x,z) \wedge 1 \leq d(x,y) \wedge 1 + d(y,z) \wedge 1.$$

Now to show that all of the sets in X, with the metrics above are bounded, it is enough to observe that for all $x, y \in X$ we have $d_1(x,y) \leq 1$ and $d_2(x,y) \leq 1$.

Summary 1.7 — Transformation of the metric. Using the proof ideas of the examples above, we can generalize it in the following proposition.

Proposition 1.1 Let (X, d) be a metric space and $\phi : [0, +\infty) \to \mathbb{R}$ be a real valued functions such that

(i) $\phi(x) = 0$ if and only if $x = 0$.

(ii) monotone increasing,

(iii) subadditive: $\forall x, y \geq 0$ we have $\phi(x+y) \leq \phi(x) + \phi(y)$. Then $\tilde{d} = \phi \circ d$ is also a metric, and $(X, \tilde{d})$ is a metric space.

Proof. Since $\phi(x) = 0$ if and only if $x = 0$, then $\tilde{d} = 0$ if and only if $d = 0$, if and only if $x = y$. So $\tilde{d}$ is positive definite. $\tilde{d}$ is also symmetric, because $\forall x, y \in X$, we have $d(x,y) = d(y,x)$ that implies $\tilde{d}(x,y) = \phi(d(x,y)) = \phi(d(y,x)) = \tilde{d}(y,x)$. For the triangle inequality, we have

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

We apply (ii) above and we will get

$$\tilde{d}(x,z) \leq \phi(d(x,y) + d(y,z)),$$

and we now apply (iii) above to get

$$\tilde{d}(x,z) \leq \tilde{d}(x,y) + \tilde{d}(y,z).$$

□

Summary 1.8 — Metric transformation examples. Using the example above, and the remark below, the followings are example functions that can transform a metric d to a new metric $\tilde{d} = f(d)$.

(i) $f_1(x) = x \wedge 1$.

(ii) $f_2(x) = x/(1+x)$.

(iii) $f_3(x) = x^\alpha/(1+x^\alpha)$ for $0 < \alpha \leq 1$.

■ **Remark 1.5** The reason that the example (iii) above works is that f_3 satisfies all the properties in the summary box above. To see the sub-additivity, it is enough to show the sub-additivity of x^α when $0 < \alpha \leq 1$. Let $\alpha = 1 - \beta$. Then

$$(x+y)^\alpha = \frac{x+y}{(x+y)^\beta} < \frac{x}{(x+y)^\beta} + \frac{y}{(x+y)^\beta} \leq \frac{x}{x^\beta} + \frac{y}{y^\beta} = x^\alpha + y^\alpha.z$$

■ **Problem 1.2.1.4** A real-valued function $\rho(x,y)$ is said to be a pseudometric on X if it satisfies conditions (M1), (M3), and (M4).

(a) Show that $\rho(x,y) \equiv 0$ is a pseudometric on any set X.

(b) Show that $\rho((x_1,x_2),(y_1,y_2)) = |x_1 - y_1|$ is a pseudometric in the plane $\mathbb{R}^2$.

Solution (a) $\rho(x,y) \equiv 0$ satisfies $(M1), (M3)$, and $(M4)$ vacuously. However, if X has only one element, then ρ is indeed a metric. But if X has more than two elements, then $\exists x, y \in X$ such that $x \neq y$ but $\rho(x,y) = 0$.

(b) Being symmetric (M3) and the triangle inequality (M4) follows from the properties of $|\cdot|$. However, for all $x, y \in \mathbb{R}^2$ that have the same first component (i.e. points on the vertical lines) have $d(x,y) = 0$. However, this will be a metric on the quotient space $\mathbb{R}^2/W$ where $W = \operatorname{Span}\{(0,1)\}$.

■ **Problem 1.2.1.5** Show that if A is nonempty, in a metric space (X,d), then $\operatorname{diam} A = 0$ if and only if A consist of a single point. Is this true in a pseudometric space?

Solution For the forward direction assume for $A \subset X$ we have $\operatorname{diam} A = 0$. So

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

This implies that $\forall x, y \in A$ we have $d(x,y) = 0$. Since d is a metric, then $A = \{x\}$ is a singleton. For the converse, let $A = \{x\}$. Then it follows immediately that $\operatorname{diam} A = 0$.

No this is not true in the case of the pseudometric spaces. In our proof above, in the forward direction, the logic breaks if d is a pseudometric. I.e. it is possible to have $\sup_{x,y\in A}\{d(x,y)\} = 0$ but A is not singleton.

1.2.2 Examples of Metric Spaces

■ **Problem 1.2.2.1** Let $X = \mathbb{R}^2$ and let

$$d(x,y) = (|x_1 - y_1|^{1/2} + |x_2 - y_2|^{1/2})^2.$$

Is (X,d) a metric space?

Solution No. Because the triangle inequality breaks. Let

$$x = (0,0), \qquad y = (0,1/4), \qquad z = (1/4,1/4).$$

It is easy to check that the triangle inequality $d(x,z) \le d(x,y) + d(y,z)$ does not hold. Because

$$1 = d(x,z) \le d(x,y) + d(y,z) = 1/4 + 1/4 = 1/2,$$

is a contradiction.

■ **Remark 1.6** With a similar idea of proof, it is easy to show that d_p fails to satisfy the triangle inequality if $0 < p < 1$.

■ **Problem 1.2.2.2** In $\mathbb{R}^2$ let $A\{x = (x_1,x_2) : (|x_1|^2 + |x_2|^2)^{1/2} < 1\}$, and

$$d(x,y) = \{|x_1 - y_1|^2 + |x_2 - y_2|^2\}^{1/2}.$$

Compute $d(x,A)$. Show that $d(x,A) = 0$ if and only if

$$|x_1|^2 + |x_2|^2 \le 1.$$

Solution Let $x = (x_1,x_2) \in \mathbb{R}^2$ (WLOG we assume that $x_1 > 0$) that does not belong to A. The ray that passes through x is given by $y = \frac{x_2}{x_1}x$, and this ray intersects the unit circle at

$$\tilde{x} = \left(\frac{1}{\sqrt{1+(x_2/x_1)^2}}, \frac{x_2/x_1}{\sqrt{1+(x_2/x_1)^2}}\right).$$

So the distance between x and the point above is

$$d(x,\tilde{x}) = \left(\left|\frac{1}{\sqrt{1+(x_2/x_1)^2}} - x_1\right|^2 + \left|\frac{x_2/x_1}{\sqrt{1+(x_2/x_1)^2}} - x_2\right|^2\right)^{1/2}.$$

To show that $d(x,A) = 0$ if and only if $x_1^2 + x_2^2 \le 1$, we first do the forward direction. Let $d(x,A) = 0$. From our explicit formula above, it is easy to see that we should have

$$x_1 = \frac{1}{\sqrt{1+(x_2/x_1)^2}}, \qquad x_2 = \frac{x_2/x_1}{\sqrt{1+(x_2/x_1)^2}}.$$

It is easy to check that $x_1^2 + x_2^2 = 1$ which implies $x_1^2 + x_2^2 \le 1$. For the converse direction, we assume $x_1^2 + x_2^2 \le 1$. When $x_1^2 + x_2^2 < 1$, then $x \in A$, and by the definition of $d(x,A)$ we see

$d(x, A) = 0$. When $x_1^2 + x_2^2 = 1$, then we can write $1 + (x_1/x_2)^2 = 1/x_1^2$. Using the fact that $x_1 > 0$ this implies that

$$d(x, A) = 0,$$

where we have used the explicit formula above.

■ **Problem 1.2.2.3** In example 5, the metric d_∞ was defined with a "sup" instead of "max". In order to see the necessity of this let $x = (x_1, x_2, \cdots), y = (y_1, y_2, \cdots)$ be given by

$$x_n = \frac{1}{n+1}, \qquad y_n = \frac{n}{n+1}.$$

And argue why "sup" should be used instead of "max"? What about Example 8 and Example 11?

Solution The reason is that if on ℓ_∞ we define d_∞ by "max", then for the presented examples $d_\infty(x, y)$ does not exist. In the case of example 8, we **can** replace "max" with "sup". The reason is that if $x, y \in C[0, T]$, then $f = |x - y| \in C[0, T]$. On the other hand, a continuous function defined on a compact set, attains its maximum and minimum. For Example 11, we can **not** use "max" in place of "sup".

■ **Problem 1.2.2.4** Sketch the set of points $x = (x_1, x_2)$ in $\mathbb{R}^2$ for which

$$d_p(0, x) = 1,$$

where $0 = (0, 0)$ and d_p is defined in Example 1. With x, y fixed show that $d_p(x, y)$ is decreasing in p. Hence

$$d_p(x, y) \leq d_q(x, y).$$

Solution It is easy to find such sketches online. For the second part, fix $x, y \in \mathbb{R}^2$. Then

$$\frac{dd_p(x, y)}{dp} = -\frac{A}{p^2},$$

where A is a positive quantity (can be calculated explicitly by using the chain rule). So $d_p(x, y)$ is decreasing.

Summary 1.9 For fixed $x, y \in \mathbb{R}^n$, $d_p(x, y)$ is decreasing in p. So

$$p \leq q \quad \Longrightarrow \quad d_p(x, y) \geq d_q(x, y).$$

■ **Problem 1.2.2.5** Let C be the unit circle in the complex plane, that is $C = \{z : |z| = 1\}$. Let X denote all complex-valued functions $f(z)$ defined on C for which

$$\int_C \left|f(z)\right| dz < \infty.$$

Show that

$$d(f, g) = \left(\int_C \left|f(z) - g(z)\right| dz\right)^{1/2} = \left(\int_0^{2\pi} \left|f(e^{2i\pi\theta}) - g(e^{2i\pi\theta})\right| d\theta\right)^{1/2}$$

is a metric on X.

Solution It is easy to check the other properties of being a metric. Thus we will show that d satisfies the triangle inequality. To see this let $f, g, h \in X$. Then

$$d(f,g) = \left(\int \left|\tilde{f} - \tilde{g}\right|^2\right)^{1/2} = \left(\int \left|\tilde{f} - \tilde{g} \pm \tilde{h}\right|^2\right)^{1/2} = \left(\int |u+v|^2\right)^{1/2},$$

where $u = \tilde{f} - \tilde{h}$ and $v = \tilde{h} - \tilde{g}$, and $\tilde{f}(\theta) = f(e^{2i\pi\theta})$. Observe that

$$\int |u+v|^2 = \int (u+v)\overline{(u+v)} = \int (|u|^2 + |v|^2 + 2\operatorname{Re}(u\overline{v})) = \int |u|^2 + \int |v|^2 + 2\int \operatorname{Re}(u\overline{v}).$$

Also

$$\int \operatorname{Re}(u\overline{v}) \leq \operatorname{Re}(\int u\overline{v}) \leq \operatorname{Re}(\|u\|\|v\|) = \|u\|\|v\|,$$

where $\|\cdot\|$ is the norm that induces the metric. So

$$\int |u+v|^2 \leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| = (\|u\| + \|v\|)^2.$$

so we can conclude

$$d(f,g) = \left(\int |u+v|^2\right)^{1/2} \leq \|u\| + \|v\| = d(f,h) + d(h,g).$$

■ **Problem 1.2.2.6** Let X denote the class of all complex-valued functions $f(z)$ that are analytic for $|z| < 1$. Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ and $g(z) = \sum_{n=0}^{\infty} b_n z^n$ be the power series expansion for f, g in X. Show that the following functions are metrics, or pseudo-metrics.

(a) $d(f,g) = \sup\{\left|f(z) - g(z)\right| \ : \ |z| \leq \rho\}$, where $0 < \rho < 1$.

(b) $d(f,g) = |a_0 - b_0|$.

(c) $d(f,g) = |a_0 - b_0| + |a_1 - b_1|$.

(d) $d(f,g) = \sum_{n=0}^{\infty} |a_n - b_n| \rho^n$, where $0 < \rho < 1$.

(e)
$$d(f,g) = \sup\{\left|\int_{|\xi|=0} \frac{f(\xi) - g(\xi)}{z - \xi} d\xi\right| \ : \ |z| < \rho\}, \quad \text{where } 0 < \rho < 1.$$

Solution Since we need to choose between a metric or a pseudometric, we don't check the triangle inequality.

(a) Is a metric.

(b) Is a pseudometric. Because the distance between functions in which in their expansion only the first term is the same will be zero.

(c) Is a pseudometric. Because of a similar argument as above.

(d) Is a metric. Because if $d(f,g) = 0$, then every term in the positive series should be zero, hence $|a_n - b_n| = 0$ for all $n \in \mathbb{N}$.

(e) Still thinking!

1.2.3 Continuity

■ **Problem 1.2.3.1** Let $C^\infty[0,T]$ denote the set of all real-valued smooth functions defined on $[0,T]$. Let D be the mapping of $C^\infty[0,T]$ to itself defined by

$$Df = \frac{df}{dt}.$$

(a) Let d_1 be the sup-metric on $C^\infty[0,T]$. Is D a continuous mapping of $(C^\infty[0,T], d_1)$ into itself?

(b) Define a metric d_2 on $C^\infty[0,T]$ by

$$d_2(x,y) = d_1(x,y) + d_1(Dx, Dy).$$

Is D a continouse mapping of $(C^\infty[0,T], d_2)$ into $(C^\infty[0,T], d_1)$?

Solution (a) No. Consider the sequence of functions $(f_n)_n$ where

$$f_n(t) = \frac{\sin(2^n x)}{n}.$$

Then $d_1(f_n, 0) \to 0$ as $n \to \infty$ but $d_1(Df_n, 0)$ becomes arbitrary large.

(b) Yes. Let $f, g \in C^\infty[0,T]$ then

$$d_1(Df, Dg) \leq d_1(Df, Dg) + d_1(f,g) = d_2(f,g).$$

So for any $\epsilon > 0$ we let $\delta = \epsilon$. Then if $d_2(f,g) < \delta$ it implies that $d_1(f,g) < \epsilon$.

■ **Problem 1.2.3.2** A delay line is a device whose output is ideally a delayed version of its input, i.e.

$$y(t) = x(t - \tau).$$

Suppose that $x, y \in L_2(-\infty, +\infty)$, where $L_2(-\infty, +\infty)$ has the usual metric. Is the mathematical model of the delay line a continouse mapping of $L_2(-\infty, +\infty)$ into itself?

Solution We know that integration, in the sense of Lebesgue, is not sensitive to shift. I.e.

$$\int_{-\infty}^{+\infty} f(t)dt = \int_{-\infty}^{+\infty} f(t-\tau)dt,$$

for all $\tau \in \mathbb{R}$. So for any $f, g \in L_2(-\infty, +\infty)$ we have

$$d(f,g) = d(Sf, Sg).$$

■ **Problem 1.2.3.3** Let $Y = C[0,T]$ be given with the sup-metric $d(x,y)$. Let

$$X = (C[0,T], d) \times (C[0,T], d),$$

be the product space. Consider the mapping F of X into Y defined by $F(x) = x_1 x_2$, where $x = (x_1, x_2)$. (That is F is a multiplier). Is F continuous? Is F uniformly continuous?

Solution Assume that the norm on X is $d_2(f,g) = d(f_1,g_1) + d(f_2,g_2)$ (we could chose other metrics on the product space as well, but we chose this one for the convenience). Then we can write

$$\begin{aligned}
d(F(f),F(g)) &= \sup_{t\in[0,T]} \big|[F(f)](t) - [F(g)](t)\big| \\
&= \sup_{t\in[0,T]} \big|f_1(t)g_1(t) - f_2(t)g_2(t)\big| \\
&= \sup_{t\in[0,T]} \big|f_1(t)g_1(t) - f_2(t)g_2(t) \pm f_1(t)g_2(t)\big| \\
&= \sup_{t\in[0,T]} \big|f_1(t)(g_1(t) - g_2(t)) + g_2(t)(f_1(t) - f_2(t))\big| \\
&\le \|f_1\|_\infty d(g_1,g_2) + \|g_2\|_\infty d(f_1,f_2) \\
&\le \max\{\|f_1\|_\infty, \|g_2\|_\infty\}(d(g_1,g_2) + d(f_1,f_2)) \\
&= \max\{\|f_1\|_\infty, \|g_2\|_\infty\} d_2(f,g).
\end{aligned}$$

So for a given $\epsilon > 0$ we can let $\delta = \epsilon/(\max\{\|f_1\|_\infty, \|g_2\|_\infty\})$. However, since the choice of δ depends on the point f, g, then the uniform continuity fails.

■ **Problem 1.2.3.4** Let (X,d) be a metric space, where X is nonempty. Let $Y = BC(X,\mathbb{R})$ denote the collection of all bounded, continuous real-valued functions defined on X.

(a) Show that the functions

$$\begin{aligned}
f_1 &: x \mapsto \frac{d(x,x_0)}{1+d(x,x_0)}, \\
f_2 &: x \mapsto d(x,x_1) - d(x,x_0), \\
f_3 &: x \mapsto 3,
\end{aligned}$$

are in Y.

(b) Show that

$$\sigma(f,g) = \sup\{\big|f(x) - g(x)\big| : x \in X\}$$

is a metric on Y.

Solution (a) Fix $x_0 \in X$. We want to show the function $f : X \mapsto \mathbb{R}$ given as $f(x) = d(x,x_0)$ is a continuous function. Because

$$d(f(x),f(y)) = \big|f(x) - f(y)\big| = \big|d(x,x_0) + d(y,y_0)\big| \le \big|d(x,y) + d(y,x_0) - d(y,x_0)\big| = d(x,y).$$

On the other hand since the function $g : \mathbb{R} \to \mathbb{R}$ given by $x \mapsto x/(1+x)$ is a bounded continouse function, then $f_1 = d/1 + d$ is also a bounded continouse function. Continuoity of f_2 follows from a similar reasoning. For the boundedness, observe that

$$d(x,x_1) - d(x,x_0) \le d(x,x_0) + d(x_0,x_1) - d(x,x_0) = d(x_0,x_1).$$

(b) Follows from the result of Exercise 13 page 56.

