MTH327 Functional Analysis and MTH427 Introduction to Hilbert Spaces

Abdullah Naeem Malik

September 15, 2015

Contents

Preface

This set of notes is meant to aid the lectures for MTH327 Functional Analysis and MTH427 Introduction to Hilbert Spaces at COMSATS Institute of Information Technology, Virtual Campus. I have tried to explain material where necessary and added some theorems which aid in the explanation of the text apart from the usual syllabus delivered. Most of the exercise questions have been solved. Where necessary, only hints have been given. The course majorly follows the text given by the Functional Analysis classic titled "Introductory Functional Analysis with Applications" by Erwin Kreyszig (January 6, 1922 – December 12, 2008), first published in 1962.

Functional Analysis is a body of knowledge and tools from calculus and algebra in which one explores the role of different spaces, their elements and operators acting on them. Mathematically, a space is any set with some structure on it. These set of notes explores such an endeavour. While I have tried to make these notes self-contained, the introductory chapters, which are not supposed to be mandatory, need to be looked at before the course begins. The first chapter is just a review of set theory to get some notation straight. The second chapter of algebra, Group Theory, goes through some basics. Of course the basics in no way mean that one can develop a mastery of the subject of algebra. Nevertheless, once the details of the chapter are kept in mind, the student should have no difficulty in understanding their use in the text. Of particular importance is the role of a binary operator. You can imagine this as a machine in which two elements belonging to a particular set are sent because of which the machine gives out usually a different element of the same set as a result. The rigorous notation for the binary function should make this clear. The next important idea of group theory used in these texts is the idea of an inverse, substructures and that of isomorphism. Be sure to keep those in mind before you proceed.

Set Topology has been given extensive attention. From a hierarchial point of view, topological spaces are the most general type of spaces there are, followed by metric spaces, vector spaces, norm spaces, inner product spaces and finally Hilbert spaces. It is thus no surprise that Set Topology is freely indulged in, since notions and ideas from this field are used in the rest of the text. Other than that, there is this notion of a topological vector space – a vector space with a topology. This has not been covered in the lectures. If it were so, one could freely indulge in the fact that geometry is more than just space – it is an additional structure on space. Even though this has been covered in these notes

but with limitations by discarding topology for the moment.

I have tried to make the text proceed just as the lectures do, starting from vector spaces. However, I made some changes to keep the flow of this monologue in line with my preferences.

Vector spaces are the subject of study in linear algebra but they are incorporated in the syllabus for the course because from them one can move on to Hilbert spaces making a tour from norm spaces and inner product spaces, exploring the mathematics of operators acting on them. Hilbert spaces, while important in their own right, are very important in mathematics from an application point of view. Sadly, the length of the course does not afford the time and space for most of these applications.

I hope the material is sufficient to give you a glimpse of the abstract of the abstract world. Just like music cannot be learned by simply watching a musician play an instrument and language cannot be learned by merely listening, so, too, can mathematics not be learned by merely reading. The reader has to indulge in it, be challenged, try to do proofs on your own, look at everything from different perspectives and see why things are the way things are. A piece of sincere advise would be thus: wherever you see a question pop or an exercise, do pause for a moment to do some scribbling.

As has been my experience, I, at times, found "left as an exercise to the reader" frustrating because the book under consideration usually did not provide sufficient details for me to do the exercise/proof on my own. However, these notes are scattered with hints and can be done if enough attention is devoted to each line (except for the first two chapters). Wherever you do see this, you may assume that I have either been lazy or know that the required exercise is routine matter/easy.

This endeavour is a single-handed production and it is bound to be flawed. I welcome any suggestions for improvements, any gaps that might need more explanation, any typos and even incorrect proofs.

Introduction

Functional analysis is an abstract branch of mathematics that originated from classical analysis. It can be viewed as a great leap from Linear Algebra and its interplay with Topology. The name is derived from the word "functional" which is a particular generalisation of the real-valued function. Technically, its name is derived from a function whose argument is a function and the name was first used in Hadamard's 1910 book on that subject. This is covered in the course MTH485 Calculus of Variation. The general concept of functional had previously been introduced in 1887 by the Italian mathematician and physicist Vito Volterra in his attempt to look at Integral Equations. The theory of nonlinear gunctionals was continued by students of Hadamard; in particular, by Fréchet and Lévy. Hadamard also founded the modern school of linear functional analysis further developed by Riesz and the group of Polish mathematicians around Stefan Banach. In modern introductory texts to functional analysis, the subject is seen as the study of vector spaces endowed with a topology, in particular infinite dimensional spaces. In contrast, linear algebra deals mostly with finite dimensional spaces, and does not use topology. An important part of functional analysis is the extension of the theory of measure, integration, and probability to infinite dimensional spaces, also known as infinite dimensional analysis.

Historically, the impetus came from problems related to ordinary and partial differential equations, numerical analysis, calculus of variations, approximation theory and integral equations. Ordinarily, one deals with limiting processes in finite dimensional vector spaces ($\mathbb{R}$ or $\mathbb{R}^n$) but problems arising in the above applications required a calculus in spaces of functions (which are infinite dimensional vector spaces), among other spaces. The theory was further refined by David Hilbert with his Hilbert spaces in his attempt to look for solutions of Integral Equations, an excellent review of which may be found at http://www.cs.umd.edu/~stewart/FHS.pdf. The applications of such spaces are found to be enormous in the fields of Quantum Mechanics, Numerical Analysis, Complex Analysis, Real Analysis, Optimisation, Fourier Analysis and Operator Theory, to name a few.

Functional analysis is a branch of mathematical analysis, the core of which is formed by the study of vector spaces endowed with some kind of limit-related structure (e.g. inner product, norm, topology, etc.) and the linear operators acting upon these spaces and respecting these structures in a suitable sense. The historical roots of functional analysis lie in the study of spaces of functions

and the formulation of properties of transformations of functions such as the Fourier transform as transformations defining continuous, unitary etc. operators between function spaces, as has been explored in the text. This point of view turned out to be particularly useful for the study of differential and integral equations. Thus, Functional Analysis is a vast generalisation of various tools used in applications of different equations.

The study of objects in a particular space is the subject of Geometry whereas Functional Analysis plays a part in the study of spaces, among other subjects such as Algebraic Topology. The idea of seperating a space independtly from its objects was first concieved by Bertrand Reimann in his seminal lecture of 1854. This subject will view some of such ideas, too.

Part I

Preliminaries

This part of the book will focus on the ground-work, building our way to the areas which Functional Analysis is famous for. We will start with set theory in order to set some notions straight, work our way with some group theory. Some Set Topology as needed has been introduced, as well. The gentle introduction continues its way slowly by turning from an introduction of norm spaces and vector spaces, regressing to familiar concepts of metric spaces and then amplifying the theory of norm and vector spaces with the addition of Cauchy Sequences and operators. We then move on to study functionals and the important Dual Spaces before skimming over Pre-Hilbert Spaces and then finally to Hilbert Spaces.

Set Theory

In this chapter, an attempt to define all of the symbols and mathematics used in this book has been undertaken and every effort has been made to ensure that the notes are self-contained. However, this portion is not intended to be a substitute for any rigorous treatment. The purpose of this section is to serve as an introductory reminder for what follows since most of the students come from a different backgrounds. Hence, this section may be skipped, if necessary. Sections where a slight introduction is necessary have been placed under appropriate headings. Proofs and extensive examples are intentionally omitted and have been left as an exercise for the reader for this chapter. The reader is encouraged to try to attempt such proofs, apart from the exercise as they will serve as "warm ups".

1.1 Basics

Definition 1 *A **set** is any well-defined collection of distinct objects.*

The words "well-defined", "collection" and "objects" may be vague and we will not indulge in commenting any further. An intuitive understanding is assumed. The word "class" will be reserved for a collection of sets.

The stated collection is denoted by capital letters while the members are listed using curly braces, separated by commas. If an element p in a set A, then the statement "p belongs to A" or "p is in A" is mathematically written as $p \in A$.

Example 2 *The empty set $\{\}$. This special case is denoted by $\varnothing$. All other sets are non-empty.*

Elements which do not belong to a set X will always belong to its complement $X^{\complement}$. The complement of a set is always defined relative to a universal set.

Example 3 (Non-example) *The collection of those collections of objects that are not members of themselves is mathematically written as $\{x \mid x \notin x\}$. This is not a set, for historical and technical reasons, which we will not mention.*

Remark 4 *We will assume that the set of natural numbers $\mathbb{N} := \{1, 2, 3, ...\}$ exists from which the set of reals can be constructed using the Dedekind Cuts.*

For a rigorous treatment on the subject, see Tom M. Apostol's Calculus. The usual (classical) logical conventions of exclusive true/false are assumed. "Iff", "if and only if" and "⇔ " will mean the usual two-way implication, implying an equivalence. If we have a statement of the form "if p then q", then its contrapositive "if not q, then not p" will be used throughout without warning. In definitions, the word "if" will refer to "iff". The symbol "∀", the inverted "A" in "all" will mean "for all" and its equivalents whereas the symbol "∃", the reflected "E" in "exists" will mean "there exists" and its equivalents. Familiarity with the principle of induction and the concept of infinity is assumed. The Axiom of Choice will be used carelessly. The symbols $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ will denote, respectively, the set of integers, rationals, real and complex numbers. Bars over elements will denote conjugates.

Definition 5 *The **intersection**, **union**, **difference** and **symmetric difference** of any two sets A and B is defined, respectively, as*

$$A \cup B = \{x | x \in A \text{ and } x \in B\}$$

$$A \cap B = \{x | x \in A \text{ or } x \in B\}$$

$$A - B = \{x | x \in A \text{ and } x \notin B\}$$

and

$$A \Delta B = \{x | x \in A or x \in B \text{ but } x \notin A \cap B\}$$

Definition 6 *Let A and B be two sets. B is called a **subset** of A if $\forall b \in B$, $b \in A$*

This is written as $B \subset A$ for proper subsets. The notation "$\subseteq$" is reserved for improper subsets i.e. $B \subset A$ or $B = A$.