■ **Problem 1.2.3.5** Let $f : X \to X$ be a continuous mapping, where X has a metric d. Let $G(f)$ denote the graph of f in $X \times X$, and Δ the diagonal set

$$\Delta = \{(x, y) : x = y\} = \{(x, x) : x \in X\}.$$

Assume that $X \times X$ has the metric

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

Define $g : \Delta \to G(f)$ by

$$g(x, x) = (x, f(x)).$$

Show that g is continuous. Show that g is invertible. Is g^{-1} continuous?

Solution g is continuous. To see this let $(x_1, x_1), (x_2, x_2) \in \Delta$. So

$$\begin{aligned} d(g(x_1, x_1), g(x_2, x_2)) &= d((x_1, f(x_1)), (x_2, f(x_2))) \\ &= d(x_1, x_2) + d(f(x_1), f(x_2)) \\ &\leq (1 + A(x_1, x_2))d(x_1, x_2) \\ &= \frac{1}{2}(1 + A(x_1, x_2))d((x_1, x_1), (x_2, x_2)). \end{aligned}$$

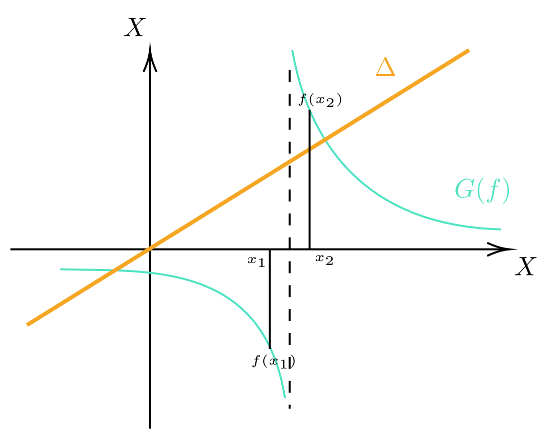
where we have used the continuity of f, i.e. $d(f(x_1), f(x_2)) \leq A(x_1, x_2)d(x_1, x_2)$. So g is a continuous mapping. Define the inverse of g as

$$h = g^{-1} : G(f) \to \Delta, \qquad g^{-1}(x, f(x)) = (x, x).$$

g^{-1} is indeed inverse of g as it is straightforward to check $g^{-1} \circ g$ and $g \circ g^{-1}$ are both identity maps. However, $h = g^{-1}$ is not necessarily continuous. To see this let $(x_1, f(x_1)), (x_2, f(x_2)) \in G(x)$. Then

$$d(h(x_1, f(x_1)), h(x_2, f(x_2))) = d((x_1, x_1), (x_2, x_2)) = d(x_1, x_2) + d(x_1, x_2) = 2d(x_1, x_2).$$

So we can come up with examples where (x_1, x_1) and (x_2, x_2) are close to each other, but $(x_1, g(x_1))$ and $(x_2, g(x_2))$ are not. For instance consider the following graph: Note that the diagram above



is just a schematic representation of the idea behind the provided example to demonstrate some cases where g^{-1} is not continuous, and there is no particular meaning behind the fact that $X \times X$ looks like the Euclidean 2 plane. The only case that this can fail is when all sets in X are bounded (see Exercise 3 page 46).

Observation 1.2.1 — Learning from a mistake. For the question above, after some discussions we found out that our answer is wrong. That is because, assuming the continuoity for f, then $d(f(x_1), f(x_2))$ can not get very large, because $d(f(x_1), f(x_2)) \leq M(x_1, x_2)d(x_1, x_2)$.

1.2.4 Convergence of Sequences

■ **Problem 1.2.4.1** Let X denote the set of all bounded piecewise continuous functions defined on $0 \leq t \leq T$, with the sup-metric d_∞.

(a) Obviously $C[0, T]$ is a subspace of X. Let x_0 be an arbitrary point in $C[0, T]$. Suppose that x_0 is to be approximated by piecewise constant functions is shows in Figure 3.6.1 in text book. That is

$$x_n(t) = x_0(j\frac{T}{n}) \text{ for } j\frac{T}{n} \leq t < (j+1)\frac{T}{n} \text{ and } j = 0, 1, \cdots, (n-1).$$

and

$$x_n(T) = x_0(\frac{n-1}{n}T).$$

Is it true that the sequence $\{x_n\}$ converges in (X, d)? If so, is it true that $x_0 = \lim_{n\to\infty} x_n$? [Hint: Use the fact that a real-valued continuous function, defined on a bounded closed interval is uniformly continuous, compare with Exercise 13, Section 17.]

(b) Suppose x_0 is not restricted to the subspace $C[0, T]$, and suppose that x_0 is approximated by function x_n as in (a). Is it true that $x_0 \lim_{n\to\infty} x_n$?

(c) Consider a different metric on X, namely the d_2 metric. Does the sequence (x_n) converges to x_0 in (X, d_2)?

Solution (a) Since the points in $C[0, T]$ are uniformly continuous, then for any choice of $\epsilon > 0$ we can choose $\delta > 0$ such that $|x - y| < \delta$ implies $\left|f(x) - f(y)\right| < \epsilon$. Let $n = T/(\delta/2)$. So on each interval $jT/n \leq tM(j+1)T/n$ we have

$$\sup_{jT/n \leq tM(j+1)T/n} \left|x_n(t) - x_0(jT/n)\right| < \epsilon.$$

This also implies

$$\sup_{0\leq t\leq T} \left|x_n(t) - x_0(jT/n)\right| < \epsilon.$$

Thus $x_n \to x_0$ in d_∞ metric.

(b) No. The uniform continuity is crucial as observed above. For instance, consider the function $x_0(t) = 0$ when $t \in [0, T)$ and $x_0(t) = 1$ when $t = 1$. With the approximation scheme as above, $x_n \equiv 0$ for all values of $n \in \mathbb{N}$, however, $d_\infty(x_n, x_0) = 1$ for all $n \in \mathbb{N}$.‘

(c) Yes. On any closed interval that misses one of the points of discontinuity, the function x_0 is uniformly continuous, hence on this interval x_n converges to x_0 uniformly (in sup norm), thus in d_2 norm. On the intervals that contains a point of discontinuity, we can bound the L_2 norm as follows. Choose n large enough that each point of discontinuity sits inside one interval $jT/n \leq t < (j+1)T/n$, and the length of each interval is less than η. Let $\|x_0\|_\infty = M$ and let $K \in \mathbb{N}$ denote the number of discontinuities. So on the intervals that we have a discontinuity point we have

$$\int_{\text{on inter. w. dis. cont.}} \left|x(t) - x_0(t)\right|^2 \leq M^2 K\eta,$$

that goes zero as $\eta \to 0$.

■ **Problem 1.2.4.2** Suppose that in a metric space (X, d) a sequence $\{x_n\}$ converges to a point x_0. Does it follow that

$$d(x_1, x_0) \geq d(x_2, x_0) \geq d(x_3, x_0) \geq \cdots \geq d(x_n, x_0) \geq \cdots ?$$

Either prove that it does, or give a counterexample.

Solution No it does not. For instance consider the sequence $\{\sin(n)\}$.

■ **Remark 1.7** However in the question above, we can say if $x_n \to x_0$ as $n \to \infty$, then there exists a subsequence $\{x_{n_k}\}$ such that

$$d(x_{k_1}, x_0) \geq d(x_{k_2}, x_0) \geq d(x_{k_3}, x_0) \geq \cdots$$

■ **Problem 1.2.4.3** Consider the metric space (X, d) given in Example 16, Section 3. Characterize the collection of all convergent sequences in (X, d)

Solution In this space a sequence is convergent if and only if it is eventually constant. Then it converges to that constant value.

■ **Problem 1.2.4.4** Consider the sequence (x_n), where

$$x_n(t) = (\cos(n!\pi t))^{2n}, \quad n = 1, 2, \cdots .$$

In the metric space $C[0, 1]$ with the sup-metric d_∞. Is (x_n) a convergent sequence?

Solution Let $p \in \mathbb{Q} \cap [0, 1]$ be a rational number. Then we can write $p = \frac{a}{b}$ for $a, b \in \mathbb{N}$, were $(a, b) = 1$. We claim $\forall n > b$ we have $x_n(p) = 1$. That is because

$$x_n(p) = (\cos\left(n!\pi\frac{a}{b}\right))^{2n} = (\cos\left(n(n-1)(n-1)\cdots(b!)\pi\frac{a}{b}\right))^{2n} = (\cos(2k\pi))^{2n} = 1,$$

where $k \in \mathbb{Z}$. On the other hand, for $q \in [0, 1]\backslash\mathbb{Q}$, since the cos function is equal to one if and only if its argument is a multiple of π, then $\cos(n!\pi q) \neq 1$ for all $n \in \mathbb{N}$. So

$$x_n(q) = (\cos(n!\pi q))^{2n} \to 0 \quad \text{as } n \to \infty.$$

So the limiting function x can be written as

$$x(t) = \begin{cases} 1 & t \in \mathbb{Q} \cap [0, 1] \\ 0 & t \in [0, 1]\backslash\mathbb{Q} \end{cases} .$$

However, $x \notin C[0, 1]$ so the sequence does not converge.

■ **Problem 1.2.4.5** If $\{x_n\}$ and $\{y_n\}$ are convergent sequences in metric space (X, d), show that the sequence of real number $\{d(x_n, y_n)\}$ converges to $d(x_0, y_0)$ where $x_0 = \lim_{n\to\infty} x_n$ and $y_0 = \lim_{n\to\infty} y_n$. (This exercise should be reconsidered after studying Section 7).

Solution By the definition of converges $x_n \to x_0$ as $n \to \infty$ is equivalent to $d(x_n, x_0) \to 0$ as $n \to \infty$. Similarly for $\{y_n\}$ we have $d(y_n, y_0) \to \infty$ as $n \to \infty$. So using the triangle inequality for the metric function we can write

$$d(x_n, y_n) \leq d(x_n, x_0) + d(y_n, x_0) \leq d(x_n, x_0) + d(y_n, y_0) + d(x_0, y_0),$$

and also

$$d(x_0, y_0) \leq d(x_n, x_0) + d(x_n, y_0) \leq d(x_n, x_0) + d(x_n, y_n) + d(y_n, y_0).$$

For any given $\epsilon > 0$ we can choose n large enough such that

$$d(x_n, x_0) < \epsilon/2, \quad \text{and} \quad d(y_n, y_0) < \epsilon/2.$$

So combining the inequalities above we will get

$$d(x_0, y_0) - \epsilon \leq d(x_n, y_n) \leq d(x_0, y_0) + \epsilon.$$

Or equivalently

$$\left| d(x_n, y_n) - d(x_0, y_0) \right| < \epsilon.$$

So $d(x_n, y_n) \to d(x_0, y_0)$ as $n \to \infty$.

■ **Problem 1.2.4.6** Let a be a real number satisfying $0 \leq a \leq 1$ and set $b = 1 - a$. Let $y_0 = 0$ and

$$y_{n+1} = \frac{1}{2}(b + y_n^2).$$

(a) Show that the sequence $\{y_n\}$ is bounded and monotone, and therefore converges. Let $y = \lim y_n$, and $x = 1 - y$. Show that $x^2 = a$.

(b) Modify part (a) for the case $a > 1$.

Solution (a) First we show the sequence being bounded. We do it by proof by induction. First, observe that $y_0 = 0 \leq 1$. Then assume $y_k \leq 1$ and we show that $y_{k+1} \leq 1$. This is true because

$$y_{k+1} = \frac{1}{2}(b + y_k^2) \leq \frac{1}{2}(1 + 1) = 1.$$

So by induction $y_n \leq 1$ for all $n \in \mathbb{N}$.

To show that the sequence is monotonically increasing, we also use the proof by induction. Since $y_0 = 0$, we have $y_1 = b/2$. So $y_1 - y_0 = b/2 \geq 0$. Now assume $y_k - y_{k+1} \geq 0$. We want to show $y_{k+1} - y_k \geq 0$. Observe that

$$y_{k+1} - y_k = \frac{1}{2}(b + y_k^2) - \frac{1}{2}(b + y_{k-1}^2) = \frac{1}{2}(y_k - y_{k-1})(y_k + y_{k+1}) \geq 0.$$

So by induction $y_{n+1} - y_n \geq 0$ for all $n \in \mathbb{N}$.

Since $\{y_n\}$ is increasing and bounded, then it converges. So taking the limit from both sides we will have

$$y = \frac{1}{2}(b + y^2).$$

By completing the square, we will have $(y - 1) = \sqrt{1 - b} = \sqrt{a}$. So $x = y - 1 = \sqrt{a}$.

(b) If $a > 1$ then we can use the following identity.

$$\sqrt{a} = \frac{1}{\sqrt{\frac{1}{a}}}$$

In other words, we calculate the square root for $a' = 1/a \leq 1$ and then we have $\sqrt{a} = 1/\sqrt{a'}$.

■ **Problem 1.2.4.7** Let $\{x_n\}$ be a sequence in a metric space (X, d) with the property that for some $\epsilon > 0$ one has $d(x_n, x_m) \geq \epsilon$ for all n, m. Show that $\{x_n\}$ is not convergent.

Solution Assume otherwise. So $\exists x_0 \in X$ such that for all $\epsilon > 0$, exists $N \in \mathbb{N}$ such that $\forall n > N$ we have $d(x_n, x_0) < \epsilon$. This contradicts the assumption.

■ **Problem 1.2.4.8** Let $X = \mathbb{R}^n$ be given with the metric $d(x, y) = \sum_{i=1}^{n} |x_i - y_i|$. Show that a sequence $\{x_n\}$ in X converges to x_0 if and only if $x_{n,i} \to x_{0,i}$ for all $i = 1, \cdots, n$.

Solution $\boxed{\Rightarrow}$ Assume $x_n \to x_0$. Then $d(x_n, x_0) \to 0$. So

$$\sum_{i=1}^{N} \left| x_{n,i} - x_{0,i} \right| \to 0.$$

So each term in the sum should go to zero, i.e. $\left| x_{n,i} - x_{0,i} \right| \to 0$ as $n \to \infty$. So $x_{n,i} \to x_{0,i}$ as $n \to \infty$.

$\boxed{\Leftarrow}$ Assume $x_{n,i} \to x_{0,i}$ for all $i = 1, \cdots, n$. So by definition $\left| x_{n,i} - x_{0,i} \right| \to 0$ as $n \to \infty$. So $\sum_{i=1}^{N} \left| x_{n,i} - x_{0,i} \right| \to 0$ as $n \to \infty$. So $x \to x_0$ in $(\mathbb{R}^N, d)$.

■ **Problem 1.2.4.9** Let $x_n(t)$ be a sequence of continuous real-valued functions where $x_n(t)$ is periodic with period $\tau_n > 0$. Assume that $x_n(t) \to x(t)$ uniformly for $\tau \in \mathbb{R}$ and $\tau_n \to \tau$. Show that $x(t)$ is periodic in t with period τ.

Solution First, observe that $x_n(t + \tau_n) \to x(t + \tau)$ as $n \to \infty$. Indeed, since $t + \tau_n \to t + \tau$ as $n \to \infty$, for any $k \in \mathbb{N}$ we have $x_k(t + \tau_n) \to x_k(t + \tau)$ that follows from continuity of x_n. Because of the uniform convergence of x_n, we can write $x_n(t + \tau_n) \to x(t + \tau)$ as $n \to \infty$. On the other hand, we want to show $x_n(t + \tau_n) \to x(t)$. To see this

$$\left| x_n(t + \tau_n) - x(t) \right| = \left| x_n(t) - x(t) \right| \to 0 \quad \text{as} \quad n \to \infty.$$

So $x_n(t + \tau_n) \to x(t)$ and $x_n(t + \tau_n) \to x(t + \tau)$ both uniformly. Then it follows that $x(t) = x(t + \tau)$.

■ **Problem 1.2.4.10** Let f and g be functions in $C[0, T]$. Define $x_0 = f$ and $x_1 = g$. Let

$$x_{n+1} = \frac{1}{2}(x_n + x_{n-1}), \quad n = 1, 2, \cdots.$$

Show that $\lim x_n = \frac{1}{3}(f + 2g)$ when $C[0, T]$ has the sup-metric.

Solution Let $t \in [0,T]$. Define $a_0 = f(t)$ and $a_1 = g(t)$, and define $a_n = x_n(t)$. Then $\{a_n\}$ is a sequence of real numbers where

$$a_{n+1} = \frac{1}{2}(a_n + a_{n-1}).$$

We can writ this recursive equation as

$$\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_n \end{bmatrix}.$$

So we will have

$$\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = A^n \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}.$$

Writing A in its Jordan block form we will have

$$A = P \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} P^{-1},$$

where

$$P = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}, \qquad P^{-1} = \frac{1}{3} \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}.$$

So

$$\lim A^n = \frac{1}{3} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}.$$

Then it follows that $\lim a_n = 1/3(f + 2g)$. So the sequence $\{x_n\}$ converges to the function $h = \frac{1}{3}(f + 2g)$ point-wise. To show the uniform convergence, i.e. the converges with the sup norm, we use the fact that the point-wise convergence of the uniform /.

1.2.5 More on open sets

■ **Problem 1.2.5.1** Let $f : X \to Y$ be continuous, where (X, d_1) and (X, d_2) are metric spaces. Let A be a connected set in X, that is the metric space (A, d_1) is connected. Show that $f(A)$ is connected in Y, that is $(f(A), d_2)$ is a connected space.

Solution We show this by contrapositive. Assume $(f(A).d_2)$ is not a connected space. So there are non-empty open sets B_1', B_2' in $\mathcal{T}_{f(A)}$ such that $B_1' \cap B_2' = \varnothing$ and $B_1' \cup B_2' = A$. Since $\mathcal{T}_A$ is the induced subspace topology, this implies there exists $B_1, B_2 \in \mathcal{T}$ such that $B_1 \cap A = B_1'$ and $B_2 \cap A = B_2'$ and $B_1 \cap B_2 = \varnothing$ and $f(A) \subset B_1 \cup B_2$. Considering the pre-image one gets

$$A = f^{-1}(f(A)) = f^{-1}(B_1) \cup f^{-1}(B_2).$$

Since f is continuous, $f^{-1}(B_1)$ and $f^{-1}(B_2)$ are both continuous and

$$C_1 = f^{-1}(B_1) \cap A, \qquad C_2 = f^{-1}(B_2) \cap A, \qquad C_1, C_2 \in \mathcal{T}_A$$

such that

$$C_1 \cup C_2 = A, \qquad C_1 \cap C_2 = \varnothing.$$

So (A, d_1) is a connected space.

■ **Problem 1.2.5.2** Suppose that F is a one-to-one mapping of a set X onto a metric space (Y, d_2).

(a) Does $d_1(x_1, x_2) = d_2(f(x_1), f(x_2))$ define a metric on X?

(b) If so, is F a homeomorphism? Anything stronger?

Solution (a) Yes. d_1 is a metric. Because $d_1(x_1, x_1) = d_2(f(x_1), f(x_1)) = 0$. Furthermore, $d_1(x_1, x_2) = d_1(x_2, x_1)$ that follows from the symmetric property of d_2. Lastly, for $x_1, x_2, x_3 \in X$ we can write

$$d_1(x_1, x_2) + d_1(x_2, x_3) = d_2(f(x_1), f(x_2)) + d_2(f(x_2), f(x_3)) \geq d_2(f(x_1), f(x_3)) = d_1(x_1, x_3),$$

where we have used the bijectivity of f.

(b) More explicitly, we want to see if $F : (X, d_1) \to (Y, d_2)$ is a homomorphism. It is. It is in fact an isometry between two metric spaces. To show the homomorphism we need to show that F is continuous with continuous inverse. Let $B_{y_0}(r)$ be an open ball in Y. Then

$$F^{-1}(B_{y_0}(r)) = F^{-1}(\{y \in Y : d_2(y, y_0) < r\}) = \{F^{-1}(y) : d_2(y, y_0) < r\} = \{x \in X : d_1(F(x), F(x_0)) < r\} = B_{x_0}(r$$

where $x_0 = F(y_0)$.

With a similar reasoning one can show that the images of open balls are open (even stronger, i.e. open balls).

■ **Problem 1.2.5.3 — Important.** Let (X, d) be a metric space, where X is nonempty, and let $Y = BC(X, R)$ denote the collection of all bounded, continuous real-valued functions defined on X. Assume that Y has the sup metric. Let x_0 be a fixed point in X and define

$$f_y(x) = d(x, y) - d(x, x_0).$$

Show that the mapping $G : y \mapsto f_y$ is an isometry from X onto a subspace of Y.

Solution First, we want to show that $f_y \in BC(X, \mathbb{R})$. Note that for any $x_0 \in X$, the mapping $d(\cdot, x_0) : X \to \mathbb{R}$ is continuous. That is because

$$\left|d(x, x_0) - d(y, x_0)\right| \leq \left|d(x, y) + d(y, x_0) - d(y, x_0)\right| = d(x, y).$$

Since f_y is sum of two continuous functions, it follows that f_y is also continuous. Furthermore, we want to show that f_y is bounded. This is true since using the triangle inequality one can write

$$\left|d(x, y) - d(x, x_0)\right| \leq \left|d(x, x_0) + d(x_0, y) - d(x, x_0)\right| = d(x_0, y).$$

So for each choice of y one has $\left|f_y(x)\right| \leq d(x_0, y)$. So it follows that

$$\sigma(f, 0) = \sup_x \{f_y(x)\} = d(x_0, y).$$

So f_y is also bounded. So G is indeed a mapping from X onto a subspace of Y. To show G is an isometry, we need to show that G is one-to-one, and also preservers the distance. To show the injectivity of G, let $y_1, y_2 \in X$ such that $G(y_1) \equiv G(y_2)$.

$$f_{y_1}(x) = d(x, y_1) - d(x, x_0), \qquad f_{y_2}(x) = d(x, y_2) - d(x, x_0).$$

Then we can write

$$f_{y_1}(x) - f_{y_2}(x) = d(x, y_1) - d(x, y_2) = 0.$$

I.e. $d(x, y_1) = d(x, y_2)$ for all $x \in X$. This implies $y_1 = y_2$. So G is injective. Lastly, we want to show that G preserves the metric. I.e. we want to show

$$d(y_1, y_2) = \sigma(G(y_1), G(y_2)).$$

One can write

$$\sigma(G(y_1), G(y_2)) = \sup\{\left|d(x, y_1) - d(x, y_2)\right| : x \in X\}.$$

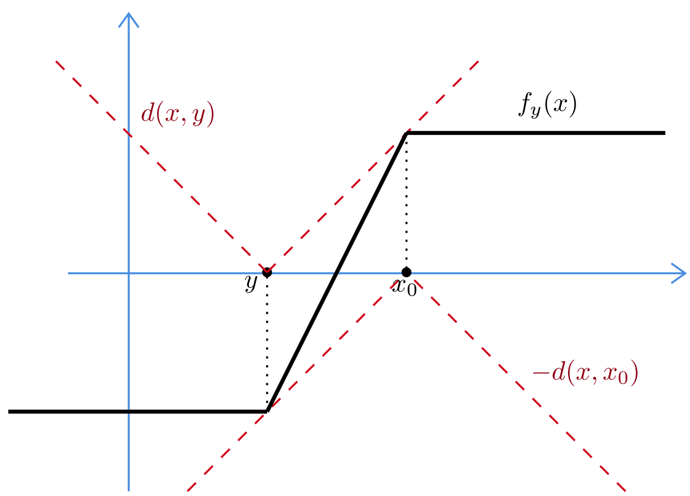
Observe that

$$d(x, y_1) - d(x, y_2) \leq d(x, y_2) + d(y_2, y_1) - d(x, y_2) = d(y_2, y_1).$$

and for $x = y_2$ one has $(d(x, y_1) - d(x, y_2))\Big|_{x=y_2} = d(y_1, y_2)$. This implies that $\sup\{\left|d(x, y_1) - d(x, y_2)\right| : x \in X\} = d(y_1, y_2)$ (because every element in the set is less than $d(y_1, y_2)$ and for $x = y_2$ this value is also achieved. So the sup of the set is $d(y_1, y_2)$).

■ **Problem 1.2.5.4** In the problem above, assume $X = \mathbb{R}$ with the usual metric. Sketch a few of the functions f_y.

Solution The graph of such a function is shown below



■ **Problem 1.2.5.5** Let $X = C(a, b)$ be the space of continuous real-valued functions defined on the open bounded interval (a, b).

2. Algebraic Structure

2.1 Summaries

Summary 2.1 Let $A : V \to W$ be a linear map between vector spaces. Then the pre-image of any linearly independent set contains a linearly independent set. Let $\{w_n\}$ be a linearly independent set of vectors in W. Let $v_n = A^{-1}(w_n)$, i.e. **one** of the pre-images of w_n (if A is not injective, then w_n can have multiple pre-images. But we choose just one pre-image. Any of them will work). Then let $\sum_n \beta_n v_n = 0$. Using the fact that $v_n = A^{-1}(w_n)$, one can write

$$0 = A(\sum_n \beta_n v_n) = \sum_n \beta_n A(A^{-1}(v_n)) = \sum_n \beta_n w_n,$$

which implies $\beta_n = 0$ for all n. So $\{v_n\}$ is linearly independent.

But note that the image of a linearly independent set is not necessarily a linearly independent set. Because a linear map can collapse some of the vectors to the origin, and the resulting collection will not be a linearly independent set of vectors.

Summary 2.2 Even in infinite dimensional vector spaces, an infinite sum of the form $\sum_{n=1}^{\infty} \alpha_n v_n$ for $\alpha_n \in F$ and $v_n \in V$ for all n, is meaningless, and one needs a topological structure on the space to make sense of such infinite sums.

Summary 2.3 — Causality and subspaces. Consider the space of functions defined on the real line (no regularity condition is necessary). Denote this space with X. Let $A : X \to X$ be a linear map between these two space, and let L_T denote the subspace that $x \in L_T$ iff $x(t) = 0$ for all $t \leq T$. L is a causal map iff L_T is invariant under A, that is $A(L_T) \subset L_T$.

Proof. Assume A is causal map and we want to show that L_T is invariant. Let $x \in L_T$. So $x(t) = 0$ for all $t \leq T$. The origin $x_0 \in L_T$ also satisfies $x_0(t) = 0$ for all $t \leq T$. From causality of A we need to have $[Ax](t) = [Ax_0](t)$ for all $t \leq T$. Since A is linear it sends the origin to the origin, i.e. $A(x_0) = 0$. This implies $[Ax](t) = 0$ for all $t \leq T$. So $Ax \in L_T$. So L_T is invariant under L.

Now assume L_T is invariant under A and we want to show that A is causal. Let $x, y \in X$

with $x(t) = y(t)$ for all $t \leq T$. Then $(x-y)(t) = 0$ for all $t \leq T$, this $x - y \in L_T$. So is $A(x-y) = Ax - Ay$. So $(Ax - Ay)(t) = 0$ for all $t \leq T$. So $[Ax](t) = [Ay](t)$ for all $t \leq T$. So A is causal. ☐

Summary 2.4 — Matrix representation of causal maps. The summary box above helps to give a characterization of causal linear maps between finite dimensional vector spaces. Let $L : V \to W$ be a causal linear map both of dimension N. Then for each $n \leq N$, if $v_i = u_i$ for all $i \leq n$, i.e. if their first n coordinates are equal, then $[Lv]_i = [Lu]_i$ for all $i \leq n$, i.e. the first i coordinates of their images is also the same. I.e. the i^{th} coordinate of $[Lv]$ only depends on the first i coordinate of v. This means that the matrix representation of the causal map is a "lower triangular" matrix. So the linear transformation in Example 6 page 65 of the text book is a causal map.

Summary 2.5 — Characterizing the causal linear and time invariant systems. The summary box above, along with the fact that a linear operator of the form

$$y(t) = \int k(t,s)x(s)ds,$$

is time invariant if and only of $k(t,s) = g(t-s)$ for some function g, then it follows that all of the causal linear time invariant operators are of the form

$$k(t,s) = \begin{cases} g(t-s) & s \leq t \\ 0 & \text{otherwise} \end{cases}.$$

Summary 2.6 Not that $\cup_\alpha B_\alpha$ can be a linear subspace and in the Problem 2.2.1.5 we did not rule out the possibility of $\cup_\alpha B_\alpha$ to be a linear subspace. For instance, consider the space of all functions defined on the real line X, and let X_T denote the subspace of functions that vanish to the left of $T \in \mathbb{R}$, i.e. $x(t) = 0$ for all $t \leq T$. In this case $\cup_T X_T$ is the space of all functions that vanishes to the left of some finite time. This is a subspace of X.

Summary 2.7 It is well know that the set of all linear transformations from a linear space to another linear space is itself a linear space. However, there is even more! Let X be an arbitrary nonempty set and let Y be a linear space. Then $\mathcal{F}$, the set of all mappings of X into Y, is a linear space (with the addition of mappings and scalar multiplies of mappings are defined as in Example 2 Section 2.

Summary 2.8 A linear transformation is one-to-one (injective) if its kernel only contains the origin.

Proof. Let L be injective, and let $a \in \ker L$. Then $La = 0$, and from linearity of L we have $L0 = 0$. From injectivity of L we must have $a = 0$.

Let $\ker L = \{0\}$. Let $x, y \in V$ such that $f(x) = f(y)$. Then from linearity of f we can write $f(x-y) = 0$. Since the kernel only contains the origin we will have $x - y = 0$, so $x = y$. ☐

Summary 2.9 The inverse of a linear transformation, if exists, is linear.

Proof. Let $L : V \to V$ be a linear transformation that its inverse exists denoted by L^{-1}. Let $y_1, y_2 \in V$. Then exists $x_1, x_2 \in V$ such that $y_1 = Lx_1$ and $y_2 = Lx_2$. Consider

$$L^{-1}(\alpha y_1 + y_2) = L^{-1}(\alpha Lx_1 + Lx_2) = L^{-1}(L(\alpha x_1 + x_2)) = \alpha x_1 + x_2 = \alpha L^{-1}(y_1) + L^{-1}(y_2).$$

So L^{-1} is linear when exists. □

Summary 2.10 Let $X = l_2(-\infty, \infty)$ and Z be the linear space made up of all complex-valued functions $f(z)$ defined on the unit circle of the complex plane such that

$$\frac{1}{2\pi i} \int_C \left| f(z) \right|^2 dz/z < \infty.$$

The usual assumption is made that if $f_1(z)$ and $f_2(z)$ differ only on a set of measure zero, then f_1 and f_2 are considered to be the same function. These two spaces are isomorphic as linear spaces with

$$(\cdots, \xi_{-1}, \xi_0, \xi_1, \cdots) \mapsto f(z)$$

with

$$f(z) = \sum_{n=-\infty}^{\infty} \xi_n z^{-n}.$$

This is the two-sided z transform.

■ **Remark 2.1** In the example above, using appropriate change of variable, and utilizing the fact that functions in Z are defined on the unit circle we can write

$$\frac{1}{2\pi i} \int_C \left| f(z) \right|^2 dz/z = \frac{1}{2\pi i} \int_0^{2\pi} \left| f(e^{i\theta}) \right|^2 d\theta.$$

In the second form, the connection to the Fourier transform is more clear.

■ **Remark 2.2** In the example above, the subspace $L \subset X$ where $x \in L$ if $x_i = 0$ for all $i < 0$ maps to the Hardy space H^2 via the constructed isomorphism. Hardy space is the space of all that are restriction of harmonic functions to the unit sphere.

Summary 2.11 — Some definitions. There are some important nuances in some definitions in linear algebra. We will highlight those definitions here.

Definition 2.1 — Linear Combination. Let V be a vector space, and $A \subset V$. Then $x \in V$ is said to be a linear combination of elements in A if there is a **finite** sum such that

$$x = \sum_{i=1}^{n} \alpha_i v_i, \quad v_i \in A \text{ for all } i.$$

Accordingly, Span$\{A\}$ is defined to be the set of all **finite** linear combinations of A.

■ **Remark 2.3** Regardless of the nature of the linear space (finite or infinite dimensional), a linear combination is always a finite sum.

Proposition 2.1 For any collection of vectors $A \subset V$, Span$\{A\}$ is **the smallest subspace** of V that **contains** A.

Proposition 2.2 — Linear Independence. Let $A \subset V$ be a collection of vectors. Then the followings are equivalent.

(a) A is a linearly independent set of vectors.

(b) For every $x \in A$ we have $x \notin \text{Span}\{A \backslash \{x\}\}$. That is x is not a linear combination of the points in $A \backslash \{x\}$.

(c) For every **finite** sub-collection $\{v_1, \cdots, v_n\} \subset A$ we have

$$\sum_{i=1}^{n} \alpha_i v_i = 0 \quad \Longrightarrow \quad v_i = 0 \quad \forall i = 1, \cdots, n.$$

(d) For $v \neq 0 \in \text{Span}\{A\}$ there exists one and only one finite subset of A, say $\{v_i\}_{i=1}^n$ and a unique n-tuple of non-zero elements such that

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

(e) A a linearly independent set of vectors if and only if there exists no proper subset $A_0 \subset A$ such that $\text{Span}\{A_0\} = \text{Span}\{A\}$. I.e. for all $B \in 2^A$ one has $\text{Span}\{B\} \subsetneq \text{Span}\{A\}$.

■ **Remark 2.4** Note that in item (d) above we have excluded the origin, and also forced the n-tuple to have non-zero elements. Because $0 = 0v_1 + 0v_2 = 0v_3 + 0v_4$, etc assuming $v_1, v_2, v_3, v_4 \in A$. Furthermore, let $v \in A$, that trivially $v \in \text{Span}\{A\}$. Then one can write $v = 1v$, $v = 1v + 0v_1$, etc. I.e. one can add more vectors from A with zero coefficients.

Summary ✎ 2.12 In Example 5, page 179, we discuss a very important example about the subspace $A_x \subseteq l_2(-\infty, \infty)$ and its connection to Z, the set of all complex valued functions defined on $\mathbb{T}$ such that

$$\int_C \left|f(z)\right|^2 dz/z = \int_0^{2\pi} \left|f(e^{i\theta})\right|^2 d\theta < \infty.$$

Z and $l_2(-\infty, \infty)$ are isomorphic via ϕ given as:

$$(\cdots, \xi_{-1}, \xi_0, \xi_1, \cdots) \mapsto \sum_{n=-\infty}^{+\infty} \xi_n z^{-n}.$$

It turns out

$$\phi[S_r^n x](t) = t^{-n} \phi[x](t),$$

i.e. shifting the sequence x to the right by n-positions, is the same as multiplying $[\phi x](t)$ be the monomial t^{-n}. So $A_x = \{\sum_{n=M}^{N} \alpha_n S_r^n x\}$ is the same as $L \subset Z$ given by

$$L = \{g(t) \cdot [\phi x](t) : g(t) \text{ is a Laurent polynomial}\}.$$

A Laurent polynomial is a polynomial in variables z and $1/z$.

■ **Remark 2.5** The importance of the monomial like $f(z) = z^n$ becomes more evident when one restricts f to the unit-circle in the complex domain, i.e. $\mathbb{T} = \{|z| = 1\}$. Because

$$f\Big|_{\mathbb{T}}(z) = f(e^{i\theta}), \quad \text{where } z = e^{i\theta}.$$

Thus restricting $f(z) = z$ to $\mathbb{T}$ we will get

$$f(e^{i\theta}) = e^{in\theta} = cos(n\theta) + i\sin(n\theta).$$

Summary 2.13 — More intuition about the basis. We know that if A is linearly independent, then every non-zero vector in $\mathrm{Span}\{A\}$ has a unique linear combination of the vectors in A. With this characterization of linear independence, and appropriate definition of spanning, we can have the following proposition.

Proposition 2.3 $A \subset X$ is a basis if and only if every non-zero vector has a unique linear combination of vectors in A.