Remark 7 *Two sets are considered equal if they are subsets of each other.*

We will denote the power set of a set X (the collection of all subsets of X) with $\mathcal{P}(X)$, instead of 2^X.

Definition 8 *Let A and B be two sets. Then, **product set** or **Cartesian product** of A and B, written $A \times B$, consists of all ordered pairs (a, b) where $a \in A$ and $b \in B$ i.e. $A \times B = \{(a, b) : a \in A, b \in B\}$*

Loosely, the Axiom of Choice states that the Cartesian product of non-empty sets is non-empty. An equivalent to this statement is that one can always find a function, called choice function, which "picks" elements from any given set.

Example 9 *If $A = \{1, 2, 3\}$ and $B = \{4, 5, 6\}$, then,*

$$A \times B = \{(1, 4), (1, 5), (1, 6), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6)\}$$

It makes sense to define $A \times A = A^2$. Furthermore, $(A \times B) \times C$ and $A \times (B \times C)$ are equal (see exercise).

It is clear that if A has n elements and B has m elements, then $A \times B$ has nm elements.

Remark 10 *An alternative way to define an* ***ordered pair*** $(a, b) := \{\{a\}, \{a, b\}\}$ *was proposed in 1921. This definition is called the Kuratowski definition and it is now the standard definition of an ordered pair in set theory.*

Definition 11 *Let A and B be two sets. Then, a* ***binary relation*** *R is a subset of $A \times B$.*

For $a \in A$, $b \in B$, if a is related to b, then $(a, b) \in R$. This is compactly written as aRb. "a is not related to b" is denoted as $a \not R b$. A **binary relation** is said to be defined on A if it is a subset of A^2.

Example 12 *For $R = \{(1, 1), (2, 2), (3, 3), ...\} \subset \mathbb{N} \times \mathbb{N}$, $aRb \Leftrightarrow a = b$*

Definition 13 *Let $<$ be a binary relation. A* ***strict weak ordering*** *is a binary relation $<$ on a set S with the following propoerties:-*

- *$<$ is transitive ($a < b$ and $b < c \Rightarrow a < c$)*
- *$<$ is irreflexive ($a < b \Rightarrow b \not< a$)*
- *$<$ is asymmetric ($a \not< a$)*

for $a, b, c \in S$

Example 14 *$<$ is said to be an order on $\mathbb{N}$, written as $(\mathbb{N}, <)$, if the elements of $\mathbb{N} \times \mathbb{N}$ satisfy all the above properties in which the former member of the ordered pair is less in magnitude than its latter member. This order is the natural order on $\mathbb{N}$.*

Example 15 *The relation "$\subset$", which is usually read as "contained in" for sets, satisfies the properties for a strict weak order and is, therefore, a strict weak order.*

Definition 16 *Let $\leq$ be a binary relation. A* ***partial order*** *is a binary relation $\leq$ over a set S if $\forall a, b, c \in S$*

- *$a \leq a$ ($\leq$ is reflexive)*
- *$a \leq b$ and $b \leq a \Rightarrow a = b$ ($\leq$ is antisymmetric)*
- *$a \leq b$ and $b \leq c \Rightarrow a \leq c$ ($\leq$ is transitive)*

Example 17 *For any set X, $\mathcal{P}(X)$ forms a partial order (is a partially ordered set) under the partial order "$\subseteq$"*

A final property called totality is added to this list to define $\leq$ as a **total order** on S. This mathematically says that $a \leq b$ or $b \leq a$ $\forall a, b \in S$. This signifies a comparison or a relation of all elements in S.

The following law, the Trichotomy law, can be proved:

Theorem 18 *For any two elements a and b in a totally ordered set, exactly one of the following is true:*

- $a < b$
- $b < a$
- $a = b$

Definition 19 *Let $(S, \leq)$ be a totally ordered set and $A \subset S$. A is said to be **bounded above** if $\exists b \in S$ such that $a \leq b$ $\forall a \in A$*

Example 20 *$\{1, 2, 3, 4\} \subset \mathbb{N}$ is bounded above by 5 whereas $\{2, 3, 4, ...\}$ is not bounded above.*

Definition 21 *Let $(S, \leq)$ be a totally ordered set and $A \subset S$ be a non-empty subset such that A is bounded above. $\beta \in S$ is said to be the **least upper bound**, or **supremum** of A if*

- *β is an upper bound of A*
- *$\alpha < \beta \Rightarrow \alpha$ is not an upper bound of A*

In such a case, $\beta = \sup A$. Such a β is unique (proof?)

Example 22 *For $A = \{1, 2, 3, 4\} \subset \mathbb{Z}$, the set of upper bounds if $\{4, 5, 6, 7, ...\}$ and the set of lower bounds is $\{1, 0, -1, -2, ...\}$ so that $\sup A = 4$ and $\inf A = 1$*

Definition 23 *Let R be a binary relation. An **equivalence relations** is an order on a non-empty set S such that for $a, b, c \in S$*

- *R is reflexive (aRa)*
- *R is symmetric (aRb implies bRa)*
- *R is transitive (aRb and $bRc \Rightarrow aRc$)*

The definition makes sense because elements of a set with an equivalence relation generate a class within the set of "equal" elements.

Example 24 *For $a, b \in \mathbb{Z}$, R is an equivalence relation if $R \subset \mathbb{Z} \times \mathbb{Z}$ such that $a - b$ is an integral multiple of 5.*

In this case, a is said to be congruent to b modulo 5, written as $a \equiv b \pmod 5$ or $a \equiv_5 b$. Thus, $1 \equiv_5 6 \equiv_5 11$ and are "equal" because they all have a remainder of 1 when divided by 5.

Definition 25 *Let X be a non-empty set. The elements of $K \subseteq \mathcal{P}(X)$ are called **pairwise disjoint** if $\forall A, B \in K$, either $A = B$ or $A \cap B = \varnothing$.*

Definition 26 *A collection Ω of proper subsets of the set X is called a **partition** of X if*

1. *The elements of Ω are pairwise disjoint and*
2. $\bigcup_{A \in \Omega} A = X$

The second bullet points to the fact that the union is taken over the subsets of X. In the exercise, you will be asked to show that this is a second way of thinking of equal objects. The example below should make this clearer.

Example 27 *For the set of rationals $\mathbb{Q} := \mathbb{Z} \times \mathbb{Z}$ such that*

$$x = (a, b) \in \mathbb{Q} \Leftrightarrow a = xb$$

We can get a partition, the elements of which are denoted by

$$[(s, t)] := \{(a, b) | a, b \in \mathbb{Z} \text{ and } bs = at\}$$

when the relation $(a, b) \sim (c, d)$ holds for $ad = bc$. Such a partition is also an equivalence class (proof?).

A set X which can be partitioned under an equivalence class $\sim$ is denoted by $X/\sim$. For the set of rationals, we have $\mathbb{Q}/\sim = \{[\left[\frac{1}{2}\right], \left[\frac{2}{3}\right], \left[\frac{3}{4}\right], ..., [1], [2], ...\}$. This is called the quotient set. This is an important concept since a set endowed with a structure, for example a space or a group, if partitioned under $\sim$, will inherit the structure. The partitioning can be thought of as a division of the set, which suggests the use of the symbol / in a quotient set, and the name. As the example illustrates, the quotient set may be thought of as a set with all the "equivalent" points identified and clumped together.

Definition 28 *Let X and Y be two sets. Then, a **function** f from X to Y is an object such that every $x \in X$ is uniquely associated with an object $f(x) \in Y$.*

This can be made more rigorous by resorting to the definition of relations. If f is a function, then $(x, y), (x, z) \in f$ implies $y = z$. Thus, we can have $f(x) = y$ since this "image" is unique. This is shortened from $xRf(x)$ where $f(x)$ is a unique association of at most one x.

This is represented by $f : X \longrightarrow Y$. The identity map will be represented by $\hat{1}$ throughout the text. This is a map such that $\hat{1}(x) = x$. Confusion with numerical one should not arise since the usage will be clear from context. The notation is intentional in the sense that it acts as the identity for group of functions (see next chapter).

Example 29 *A specific type of function called Boolean function used in Boolean algebra may be defined as* $f : \{0,1\} \to \{0,1\}$. *For* $m \in \mathbb{N}$, *this definition may be extended by using an* m*-tuple input formed by taking the Cartesian product of* $\{0,1\}$ m*-times i.e.* $\{0,1\}^m$

The set of values at which a function is defined is called its domain. The **domain** of f, X, will be reserved by the symbol $\mathcal{D}(f)$. The set of values that the function can produce is called its range or image. Such a set Y is called the **codomain**. The set $f(X) \subseteq Y$ is called the **image** of X under f and will be denoted by $\mathcal{R}(f)$. Thus,

$$\mathcal{R}(f) = f(X) = \{f(x) : x \in X\}$$

This image is a subset of the codomain Y. For $B \subset Y$, the set

$$f^{-1}(B) := \{x \in X : f(x) \in B\}$$

is called the inverse image of f and does not require for the function to have an inverse (see exercise). The **graph** of f is defined as

$$G(f) = \{(x, y) : x \in X \text{ and } y = f(x)\}$$

If you recall your high school mathematics, you had to draw graphs of a function on a piece of graph paper. Thus, this set is actually a more compact way then giving out a pictorial representation of a particular function, hence the name. Caution: both are used interchably in the course.

A function is therefore a many-to-one or sometimes one-to-one relation. f is **one-to-one** if $f(x) = f(y)$ implies $x = y$ for $x, y \in X$. A function is **onto** if $f(X) = Y$. A function is well-defined if $x = y$ implies $f(x) = f(y)$.

For $f : X \longrightarrow Y$, if $A \subset X$ and then the function $g : A \longrightarrow Y$ is the **restriction** of f to A. This is denoted by $g = f|_A$. Dually, f is called the extension of g.