Summary 2.14 Let $X = \ell(\mathbb{R})$ be the space of real sequences. Then one might tend to think that the set $A = \{e_1, e_2, \cdots\}$ where $e_i[j] = \delta_{ij}$, i.e. the i^{th} position of e_i is non-zero (equal to one). But this set is not a basis. Because one can not write the sequence with infinite support as a finite linear combination of elements in this collection. In fact $\mathrm{Span}\{A\}$ is a proper subspace of X consisting of all sequences with finite support.

Summary 2.15 Some useful facts

(a) Cardinality of the Hamel basis of a vector space is the dimension of the space.

(b) Two Hamel bases for a fixed space have the same Cardinal numbers.

(c) Two vector spaces are isomorphic if and only if the have the same dimension.

Proof. (a) It is a definition!

(b) See Appendix B in the text book

(c) Let B_1 be a basis for V_1 that is isomorphic to V_2 via ϕ. Then $\phi(B_1)$ is a basis for V_2 and since ϕ is a bijection, then B_1 and $\phi(B_1)$ both has the same cardinality. For the converse, assume B_1 and B_2 has the same cardinality so there is a bijection between these two sets. Let ϕ be a linear map that maps the bases to bases according to the bijection. This is a linear map and is a bijection by construction. So V_1 and V_2 are isomorphic.

□

Summary 2.16 — Facts about subspaces. Consider the following two propositions.

Proposition 2.4 Let V be a linear space and $A \subset V$ be a set of vectors. Then $a \implies b$ and b, c are equivalent.

(a) $\text{Span}\{A\}$ is the set of all linear combinations of vectors in A.

(b) $\text{Span}\{A\}$ is the smallest subspace containing A.

(c) $\text{Span}\{A\}$ is the intersection of all subspaces containing A.

Proof. The equivalence between (b), and (c) is merely the definition of intersection. Now we want to establish $a \implies b$. The fact that $\text{Span}\{A\}$ contain A and $\text{Span}\{A\}$ are easy. We will show that $\text{Span}\{A\}$ is the smallest such space. So we prove that $\forall W \subset V$ a linear subspace, if $A \subset W$ then $\text{Span}\{A\} \subset W$. Let W be any subspace that contains A, and let $x \in \text{Span}\{A\}$. Then x is a linear combination of the points in A, and since W also contains A and is a subspace, then it also contains the linear combination above, thus it contains x. So $\text{Span}\{A\} \subset W$. □

One important side effect of the proposition above is the following corollary:

Corollary 2.1 A is a linear subspace if and only if $A = \text{Span}\{A\}$.

Proof. $\text{Span}\{A\}$ is the smallest subspace that contains A. Since A is subspace itself, the it follows then the smallest subspace containing A is itself. Thus $\text{Span}\{A\} = A$. The converse is also immediate using the fact that $\text{Span}\{A\}$ is a linear space. □

There is an interesting similarity between the notions above and the notion of the set closure in the topological setting. Closure of a set A in a topological space is the smallest closed set that contains A. Equivalently, it is the intersection of all closed sets that contains A. Similar to the fact that the linear combination of any points (finite) is the span, the topological equivalence is that if there is a sequence is a closed set, then the limit of the sequence (if exists) is also in the set.

Summary 2.17 Let L be a linear transformation of X into Y where X and Y are both finite dimensional.

(a) L maps X onto Y if and only if $\dim \mathcal{R}(L) = \dim Y$.

(b) L is one-to-one if any only if $\dim \mathcal{R}(L) = \dim X$.

(c) L is invertible if and only if $\dim X = \dim Y = \dim \mathcal{R}(L)$.

Summary 2.18 — Some scattered connections between causality and adaptive processes. I have been thinking about the causal maps. A linear causal map $\phi : V \to W$ is a map between two vector spaces (function spaces in this case) such that if $x(t) = y(t)$ for $t \leq T$, then $[\phi x](t) = [\phi y](t)$ for $t \leq T$, which this holds for each T. One can look at this from a slightly different angle! For each T define the equivalence relation $x =_T y$ if x and y are the same "upto" time T. So for each T one gets a partition of the space, say P_T. With this point of view, the definition of causality simply forces the image of each equivalence class of P_T to be fully contained in an equivalence class $\hat{P}_T$ of W. This is equivalent to the statement that the pre-image of each equivalence class in $\hat{P}_T$ is a union of equivalence classes in P_T. Which is the same as saying the pre-image of measurable sets (belonging to the sigma algebra generated by $\hat{P}_T$) is measurable (that belongs to the sigma algebra generated by P_T). So a map is causal is the same thing as saying it is $\mathcal{F}_T$ adaptive process.

2.1.1 Direct Sums and Sums

Direct sums, sums, inner sums, outer sums, etc, are in the Category of notions that because they are very important it has been discussed by different people and different fields, and this has lead to many wrong interpretations about these concepts. So it is very easy for someone to get lost in the subtle nuances between these notions. In this section I will keep everything simple.

Let X, Y be any two vector spaces (maybe subspaces of a same containing vector space). Then $X \oplus Y$ is a vector space whose underlying set is $X \times Y$. It is also possible to view this from the universal property point of view, which we will not do here. See Roman for more on this.

■ **Example 2.1** Let V be a vector space and let $B = \{e_1, e_2, e_3\}$ be a basis for this space. Let $X_1 = \text{Span}\{e_1\}$ and $X_2 = \text{Span}\{e_1, e_2\}$ be two subspaces. Then $X_1 \oplus X_2$ is the set of all vectors of the form $(\alpha_1 e_1, \alpha_2 e_1 + \alpha_3 e_3)$. One can write

$$(\alpha_1 e_1, \alpha_2 e_1 + \alpha_3 e_3) = \alpha_1 (e_1, 0) + \alpha_2 (0, e_2) + \alpha_3 (0, e_3).$$

■

In contrast, $X_1 + X_2$ is the smallest subspace that contain X_1 and X_2 and is only defined when X_1 and X_2 are subspaces of a containing linear space. Equivalently, $X_1 + X_2$ is the set of all vectors that are linear combination of elements in X_1 or X_2.

The following proposition is very important.

Proposition 2.5 Let X_1, X_2 be subspaces of X. Then the followings are equivalent:

(a) $X_1 \cap X_2 = \{0\}$.

(b) Every $x \in X_1 + X_2$ has a unique representation $x = x_1 + x_2$ with $x_1 \in X_1$ and $x_2 \in X_2$.

(c) If $\alpha_1 x_1 + \alpha_2 x_2 = 0$ then $\alpha_1 = \alpha_2 = 0$. I.e. 0 has a unique representation.

When X_1 and X_2 are both subspaces of a containing vector space, then there is a natural mapping between $X_1 \oplus X_2$ and $X_1 + X_2$ given as

$$\Phi(x_1, x_2) = x_1 + x_2.$$

Proposition 2.6 Φ is isomorphism of vector spaces if $X_1 \cap X_2 = \varnothing$.

Proof. Every (x_1, x_2) determines a unique $x_1 + x_2 \in X_1 + X_2$. For the converse, from the proposition above if $X_1 \cap X_2 = \varnothing$, then every point in $X_1 + X_2$ is uniquely determined via (x_1, x_2). So Φ is an isomorphism. □

Proposition 2.7 In general ones has

$$\dim(X \oplus Y) = \dim(X) + \dim(Y), \qquad \dim(X + Y) = \dim(X) + \dim(Y) - \dim(X \cap Y).$$

2.1.2 Linear Functionals

The 0 level set of every non-zero linear functional is a hyperplane. So the null-set of a linear functional is maximal in the sense that the only subspace that contains it is the whole space. The following theorem makes this precise.

Proposition 2.8 Let $\ell : V \to \mathbb{R}$ be a linear functional. Then $\dim \mathcal{N}(\ell) = n-1$. Furthermore, for any subspace A that contains $\mathcal{N}(\ell)$ but is not equal to it, we have $A = X$.

Proof. Since ℓ is a non-zero functional, $\exists x_0 \in X$ such that $\ell(x_0) \neq 0$. Let $x \in X$. One can write

$$x = x \pm \frac{\ell(x)}{\ell(x_0)} x_0 = (x - \frac{\ell(x)}{\ell(x_0)} x_0) + \frac{\ell(x)}{\ell(x_0)} x_0.$$

Observe that the first term belongs to $\mathcal{N}(\ell)$ (apply ℓ and utilize linearity). Call $M = \mathrm{Span}\{\{\}\}x_0$ and since any arbitrary $x \in X$ can be written as a linear combination of a point in $\mathcal{N}(\ell)$ and a point in M it follows that

$$X = \mathcal{N}(\ell) + M.$$

Using the fact that $\mathcal{N}(\ell) \cap M = \{0\}$ it follows that the sum above is direct and in fact $X = \mathcal{N}(\ell) \oplus M$. So M is the algebraic complement of $\mathcal{N}(\ell)$ and thus $\mathcal{N}(\ell)$ has co-dimension 1. Thus it is a hyperplane.

For the second part of the proposition, let A be a subspace that contains $\mathcal{N}(\ell)$. Since it is not equal to $\mathcal{N}(\ell)$, then we can choose $x_0 \in A \backslash \mathcal{N}(\ell)$, and the rest of proof follows. □

2.2 Problems

2.2.1 Linear Spaces and Linear Subspaces

■ **Problem 2.2.1.1** Let X be the linear space $\mathbb{R}^4$, For what values of r, if any, is the set

$$A_r = \{x \in \mathbb{R}^4 : x_1 + x_2 + x_3 + x_4 = r\},$$

where r is a real number, a linear subspace of X? For what values of r, if any, is the set

$$B_r = \{x \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 = r^2\}$$

a linear space of X?

Solution Let $x, y \in A_r$. Then we need to have

$$\alpha x + y \in A_r \qquad \forall \alpha \in F$$

Thus we need to have $r\alpha + r = r$ for all α in the underlying field. Thus we need to have

$$r = 0.$$

Let $x \in B_r$. Since we need to have $\alpha x \in B_r$ for all $\alpha \in F$, it follows that $\alpha r^2 = r^2$, thus $r = 0$. So B_0, i.e. the origin, is the only case that B_r can be a linear subspace of $\mathbb{R}^4$.

■ **Problem 2.2.1.2** Let X be the linear space made up of all complex-valued functions $T(s)$ defined on the imaginary axis of the complex plane such that

$$\Big| \int_{-i\infty}^{i\infty} \big|T(s)\big| \, ds \Big| < \infty,$$

where the integral is along the imaginary axis. That is, $X = L_2(-i\infty, i\infty)$. Let A be the set made up of all rational functions, that is all functions of the form

$$T(s) = \frac{a_0 s^m + \cdots + a_m}{b_0 s^n + \cdots + b_n},$$

where $a_0 \neq 0$ and $b_0 \neq 0$, and m, n are integers. Is the set A a linear subspace of X? Next consider the subset B of A made up of all rational functions with $n > m$. Is B a linear subspace of X? What about the subset C of B made up of all functions with all their finite poles in the left hand plane?

Solution The first part of the question is quite vague. Because $T(s) = 1/s$ belongs to A but not to X as

$$\int_{-\infty}^{\infty} \frac{1}{|x|^2} dx = \infty$$

So A is not a subset of X. However, if we want to consider $A \cap X$, then it is a linear subspace of X. That is because $T, L \in A$, then $\alpha T + \beta L$ still has the rational function form, and since $T, L \in X$, and X is a linear space, then $\alpha T + \beta L$ still satisfies the integral inequality, hence belongs to $A \cap X$.

STILL THINKING ON THE OTHER PARTS OF THE QUESTION.

■ **Problem 2.2.1.3** Show that the set of all $n \times m$ matrices can be viewed as a linear space.

Solution We exhibit an homomorphism between $M_{n \times m}$ and $\mathbb{R}^{mn}$. Let $A \in M_{n \times m}$ and $v \in R^{mn}$. Let $\phi : M_{n \times m} \to \mathbb{R}^{mn}$ be defined as $\phi(A) = v$, where

$$v_{ni+j} = a_{ij}.$$

It is straightforward to show that ϕ is bijection. Using this function we can transfer all of the properties of $\mathbb{R}^{mn}$ to $M_{m \times n}$. So $M_{m \times n}$ is a linear space, and ϕ is in fact an homomorphism of vector spaces. So $M_{m \times n}$ has dimension mn.

■ **Problem 2.2.1.4** Let X be the linear space $C[0, T]$. Which, if any, of the following subsets of X are linear subspaces?

(a) $B_1 = \{x \in C[0, T] : x(0) = x(T)\}$,

(b) $B_2 = \{x \in C[0, T] : x(0) = x(T) = 0\}$,

(c) $B_3 = \{x \in C[0, T] : x(t_1) = x(t_2)$ for all t_1, t_2 such that $t_1 + t_2 = T\}$,

(d) $B_4 = \{x \in C[0, T] : x(0) = 1\}$,

(e) $B_5 = \{x \in C[0, T] : \int_0^T x(\tau) d\tau = 1\}$,

(f) $B_6 = \{x \in C[0, T] : \left|x(t_1) - x(t_2)\right| \leq 10 \left|t_1 - t_2\right|$ for all $t_1, t_2 \in [0, T]\}$.

Solution

Yes. $x, y \in B_1$, then $x(0) = x(T)$ and $y(0) = y(T)$. Then it follows that $(\alpha x + y)(0) = \alpha x(0) + y(0) = \alpha x(1) + y(1) = (\alpha x + y)(1)$. So $\alpha x + y \in B_1$.

Yes. Special case of above.

Yes. Let $x, y \in B_3$. Then $x(t) = x(T - t)$ and $y(t) = y(T - t)$. Then one can write

$$(\alpha x + y)(t) = \alpha x(t) + y(t) = \alpha x(T - t) + y(T - t) = (\alpha x + y)(T - t).$$

So $\alpha x + y \in B_3$.

No. $x \equiv 1$ is in B_4, but $2x$ is not.

No. Let $x \in B_5$. Then $\int_0^T (2x)(\tau)d\tau = 2 \neq 1$, so $2x \notin B_5$.
No. $x(t) = 10t$ is in B_6 but $2x$ is not.

■ **Problem 2.2.1.5** Show that if $\{B_\alpha\}$ is a family of linear subspaces of a linear space X, then $B = \cap_\alpha B_\alpha$ is a linear subspace of X. What about $\cup_\alpha B_\alpha$?

Solution Let $x, y \in B$. Then $\forall \alpha$ we have $x, y \in B_\alpha$. Since B_α is a linear subspace, it follows that $ax + by \in B_\alpha$, which is true for all α. Thus $ax + by \in B$. The same is not true in the case of $\cup_\alpha B_\alpha$. For instance in the case of $\mathbb{R}^2$, consider the subspaces $B_1 = \{(\alpha, 0) : \alpha \in \mathbb{R}\}$ and $B_2\{(0, \beta) : \beta \in \mathbb{R}\}$. Then $(1, 0) \in B_1$ and $(0, 1) \in B_2$ but $(1, 1) \notin B_1 \cup B_2$ (in fact $(1, 1) \in B_1 + B_2$, where $B_1 + B_2$ is the smallest subspace that contains B_1 and B_2).

■ **Remark 2.6** Not that $\cup_\alpha B_\alpha$ can be a linear subspace and in the problem above we did not rule out the possibility of $\cup_\alpha B_\alpha$ to be a linear subspace. For instance, consider the space of all functions defined on the real line X, and let X_T denote the subspace of functions that vanish to the left of $T \in \mathbb{R}$, i.e. $x(t) = 0$ for all $t \leq T$. In this case $\cup_T X_T$ is the space of all functions that vanishes to the left of some finite time. This is a subspace of X.

■ **Problem 2.2.1.6** Let X be the linear space made up of all real-valued sequences. Show that A_1, the set of all sequences that have a finite number of nonzero entries only, is a linear subspace of X. Show that A_2, the set of all sequences that have an infinite number of nonzero entries, is not a linear subspace of X.

Solution Let $a, b \in A_1$ with their number of non-zero elements as n_1, n_2 respectively. Then $\alpha a + \beta b$ has at most $n_1 + n_2$ number of non-zero elements, hence in A_1. A_2 is not a linear subspace of X because it does not contain the origin.

■ **Problem 2.2.1.7** Often in systems theory the linear space $L_2(-\infty, \infty)$ is a good mathematical model for the set X of all inputs to a system as well as the set Y contains the range. Let $\mathcal{A}$ be the set of all mappings (linear and nonlinear) of X into Y. Show that $\mathcal{A}$ can be viewed as a linear space. Show that the subset $\mathcal{L} \subset \mathcal{A}$ of all mappings representing causal (Section 2.8) systems is a linear subspace of $\mathcal{A}$.

Solution Let $\phi, \psi \in \mathcal{A}$, and define the addition and scalar multiplication in this space as

$$(\alpha\phi + \beta\psi)(f) = \alpha\phi(f) + \beta\psi(f).$$

It is straightforward to check that with these definitions, $\mathcal{A}$ is a vector space.

Let $F, G \in \mathcal{L}$ be causal maps. Then for all $T \in \mathbb{R}$, $x(t) = y(t)$ for $t < T$ implies $[Fx](t) = [Fy](t)$, and $[Gx](t) = [Gy](t)$ for all $t < T$. Then for $t \leq T$ one can write

$$[\alpha F + G](x)(t) = \alpha[Fx](t) + [Gx](t) = \alpha[Fy](t) + [Gy](t) = [\alpha F + G](y)(t).$$

Thus $\mathcal{L}$ is a linear subspace.

■ **Remark 2.7 — Review of the causal maps.** $\Phi : L_2(-\infty, \infty) \to L_2(-\infty, \infty)$ is causal, if f1or all $T \in \mathbb{R}$, $x(t) = y(t)$ for all $t \leq T$, implies $(\Phi(x))(t) = (\Phi(t))(t)$ for all $t \leq T$.

■ **Problem 2.2.1.8** Let X be the linear space made up of all absolutely convergent sequences of real numbers. Show that B, the set of all absolutely convergent sequences of real numbers with limit zero, is a linear subspace of X.

Solution I can not understand what does absolutely convergent **sequence** means (in contrast to the absolutely convergent series). Also, not sure about the meaning of the abs. conv. sequences with limit zero.