The term "map" is synonymous with function. However, every function is a mapping but not every mapping (or map) is a function since a mapping might also take one element and map it to many others. The notation $\circ$ is reserved for function composition. A function will be called bijective or as having a one-one correspondence if it is one-to-one and onto.

Assuming that the underlying sets have some structure on them, we then use the term **support** of a function supp$f$ to mean the collection of all those values of the domain which are non-zero. That is, supp$f = \{x \in \mathcal{D}(f) : f(x) \neq 0\}$

Definition 30 *A set is* ***countable*** *if it can be placed in a one-to-one correspondence with a subset of the natural numbers, even* $\mathbb{N}$ *itself.*

The indexing set $I_n := \{1, 2, 3, ..., n\}$ will be used for convenience. In cases where an infinite number of items have to be indexed, the set $I = \mathbb{N}$ will be used. This is particularly important for using countable sets.

One-to-one correspondence refers to a bijective function.

The set of natural numbers (trivially), integers and rationals are countable whereas the set of reals are not. For a beautiful proof of the usage of the concept, see Cantor's diagonal argument in any standard text book of Set Theory. In notation, this correspondence is denoted by $\mathbb{N} \sim \mathbb{Z}$, $\mathbb{N} \sim \mathbb{Q}$. We shall make little use of this, however.

Technically, a sequence is any function from the natural numbers to any given set. So, if we have $f(n) = n$, we can have a sequence of natural numbers, or an arithmetic sequence of unit difference. Any such range is called the range of the sequence. As seen, this range can be an integer, a real number, a natural number, a complex number, a function or even a set. It is not yet meaningful to talk about convergence of a sequence since we don't have the idea of distance between two points. This is explored in metric and norm spaces in detail in the upcoming chapters. For now, the ideas of Calculus will suffice i.e. a sequence x_n converges to a point x if $\lim_{n \longrightarrow \infty} x_n = x$. We will call a function continuous if $\lim_{x \longrightarrow x_0} f(x) = x_0$ or equivalently, $\lim_{n \longrightarrow \infty} f(x_n) = x_0$ provided $\lim_{n \longrightarrow \infty} x_n = x_0$, as was probably expounded in your Real Analysis course. Of course there is more rigour later on. We will also do some mathematics with series i.e. partial and infinite sums of elements of a sequence. One such particular example, called the Basel problem, is listed below, since it is made use of in these notes (and it's beautiful)

Example 31 $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$

Since $\sin x = x - x^3/3! + x^5/5! - x^7/7! + ...$ *we have*

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + ...$$

Now, the roots of this polynomial occur at $\pm n\pi$ *for* $n \in \mathbb{N}$ *so that we can factor this polynomial as follows:*

$$\begin{aligned} \frac{\sin x}{x} &= \left(1 - \frac{x}{\pi}\right)\left(1 + \frac{x}{\pi}\right)\left(1 - \frac{x}{2\pi}\right)\left(1 + \frac{x}{2\pi}\right)... \\ &= \left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{(2\pi)^2}\right)\left(1 - \frac{x^2}{(3\pi)^2}\right)... \end{aligned}$$

Now, we can multiply these infinite terms and get coefficients in x^0, x^2, x^4 *and so on. The coefficient for* x^0 *is only* 1*. The coefficient for* x^2 *is*

$$-\frac{1}{\pi^2}\left(\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^3} + ...\right)$$

Comparing this with $-1/3!$*, we have* $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/3!$

We can move ahead and compare the coefficients for x^4 and relate this heavily with the Reimann-Zeta function but this beyond our requirements.

1.2 Ordinals and Cardinals

In informal use, a cardinal number is what is normally referred to as a counting number. Technically, the name given to a size for the set is **cardinality of a set**. Cardinality is defined by constructing a bijective map from the set under observation to another well-known set, usually the natural numbers. Thus, the cardinality of a finite set X is always a natural number n if and only if there exists a bijective mapping from the set X to a finite subset $\{1, 2, ..., n\}$ of the natural numbers $\mathbb{N}$. This is denoted by $|X| = n$.

A set X is countably infinite if and only if there exists a bijective mapping between X and the natural numbers. In fact, we can construct a bijective mapping from $\mathbb{N}$ to $\mathbb{Z}$ and to $\mathbb{Q}$ so that $|\mathbb{Z}| = |\mathbb{Q}|$. This leads to interesting paradoxes, which are well covered in the Hilbert's Paradox of the Grand Hotel. Galileo had come close to the idea but basing his intuition that the sum of the whole is always less than the whole, rejected the fact that $|2\mathbb{N}| = |\mathbb{N}|$. It was Cantor who formalised the idea in his ground-breaking work of Set Theory in 1874–1884. He was the first to denote $|\mathbb{N}| = \aleph_0$ (read aleph-nought).

Cardinals are a generalization of the natural numbers used to measure the cardinality (size) of sets. Thus, 0,1 and 2 are all finite cardinals whereas the first infinite cardinal number is $\aleph_0$. A fundamental theorem due to Georg Cantor shows that it is possible for infinite sets to have different cardinalities and in particular the cardinality of the set of real numbers is greater than the cardinality of the set of natural numbers. The next infinite cardinal is $\aleph_1$ and so on, ordered by the usual "order" (ordinal) numbers. A transfinite sequence, strengthened with regards to order by the Axiom of Choice, is as follows:

$1, 2, 3, ..., \aleph_0, \aleph_1, ...$

Note that in the above definition, only functions are needed without regard to the nature of the elements of the set. In particular, the order of the elements is immaterial.

A non-zero number can be used for two purposes: to describe the size of a set, or to describe the position of an element in a sequence. For finite sets and sequences it is easy to see that these two notions coincide, since for every number describing a position in a sequence we can construct a set which has exactly the right size, e.g. 2 describes the position of b in the sequence $a, b, c, d, ...$ and we can construct the set $\{a, b\}$ which has 2 elements. Notice the place (order) of b. Thus, in the finite case, the ordinals and the cardinals are the same. When dealing with infinite sets it is essential to distinguish between the two — the two notions are in fact different for infinite sets. To motivate the definition of ordinal numbers, we need another definition: The posets $(P, \leq)$ and $(P', \leq')$ are order isomorphic if there is a bijection f such that $f(a) \leq' f(b)$ if and only if $a \leq b$. That is, both f and its inverse must be order preserving. In such a case, the two sets are said to have the same order type. This also happens to be an equivalence relationship. Thus, the set of integers and the set of even integers have the same order type under the bijection $f(n) = 2n$ but the set of integers and the rationals are not because there does not exist any order preserving map between them, even though both have the same cardinality. As remarked, in the

finite case, the distinction between the cardinals are ordinals is blurred: any two finite, well-ordered sets with the same cardinality are order-isomorphic, as can be seen from the above example. We are now in a position to apply a notation:

$\text{ord}(A, \leq) = 0$ if and only if $A = \varnothing$ and $\text{ord}(A, \leq) = n$ if and only if $|A| = n$ where A is a well-ordered set (the empty set is vacuously well-ordered). An ordinal number or just ordinal is, therefore, the order type of a well-ordered set. This is in line with the intuition of order. It is clear that ordinals are different from cardinals and, therefore, serve as an extension of the natural numbers. The least infinite ordinal is ω, which is identified with the cardinal number $\aleph_0$. That is, $\text{ord}(\mathbb{N}, \leq) = \omega$. However in the transfinite case, beyond ω, ordinals draw a finer distinction than cardinals on account of their order information. To each well-ordered set $(A, \leq)$, an ordinal number is assigned denoted by $\text{ord}(A, \leq)$ and if α is an ordinal number, then there is a well-ordered set $(A, \leq)$ such that $\text{ord}(A, \leq) = \alpha$. Also, $\text{ord}(A, \leq) = \text{ord}(B, \leq)$ if and only if A and B are order isomorphic. A given well-ordered set has only one cardinal number but it is possible to obtain a different well-ordering on the same set and, therefore, to yield a distinct ordinal number. Since order-type is an equivalence relationship, the ordinal numbers are taken to be the canonical representatives of their classes and so the order type of a well-ordered set is usually identified with the corresponding ordinal. Given a class of ordinals, one can identify the α-th member of that class, i.e. one can index (count) them.

In summary, an order type categorizes totally ordered sets in the same way that a cardinal number categorizes sets.

1.3 Exercise

1. For any subsets A, B and C of a set X, prove the following:

   (a) $A \cup B = B \cup A$
   (b) $A \cap B = B \cap A$
   (c) $(A \cup B) \cup C = A \cup (B \cup C)$
   (d) $(A \cap B) \cap C = A \cap (B \cap C)$
   (e) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
   (f) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
   (g) $\mathcal{P}(A) \cup \mathcal{P}(B) \subseteq \mathcal{P}(A \cup B)$
   (h) $\mathcal{P}(A) \cap \mathcal{P}(B) = \mathcal{P}(A \cap B)$
   (i) $A - B = A \cap B^{\complement}$
   (j) $A \subseteq B \Leftrightarrow B^{\complement} \subseteq A^{\complement}$
   (k) $(A \cup B)^{\complement} = A^{\complement} \cap B^{\complement}$
   (l) $(A \cap B)^{\complement} = A^{\complement} \cup B^{\complement}$
   (m) $A \cup C = C \Leftrightarrow A \subset C$
   (n) $A \cap C = A \Leftrightarrow A \subset C$

2. For any set X, prove that $\mathcal{P}(X)$ is unique. Do this by showing that there are two such classes and then show that they are the same by employing set-theoretic arguments

3. For any sets A, B prove that $A \times B = B \times A \Leftrightarrow B = A$ hence the use of $A \times A = A^2$ is justified.