■ **Problem 2.2.1.9** Let X be the set of all convergent sequences of real numbers. Is X a linear space?

Solution Yes. Viewing $\mathbb{R}$ is a linear space, and using the continuity of the addition and scalar multiplication, if $x_n \to x$ and $y_n \to y$, then $\alpha x_n + \beta y_n \to \alpha x + \beta y$.

■ **Problem 2.2.1.10** Let X denote the collection of all real-valued Lipschitz-continuous functions $x(t)$ defined for $-\infty < t < \infty$. That is $x(t)$ satisfies $\left|x(t) - x(s)\right| \leq k\,|t-s|$ for some constant k which depends on x and for all t, s. Show that X is a real linear space.

Solution Let $x, y \in X$. Then we claim that $\alpha x + \beta y$ is also in X with Lipschitz constant $k \leq \alpha K_x + \beta K_y$. Because

$$\left|(\alpha x(t) + \beta y(t)) - (\alpha x(s) - \beta y(s))\right| \leq k_x\,|t-s| + k_2\,|t-s|\,.$$

2.2.2 Linear Transformation

■ **Problem 2.2.2.1** Let X and Y be linear spaces over the same scalar field. Show that $I : X \to X$ the identity transformation, and $0 : X \to Y$ the zero transformation are linear.

Solution Let $x, y \in X$. The $I(\alpha x + y) = \alpha x + y = \alpha I(x) + I(y)$, so I is linear. The zero transformation is trivially linear.

■ **Problem 2.2.2.2** Let X, Y, and Z be linear spaces over the same scalar field, and let $L_1 : X \to Y$ and $L_2 : Y \to Z$ be linear. Show that the composition $L_2 L_1 : X \to Z$ is linear.

Solution Let $x, y \in X$. Then

$$L_2(L_1(\alpha x + y)) = L_2(\alpha(L_1 x) + L_1 y) = \alpha L_2(L_1(x)) + L_2(L_1(y)) = \alpha(L_2 L_1)x + L_2 L_1 y.$$

So $L_2 L_1$ is linear.

■ **Problem 2.2.2.3** Suppose that we consider a system whose output is a delayed version of the input. That is, if $x(t)$ is the input, then the output $y(t) = x(t - \tau)$, where τ is a constant. Let X be the linear space $C(-\infty, \infty)$ of continuous real-valued functions defined on $(-\infty, \infty)$. Let D denote the system operation. Is D a linear transformation of X into itself? suppose that instead of being constant the delay τ is given by $\tau = e^{-t}$. Do we have a linear transformation? Then suppose that $\tau = \exp\left[-\int_{-\infty}^{t}\left|x(\xi)d\xi\right|\right]$, where of course the linear space X must be selected so that the integral exists. Do we have a linear transformation?

Solution In all of the cases D is a linear operator. Assume the delay is given as a function $\tau(t)$. Let $x, y \in C(-\infty, \infty)$. Then one can write

$$[D(\alpha x + y)](t) = (\alpha x + y)(t - \tau(t)) = \alpha x(t - \tau(t)) + y(t - \tau(t)) = \alpha[Dx](t) + [Dy](t).$$

■ **Problem 2.2.2.4** Let $Y = C([0m\infty], \mathbb{R}^n)$ be the linear space made up of all continuous mappings of $[0, \infty)$ into $\mathbb{R}^n$, that is, each component is continuous. Let $X = C^1([0, \infty], \mathbb{R}^n)$ be the linear subspace of Y made up of all elements of Y with continuous derivatives, that is, each component has a continuous derivative. Does the expression $y = Tx$, where

$$Tx = \frac{dx}{dt} - Ax$$

and A is a real $n \times n$ matrix, represent of linear transformation of X into Y?

Solution Yes. Let $x, y \in X$. Then

$$T(\alpha x + y) = \alpha\frac{dx}{dt} + \frac{dy}{dt} - \alpha Ax - Ay = \alpha Tx + Ty,$$

where we have used the linearity of the differentiation operator and the linearity of multiplication by a matrix.

■ **Problem 2.2.2.5** Let $Y = BC(-\infty, \infty)$ denote the space of all bounded real-valued continuous functions $y(t)$ defined for $-\infty < t < \infty$ and let X denote the space of all Lipschitz continuous functions. Define $x = Ly$ be $x(t) = \int_0^t y(s)ds$. Show that L is a linear mapping of Y into X.

Solution Follows from the properties of integration: for $y_1, y_2 \in Y$

$$\int_0^t (\alpha y_1 + y_2)ds = \alpha \int_0^t y_1(s)ds + \int_0^t y_2(s)ds.$$

■ **Remark 2.8** Note that in the example above, integration improved the regularity of the functions. The input of the operator was $BC(-\infty, \infty)$ and its output is Lipschitz continuous functions.

2.2.3 Inverse Transformation

■ **Problem 2.2.3.1** Show that the linear transformation $y = Lx$ on $L_2(-\infty, \infty)$ given by

$$y(t) = \int_{-\infty}^t a^{-1}e^{a(t-\tau)}x(\tau)d\tau$$

is one-to-one. *Hint: Show that* $Lx = 0$ *reduces to* $\int_{-\infty}^t e^{a\tau}x(\tau)d\tau = 0$. *Then differential and use theorem D.13.3.*

Solution The transformation L is linear. So L is injective iff $\ker L = 0$. Let $x \in \ker L$. Then one can write

$$\int_{-\infty}^t a^{-1}e^{-a(t-\tau)}x(\tau)d\tau = a^{-1}e^{-at}\int_{-\infty}^t e^{a\tau}x(\tau)d\tau = 0 \quad \forall t.$$

This implies

$$\int_{-\infty}^t e^{at}x(\tau)d\tau = 0 \quad \forall t.$$

Differentiating with respect to t one gets

$$e^{a\tau}x(\tau) \equiv 0.$$

This implies $x(\tau) \equiv 0$.

■ **Problem 2.2.3.2** Let $k(t, s)$ be continuous for $0 \leq s \leq t \leq T$ and consider

$$y(t) = x(t) + \int_0^t k(t, s)y(s)ds.$$

The following steps will lead to a proof that the relationship $y = Fx$ implicitly given above, does define a linear mapping F on $C[0, \alpha]$ provided α is a sufficiently small positive number.

(a) Define G by $x = Gy$, where

$$x(t) = y(t) - \int_0^t k(t, s)y(s)ds.$$

Show that G is linear.

(b) Assume that for some $\alpha > 0$, G maps $C[0, \alpha]$ onto itself. Show that if G is one-to-one, then G^{-1} exists and is F, so F is linear.

(c) Let M satisfy $\left|k(t, s)\right| \leq M$ for $0 \leq s \leq t \leq T$, then show that $Gy = 0$ reduces to

$$\left|y(t)\right| \leq M \int_0^t \left|y(s)\right| ds.$$

Now use the Gronwall inequality, see Cesari [1,p.35], to show that G is one-to-one.

Solution (a) Both terms in the RHS of

$$x(t) = y(t) - \int_0^t k(t, s)y(s)ds$$

with respect to the position of y (you know what I mean!!). So the linearity of G follows immediately.

(b) By assumption G is onto and one-to-one. So G^{-1} exists. It remains to show that G^{-1} is F. So we need to show $GF = FG = I$. STILL THINKING ON THIS.

(c) STILL THINKING ON THIS.

2.2.4 Isomorphisms

■ **Problem 2.2.4.1** Let χ denote the set of all linear spaces over a scalar field F. Show that the relation $X \sim Y \Leftrightarrow (X, Y$ are isomorphic) is an equivalent relation and, therefore, induces a partition on χ.

Solution Every linear space is isomorphic with itself with the identity operator. If U is isomorphic to Y via $\phi : U \to V$, then V is also isomorphic to U via $\phi^{-1} : V \to U$. Lastly, let $\phi : U \to V$ and $\psi : V \to W$ be isomorphism between spaces. Then U and W are also isomorphic via the isomorphism $\phi \circ \psi$.

■ **Problem 2.2.4.2** Referring to the Example 2 in Page 174 of the text book, under what conditions on real numbers $c_{11}, c_{12}, c_{21}, c_{22}$ does

$$T_2(x) = (c_{11}x_1 + c_{12}x_2) + (c_{21}x_1 + c_{22}x_2)t,$$

where $x = (x_1, x_2)$, define an isomorphism from $\mathbb{R}^2$ to Z?

Solution Let $v = (v_1, v_2)$ denote the function $v_1 + v_2 t \in Z$. Then we want the following matrix to be invertible

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

So the matrix should have a non-zero determinant. I.e.

$$c_{11}c_{22} - c_{12}c_{21} \neq 0.$$

■ **Problem 2.2.4.3** Show that the real linear space X made up of all functions of the form $x = a\cos(\omega t + \phi)$, where ω is fixed, is isomorphic to the complex numbers considered as a real linear space. This fact is the cornerstone of the so-called phasor method of analyzing alternating current electrical networks.

Solution Given $a\cos(\omega t + \phi)$ observe that

$$a\cos(\omega t + \phi) = \mathrm{Re}(a\cos(\omega t + \phi) + ia\sin(\omega t + \phi)) = \mathrm{Re}(ae^{i\phi}e^{i\omega t}).$$

So let the map between these to spaces be

$$a\cos(\omega t + \phi) \mapsto z = ae^{i\phi}.$$

Showing the linearity of the mapping above is very dense with different trigonometric identities. So we exhibit the mapping that is the other way around. Consider the following map

$$z \mapsto \mathrm{Re}(ze^{i\omega t}).$$

This map is linear and onto. Linearity follows from linearity of the $\mathrm{Re}(\cdot)$ map and it is onto because for any $a\cos(\omega t + \phi)$ we have a preimage given by $z = ae^{i\phi}$. It is also bijective because it is linear and its kernel contains only the origin. So The mapping above is isomorphism between vector spaces.

■ **Problem 2.2.4.4** Let $L_2^\sigma(-\infty, \infty)$ denote the linear space made up do all complex-valued functions x defined on $(-\infty, \infty)$ such that

$$\int_{-\infty}^{\infty} \left|x(t)\right|^2 e^{-2\sigma t} dt < \infty,$$

where σ is a real number. Show that $L_2^\sigma(-\infty, \infty)$ is isomorphic to $L_2^\tau(-\infty, \infty)$ for any σ and τ.

Solution Let $x \in L_2^\sigma$. Consider the mapping ϕ

$$x(t) \mapsto x(t)e^{(\sigma - \tau)t}.$$

This map is linear because

$$\phi(\alpha x + y) = (\alpha x(t) + y(t))e^{(\sigma-\tau)t} = \alpha\phi(x)(t) + \phi(y)(t).$$

The inverse map is ϕ^{-1} given by

$$\phi^{-1}(y)(t) = y(t)e^{-(\sigma-\tau)t}.$$

It is straightforward to check that $\phi\phi^{-1}$, and $\phi^{-1}\phi$ are the identities maps of the corresponding spaces. It only remains to show that for $x \in L_2^\sigma$, ϕx really belongs to L_2^τ. This is true because

$$\int_{-\infty}^{\infty} \left|[\phi x](t)\right|^2 e^{-2\tau t} dt = \int_{-\infty}^{\infty} \left|x(t)e^{(\sigma-\tau)}\right|^2 e^{-2\tau t} dt = \int_{-\infty}^{\infty} \left|x(t)\right|^2 e^{-2\sigma t} dt < \infty.$$

■ **Problem 2.2.4.5** Show that $L_2(-\infty, \infty)$ is isomorphic to $L_2[0, \infty)$.

Solution Consider the following linear functions

$$\phi : L_2(-\infty, +\infty) \to L_2[0, +\infty), \qquad [\phi x](t) = x(\log(t))/\sqrt{t},$$

and

$$\psi : L_2[0, +\infty) \to L_2(-\infty, +\infty), \qquad [\psi y](t) = y(e^t)e^{t/2}.$$

It is easy to show that $\psi\phi$ and $\phi\psi$ are the identity maps of the corresponding spaces, so are inverses of each other. Now we need to show that for $x \in L_2(-\infty, +\infty)$, ϕx belongs to $L_2[0, +\infty)$. Consider the integral

$$\int_0^{\infty} \left|[\phi x](t)\right|^2 dt = \int_0^{\infty} \left|x(\log(t))\right|^2 dt/t.$$

We do the change of variable $\tau = \log(|t|)$.

$$\int_0^{\infty} \left|[\phi x](t)\right|^2 dt = \int_{-\infty}^{\infty} \left|x(\tau)\right|^2 d\tau = \int_{-\infty}^{\infty} \left|x(\tau)\right|^2 d\tau < \infty.$$

A similar approach can be used to show that for $y \in L_2[0, \infty)$ we have $\phi y \in L_2(-\infty, \infty)$.

■ **Problem 2.2.4.6** Show that $l_2(-\infty, \infty)$ is isomorphic with $l_2(0, \infty)$.

Solution Map every negative indexed element of the sequence in $l_2(-\infty, \infty)$ to an odd index, and every positive indexed element of the sequence to an even index. Since the sequence is absolutely summable, so is the new sequence, hence in $l_2(0, \infty)$.

2.2.5 Linear Independence and Dependence

■ **Problem 2.2.5.1** Let T be an isomorphic mapping of a linear space X onto a linear space Y. Show that $A \subset X$ is linearly independent if and only if its image $T(A) \subset Y$ is linearly independent.

Solution Assume $\{T(A)\}$ is linearly independent. We want to prove that A is linearly independent. To see this consider a sub-collection $\{x_i\} \subset A$. Let the following be any linear combination of these vectors that is equal to zero

$$\sum_{i=1}^{n} \alpha_i x_i = 0$$

One can apply T and using its linearity

$$0 = T(\sum_{i=1}^{n} \alpha_i x_i) = \sum_{i=1}^{n} \alpha_i T(x_i).$$

Since $\{T(x_i)\}$ is linearly independent, it follows that $\alpha_i = 0$ for all i. So $\{x_i\}$.

For the converse let A be a linearly independent set of vectors. We want to show $\{T(A)\}$ is also linearly independent. Consider the following linear combination of $\{T(x_i)\}_{i=1}^n$ where

$$\sum_{i=1}^{n} \alpha_i T(x_i) = 0.$$

Applying T^{-1} to both sides and using its linearity we have

$$\sum_{i=1}^{n} \alpha_i x_i = 0.$$

Since $\{x_i\}$ is linearly independent, it follows that $\alpha_i = 0$ for all i. Since this is true for any sub-collection of vector in $\{T(A)\}$ it follows that $\{T(A)\}$ is linearly independent.

■ **Remark 2.9** Note that one $\{T(A)\}$ being linearly independent (L.I.) implies $\{A\}$ being L.I., however, the converse is not true. Because a map might not be injective and can send some of the vectors to the origin, thus the image will not be linearly independent set of vectors.

■ **Problem 2.2.5.2** Let $X = L_2(\Omega, \mathcal{F}, \mathbb{P})$, the linear space made up of all complex-valued random variables x defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$\mathbb{E}\Big[x\overline{x}\Big] = \int_\Omega = \int_\Omega |x|^2 \, d\mathbb{P},$$

where $\overline{x}$ denotes the complex conjugate of x. Let $A \subset X$ be the set containing two random variables x_1 and x_2 such that

$$\mathbb{E}\Big[x_1\overline{x_1}\Big] = 1, \qquad \mathbb{E}\Big[x_2\overline{x_2}\Big] = 1, \qquad \mathbb{E}\Big[x_1\overline{x_2}\Big] = 0.$$

Sow that $y \in X$ is a linear combination of points in A if and only if

$$\mathbb{E}\Big[y\overline{y}\Big] = \Big|\mathbb{E}\Big[y\overline{x}\Big]\Big|^2 + \Big|\mathbb{E}\Big[y\overline{x_2}\Big]\Big|^2.$$

Moreover, show that A is linearly independent.

Solution We are dealing with a linear space that we have a notion of inner product, hence projection, etc. So I will use that notation here. For two random variables x, y define

$$\langle x, y \rangle = \mathbb{E}\Big[x\overline{y}\Big].$$

So with this language the properties in the question statement can be written as

$$\langle x_1, x_1 \rangle = 1, \qquad \langle x_2, x_2 \rangle = 1, \qquad \langle x_1, x_2 \rangle = 0.$$

To show the forward direction let $y = \alpha x_1 + \beta x_2$. Then

$$\mathbb{E}\Big[y\overline{y}\Big] = \langle y, y \rangle = |\alpha|^2 + |\beta|^2 = \Big|\mathbb{E}\Big[y\overline{x_1}\Big]\Big|^2 + \Big|\mathbb{E}\Big[y\overline{x_2}\Big]\Big|^2.$$

For the converse, assume

$$\mathbb{E}\Big[y\overline{y}\Big] = \Big|\mathbb{E}\Big[y\overline{x_1}\Big]\Big|^2 + \Big|\mathbb{E}\Big[y\overline{x_2}\Big]\Big|^2$$

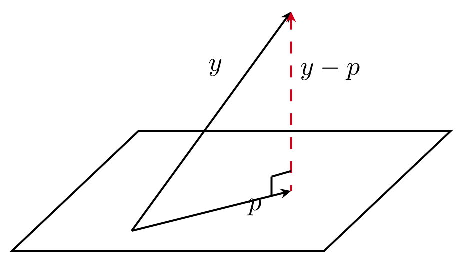
holds. We want to show $y = \alpha x_1 + \beta x_2$ for some α, β. Let $p = \alpha x_1 + \beta x_2$. Then

$$\langle y - p, x_1 \rangle = \langle y, x_1 \rangle - \langle p, x_1 \rangle = |\alpha|^2 - |\alpha|^2 = 0.$$

And for a similar reason one can write

$$\langle y - p, x_2 \rangle = 0.$$

So $y - p \perp \text{Span}\{x_1, x_2\}$. See the following figure



So we can write

$$\|y - p\|^2 = \|y\|^2 + \|p\|^2.$$

Using the fact that $\|y\|^2 = |\alpha|^2 + |\beta|^2$ and $\|p\|^2 = |\alpha|^2 + |\beta|^2$ one gets

$$\|y - p\| = 0.$$

Using the properties of the norm we have

$$y = p.$$

To show that A is linearly independent, consider the following linear combination

$$\alpha x_1 + \beta x_2 = 0.$$

By applying $\langle x_1, \cdot \rangle$ and $\langle x_2, \cdot \rangle$ to both sides one gets

$$\alpha = 0, \qquad \beta = 0.$$

So A is a linearly independent set of vectors.