4. If A, B and C are non-empty sets, prove that there exists a one-to-one correspondence between

   (a) $A \times B$ and $B \times A$
   (b) $(A \times B) \times C$ and $A \times (B \times C)$
   (c) $(A \times B) \times C$ and the ordered triples (a, b, c) where $a \in A$, $b \in B$ and $c \in C$

5. Prove that an equivalence relation yields a partition of a set X and conversely.

6. Show that the mapping $f : X \longrightarrow Y$ is bijective if and only if there exists a mapping $f : Y \longrightarrow X$ such that $g \circ f = f \circ g = 1$

7. Let Ω be any indexing set. For $f : X \longrightarrow Y$, $A \subset X$ and $B \subset Y$, prove the following: (remember, the following are sets, so you'll have to use set-theoretic arguments)

(a) $f(f^{-1}(B)) \subseteq B$.

(b) $A \subseteq f^{-1}(f(A))$

(c) $f(\bigcup_{i\in\Omega} A_i) = \bigcup_{i\in\Omega} f(A_i)$

(d) $f(\bigcap_{i\in\Omega} A_i) \subseteq \bigcap_{i\in\Omega} f(A_i)$

(e) Equality holds in d if f is one-to-one

(f) $f^{-1}(B^c) = \big(f^{-1}(B)\big)^c$

(g) $f^{-1}(\bigcup_{i\in\Omega} B_i) = \bigcup_{i\in\Omega} f^{-1}(B_i)$

(h) $f^{-1}(\bigcap_{i\in\Omega} B_i) = \bigcap_{i\in\Omega} f^{-1}(B_i)$

8. For $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$, show that

   (a) f and g are one-to-one implies $g \circ f$ is one-to-one

   (b) f and g are onto implies $g \circ f$ is onto

   (c) $g \circ f$ is one-to-one implies f is one-to-one

   (d) $g \circ f$ is onto implies g is onto

9. Show that $f : X \longrightarrow Y$ is one-to-one if and only if $\exists g : Y \longrightarrow X$ such that $g \circ f = \hat{1}_X$

10. Show that $f : X \longrightarrow Y$ is onto if and only if $\exists g : Y \longrightarrow X$ such that $f \circ g = \hat{1}_Y$

11. A permutation on any set X is a mapping $\varsigma : X \longrightarrow X$. Let $\sigma(X)$ be a set of permutations on X. Show that, for $\alpha, \beta, \gamma \in \sigma(X)$,

    (a) $\alpha \circ \beta \in \sigma(X)$

    (b) $\alpha \circ (\beta \circ \gamma) = (\alpha \circ \beta) \circ \gamma$

    (c) $\exists \hat{1} \in \sigma(X)$ such that $\alpha \circ \hat{1} = \hat{1} \circ \alpha = \alpha$

    (d) $\forall \alpha \in \sigma(X)$, $\exists \alpha^{-1} \in \sigma(X)$ such that $\alpha \circ \alpha^{-1} = \alpha^{-1} \circ \alpha = \hat{1}$

Abstract Algebra

This chapter introduces most of the common notions of Algebra. If the reader is already familiar with the topics, a casual reading will suffice.

1.4 Groups

Definition 32 *Let S be a set. Then, a **binary operation** $*$ on S is a function such that $* : S \times S \longrightarrow S$.*

This may be extended to include more than two elements but is beyond the scope of these notes.

Definition 33 *Let G be a set. Then, G, together with a binary operation $*$, written as $(G, *)$, is a **group** if it satisfies the following axioms:-*

- $\forall \ \ a, b, c \in G, (a * b) * c = a * (b * c)$ *i.e. $*$ is associative.*
- $\exists \ \ e \in G$ *such that* $a * e = a, \ \forall \ a \in G$
- $\forall \ a \in G, \ \exists \ b \in G$ *such that* $a * b = e$

Remark 34 *The third axiom is possible only when the uniqueness of e is justified.*

Proof. $e_1 * e_2 = e_1 = e_2$ ■

$(G, *)$ will be shortened to G, where possible. Also, the "$*$" symbol will be skipped and juxtaposition will be used instead. An **Abelian group** is a group in which $*$ is commutative i.e. $\forall a, b \in G, \ a * b = b * a$.

Strictly speaking, it is not generally true that the existence of a right identity and right inverse implies the existence of a left identity and left inverse in a binary relation unless the relation is associative. This explains why most definitions of a group assume that both left and right inverse exist and the left and right identities hold.

Example 35 *One of the most familiar groups is the set of integers $\mathbb{Z}$ which consists of the numbers..., -4, -3, -2, -1, 0, 1, 2, 3, 4, ..., together with addition $+ : \mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}$ where for $a, b \in G$, $+(a, b) = a + b$. The following properties*

of integer addition serve as a model for the abstract group axioms given in the definition below. For any two integers a and b, the sum $a+b$ is also an integer. Thus, adding two integers never yields some other type of number, such as a fraction. This property is known as closure under addition. For all integers a,b and c, $(a+b)+c=a+(b+c)$. Expressed in words, adding a to b first, and then adding the result to c gives the same final result as adding a to the sum of b and c, a property known as associativity. If a is any integer, then $0+a=a+0=a$. Zero is called the identity element of addition because adding it to any integer returns the same integer. For every integer a, there is an integer b such that $a+b=b+a=0$. The integer b is called the inverse element of the integer a and is denoted $-a$.

A group is said to be finite if it has a finite number of elements. Such a group is said to be of finite order. Otherwise, it is of finite order.

Lemma 36 *For a group G and for $g_1, g_2, ..., g_n \in G$,*

$$g_1 g_2 ... g_n = \prod_{i=1}^{n} g_i$$

This lemma might sound trivial but the reader is reminded that such an impression comes from taking for granted that $(g_1 g_2)\, g_3 = g_1\, (g_2 g_3)$. For example, $(5-3)-2 \neq 5-(3-2)$. Of course, in a group, this has to be true. This theorem basically states that the brackets can be removed without any ambiguity. It thus makes sense to define $g^n = gg...g$ n-times. Note that the above is valid for a finite n.

Proof. We use induction to show that for every n any meaningful product $g_1 g_2 ... g_n = \prod_{i=1}^{n} g_i$. Since the axioms of a group dictate that this is true for $n = 1, 2$ and 3 we move on to consider a general n. Let $m < n$. Then, $g_1 g_2 ... g_n = (g_1 g_2 ... g_m)\,(g_{m+1} g_{m+2} ... g_n)$

$$\begin{aligned}
&= \left(\prod_{i=1}^{m} g_i\right)\left(\prod_{i=1}^{n-m} g_{m+i}\right) \\
&= \left(\prod_{i=1}^{m} g_i\right)\left(\left(\prod_{i=1}^{n-m-1} g_{m+i}\right) g_n\right) \\
&= \left(\left(\prod_{i=1}^{m} g_i\right)\left(\prod_{i=1}^{n-m-1} g_{m+i}\right)\right) g_n \\
&= \left(\prod_{i=1}^{n-1} g_i\right) g_n \\
&= \prod_{i=1}^{n} g_i \quad \blacksquare
\end{aligned}$$

Definition 37 *A **subgroup** H of a group $(G, *)$ is a subset of G together with all the axioms of the group G.*

In short, if we restrict the action of the binary operators with the axioms mentioned to a set H, we get a subgroup. Needless to say, the binary operation induced on H is that of G. Mathematically, a subgroup is represented by $H \leq G$. In analogy to set theory, the improper subgroups of a group are $\{e\}$ and G. The identities and inverses of the group and subgroup should necessarily coincide (proof?).

Example 38 *Let $n \in \mathbb{Z}$. Let $n\mathbb{Z} = \{nx \mid x \in \mathbb{Z}\}$. Then $n\mathbb{Z}$ is a subgroup of $\mathbb{Z}$, the group of integers under addition. $n\mathbb{Z}$ consist of all multiples of n. First, we have to show that $n\mathbb{Z}$ is closed under addition. If $nx, ny \in n\mathbb{Z}$, then $nx + ny = n(x + y) \in n\mathbb{Z}$. Therefore, $n\mathbb{Z}$ is closed under addition. Next, the identity element of $\mathbb{Z}$ is 0. Now $0 = n \cdot 0$, so $0 \in n\mathbb{Z}$. Finally, suppose $nx \in \mathbb{Z}$. The additive inverse of nx in $\mathbb{Z}$ is $-nx$, and $-nx = n(-x)$. This is n times something, so it's in $n\mathbb{Z}$. Thus, $n\mathbb{Z}$ is closed under taking inverses. Therefore, $n\mathbb{Z}$ is a subgroup of $\mathbb{Z}$*

In this case, the binary operation of G is carried over. Associateive law, therefore, trivially follows. The identity for a subgroup H is the same as that for the group G (proof?).

One way to check whether we have a bonafide subgroup is by checking that it satisfies the axioms for a group in its own right. Another way is to use the following:

Theorem 39 *H is a subgroup of $G \iff$ for any $a, b \in H$, $ab^{-1} \in H$.*

Proof. We should first prove that H is non-empty. If it is a subgroup, then it will at least contain $\{e\}$. Since H is a subgroup, if $b \in H$, then $b^{-1} \in H$. Also, if $a, b^{-1} \in H$, then $ab^{-1} \in H$. Conversely, suppose that for any $a, b \in H$, $ab^{-1} \in H$. First we prove associativity. For $a, b, c \in H$, we know that $a, b, c \in G$, so associativity is trivially proved for any three elements in H. Since $a, b \in H$ implies $ab^{-1} \in H$, we can reverse the roles of the elements and can conclude that $ba^{-1} \in H$. Since we know that these two elements belong to the group G, from this we have $\left(ab^{-1}\right)\left(ba^{-1}\right) = aea^{-1} \in H$. This step is valid since b is an element of a group and we've already established associativity. Hence, we have $aea^{-1} = aa^{-1} = e \in H$. Finally, $e, a \in H$ implies $ea^{-1} = a^{-1} \in H$. Lastly, $a \in H$ and $b^{-1} \in H$ implies $a\left(b^{-1}\right)^{-1} \in H$ or that $ab \in H$, establishing that H is closed under the binary operation of G. ■

Definition 40 *For any subgroup H, K of G,*

$$HK := \{x \in G \mid x = hk, h \in H, k \in K\}$$

Lemma 41 *For any subgroup H, K of G, HK is a subgroup iff $HK = KH$*

Theorem 42 *For any subgroup H, $HH = H$*

Proof. Since $h = he \in HH$ implies $H \subset HH$

On the other hand, $HH \ni h_1h_2 = h \in H$ ■

1.4.1 Words on the cancellation law

In a group $(G, *)$, the following are called the cancellation laws:

1. if $ab = ac$ then $b = c$ **(left cancellation law)**
2. if $ba = ca$ then $b = c$ **(right cancellation law)**

Proof. We only prove the left cancellation law. The second proof is similar. Let $ab = ac$

$\implies a^{-1}(ab) = a^{-1}(ac)$

$\implies (a^{-1}a)b = (a^{-1}a)c$

$\implies eb = ec$

$\implies b = c$ ■

This property makes sense because the inverse of a, if it exists, is being "multiplied" on both sides on the left. In simple cases of the integers, rationals, reals and complex numbers, this is very clear and usually applied without the need for justification as a part of school training. Some situations can be constructed in which there is no clear answer. For instance, if $\mathbf{a} \times \mathbf{b} = \mathbf{a} \times \mathbf{c}$, it does not at all necessarily follow that $\mathbf{b} = \mathbf{c}$ for vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ with the usual cross product definition.

The distinction between left and right cancellation is important in non-commutative algebra. For instance, for matrices A, B, C, $AB = CA$ does again not necessarily mean that $B = C$, even if $\det(A) \neq 0$

The cancellation law is one simple way to state divisibility and invertibility combined. If inverses exists, then clearly, the cancellation property holds. In fact, every group is therefore a semi-group (set with binary operation and associative law) in which the cancellation law holds. In the finite case, the existence of inverses and the cancellation law coincide but in the infinite case, this is not so. Can you come up with an example?

1.4.2 Quotient Groups

Let R be an equivalence relation on a set X. Then, we can have for ourselves a quotient set X/R. In a similar vein, we can have for ourselves a quotient group G/R but for that, we need to be able to define an equivalence relation in light of the idea of a binary relation.

Definition 43 *Suppose H is a subgroup of a group G. A **left (right) coset** if the set $aH = \{ah \mid h \in H\}$ (Ha).*

Such a set is simply called a coset if $aH = Ha$. Note that this implies $ah_1 = h_2a$ where h_1 is not necessarily equal to h_2. It will be shown below that H partitions G into right cosets. It also partitions G into left cosets, and in general these partitions are distinct.

Definition 44 *Let G be a group, H be a subgroup of G. For $a, b \in G$, we say that $a \equiv b \bmod H$ iff $ab^{-1} \in H$.*

This relation is an equivalence relation.

Proof. Clearly, $e = aa^{-1} \in H$ so that $a \equiv a \bmod H$

Second, $ab^{-1} \in H$ implies $ba^{-1} = \left(ab^{-1}\right)^{-1} \in H$ so that $a \equiv b \bmod H$ implies $b \equiv a \bmod H$

Finally, if $a \equiv b \bmod H$ and $b \equiv c \bmod H$, then $ab^{-1} \in H$ and $bc^{-1} \in H$ so that $ab^{-1}bc^{-1} = ac^{-1} \in H$, satisfying transitivity. ■

Notice that $a \equiv b \bmod H$ is defined for the subgroup relation.

We can have that $Ha = H \;\; iff \;\; a \in H$.

Proof. $Ha = H$

$\implies h_1 a = h_2$

$\implies a = h_1^{-1} h_2 \in H$

Conversely, if $a \in H$

then $ha \in Ha$ for some h

In particular, for $h = e$, we have $ea = a \in Ha$

Thus, $H \subseteq Ha$

For the other containment, we have $ha \in Ha$

Take $h = a$, then $a^2 \in H$

$\implies Ha \subseteq H$ ■

Theorem 45 *Suppose H is a subgroup of a multiplicative group G. If $a \in G$, define the right coset containing a to be $Ha = \{ha : h \in H\}$, then the following hold:*

1. If $b \in Ha$, then $Hb = Ha$ i.e. if $b \in H$, then $H(ba) = (Hb)a = Ha$.
2. If $Hc \cap Ha \neq \varnothing$ then $Hc = Ha$.
3. Elements a and b belong to the same right coset iff $ab^{-1} \in H$ or if $ba^{-1} \in H$
4. The right cosets form a partition of G, i.e., each a in G belongs to one and only one right coset.

Proof.

1. $b \in Ha$

   $\implies b = h_1 a$

   $\implies h_2 b = h_2 h_1 a$

   $\implies Hb = Ha$

2. Let $b \in Hc \cap Ha$

   Since $Hc \cap Ha \neq \varnothing$

   $\implies b \in Hc$ and $b \in Ha$

   $\implies Hb = Hc$ and $Hb = Ha$ from 1.

   $\implies Hc = Ha$

3. Since $b \in Ha$ therefore

   $Ha = Hb$

   $\Longrightarrow b \in Hb$

   $\Longrightarrow ba^{-1} \in (Ha)a^{-1}$

   $\Longrightarrow ba^{-1} \in H(aa^{-1})$

   $\Longrightarrow ba^{-1} \in (He)$

   $\Longrightarrow ba^{-1} \in H$

   Conversly, $ba^{-1} \in H$

   $\Longrightarrow Hba^{-1} = H$

   $\Longrightarrow Hba^{-1}a = Ha$

   $\Longrightarrow Hbe = Ha$

   $\Longrightarrow Hb = Ha$

   Similarly we can prove that $aH = bH \Leftrightarrow ab^{-1} \in H$

   Since H is a subgroup. Therefore if $ba^{-1} \in H$

   $(ba^{-1})^{-1} = (a^{-1})^{-1}b^{-1} = ab^{-1} \in H$

4. Suppose H is a subgroup of G. H itself is a right coset. If there is an element $a \in G$ such that $a \notin H$, then Ha will be another distinct right coset. Again if there is another element $b \in G$ such that $b \notin H$ and so $b \notin Ha$ then Hb will be another distinct right coset. Proceeding in this way we can get all distinct right cosets of H in G.

Then we shall have

$$G = H \cup Ha \cup Hb \cup Hc \cup ...$$

where $a, b, c, ...$ are elements of G so chosen that all right cosets are distinct.

No right coset of H in G is empty since $e \in H$. Any two right cosets of H in G are either disjoint or identical. The union of all right cosets of H in G equal to G. Therefore the union of all right cosets of H in G gives us a partition of G. We can therefore have $Ha = \{x \in G \mid a \equiv x \bmod H\} := [a]$. ■

Lemma 46 *Let Ha, Hb be any two right cosets. Then, $f : Ha \longrightarrow Hb$ such that $f(ha) = hb$ is one-to-one*

Proof. $f(h_1a) = f(h_2a)$

$\Longrightarrow h_1b = h_2b$

$\Longrightarrow h_1 = h_2$

$\Longrightarrow h_1a = h_2a$

From any hb, we can construct $ha = f(ha)\, b^{-1}a$, implying surjectivity. ■

Thus, in the finite case, any two right cosets of the same subgroup H have the same number of elements. In the commutative case, the left and right cosets agree. Thus, we have proved the famous

Theorem 47 (Lagrange's theorem) *The order of a subgroup divides the order of the group*

If the group itself is not commutative but if every left coset of subgroup N of G is a right coset, then such a subgroup is called **normal**. This will be denoted by $N \trianglelefteq G$. If N is a proper subgroup, then $N \triangleleft G$. For any group G, G and e are normal subgroups.

Theorem 48 *If H is a subgroup of a group G, then the following are equivalent.*

1. *If $a \in G$, then $aHa^{-1} = H$ Normal Test Condition (NTC)*
2. *If $a \in G$, then $aHa^{-1} \subset H$*
3. *If $a \in G$, then $aH = Ha$*
4. *Every right coset is a left coset, i.e., if $a \in G$, $\exists\ b \in G$ with $Ha = bH$.*

Corollary 49 *Subgroups of an Abelian group are normal*

Proof. For Abelian groups, $Ha = aH$ ■

If H satisfies any of the four conditions above, then H is said to be a normal subgroup of G.

Example 50 *Let G be a group of all 2×2 non-singular matrices. Let S be the set of nonsingular matrices of the form*

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

Notice that a cannot be zero, for otherwise the determinant would be zero and the inverse would not exist. Let's find out if S can be a normal subgroup of G. First we need to find out that S is a subgroup.

1. ***Closure:*** *Let $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \in S$ and $\begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix} \in S$*
   $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}\begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix} = \begin{bmatrix} ab & 0 \\ 0 & ab \end{bmatrix}$ *since a and b cannot be zero, $ab \neq 0$, and this result is clearly in S.*
2. ***Inverse:*** *Let $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \in S$. Then, $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{a} \end{bmatrix}$ which is clearly in S. This implies that S is a subgroup of G.*
3. ***Normal Test Condition (NTC):*** *Let $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \in S$, Let $\begin{bmatrix} x & y \\ z & w \end{bmatrix}$ be an arbitrary nonsingular 2×2 matrix. Let $D = xw - zy$ be the determinant of $\begin{bmatrix} x & y \\ z & w \end{bmatrix}$ and $D \neq 0$. Consider $\begin{bmatrix} x & y \\ z & w \end{bmatrix}\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}\begin{bmatrix} x & y \\ z & w \end{bmatrix}^{-1}$*

$$= \begin{bmatrix} xa & ya \\ za & wa \end{bmatrix} \begin{bmatrix} \frac{W}{D} & -\frac{y}{D} \\ -\frac{z}{D} & \frac{x}{D} \end{bmatrix} = \begin{bmatrix} \frac{xaw-yaz}{D} & \frac{-xay+xay}{D} \\ \frac{zaw-zaw}{D} & \frac{-zay+wax}{D} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{aD}{D} & 0 \\ -0 & \frac{aD}{D} \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \in S$$
$\Longrightarrow S$ *is a Normal subgroup.*

Lemma 51 *A subgroup of N in G is normal iff the product of two right cosets of N in G is again a right coset of N in G.*

Proof. ($\Longrightarrow$)
$NaNb = N(aN)b = NNab = Nab = Nc$
($\Longleftarrow$) $NaNb = Nc$, then
$n_1an_2b = n_3ab$ where $c = ab$
$\Longrightarrow an_2 = n_1^{-1}n_3abb^{-1}$
$\Longrightarrow an_2 = na$
$\Longrightarrow aN = Na$ ■

Note that in this proof, it is inherently stated that the operation

$$NaNb = Nab \tag{1.1}$$

is well-defined because the binary operation itself is well-defined.