■ **Problem 2.2.5.3** Let X be the linear space $L_2[0, 2\pi]$ and A the set of all functions $x_n(t) = e^{int}$, $n = 0, 1, 2, \cdots$. Show that A is linearly independent.

Solution In order not to deal with messy sums, we demonstrate the linear independence of the subset $\{1, e^{it}, e^{2it}\}$. The linear independence for arbitrary finite subset of A can be shown in a similar way. Consider

$$\alpha_0 + \alpha_1 e^{it} + \alpha_2 e^{2it} = 0.$$

It follows that $\alpha_0 = 0$. Furthermore, differentiating both sides and dividing by ie^{it} one gets

$$\alpha_1 + 2\alpha_2 e^{2it} = 0.$$

This time we get $\alpha_1 = 0$. Differentiating again and dividing by e^{2it} one gets $\alpha_2 = 0$. So $\{1, e^{it}, e^{2it}\}$ is linearly independent.

■ **Problem 2.2.5.4** Show that a finite set $A = \{x_1, \cdots, x_n\}$ is a linear space X is linearly independent if and only if the only n-tuple of scalars satisfying the equation

$$a_1x_1 + \cdots + a_nx_n = 0$$

is $a_1 = \cdots = a_n = 0$..

Solution Assume A is linearly independent. We want to show $a_1x_1 + \cdots + a_nx_n = 0$ implies $a_1 = \cdots = a_n = 0$. We show this by contrapositive. Let $a_1x_1 + \cdots + a_nx_n = 0$ with not all of a_i equal to zero. Then one can write

$$x_i = \beta_1x_1 + \cdots + \beta_nx_n$$

, where x_i is the vector that its corresponding coefficient α_i was not zero. This implies that the set A is not linearly independent.

For the converse we want to show $a_1x_1 + \cdots + a_nx_n = 0$ implying $a_i = 0$ for all i implies A is linearly independent. We do it by proof by contrapositive. Assume A is not linearly independent. WLOG we can assume x_1 is linear combination of others:

$$x_1 = \sum_{i=2}^{n} \alpha x_i.$$

Rearranging this sum one gets

$$x_1 - \alpha_2x_2 - \cdots - \alpha_nx_n = 0$$

with not all of coefficients equal to zero. This finishes the proof.

■ **Problem 2.2.5.5** Let X be the linear space $\mathbb{R}^n$, and let A be the set containing the n vectors

$$x_i = (x_{1i}, x_{2i}, \cdots, x_{ni}) \quad \text{for } i = 1, \cdots, n.$$

Show that this set is linearly independent if and only if

$$\det \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} \neq 0.$$

Solution Assume the determinant is non zero and we want to show that A is linearly independent. We do this by contrapositive. Assume A is not linearly independent. So we can write one column of the matrix as the linear combination of the other column. Then by elementary column operations (that keeps the determinant intact) one can reduce that column to zero column and hence the determinant will be zero.

Now assume A is linearly independent and we want to show that the determinant is non-zero. Let $c = (c_1, \cdots, c_n)$ be a vector of coefficients. Then columns of A being linearly independent is the same as saying the following system of equations has the only zero answer

$$Ac = 0 \quad \text{only when} c = (0, \cdots, 0).$$

In other words, $\ker A = \{0\}$. This implies A is invertible (because it is automatically onto), and thus $\det(A) \neq 0$.

■ **Problem 2.2.5.6** Let the state of a dynamical system be a point in the linear space $X = \mathbb{R}^n$. Let the state at time $k = 0, 1, 2, \cdots$ be denoted by x_k. Further, suppose that the evolution of the system is characterized by a linear transformation $T : X \to X$. In particular, $x_k = Tx_{k-1}$. Let x_0 be an arbitrary initial state, and consider the set

$$A = \{x_0, Tx_0, T^2x_0, \cdots\}.$$

Show that there exists an integer p such that for

$$A_p = \{x_0, Tx_0, \cdots, T^p x_0\}$$

one has $V(A) = V(A_p)$, that is, A and A_p span exactly the same linear spaces.

Solution Let $A = \{x_1, x_2, \cdots, x_N\}$ be the first linearly independent vectors that appear in A. Note that $N \leq n$, as the space X is n dimensional. Let $p = N$. Then $\text{Span}\{A\} = \text{Span}\{A_p\}$.

■ **Problem 2.2.5.7** Let X be the linear space made up of all real-valued random variables defined on some probability space. Let A be a set of random variables in X. Show that if a random variable $z \in X$ is stochastically independent of each random variable in A, then z is ot in $V(A)$.

Solution STILL THINKING ON THIS.

■ **Problem 2.2.5.8** Let X be a linear space. A set $K \subset X$ is said to be convex if

$$\lambda x + (1 - \lambda)y \in K \qquad (0 \leq \lambda \leq 1),$$

whenever $x, y \in K$. Let K_1 and K_2 be two convex sets in X.

(a) Show that $K_1 \cap K_2$ is convex.

(b) Is $K_1 \cup K_2$ convex?

Solution (a) Let $x, y \in K_1 \cap K_2$. Then since K_1 and K_2 are convex we have

$$\lambda x + (1 - \lambda)y \in K_1 \quad \lambda x + (1 - \lambda)y \in K_2 \quad (0 \leq \lambda \leq 1).$$

Thus $\lambda x + (1 - \lambda)y \in K_1 \cap K_2$ for $\lambda \in [0, 1]$. Thus $K_1 \cap K_2$ is convex.

(b) No. Let $x_1, x_2 \in X$ be two points in the vector space. Each of the sets $\{x_1\}$ and $\{x_2\}$ are convex, but $\{x_1, x_2\}$ is not.

■ **Problem 2.2.5.9** Let $y = f(x)$ be a C^2 function defined for $-\infty < x < \infty$. Find a condition on d^2f/dx^2 in order that

$$K = \{(x, y) \in \mathbb{R}^2 : y \geq f(x)\}$$

be convex.

Solution We use the fact that $f'' > 0$ implies f is convex, i.e.

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2), \qquad \lambda \in [0, 1]$$

Choose $(x_1, y_1), (x_2, y_2) \in K$. So by definition $y_1 \geq f(x_1)$ and $y_2 \geq f(x_2)$. Thus

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda y_1 + (1 - \lambda)y_2, \qquad \lambda \in [0, 1].$$

So $\lambda y_1 + (1 - \lambda)y_2 \in K$ for all λ. Thus K (the epigraph of f) is convex if and only if $f'' > 0$.

2.2.6 Hamel Basis

■ **Problem 2.2.6.1** Prove the Rank-Nullity Theorem. I.e. for a linear map $L : V \to W$ we have

$$\mathcal{R}(L) + \mathcal{N}(L) = \dim V,$$

where

$$\mathcal{R}(L) = \dim \mathrm{Image}(L), \qquad \mathcal{N}(L) = \dim \ker(L)$$

Solution Let $B_{\mathcal{N}}$ be a basis for $\mathcal{N}(L)$. We can extend this basis to a basis to the whole space:

$$B = B_{\mathcal{R}} \cup B_{\mathcal{N}}.$$

So we have

$$\dim V = |B| = |B_{\mathcal{R}}| + |B_{\mathcal{N}}|.$$

We already have $|B_{\mathcal{N}}| = \dim \ker(L) = \mathcal{N}(L)$. Thus it remains to show $B_{\mathcal{R}} = \mathcal{R}(L) = \dim \mathrm{Image}(L)$. To see this we need to show that $L(B_{\mathcal{R}})$ is linearly independent. Consider

$$0 = \sum_{i=1}^{n} \alpha_i L(x_i), \qquad x_i \in B.$$

Applying L and using linearity one gets

$$L(\sum_{i=1}^{n} \alpha_i x_i) = 0.$$

However, $x_i \notin B_{\mathcal{N}}$ for all i. So this implies that $\alpha_i = 0$ for all i. Hence $L(B_{\mathcal{R}})$ is linearly independent and using that fact that it spans $\mathrm{Image}(L)$ we conclude that it is a basis for the image of L, thus $\dim \mathrm{Image}(L) = |B_{\mathcal{R}}|$.

■ **Problem 2.2.6.2** Let X be a finite-dimensional space with $\dim(X) = n$. Show that every set containing $n+1$ points is linearly dependent.

Solution Assume otherwise. If there are $n+1$ vectors linearly independent in the set, then there is a basis that contains those linearly independent vectors (or in other words, this set can be extended to a basis for the whole space). So that dimension of the space should be at least $n+1$, which a contradiction.

■ **Problem 2.2.6.3** Let A be a linear subspace of a linear space X. Show that $\dim(A) \leq \dim(X)$. Moreover, if X is finite dimensional and A is a proper linear subspace of X, show that $\dim(A) < \dim(X)$.

Solution There exists a basis for the subspace A that can further be extended to a basis for the whole space.

1. In the case that X is finite dimensional, then it implies that the cardinality of the basis for X is at least the cardinality of A, thus $\dim A \leq \dim X$.
2. In the case that X is infinite dimensional (countable), then if X is finite dimensional then we automatically have $\dim A \leq \dim X$ and if X is also infinite dimensional (countable), then $\dim A = \dim X$.

3. In the case that X infinite dimensional (uncountable, equal to $\mathfrak{c}$), then if A either finite or countable we will have $\dim A \leq \dim X$ and if A is also infinite dimensional (uncountable, equal to $\mathfrak{c}$) then we will have $\dim A = \dim X$.

For the case that X is finite dimensional and A is a proper subset of X, let B_A be a basis for A and extend it to a basis for X called B_X. Then $B_X \backslash B_A \neq \varnothing$ (otherwise A is not a proper subspace. That is because $\mathrm{Span}\{B_A\}$ is the smallest subspace that contains A and if it contains all the elements that B_X has then it should be at least the whole space $A \supseteq X$. And using $A \subset X$ one reaches to the contradiction). This implies there are vectors in B_X that is not in B_A. So $\dim A \leq \dim X$.

■ **Problem 2.2.6.4** Consider the following differential equation defined on $C^2[0,\infty)$

$$\frac{d^2x}{dt^2} + b\frac{dx}{dt} + cx = 0.$$

If X denotes the set of all solutions of the equation above, show that X is a linear subspace of $C^[0,\infty)$ and that $\dim(X) = 2$.

Solution Let $X \subset C^2[0,\infty)$ be the set of all solutions to the differential equation. Observe that one can write the equation as

$$(\frac{d^2}{dt^2} + b\frac{d}{dt} + c)x = \mathcal{L}x = 0,$$

where $\mathcal{L} : C^2[0,\infty) \to C^2[0,\infty)$ is a linear operator. Then the solution set X is in fact the pre-image of the origin $0 \in \mathbb{R}$, i.e. the kernel of the linear operator. The kernel of any linear operator is a linear subspace. So X is a linear subspace. One can transform this second order equation into a system of first order equations by letting:

$$\phi_0 = x, \qquad \phi_1 = x'$$

Thus the equivalent equation (that its solutions are in one-to-one correspondence with the solutions of $\mathcal{L}x = 0$) will be

$$\Phi' = A\Phi$$

where A is the companion matrix of the polynomial $t^2 + bt + c = 0$. The ODE above has the solution

$$\Phi(t) = e^{At}\Phi(0).$$

Thus every solution is characterized by $\Phi(0) = (\phi_0(0), \phi_1(0)) = (x(0), x'(0))$.

■ **Remark 2.10** One can solve this problem more "manually", the way that Braun did in Theorem 2, Page 132. I.e. to show that $x = \alpha_1 y_1 + \alpha_2 y_2$ is a solution (where $y_1, y_2 \in \ker \mathcal{L}$), and also showing that any solution is of this form (by utilizing uniqueness and existence theorem). But note that one must argue that $\ker \mathcal{L}$ contains at least two elements, which can be done by transforming the second order into two first order ODEs and arguing there exist a solution for each, and utilizing those solutions to construct two solutions, i.e. y_1, y_2

■ **Problem 2.2.6.5** Show that if A is a set in a linear space X with $V(A) = X$, then A contains a Hamel basis of X.

Solution Let $B \subset A$ be a maximal linearly independent subset of A (exist by Zorn's lemma). Then $\mathrm{Span}\{B\} = \mathrm{Span}\{A\}$. Because otherwise, considering the fact $B \subset A$ thus $\mathrm{Span}\{B\} \subset \mathrm{Span}\{A\}$, if there exists $a \in \text{Span}\{A\}$ such that $a \notin \text{Span}\{B\}$, then $B \cup \{a\}$ is a larger linearly independent subset of A, contradicting the maximality of B. Since $\text{Span}\{B\} = \text{Span}\{A\} = X$, it follows that B is a basis for X.

■ **Problem 2.2.6.6** Let X be the real linear space made up of all functions of the form $x(t) = a\cos(\omega t + \phi)$, where ω is fixed. Show that $B = \{\cos\omega t, \sin\omega t\}$ is a basis for X.

Solution We want to show that B is linearly independent, and spans X. For the linear independence consider

$$a\cos\omega t + b\sin\omega t = 0.$$

Since both $\cos\omega t$ and $\sin\omega t$ has countably many zeros, then the equality above is true for all t only when $a, b = 0$. So B is linearly independent. For the spanning, we want to show that every $a\cos(\omega t + \phi)$ is a linear combination of $\sin\omega t$ and $\cos\omega t$ (uniqueness follows from linear independentness show above). We use

$$\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y).$$

So one can write

$$a\cos(\omega t + \phi) = a(\cos\phi\cos(\omega t) - \sin(\phi)\sin(\omega t)).$$

So B is a basis for the space X.

■ **Problem 2.2.6.7** Show that $C[0, T]$ is infinite dimensional. *Hint: Construct a linearly independent set of dimension n, where n is an arbitrary integer.*

Solution The set $\{1, t, t^2, t^3, \cdots\}$ is a linearly independent set of vectors (any attempt to write some t^k as a linear combination (finite) of other vectors results in the conclusion that a polynomial has infinitely many zeros, which is not correct). So the dimension of X is at least $\aleph_0$, thus infinite.

■ **Problem 2.2.6.8** Let L be a linear transformation of X into Y where X and Y are both finite dimensional.

a Show that L maps X onto Y if and only if $\dim\mathcal{R}(L) = \dim Y$.

b Show that L is one-to-one if any only if $\dim\mathcal{R}(L) = \dim X$.

c Show that L is invertible if and only if $\dim X = \dim Y = \dim\mathcal{R}(L)$.

d What can one say about infinite-dimensional spaces?

Solution TODO: to add solution

2.2.7 Matrix Representation of Operators

■ **Problem 2.2.7.1** Let $[T]$ be an $m \times n$ matrix of scalars, and let X and Y be linear spaces where $\dim(X) = n$ and $\dim(Y) = m$. Show that there exists a linear transformation $T : X \to Y$ such that $[T]$ represents T relative to some Hamel basis.

Solution Let $\{e_i\}_{i=1}^n$ and $\{\hat{e}_j\}_{j=1}^m$ be a basis for X and Y. Let $L : X \to Y$ be a linear transformation that

$$Te_i = \sum_{j=1}^{m} [T]_{ji} \hat{e}_j.$$

I.e. maps each e_i to a vector in Y that its coordinates are the entries in the i^{th} column of the matrix.

■ **Problem 2.2.7.2** Let X and Y be finite-dimensional linear spaces with $\dim X = n$ and $\dim Y = m$. Consider the linear space $lt[X, Y]$ of all linear transformation $T : X \to Y$ and the linear space M_{mn} of all $m \times n$ matrices of scalars. The relation "$[T]$ represents T relative to B_1 and B_2" is a mapping of $lt[X, Y]$ into M_{mn}. Show that it is an isomorphism. In other words, $lt[X, Y]$ and M_{mn} are isomorphic.

Solution Let B_X and B_Y be two basis for X and Y respectively. Let $\Phi : \mathcal{L}[X, Y] \to M_{mn}$ such that

$$[\Phi T]_i = (\alpha_{1i}, \alpha_{2i}, \alpha_{3i}, \cdots, \alpha_{mi})^T.$$

Where $[\Phi T]_i$ is the i^{th} column of the matrix $[\Phi T]$. Furthermore let $\Psi : M_{mn} \to \mathcal{L}[X, Y]$ such that

$$([\Psi M]x)_i = (M_{1i}, M_{2i}, \cdots, M_{mi}).$$

It is easy to check that Ψ and Φ are inverses of each other and are both linear. So Φ is an isomorphism between these spaces.

■ **Problem 2.2.7.3** Let X be a finite-dimensional linear space and let $B_1 = \{x_1, \cdots, x_n\}$ and $B_2 = \{y_1, \cdots, y_n\}$ be two bases for X. Thus any $x \in X$ can be expressed uniquely as

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n,$$

and

$$x = \beta_1 x_1 + \cdots + \beta_n x_n.$$

Show that an $n \times n$ matrix can be used to represent the transformation from the B_1 to B_2 coordinate system. Moreover, show that the same $n \times n$ matrix is the representation of the identity transformation $I : X \to X$ relative to B_1 and B_2.