Suppose N is a normal subgroup of G, and C and D are cosets. We wish to define a coset E which is the product of C and D. If $c \in C$ and $d \in D$, define E to be the coset containing cd, i.e. $E = N(cd)$. The coset E does not depend upon the choice of c and d since there are other elements belonging to the respective equivalence class. Such a collection forms a group, denoted by G/N.

Proof. Suppose G is a group, N is a normal subgroup, and G/N is the collection of all cosets. Then $(Na)(Nb) = N(ab)$ is a well-defined multiplication (binary operation) on G/N as shown above, and with this multiplication, G/N is a group. Its identity is N and $(Na)^{-1} = (Na^{-1})$. Multiplication of elements in $G = N$ is multiplication of subsets in G.

$(Na)(Nb) = N(aN)b = N(Na)b = N(a \cdot b)$.

Once multiplication is well defined, the group axioms are immediate. ■

In the left side of the equality of **1.1**, we use the binary operation on the set G/N whereas on the right side of the equality, we make use of the binary operation on the set G.

If G is finite, then Lagrange's theorem tells us that $|G/N| = |G| \,/\, |N|$

Example 52 *For example, consider the group with addition modulo 6:*

$$G = \mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$$

Let $N = \{0, 3\}$. *Then,* $G/N = \{aN \mid a \in G\}$
$= \{a +_6 \{0, 3\} \mid a \in \{0, 1, 2, 3, 4, 5\}\}$
$= \{0 +_6 \{0, 3\}, 1 +_6 \{0, 3\}, 2 +_6 \{0, 3\}, 3 +_6 \{0, 3\}, 4 +_6 \{0, 3\}, 5 +_6 \{0, 3\}\}$
$= \{\{0, 3\}, \{1, 4\}, \{2, 5\}, \{3, 0\}, \{4, 1\}, \{5, 2\}\}$
$= \{\{0, 3\}, \{1, 4\}, \{2, 5\}\}$

1.4.3 Homomorphisms of Groups

Homomorphisms are functions between groups that respect the group operations. It follows that they honor identities and inverses.

Definition 53 *If $(G, \cdot)$ and $(G', \cdot')$ are groups, a function $f : G \longrightarrow G'$ is a **homomorphism** if, for all $a, b \in G$, $f(a \cdot b) = f(a) \cdot' f(b)$.*

In the case $G = G'$, f is called an **automorphism**. The kernel of f is defined by $\ker(f) = f^{-1}(e') = \{a \in G : f(a) = e'\}$. In other words, the kernel is the set of solutions to the equation $f(x) = e'$. Let a function $f : G \longrightarrow H$ be a homomorphism. If f is also a one-one correspondence, then f is called an **isomorphism**. Two groups G and H are called isomorphic, denoted by $G \cong H$, if there exists an isomorphism between them. A group isomorphism is, therefore, a function between two groups that sets up a one-to-one correspondence between the elements of the groups in a way that respects the given group operations. If there exists an isomorphism between two groups, then the groups are called isomorphic. From the standpoint of group theory, isomorphic groups have the same properties and need not be distinguished.

Lattices are isomorphic when orders between elements are preserved. Graphs are isomorphic when they both have a similar structure, with vertices and nodes being the same. Metric spaces are isomorphic when distance is preserved. Topological spaces are isomorphic when arbitrarily small distances are preserved between the two topological spaces. In short, two mathematical structures are isomorphic or the same if their underlying structures are similar, with a disregard to nature of the elements.

The definition introduced here is only to make the reader familiar with the rigour of this central concept. Group isomorphism is hardly used afterwards. We shall make use of the concept of isomorphism for metric spaces, vector spaces, norm spaces and inner product spaces. However, since any vector space is also a group in its own right, it is for the benefit of the reader that the following be read carefully enough.

Example 54 *The constant map $f : G \longrightarrow G'$ defined by $f(a) = e'$ is a homomorphism.*

Example 55 *If H is a subgroup of G, the inclusion $i : H \hookrightarrow G$ is a homomorphism.*

Example 56 *The function $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ defined by $f(t) = 2t$ is a homomorphism of additive groups, while the function defined by $f(t) = t + 2$ is not a homomorphism.*

Example 57 *The function $h : \mathbb{Z} \longrightarrow \mathbb{R}\setminus\{0\}$ defined by $h(t) = 2^t$ is a homomorphism from an additive group to a multiplicative group.*

Example 58 *To show that $(\mathbb{R}, +) \cong (\mathbb{R}_{>0}, \times)$. let $f(x) = e^x$. To prove that this is an isomorphism, we should check tha $f : \mathbb{R}^+ \longrightarrow \mathbb{R}^\times_{>0}$ is one-one*

correspondence and that $f(x+y) = f(x)f(y)$ *for all* $x, y \in \mathbb{R}$. *The first part is trivial, since* $f(x) = e^x$ *is defined for all* $x \in \mathbb{R}$ *and its inverse* $g(x) = \ln x$ *is also defined for all* $x \in \mathbb{R}_{>0}$. *The second part is also true, since* $f(x+y) = e^{x+y} = e^x e^y = f(x)f(y)$.

Example 59 *The group* $\mathbb{Z}$ *of integers (with addition) is a subgroup of* $\mathbb{R}$, *and the factor group* $\mathbb{R}/\mathbb{Z}$ *is isomorphic to the group* S^1 *of complex numbers of absolute value* 1 *(with multiplication). That is,* $\mathbb{R}/\mathbb{Z} \cong S^1$. *An isomorphism is given by* $f(x+\mathbb{Z}) = e^{2\pi x i}$ *for every* x *in* $\mathbb{R}$

Example 60 *Let* G *be an infinite cyclic group. Then* G *is isomorphic to the additive group of integers:* $G \cong (\mathbb{Z}, +)$. *From infinite cyclic group, we have* $G = \langle a \rangle = \{a^k : k \in \mathbb{Z}\}$. *Let us define* $\varphi : \mathbb{Z} \longrightarrow G$ *such that* $\varphi(k) = a^k$. *First we show that* φ *is a homomorphism. Let* $k, l \in \mathbb{Z}$.. *Then,* $\varphi(k+l) = a^{k+l} = a^k a^l = \varphi(k)\varphi(l)$. *Now we show that* φ *is a surjection. As* G *is cyclic, every element of* G *is a power of* a *(for some* $a \in G$ *such that* $G = \langle a \rangle$*). Thus,* $\forall x \in G : \exists\ k \in \mathbb{Z} : x = a^k$. *By the definition of* φ, $\varphi(k) = a^k = x$. *Thus* φ *is surjective. As* G *is cyclic, every element of* G *is a power of* a *(for some* $a \in G$ *such that* $G = \langle a \rangle$*). Thus,* $\forall x \in G : \exists k \in \mathbb{Z} : x = a^k$. *By the definition of* φ, $\varphi(k) = a^k = x$. *Thus* Φ *is surjective. Now we show that* φ *is an injection. This follows directly from Powers of Infinite Order Element proved above, where* $\forall m, n \in \mathbb{Z} : m \neq n \Longrightarrow a^m \neq a^n$ *Thus* φ *is an injective, surjective homomorphism, thus* $G \cong (\mathbb{Z}, +)$ *as required.*

Suppose G and G' are groups and $f : G \longrightarrow G$ is a homomorphism. Then, the following properties can be proved and are left to the reader as an exercise.

1. $f(e) = e'$.
2. $f(a^{-1}) = f(a)^{-1}$ where the first inverse is in G, and the second is in G'.
3. $\ker f$ is a Normal subgroup.
4. f is injective $\Leftrightarrow \ker(f) = \{e\}$.
5. If H is a subgroup of G, $f(H)$ is a subgroup of G' In particular, image (f) is a subgroup of G'.
6. If H' is a subgroup of G', $f^{-1}(H')$ is a subgroup of G, Furthermore, if H' is normal in G', then $f^{-1}(H')$ is normal in G..
7. The composition of homomorphisms is a homomorphism, i.e., if

   $$h : G' \longrightarrow G"$$

   is a homomorphism, then $h \circ f : G \longrightarrow G''$ is a homomorphism.
8. If $f : G \longrightarrow G'$ is a bijection, then the function $f^{-1} : G' \longrightarrow G$ is an isomorphism.

Isomorphisms preserve all algebraic properties. For example, if f is an isomorphism and $H \subset G$ is a subset, then H is a subgroup of G iff $f(H)$ is a subgroup of G'. H is normal in G iff $f(H)$ is normal in G', G is cyclic iff G' is cyclic. Orders of elements are also preserved.

9. Suppose H is a normal subgroup of G. If $H \subset \ker f$, then $f' : G/H \longrightarrow G'$ defined by $f'(Ha) = f(a)$ is a well-defined homomorphism.

10. Suppose K is a group. Then K is an infinite cycle group iff K is isomorphic to the integers under addition, i.e., $K \cong \mathbb{Z}$. K is a cyclic group of order n iff $K \cong \mathbb{Z}_n$.

1.5 Rings

Definition 61 *Let $R \neq \varnothing$ with two binary operations $+ : R \times R \longrightarrow R$ and $. : R \times R \longrightarrow R$. Then, R is called a **ring** if*

1. *$(R, +)$ is an Abelian group*
2. *$.$ is associative*
3. *$\forall a, b, c \in R$, $a.(b + c) = a.b + a.c$ and $(a + b).c = a.c + b.c$*

A subset S of R is called a subring if it obeys the above properties.

Lemma 62 *Let R be a ring. Then*

1. *$--a = a$*
2. *$0a = a0 = 0$ for all $a \in R$.*
3. *$(-a)b = a(-b) = -(ab)$ for all $a, b \in R$*
4. *$(-a)(-b) = ab$ for all $a, b \in R$.*
5. *if R has an identity 1, then the identity is unique and $-a = (-1)a$.*

Proof. 1) Since $-a + a = 0$, therefore a is the inverse of $-a$ or $a = --a$
2) $0a = (0 + 0)\, a = 0a + 0a$
$\implies 0a - 0a = 0a$
$\implies 0 = 0a$
3) $0 = 0b = (a - a)\, b = ab + (-a)\, b$. That is, $-(ab) = (-a)\, b$
Similary, $0 = a0 = a\,(b - b) = ab + a\,(-b)$. That is, $-(ab) = a\,(-b)$
4)$(-a)\,(-b) = a\,(--b) = ab$ from 3 and 1
5) Assume there are two identities 1 and $1\prime$. Then, $1 = 11\prime = 1\prime$.
Put $b = -1$ in 4 to get $(-a)\,(--1) = (-a)\,1$ which equals $a\,(-1)$ by 3 ■

Remark 63 *R has identity element if there exists an element $1_R \in R$ such that $a.1_R = 1_R.a = a$ $\forall a \in R$. From now on, all rings will be assumed to have the identity element whereas the $+$ binary operation will be used*

Remark 64 *$2\mathbb{Z}$ is a "ring" without the identity. Thus, a "ring" may not necessarily have an identity. If R does not have the identity element, the jocular term "rng" will be used.*

Remark 65 *Henceforth, the . will be dropped and a juxtaposition will be used instead.*

Remark 66 *R itself is commutative if $ab = ba \ \forall a, b \in R$*

Remark 67 (Warning) *From now on, reference will not be given as to whether or not a structure is commutative or not.*

Definition 68 *An element a of a ring R is called a **left zero divisor** if there exists a nonzero x such that $ax = 0$.*

Similarly, for the right zero divisor. In commutative algebra, this distinction is blurred.

Lemma 69 *Let R be a ring and let $\psi : R \longrightarrow R$ such that $\psi(x) = ax$ for some $a \in R$. Then, a is a zero divisor if and only if ψ is not injective.*

Proof. ($\Longrightarrow$) suppose a is a zero divisor and $\psi(x)$ is injective. Then,

$$\psi(x) = \psi(y)$$

or $ax = ay$ implies $x = y$. That is, the cancellation laws hold, implying the existence of the inverse of a, a^{-1}. But if $ax = 0$, then we can apply a^{-1} on both sides to get $x = 0$, a contradiction.

($\Longleftarrow$) Suppose ψ is not injective but that a is also not a zero divisor. Then, $\psi(x) = \psi(y)$ or $ax = ay$ does not imply $x \neq y$. But if $ax = ay$, then $ax - ay = 0$ $\Longrightarrow a(x - y) = 0$. Since a is not a zero divisor, $(x - y) = 0$ from which we have the contradiction that $x = y$ ■

Corollary 70 *Let R be a ring and let $\psi : R \longrightarrow R$ such that $\psi(x) = ax$ for some $a \in R$. Then, a is a not a zero divisor if and only if ψ is injective.*

Invertible elements will be called **units**. Note that it is strictly not true that if an element is not a zero divisor, then it is necessarily a unit. In fact, it will be seen later, a nonzero divisor has an inverse but only in a larger ring.

Lemma 71 *Set of units of R, $U(R)$ form a group under multiplication*

Proof. Let $a, b \in U(R)$. Then, b^{-1} and a^{-1} exist. Since b^{-1} and a^{-1} are members of the larger ring, therefore the product $b^{-1}a^{-1}$ exists. Hence,

$$ab\left(b^{-1}a^{-1}\right) = 1$$

implying that the inverse of ab exists or that $ab \in U(R)$. Associativity follows from the structure of the ring. The multiplicative identity is the inverse of itself and hence trivially a unit. Suppose that this is not true and that $1^{-1}1 = 1$. But then multiplying by 1^{-1} on both sides and we get $1^{-1} = 1$ implying that $1 \in U(R)$. Finally, for any $a \in U(R)$, a^{-1} exists and that $a = \left(a^{-1}\right)^{-1}$ hence $a^{-1} \in U(R)$ ■

Definition 72 *An element x of a ring R is called* ***nilpotent*** *if there exists some positive integer n such that $x^n = 0$. If there is a smallest positive integer n such that $n.a = 0$ for all $a \in R$, then such a positive integer is called the* ***characteristic*** *of R.*

This is denoted as $\text{char}R = n$. If there is no such positive integer, then R is said to be of characteristic zero.

The matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ is a nilpotent element of order 3. In the factor ring $\mathbb{Z}/9\mathbb{Z}$, the equivalence class of 3 is nilpotent because 3^2 is congruent to 0 modulo 9.

Lemma 73 *A nilpotent element a of a nonzero ring is always a two-sided zero divisor. In particular, an idempotent element (a nilpotent element of order $n = 2$) is always a two-sided zero divisor.*

Proof. *$a^n = 0 \implies aa^{n-1} = 0 \implies ax = 0$ for $x = a^{n-1}$.*

Similarly, $xa = 0$ ■

Definition 74 *A commutative ring with no zero divisor is called an* ***integral domain****.*

Example 75 *$\mathbb{Z}_c$ for any $c \in \mathbb{Z} - \mathbb{P}$ is not an integral domain.*

Example 76 $(n\mathbb{Z}, +, .)$

An element of a ring that is not a zero divisor is called regular, or a non-zero-divisor. A zero divisor that is nonzero is called a nonzero zero divisor or a nontrivial zero divisor

Lemma 77 *R is a division ring $\iff U(R) = R \backslash \{0\}$*

Definition 78 *A* ***Boolean ring*** *B is a ring with identity in which $x^2 = x$ for all $x \in B$*

Theorem 79 *A Boolean ring B is commutative and has a characteristic of 2*

Solution 80 $2x = x + x = (x + x)^2 = 2x^2 + 2x^2 = 2x + 2x$

Hence, $2x = 0 \ \forall x \in B \backslash \{0\}$

Therefore, $char(B) = 2$

$x + y = (x + y)^2 = x^2 + y^2 + xy + yx = x + y + xy + yx$

Hence, $xy + yx = 0$. Since $2xy = 0$,

or, $xy = xy \implies xy + 0 = xy \implies xy + xy + yx = xy \implies 2xy + yx = xy$ or $xy = yx$

1.5.1 Ideals

Definition 81 *Let R be a ring and I be a non-empty subset of R. Then, I is said to be a* ***left ideal*** *(right ideal) if I is a subgroup under $+$ and if $\forall r \in R$ and $\forall x \in I$, $rx \in I$ ($xr \in I$)*

I is two-sided ideal if it is both left and right ideal. In such a case, we will simply refer to the resulting structure as an ideal. In the commutative case, of course this distinction is blurred. $I \neq \{0\}$ is a proper ideal if it is a proper subset of R, that is, I does not equal R. Therefore, to prove a subset I of a ring R is an ideal, it is necessary to prove that I is nonempty, closed under subtraction and closed under multiplication by all the elements of R. This naturally also includes the elements of I.

Trivially, $\{0\}$ and R are ideals. In the case of integers, $2\mathbb{Z}$ is an ideal because addition and subtraction of even numbers preserves evenness, and multiplying an even number by any other integer results in another even number. Similarly, the set of all integers divisible by a fixed integer n is an ideal denoted $n\mathbb{Z}$. The set of all $n \times n$ matrices whose last row is zero forms a right ideal in the ring of all $n \times n$ matrices. It is not a left ideal. The set of all $n \times n$ matrices whose last column is zero forms a left ideal but not a right ideal.

1.5.2 Quotient Ring

Similar to quotient groups, quotient rings can be constructed. One starts with a ring R and a two-sided ideal I in R, and constructs a new ring, the quotient ring R/I, whose elements are the cosets of I in R.

Definition 82 *Let R be a ring, I be an ideal of R. For $a, b \in I$, we say that $a \equiv b \bmod I$ if $a - b \in I$.*

This relation is an equivalence relation.

Proof. Clearly, $0 = a - a \in I$ so that $a \equiv a \bmod I$

Second, $a - b \in I$ implies $b - a = -(a - b) \in I$ so that $a \equiv b \bmod I$ implies $b \equiv a \bmod I$

Finally, if $a \equiv b \bmod I$ and $b \equiv c \bmod I$, then $a - b \in I$ and $b - c \in I$ so that $a - b - (b - c) = a - c \in I$, satisfying transitivity. ■

Corollary 83 $I + a = \{x + a \mid x \in I\} = \{x \in R \mid a \equiv x \bmod I\} = [a]$

Proof. Let $y \in I + a$. Then, $y = y_0 + a$ for some $y_0 \in I \implies y_0 = y - a \in I$
$\implies a \equiv y \bmod I$
$\implies y \in [a]$
Conversely, $y \in [a]$
$y - a \in I$
$\implies x = y - a \in I$
$\implies y = x + a \in I + a$ ■

This is called the residue class of a modulo I. Thus, any two such classes $I + a$ and $I + b$ of R are identical or have no two elements in common.