Solution There are Two ways to look at this problem. T as a map from (X, B_1) to (X, B_2), where B_1 and B_2 are two bases for the space X. In this case one has

$$Tx_1 = y_1, \quad Tx_2 = y_2 \quad , \cdots, Tx_n = y_n.$$

So T is represented by the identity matrix. However, this is not a useful formulation. We are interested in relating the expansion coefficients $(\alpha_1, \cdots, \alpha_n)$ to $(\beta_1, \cdots, \beta_n)$. For this purpose, we consider the map T from (X, B_1) to (X, B_1). I.e. T maps x_i to y_i as an element in (X, B_1) (thus expanded in terms of B_1), not in terms of B_2. So

$$Tx_i = y_i = \sum_{j=1}^{n} T_{ji} x_j.$$

Recording these coefficients in a matrix T one can write

$$[T] = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{21} & T_{22} & \cdots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1} & T_{n2} & \cdots & T_{nn} \end{pmatrix},$$

Then $[T]$ transforms $(1, 0, \cdots, 0)^T$ to $(T_{11}, \cdots, T_{n1})$, i.e. the coordinates of y_i expanded in terms of B_1. So for

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n,$$
$$x = \beta_1 x_1 + \cdots + \beta_n x_n,$$

one can write

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = [T] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

■ **Problem 2.2.7.4** Let X and Y be linear spaces over the same scalar field, and let B_1 and B_2 be countable Hamel bases for X and Y, respectively. Further, let $T : X \to Y$ be linear. Show that an infinite matrix can be used to represent T./

Solution Let $e_i \in B_1$. Then one can write

$$Te_i = \sum_{j \in I_i} \alpha_{ji} \bar{e}_j,$$

where $I_i \subset \mathbb{N}$ is finite (because expansions in Hamel basis are finite). Let $[T]$ be an infinite matrix that $[T]_{ij} = 0$ if $j \notin I_i$ and $T_{ji} = \alpha_{ji}$ otherwise. Observe that the columns of $[T]$ are finitely supported.

■ **Problem 2.2.7.5** Let X be the linear space made up of all functions $\Phi(t, x)$ of the form

$$\Phi(t,x) = \alpha_1 + (\alpha_{21} t + \alpha_{22} x) + (\alpha_{31} t^2 + \alpha_{32} tx + \alpha_{33} x^2) + \cdots + (\alpha_{n1} t^{(n-1)} + \alpha_{n2} t^{(n-2)} x + \cdots + \alpha_{nn} x^{(n-1)}),$$

where $(t, x) \in [0, T] \times [0, L]$, n is a fixed positive integer, and the α's are scalars. What is the dimension of X? Show that the equation

$$\frac{\partial \Phi}{\partial t} - k \frac{\partial^2 \Phi}{\partial x^2} = \Psi$$

where k is a constant, represents a linear transformation of X into itself. Represent this linear transformation with a matrix. Is the transformation one-to-one? Does it map X onto itself?

Solution There are $1 + 2 + \cdots + n$ coefficients. So $\dim X = (n)(n + 1)/2$. For the matrix representation of the partial differential equation, we let $n = 5$ for a more concrete example. When $n = 5$ then Φ will be of the form

$$\begin{aligned} \Phi(t,x) = \alpha_1 &+ (\alpha_{21} t + \alpha_{22} x) \\ &+ (\alpha_{31} t^2 + \alpha_{32} xt + \alpha_3 x^2) \\ &+ (\alpha_{41} t^3 + \alpha_{42} xt^2 + \alpha_{43} x^2 t + \alpha_{44} x^3) \\ &+ (\alpha_{51} t^4 + \alpha_{52} xt^3 + \alpha_{53} x^2 t^2 + \alpha_{54} x^3 t + \alpha_{55} x^4). \end{aligned}$$

So the mapping $\frac{\partial}{\partial t} : X \to X$ is given by

$$\begin{aligned}
&1 \mapsto 0, \quad t \mapsto 1, \quad x \mapsto 0, \quad t^2 \mapsto 2t\\
&tx \mapsto x, \quad x^2 \mapsto 0, \quad t^3 \mapsto 3t^2, \quad t^2x \mapsto 2tx\\
&tx^2 \mapsto x^2, \quad x^3 \mapsto 0, \quad t^4 \mapsto 4t^3, \quad xt^3 \mapsto 3xt^2\\
&x^2t^2 \mapsto 2x^2t, \quad x^3t \mapsto x^3, \quad x^4 \mapsto 0.
\end{aligned}$$

Similarly, for $\partial^2/\partial x^2 : X \to X$ one can write

$$\begin{aligned}
&1 \mapsto 0, \quad t \mapsto 0, \quad x \mapsto 0, \quad t^2 \mapsto 0\\
&tx \mapsto 0, \quad x^2 \mapsto 2, \quad t^3 \mapsto 0, \quad t^2x \mapsto 0\\
&tx^2 \mapsto 2t, \quad x^3 \mapsto 6x, \quad t^4 \mapsto 0, \quad xt^3 \mapsto 0\\
&x^2t^2 \mapsto 2t^2, \quad x^3t \mapsto 6xt, \quad x^4 \mapsto 12x^2.
\end{aligned}$$

So with respect to the ordered basis

$$\{1, t, x, t^2, xt, x^2, t^3, xt^2, x^2t, x^3, t^4, xt^3, x^2t^2, x^3t, x^4\}$$

can be constructed using the value of the mapping on each basis vector. This map is not one-to-one because its kernel has other vectors than the origin. For instance $x \in \ker L$. And also, it is not onto. Because, for instance $x_4 \in X$ has not pre-image.

■ **Problem 2.2.7.6** Consider the linear operator $K : X \to X$ given by $y = Kx$, where

$$y(t) = \int_0^{2\pi} k(t,s)x(s)ds, \qquad k(t,s) = 4\cos\big(2(t-s)\big),$$

and X is the linear space spanned by $\{1, \cos s, \cos 2s, \sin s, \sin 2s\}$.

(a) Express K as a matrix.

(b) Is K one-to-one?

(c) Does it map X onto itself?

Solution (a) In order to write K as a matrix, we need to calculate the value of K when applied on the basis functions. It turns out

$$1 \mapsto 0, \quad \sin s \mapsto 0, \quad \cos s \mapsto 0, \quad \cos(2s) \mapsto \pi\cos(2s), \quad \sin(2s) \mapsto \pi\sin(2s).$$

So with respect to the ordered basis $\{1, \sin s, \cos s, \sin 2s, \cos 2s\}$ the matrix representation of the operator will be

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & \pi & 0\\ 0 & 0 & 0 & 0 & \pi \end{bmatrix}$$

(b) No. Its kernel is non-trivial.

(c) No. $2 = \operatorname{rank} K \neq \dim X = 4$.

■ **Problem 2.2.7.7** Let L be a linear transformation of X onto Y where X is finite dimensional. Let $[L]$ be a matrix representing L. Show that if L is one-to-one, then $[L]$ is a square matrix and $\det[L] \neq 0$.

Solution Since L is one-to-one, then its kernel is trivial, so $\mathcal{N}L = \{0\}$. Using the fact that $\dim X$ finite and L is onto, then $\operatorname{rank} L = \dim Y$. Using the rank-nullity theorem one gets $\dim X = \operatorname{rank} L$. So $\dim X = \dim Y$ and the matrix representation of T is an square matrix.

L being one-to-one means its kernel is trivial, i.e. the only solution to $Ax = 0$ is $x = 0$. So A should be invertible, thus we need to have $\det(A) \neq 0$.

■ **Remark 2.11** Since L is onto, then $\dim X \geq \dim Y$ (otherwise, the extra bases in Y has no pre-images, thus the map is not onto)

■ **Problem 2.2.7.8** Let $X = Y$ denote the space of all fourth degree polynomials in t and define $L : X \to Y$ by $L = D^2 + 2D + I$ where D is the differential operator, that is

$$Lx = \frac{d^2x}{dt^2} + 2\frac{dx}{dt} + x.$$

(a) Represent L by a matrix L in terms of the ordered basis $\{1, t, t^2, t^3, t^4\}$ on X and Y.

(b) Represent $L^2 = LL$ in terms of this basis. Show that $[L^2] = [L][L]$ in terms of the usual matrix product.

(c) Repeat steps (a) and (b) for $M = \alpha D^2 + \beta D + \gamma I$.

Solution (a) We need to evaluate the effect of L on the basis vectors. It is straightforward to see

$$1 \mapsto 1, \quad t \mapsto t + 2, \quad t^2 \mapsto t^2 + 4t + 2, \quad t^3 \mapsto t^3 + 6t^2 + 6t, \quad t^4 \mapsto t^4 + 8t^3 + 12t^2.$$

So with respect to the ordered basis $\{1, t, t^2, t^3, t^4\}$ one can write

$$[L] = \begin{bmatrix} 1 & 2 & 2 & 0 & 0 \\ 0 & 1 & 4 & 6 & 0 \\ 0 & 0 & 1 & 6 & 12 \\ 0 & 0 & 0 & 1 & 8 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

(b) With the map $L^2 = LL$ one reads

$$1 \mapsto 1, \quad t \mapsto t + 4, \quad t^2 \mapsto (t^2 + 4t + 2) + 4(t + 2) + 2$$
$$t^3 \mapsto (t^3 + 6t^2 + 6t) + 6(t^2 + 4t + 2) + 6(t + 2), \quad t^4 \mapsto (t^4 + 8t^3 + 12t^2) + 8(t^3 + 6t^2 + 6t) + 12(t^2 + 4t + 2).$$

writing this operator with respect to the ordered basis as above one verifies

$$[L^2] = [L][L]$$

(c) Similar to above!

■ **Problem 2.2.7.9** Let X_n denote the space of all polynomials in t of degree less than n.

(a) Consider the operator

$$y = Lx \quad \Longleftrightarrow y(t) = \frac{d}{dt}(t^2-1)\frac{dx}{dt}$$

on X_4. Find a representation of $L : X_4 \to X_4$ with respect to the basis

$$A = \{1, t, 3/2t^2 - 1/2, 5/2t^3 - 3/2t\}.$$

(b) Find a basis B for X_4 such that the operator

$$y = Hx \quad \Longleftrightarrow y(t) = \frac{d^2x}{dt^2} - 2t\frac{dx}{dt}$$

can be represented by a diagonal matrix.

Solution (a) We express $[L]$ in the ordered basis $B = \{1, t, t^2, t^3\}$.

$$1 \mapsto 0, \quad t \mapsto 2t, \quad t^2 \mapsto 6t^2 - 2, \quad t^3 \mapsto 12t^3 - 6t.$$

So the matrix representation will be

$$[L]_B = \begin{pmatrix} 0 & 0 & -2 & 0 \\ 0 & 2 & 0 & -6 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 12 \end{pmatrix}.$$

In order to express $[L]$ with respect to the basis, consider the change of basis matrix

$$M = \begin{pmatrix} 1 & 0 & -1/2 & 0 \\ 0 & 1 & 0 & -3/2 \\ 0 & 0 & 3/2 & 0 \\ 0 & 0 & 0 & 5/2 \end{pmatrix}$$

So we will have

$$[L]_A = M[L]_B M^{-1}.$$

So

$$[L]_A = \frac{1}{30}\begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 2 & 0 & -3 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}\begin{pmatrix} 0 & 0 & -2 & 0 \\ 0 & 2 & 0 & -6 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 12 \end{pmatrix}\begin{pmatrix} 15 & 0 & 5 & 0 \\ 0 & 15 & 0 & 9 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}$$

(b) Let $B = \{1, t, t^2, t^3\}$ be an ordered basis. Then the effect of H on the basis vectors is

$$1 \mapsto 0, \quad t \mapsto -2t, \quad t^2 \mapsto 2 - 4t^2, \quad t^3 \mapsto 6t - 6t^3.$$

Then H can be represented by

$$A = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & -2 & 0 & 6 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & -6 \end{pmatrix}$$

We find the eigenvalues and eigenvectors of this matrix by solving all λ_i and v_i that satisfies

$$\det(A - I\lambda) = 0, \qquad (A - I\lambda)v = 0.$$

The characteristic polynomial is

$$\lambda(\lambda + 2)(\lambda + 4)(\lambda + 6) = 0.$$

Then

$$\lambda_1 = 0, \quad \lambda_2 = -2, \quad \lambda_3 = -4, \quad \lambda_4 = -6.$$

The corresponding eigenvectors are

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 \\ -3 \\ 0 \\ 2 \end{bmatrix}.$$

So the basis under which the operator is diagonal is:

$$\{1, t, 1 - 2t^2, -3t + 2t^3\}.$$

2.2.8 Equivalent Linear Transformation

■ **Problem 2.2.8.1** Let X be the sequence space $\ell_2(-\infty, \infty)$ and let Z be the same as the linear space Z in Example 4, Section 5. Let $S_r : X \to X$ be the right shift operator. Let $T : Z \to Z$ be defined by $(Ty)(z) = y(z)/z$ for $|z| = 1$. Show that T and S_r are similar.

Solution We want to show that the following diagram commutes.

$$\begin{tikzcd} \ell_2(-\infty,\infty) \arrow[r, "S_r"] \arrow[d, "\Phi"'] & \ell_2(-\infty,\infty) \arrow[d, "\Phi"] \\ Z \arrow[r, "T"'] & Z \end{tikzcd}$$

where Φ is the Z-transformation. Recall that

$$[\Phi\xi](z) = \sum_{n=-\infty}^{\infty} \xi_n z^{-n},$$

where $\xi = (\cdots, \xi_1, \xi_0, \xi_1, \cdots)$. So

$$(T \circ \Phi)(z) = 1/z \sum_{n=-\infty}^{\infty} \xi_n z^{-n} = \sum_{n=-\infty}^{\infty} \xi_n z^{-n-1}.$$

On the other hand

$$(\Phi \circ T)(z) = \sum_{n=-\infty}^{\infty} \xi_n z^{-n-1}.$$

So the diagram above commutes and T and Φ are similar.

■ **Problem 2.2.8.2** Let $X, Y, \mathcal{X}, \mathcal{Y}$ be finite dimensional linear spaces over the same scalar field, where X and $\mathcal{X}$ are isomorphic and Y and $\mathcal{Y}$ are isomorphic. Let $T : X \to Y$ and $\mathcal{T} : \mathcal{X} \to \mathcal{Y}$ be linear. Show that T and $\mathcal{T}$ are isomorphically equivalent if and only if $\text{rank}(T) = \text{rank}(\mathcal{T})$.

Solution Let $\Phi : X \to \mathcal{X}$ and $\Psi : Y \to \mathcal{Y}$ be the isomorphism. For the forward direction we want to show that similarity of T and $\mathcal{T}$ implies $\operatorname{rank} T = \operatorname{rank} \mathcal{T}$. We show this by proof by contrapositive. Assume $\operatorname{rank} T \neq \operatorname{rank} \mathcal{T}$. So $\dim T(X) = m_1$ and $\dim \mathcal{T}\mathcal{X} = m_2$ where $m_1 \neq m_2$. Using the fact that Φ and Ψ do preserve the dimension, one can write

$$\dim \Psi T(X) = m_1, \qquad \dim \mathcal{T}\Phi(X) = m_2.$$

Since $m_1 \neq m_2$, thus $\Psi T \neq \mathcal{T}\Phi$.

For the converse, we want to show that $\operatorname{rank} T = \operatorname{rank} \mathcal{T}$ implies T and $\mathcal{T}$ are similar. We use proof by contrapositive. Assume $\operatorname{rank} T \neq \operatorname{rank} \mathcal{T}$. Since Φ and Ψ both preserves the dimension, it follows that

$$\operatorname{rank} \Psi T \neq \operatorname{rank} \Phi \mathcal{T}.$$

So $\Psi T \neq \Phi \mathcal{T}$.

■ **Remark 2.12** In short, the proof idea of the question above is to use the fact that the isomorphism preserves the dimension. So to prove equal rank implies similarity, by contrapositive we need to show non similarity implies non-equal rank. So if

$$\Psi T \neq \Phi \mathcal{T},$$

then it follows that $\operatorname{rank} \Psi T \neq \operatorname{rank} \Phi \mathcal{T}$ which furthermore implies that $\operatorname{rank} T \neq \operatorname{rank} \mathcal{T}$. And to show that similarity implies equal rank, by contrapositive we need to show that non-equal rank implies non-similarity. So if

$$\operatorname{rank} T \neq \operatorname{rank} \mathcal{T}$$

then it follows that

$$\operatorname{rank} \Psi T \neq \operatorname{rank} \mathcal{T}\Phi.$$

So $\Psi T \neq \mathcal{T}\Phi$.

■ **Problem 2.2.8.3** Let S and T be linear operators on $\mathbb{R}^2$ and assume that there is a basis for $\mathbb{R}^2$ in which S and T have the following representation.

$$S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad T = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}.$$

(a) Show that S and T are not similar.

(b) Are S at T ever isomorphically equivalent?

Solution (a) The matrix S has one eigenvalue $\lambda = 1$ with algebraic multiplicity equal to 2. However, the geometric multiplicity (dimension of the null space of $A - \lambda I$) is 1. So this matrix is deficient and can not be diagonalized. So S and T are not similar.

(b) No. Because of the reason above.

■ **Problem 2.2.8.4** Let M and N be linear operators on C^n that can be represented as diagonal matrices

$$[M] = \text{diag}\{\mu_1, \cdots, \mu_n\}, \qquad [N] = \text{diag}\{\nu_1, \cdots, \nu_n\},$$

in terms of some basis. Show that M and N are similar if and only if the two sets $\{\mu_1, \cdots, \mu_n\}$ and $\{\nu_1, \cdots, \nu_n\}$ are the same.

Solution Assume the sets $\{\mu_1, \cdots, \mu_n\}$ and $\{\nu_1, \cdots, \nu_n\}$ are the same. So there is a permutation $\sigma \in S_n$ that maps the ordered tuple $(\mu_1, \cdots, \mu_n)$ to $(\nu_1, \cdots, \nu_n)$. If $\sigma\mu_i = \mu_j$, then let S be a matrix that in its i^{th} row, all of the columns are zero except for the j^{th} row. So

$$M = S^{-1}NS.$$

For the converse, note that M and N commutes with any other matrix. So assuming M, N are similar, then $M = SNS^{-1}$. One can write

$$M = SNS^{-1} = NSS^{-1} = N.$$

So $\{\mu_1, \cdots, \mu_n\} = \{\nu_1, \cdots, \nu_n\}$.

■ **Remark 2.13** To prove that N commutes with any other matrix, let A be any matrix. Then

$$[AN]_{ij} = \sum_k A_{ik}N_{kj} = A_{ij}N_{jj}.$$

And

$$[NA]_{ij} = \sum_k N_{ik}A_{kj} = N_{ii}A_{ij}.$$

Since $N^T = T$, then it follows that $[AN]_{ij} = [NA]_{ij}$.

2.2.9 Projection

■ **Problem 2.2.9.1** Let P be a projection on a linear space X. Show that the range of P is given by

$$\mathcal{R}(P) = \{x \in X : Px = x\}.$$

Solution Let $y \in \mathcal{R}(P)$. Then $\exists x \in X$ such that $Px = y$. Applying P to both sides one gets $P^2x = Py$. Since $P^2 = P$ one gets $Px = Py$ and since $Px = y$ one gets $y = Py$. So $y \in \{x \in X : Px = x\}$. For the converse, let $x \in \{x \in X : Px = x\}$. By definition of the range we have $x \in \mathcal{R}(P)$.

■ **Problem 2.2.9.2** Let $X = L_2[-\pi, \pi]$ and show that

$$(Px)(t) = \int_{-\pi}^{\pi} K(t, \tau)x(\tau)d\tau,$$

where

$$K(t, \tau) = \frac{1}{2\pi} \sum_{n=-10}^{10} e^{in(t-\tau)}$$

represents a projection on X.

Solution To see this observe that

$$e^{in(t-\tau)} = \cos\big(n(t-\tau)\big) + i\sin\big((t-\tau)\big).$$

Since the sum in $K(t,\tau)$ runs from -10 to $+10$, then the sine terms will cancel out (because they are even functions). So we will have

$$K(t,\tau) = \frac{1}{2\pi}\sum_{n=-10}^{10}\cos\big(n(t-\tau)\big).$$

Furthermore, it is easy to check that

$$\int_{-\pi}^{\pi}\cos\big(n(t-s)\big)\cos\big(m(s-\tau)\big)ds = 2\pi\delta_{n,m}\cos\big(n(t-\tau)\big).$$

So

$$\int_{-\pi}^{\pi}K(t,s)K(s,\tau) = K(t,\tau).$$

So one can write

$$P(Px)(t) = \int_{-\pi}^{\pi}K(t,s)\int_{-\pi}^{\pi}K(s,\tau)x(\tau)d\tau ds = \int_{-\pi}^{\pi}K(t,\tau)x(\tau)d\tau = (Px)(t).$$

So since P is linear and $P^2 = P$, P is a projection map.

■ **Problem 2.2.9.3** Let $X = C[0,T]$ and define P by

$$(Px)(t) = x(0)(1-t), \qquad \text{for } 0 \le t \le T.$$

Show that P is a projection.

Solution P is a linear map. Further more, observe that

$$P(Px)(t) = (Px)(0)(1-t) = x(0)(1-t).$$

So $P^2 = P$ and P is a projection map.

■ **Remark 2.14** In general, for a nice enough function

$$(Px)(t) = x(0)f(t),$$

where $f(0) = 1$ is a projection map.

3. Combined Topological and Algebraic Structures

Summary 3.1 — A useful trick. Let $x, y \in \mathbb{R}$. Then we want to show for $p \in (0,1)$ we have

$$|x+y|^p \leq |x|^p + |y|^p \,.$$

To see this let $\beta = 1 - p$. Then we can write

$$|x+y|^p = |x+y|^{1-\beta} = \frac{|x+y|}{|x+y|^\beta} \leq \frac{|x|}{|x+y|^\beta} + \frac{|y|}{|x+y|^\beta} \leq \frac{|x|}{|x|^\beta} + \frac{|y|}{|y|^\beta} = |x|^p + |y|^p \,.$$

Summary 3.2

$$\|x\| = \int_0^1 |x(t)|^p \, dt < \infty$$

is a not a norm but $d(x,y) = \|x-y\|$ is a metric.

Summary 3.3 — Not norm but metric!. There are functions $\|\cdot\| : V \to \mathbb{R}$ defined on a vector space, that are not norm, but $d(x,y) = \|x-y\|$ is a metric. It is enough for a function $\|\cdot\|$ to be non-negative, and positive definite, and satisfy the triangle inequality and have $\|x\| = \|-x\|$. Then $d(x,y) = \|x-y\|$ will be a metric. There is no need for homogeneity condition.

Summary 3.4 Any metric d on a vector space V comes from a function $\|x\|$ that need to satisfy all norm properties except for the homogeneity condition. And it also need to satisfy $\|x\| = \|-x\|$. See Problem 3.1.1.7.

3.1 Problems

3.1.1 Normed Linear Spaces

■ **Problem 3.1.1.1** Show that the norm $\|\cdot\|$ considered as a mapping of a normed linear space X into the reals is continues.

Solution We use the reverse triangle inequality, i.e. for $x, y \in X$ we have

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

Let $\epsilon > 0$ given and let $d(x, y) = \|x - y\| < \epsilon$. Then it implies that

$$d(\|x\|, \|y\|) = \big|\|x\| - \|y\|\big| \leq \|x - y\| < \epsilon.$$

So $\|\cdot\|$ is continuous.

■ **Remark 3.1** In the problem above we implicitly assume that the real line (co-domain of the norm function) is has the topology induced by $|\cdot|$.

■ **Problem 3.1.1.2** Show that the addition and scalar multiplications in a vector space are continuous.

Solution Let $\cdot + \cdot : X \times X \to X$ denote the addition operation in a vector space. We assume that X has the topology induced by $\|\cdot\|$ and $X \times X$ has the topology induced by $\big\|(x, y)\big\| = \|x\| + \|y\|$. Let $(a, u), (b, v) \in X \times X$. Let $\epsilon > 0$ such that $d((a, u), (b, v)) = \big\|(a - b, v - u)\big\| = \|a - b\| + \|v - u\| < \epsilon$. Then we can write

$$d(a + u, b + v) = \|a + u - b - v\| \leq \|a - b\| + \|u - v\| = d((a, u), (b, v)) < \epsilon.$$

For the scalar multiplication, let $\cdot : F \times X$ denote the scalar multiplication. We assume F is $\mathbb{R}$ with the usual topology. Then

$$\|\alpha u - \beta v\| = \|\alpha u \pm \alpha v - \beta v\| \leq \big\|\alpha(u - v)\big\| + \big\|v(\alpha - \beta)\big\| = |\alpha| \, \|u - v\| + |\alpha - \beta| \, \|v\|.$$

Let $\epsilon > 0$ given and choose $\|u - v\| < \epsilon/\alpha$ and $|\alpha - \beta| < \epsilon/\|v\|$. So when $d((\alpha, u), (\beta, v)) = \|u - v\| + |\beta - v| < \delta$ where $\delta = \epsilon/\alpha + \epsilon/\|v\|$ we will have

$$\|\alpha u - \beta v\| < \epsilon.$$

So the scalar multiplication is continuous. Note that the choice of δ depends on v. So the scalar multiplication is **not** uniformly continuous.

■ **Problem 3.1.1.3** Characterize all possible norms on the real line $\mathbb{R}$, where $\mathbb{R}$ is considered a real linear space. Ont eh complex plane $\mathbb{C}$, where $\mathbb{C}$ is considered as a complex linear space. Show that $\mathbb{C}$ may have other norms when it is considered as a real linear space.

Solution We know that $|\cdot| : \mathbb{R} \to \mathbb{R}$ is a norm on $\mathbb{R}$. Let $f(x)$ be any other norm on this space. Then define $g(x) = f(x)/|x|$ for $x \neq 0$. For any $\alpha \in \mathbb{R}$ we will have

$$g(\alpha x) = g(x) \quad \forall \alpha, x \in \mathbb{R}.$$

This implies that g is a constant function. Because, let $x, y \in \mathbb{R}\backslash\{0\}$ then $g(x) - g(y) = g(x) - g(\beta x) = g(x) - g(x) = 0$ where $\beta = y/x$. So $f(x) = \alpha |x|$ where $g(x) \equiv \alpha$. So all of the norms on $\mathbb{R}$ are of the form $\|x\|$.

On the complex plane $\mathbb{C}$ as a vector space over $\mathbb{C}$ we define the norm as

$$\|u\| = \sqrt{u\overline{u}}.$$

Similar to above, define $g(u) = f(u)/\|u\|$ where $f(u)$ is any other norm. Then we will have $g(\alpha u) = g(u)$ for all $u, \alpha \in \mathbb{C}$. Let $u, v \in \mathbb{C}$. Then $\exists \alpha \in \mathbb{C}$ such that $u = \alpha v$ (this is only true if $\mathbb{C}$

is a vector space over $\mathbb{C}$). So $g(u) - g(v) = g(u) - g(\alpha u) = 0$. So g is a constant function. So we conclude that every norm in $(\mathbb{C}, \mathbb{C})$ is in the form of $|\alpha| \|\cdot\|$ for $\alpha \in \mathbb{C}$.

However, if we consider $\mathbb{C}$ as a vector space over $\mathbb{R}$, then there are other norms different from the form above. For instance $\|u\| = \sqrt{\mathrm{Im}(u)^2 + \mathrm{Re}(u)^2}$ is a norm.

■ **Problem 3.1.1.4** Let $(X, \|\cdot\|)$ be a normed linear space and let $S_r = \{x : \|x\| = r\}$ where $r > 0$. Assume that $X \neq \{0\}$. Show that $(X, \|\cdot\|)$ is a Banach space if and only if the metric space $(S_r, \|\cdot\|)$ is complete for some $r > 0$.

Solution First we show if S_r is complete then X is complete. Let $\{x_n\}$ be a Cauchy sequence in X. There are two cases: The first case is that x_n gets arbitrarily close to the origin, and the second case is that for large enough n, x_n is uniformly away from the origin. Since the sequence is Cauchy, these are the only two possible cases. For the first case we have $\forall \epsilon > 0$, $\exists N > 0$ such that $\forall n > N$ we have $\|x_n\| < \epsilon$, we have $x_n \to 0 \in X$, thus converges. The second case is that for n large enough, the sequence is uniformly away from the origin. For this case case, where $\exists \epsilon > 0$ and $\exists N > 0$ such that $\|x_n\| > 0$. We map the sequence $\{x_n\}$ to a sequence $\{y_n\} \subset S_r$ via

$$y_n = \frac{r x_n}{\|x_n\|}.$$

Since this map is continuous (using the fact that every Cauchy sequence is bounded, so is $\{x_n\}$), y_n is also Cauchy, thus it converges: $y_n \to y \in S_r$. Using the continuity of $\|\cdot\|$, the sequence $\{\|x_n\|\} \subset \mathbb{R}$ is Cauchy, thus it converges: i.e. $\|x_n\| \to a \in \mathbb{R}$. We claim $x_n \to x = ay/r$. Because if we assume other wise, then $\{y_n\}$ is not converging to y, which is against our hypothesis.

For the other direction, we assume X is Banach and we want to conclude S_r is complete. Let $\{x_n\} \subset S_r$ be a Cauchy sequence. Since $S_r \subset X$, then $\{x_n\}$ is also a Cauchy sequence in X, hence converges to $x \in X$. We want to show that $\|x\| = r$, thus belongs to S_r. To show this we use the fact that $\|\cdots\|$ is a continuous map. So since $\|x_n\| = r$ for all n, it follows that $\|x\| = 1$ as well.

■ **Problem 3.1.1.5** Let p satisfy $0 < p < 1$ and consider the space $L_p[0,1]$ of all functions with

$$\|x\| = \int_0^1 |x(t)|^p \, dt < \infty.$$

Show that $\|x\|$ is not a norm on $L_p[0,1]$. Show that $d(x,y) = \|x - y\|$ is a metric on $L_p[0,1]$. Hint: Note that if $0 \leq \alpha \leq 1$ then $\alpha < \alpha^p \leq 1$.

Solution The fact that $\|\cdot\|$ as defined above is not a norm follows from the fact that it is not homogeneous. I.e. for $\alpha \in \mathbb{R}$ we have $\|\alpha x\| = |\alpha|^p \|x\|$. But $d(x,y) = \|x - y\|$ is a metric. That is because it satisfies the metric properties, in particular $d(x,y) \leq d(x,z) + d(z.y)$. See the following remark for the triangle inequality.

■ **Remark 3.2** Let $x, y \in \mathbb{R}$. Then we want to show for $p \in (0,1)$ we have

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

To see this let $\beta = 1 - p$. Then we can write

$$|x + y|^p = |x + y|^{1-\beta} = \frac{|x + y|}{|x + y|^\beta} \leq \frac{|x|}{|x + y|^\beta} + \frac{|y|}{|x + y|^\beta} \leq \frac{|x|}{|x|^\beta} + \frac{|y|}{|y|^\beta} = |x|^p + |y|^p.$$

■ **Problem 3.1.1.6** Define

$$\begin{aligned}\sigma_n(f) &= \sup\{\left|f(t)\right| : |t| \le n\}, \\ \rho_n(f) &= \min(1, \sigma_n(f)), \\ \|f\| &= \sum_{n=1}^{\infty} 2^{-n}\rho_n(f),\end{aligned}$$

where $f \in C(-\infty, \infty)$.

(a) Show that $\sigma_n(f)$ is a pseudo-norm on on $C(-\infty, \infty)$.

(b) Show that $\rho_n(f)$ and $\|f\|$ are not norms.

(c) Show that $d(f, g) = \|f - g\|$ is a metric on $C(-\infty, \infty)$.

(d) Show that $d(f_n, f) \to 0$ as $n \to \infty$ if and only if $f_n(t) \to f(t)$ uniformly on compact sets in $-\infty < t < \infty$.

Solution (a) For any choice of n, the function

$$f(t) = \begin{cases} 0 & |t| < n, \\ x - n & t \ge n, \\ -x - n & t \le n \end{cases},$$

and the origin both has norm zero.

(b) Let $f \equiv 5$. Then $\sigma_n(f) = 5$ for all $n \in \mathbb{N}$ and $\rho_n(f) = 1$ for all $n \in \mathbb{N}$. For all $\alpha \ge 1/5$ we also have $\rho_n(\alpha f) = \rho_n(f)$. Thus ρ is not homogeneous. So it is not a norm. For the same reason, $\|\cdot\|$ is not a norm.

(c) $d(x, y) = \|x - y\|$ satisfies all of the metric properties. In particular, for the triangle inequality we have

$$\sigma_n(f + g) = \sup\{|f + g|\} \le \sup\{|f| + |g|\} \le \sup\{f\} + \sup\{g\} = \sigma_n(|f|) + \sigma_n(|g|),$$

where $\sup\{|f|\}$ is used as a short notation for $\sup\{\left|f(t)\right| : |t| < n\}$.. Furthermore, we know that $\min(1, a + b)$ for $a, b > 0$ we have $\min(1, a + b) \le \min(1, a) + \min(1, b)$. Because there are 4 cases: $a, b < 1$ with their sum $a + b < 1$, $a, b < 1$ with $a + b > 1$, $a, b > 1$, and $a < 1 < b$. In all of these cases the inequality holds. Furthermore, we have

$$\|f + g\| = \sum_{n=1}^{\infty} 2^{-n}\rho_n(f + g) \le \sum_{n=1}^{\infty} 2^{-n}\rho_n(f) + \sum_{n=1}^{\infty} 2^{-n}\rho_n(g).$$

This is true because $\rho_n(f+g) \le 1$ and $\rho_n(f+g) \le \rho_n(f)+\rho_n(g) \le 2$. So the $\sum 2^{-n}\rho_n(f+g)$ is absolutely convergent. This $\sum 2^{-n}\rho_n(f + g) \le \sum 2^{-n}(\rho_n(f) + \rho_n(g)) \le \sum 2^{-n}\rho_n(f) + \sum 2^{-n}\rho_n(g)$.

(d) Still thinking!

■ **Problem 3.1.1.7** Let X be a linear space and let $\|x\|$ be a real-valued function defined on X. Show that $d(x, y) = \|x - y\|$ is a metric if and only if $\|x\|$ satisfies (N1), (N2), (N4), and $\|x\| = \|-x\|$ for all $x \in X$.

Solution Assume $\|x\|$ satisfies (N1), (N2), (N4), and $\|x\| = \|-x\|$. Then $d(x, y) = \|x - y\|$ is positive (from N1) and $d(x, y) = 0$ iff $x = y$ (from positive definiteness of $\|\cdot\|$), and $d(x, y) = d(y, x)$ from $\|x\| = \|-x\|$. $d(x, y) \le d(x, z) + d(z, y)$ from the triangle inequality for $\|\cdot\|$.

Now assume $d(x, y) = \|x - y\|$ is a metric. $\|x\| = d(x, 0) \le 0$, so $\|\cdot\|$ is non-negative. Also $d(x, y) = d(y, x)$ implies $\|x\| = \|-x\|$. Furthermore, in $d(x, y) \le d(x, z) + d(z, y)$, let $z = 0$, and we will recover $\|x - y\| \le \|x\| + \|-y\|$. Let $\tilde{y} = -y$. Then $\|x + \tilde{y}\| \le \|x\| + \|\tilde{y}\|$, hence the triangle inequality for the norm. Lastly, for the positive definiteness, Let $\|x\| = 0$. Since $\|x\| = \|x - 0\| = d(x, 0)$, it follows that $d(x, 0) = 0$. Thus $x = 0$ from metric properties. Also if $x = 0$, then $\|0\| = \|0 - 0\| = d(0, 0) = 0$.

■ **Problem 3.1.1.8** Let $(X, \|\cdot\|)$ be a normed linear space and let $\{x_n\}$ be a sequence in X with $x = \lim_{n\to\infty} x_n$. Assume that $\|x_n - y\| \le a$ for all n. Show that $\|x - y\| \le a$.

Solution Since $x_n \to x$, from the continuity of the addition in vector space, it also follows that $x_n - y \to x - y$. From continuity of $\|\cdot\|$ it follows that the real sequence $\{\|x_n - y\|\}$ converges to $\|x - y\|$. Since all elements of the real sequence $\|x_n - y\|$ is less than or equal to a, then the for the limit will will have $\|x - y\| \le a$. This property for the real sequences can be shown with a simple proof by contradiction.

■ **Problem 3.1.1.9** Show that a Hamel basis for a Banach space is either finite or uncountably infinite. (Completeness is important for this. Use Exercise 17, section 3.13).