Lemma 84 *Let $I+a$, $I+b$ be any two modulo classes. Then, $f : I+a \longrightarrow I+b$ such that $f(r+a) = r+b$ is one-to-one*

Proof. $f(r_1+a) = f(r_2+a)$
$\implies r_1+b = r_2+b$
$\implies r_1 = r_2$
$\implies r_1+a = r_2+a$
From any $r+b$, we can construct $r+a = f(r+a) - b + a$ ■

Thus, in the finite case, any two right residue classes of the same ideal I have the same number of elements.

From hereon, we will assume that R is abelian.

Definition 85 *$R/I := \{r+I \mid r \in R\}$ is called the **quotient ring***

Now we define two operations for R/I, the addition and multiplication

$$\cdot : R/I \times R/I \longrightarrow R/I$$

such that $(r+I)\cdot(s+I) = rs+I$ and

$$+ : R/I \times R/I \longrightarrow R/I$$

such that $(r+I)+(s+I) = (r+s)+I$. We now show that this is well-defined.

Since I is a commutative subgroup under addition, it is a normal subgroup hence the need to justify the plus operation can be foregone. Let $(r+I)+(s+I) = (r+s)+I$ is well-defined, as was proved for normal subgroups above. For multiplication, let $(r_1+I, s_1+I) = (r_2+I, s_2+I)$. Then, $r_1+I = r_2+I$ and $s_1+I = s_2+I$. Thus, $r_1 \in r_2+I$, $r_2 \in r_1+I$, $s_2 \in s_1+I$ and $s_1 \in s_2+I$
$\implies r_1 - r_2 \in I$ and $s_1 - s_2 \in I$. Let $r_1 - r_2 = a$ and $s_1 - s_2 = b$. Then, $r_1 s_1 = (a+r_2)(b+s_2) = r_2 s_2 + b s_2 + ab + a s_2 \in r_2 s_2 + I$
$\implies r_1 s_1 + I = r_2 s_2 + I$.

Another way to see this is as follows: If $I+a = I+c$ and $I+b = I+d$, what must be true about I so that we can be sure $I+(ab) = I+(cd)$? Using the definition of cosets: If $a-c, b-d \in I$, what must be true about I to assure that $ab-cd \in I$? We need to have the following element always end up in I: $ab-cd = ab-ad+ad-cd = a(b-d)+(a-c)d \in I \implies ab+I = cd+I$. Because $b-d$ and $a-c$ can be any elements of I (and either one may be 0), and a, d can be any elements of R, the property required to assure that this element is in I, and hence that this multiplication of cosets is well-defined, is that, for all s in I and r in R, sr and rs are also in I. If this condition holds, then we don't need to assume separately that the product of two elements of I is again in I.

1.6 Fields

Definition 86 *Let R be a ring. Then, R is called a **divison ring** or a **skew field** if $\forall a, b \in R \setminus \{0\}$, $\exists b$ such that $ab = ba = 1_R$*

In other words, a divison ring is a ring in which division is possible. The set of rational numbers, real numbers, complex numbers and even the modulo set $(\mathbb{Z}_p, +, .)$ are division rings. The real quaterions $(\mathbb{H}, +, .)$ are formed in the following way: start with a vector space over $\mathbb{R}$ (this could be the rationals, too) with basis vectors $1, i, j, k$. In other words, an element of $\mathbb{H}$ looks like $a_1 + a_2 i + a_3 j + a_4 k$. Products are determined by the relationships that $i^2 = j^2 = k^2 = -1$, $ij = k$, $jk = i$, and $ki = j$. Therefore, the product of two elements is $(a_1 + a_2 i + a_3 j + a_4 k)(b_1 + b_2 i + b_3 j + b_4 k)$

$= (a_1 b_1 - a_2 b_2 - a_3 b_3 - a_4 b_4) + (a_1 b_2 + a_2 b_1 + a_3 b_4 - a_4 b_3) i$

$+(a_1 b_3 + a_3 b_1 + a_4 b_2 - a_2 b_4) j + (a_1 b_4 + a_4 b_1 + a_2 b_3 - a_3 b_2) k.$

Choosing $a_1 = 1$, $a_2 = 0$, $a_3 = 0$,and $a_4 = 0$, we have ourselves an identity. It is not commutative since $ij = k$, but

$$ji = ji(j)(-j) = -j(ij)j = -jkj = -ij = -k$$

Elements have inverses, and the inverse of is

$$\frac{a_1}{a_1^2 + a_2^2 + a_3^2 + a_4^2} - \frac{a_2}{a_1^2 + a_2^2 + a_3^2 + a_4^2} i - \frac{a_3}{a_1^2 + a_2^2 + a_3^2 + a_4^2} j - \frac{a_3}{a_1^2 + a_2^2 + a_3^2 + a_4^2} k$$

Finally, there are no zero divisors since every non-zero element has an inverse. Therefore, $(\mathbb{H}, +, .)$ is a division ring which is not commutative.

In summary, we have

Definition 87 *Let $\mathbb{F}$ be a set. $\mathbb{F}$ is **field** with two binary operations $+$ and $\cdot$, denoted by $(\mathbb{F}, +, \cdot)$, satisfying the following axioms:-*

- *$(\mathbb{F}, +)$ is an abelian group*
- *$(\mathbb{F}^*, \cdot)$ is also an abelian group.*
- *$\forall \ \ a, b, c \in \mathbb{F}$, the left distributive law $a \cdot (b + c) = a \cdot b + a \cdot c$ and the right distributive law $(b + c) \cdot a = b \cdot a + c \cdot a$ hold*

Here, $\mathbb{F}^* = \mathbb{F} - \{0\}$. That is, inverse for the additive identity does not exist.

In simple words, $1/0$ is undefined.

This definition should not be surprising, given what we just covered. *This* definition needs to be remembered from here on.

$\mathbf{x} \in \mathbb{F}^n$ is a vector defined by an n-ordered pair for $n \in \mathbb{N}$. The same rules as for the field are applied to each jth-tuple for $1 \leq j \leq n$. This is one way of obtaining for ourselves a vector space, as we will see in the next chapter.

Example 88 ***Rational Numbers:*** *A simple example of a field is the field of rational numbers, consisting of numbers which can be written as fractions a/b, where a and b are integers, and $b \neq 0$. The additive inverse of such a fraction is simply $-a/b$, and the multiplicative inverse (provided that $a \neq 0$) is b/a. For any two rational numbers $a/b, c/d \epsilon \mathbb{Q}$, the sum as well as the product of a/b and c/d is again a rational number. Associativity holds for rational numbers as*

well as commutativity with respect to addition and multiplication. "0" is called the additive identity and "1" is called the multiplicative identity for the set of rational numbers, that is, $0 + a/b = a/b + 0 = a/b$ *and* $1 \times a/b = a/b \times 1 = a/b$. *Distribution law of multiplication over addition also holds for the set of rational numbers i.e. for all* a/b, c/d *and* $e/f \in \mathbb{Q}$, *we have*

$$\frac{a}{b} \cdot \left(\frac{c}{d} + \frac{e}{f} \right) = \left(\frac{a}{b} \cdot \frac{c}{d} \right) + \left(\frac{a}{b} \cdot \frac{e}{f} \right)$$

Example 89 *Such a reasoning can be extended to include the reals and the complex numbers, use of which will be made extensively. The reader is invited to show that these are indeed fields.*

Example 90 $\mathbb{F}_p$ *forms a field for any prime* p. *This is because for any*

$$p \neq a \in \mathbb{F}_p$$

and $a \neq 0$, *we have* $\gcd(a, p) = 1 \implies a^{-1}$ *exists. We've already seen that* $\mathbb{F}_p$ *forms a group under addition.*

Proof. $\gcd(a, p) = 1$
$\implies \exists m, n$ such that $ma + np = 1$
$\implies ma \equiv 1 \bmod p$ ■

This number could very well be a prime power and the results still hold.

We can go on to show how polynomials are constructed but that divert us from our main focus. Let us assume that the reader is familiar with polynomials generally. A high school mathematical book will tell you that the complex numbers are algebraiclly closed whereas the reals are not. That is, the solution to the equation $x^2 + 1 = 0$ exists in the complex numbers but not in the reals. The complex numbers are, in some sense, stronger than the reals. However, there is a great deal more of properties the reals enjoy which the complex numbers do not.

Definition 91 *A field* $(\mathbb{F}, +, \cdot)$ *together with a total order* $\leq$ *on* $\mathbb{F}$ *is an* ***ordered field*** *if the order satisfies*

1. *if* $a \leq b$ *then* $a + c \leq b + c$
2. *if* $0 \leq a$ *and* $0 \leq b$ *then* $0 \leq ab$

It follows from these axioms that for every a, b, c, d in $\mathbb{F}$:
Either $-a \leq 0 \leq a$ or $a \leq 0 \leq -a$.
We are allowed to "add inequalities": If $a \leq b$ and $c \leq d$, then $a + c \leq b + d$
We are also allowed to "multiply inequalities with positive elements": If $a \leq b$ and $0 \leq c$, then $ac \leq bc$.

Example 92 *The rational numbers* $(\mathbb{Z}, \mathbb{Z}^+)$ *form an ordered field, where* $\mathbb{Z}^+$ *denotes the familiar set of positive integers so do the reals but not the complex numbers*

Definition 93 *An **Archimedean ordered field** is an ordered field $\mathbb{F}$ which obeys the Archimedean Property: $\forall x \in \mathbb{F}$, $\exists n \in \mathbb{Z}$ such that $x \leq n$*

If $\mathbb{F}$ is an Archimedean ordered field we can define a bracket function $[x]$ to be $n-1$ where n is the least $n \in \mathbb{Z}$ such that $x \leq n$. The Archimedean Property guarantees there is such an n . The least number principle for integers or the well ordering principle guarantees there is a least such n. Examples of Archimedean ordered fields include the reals $\mathbb{R}$ and the rationals $\mathbb{Q}$. The intuition needed to be able to say that we can have a bracket function will come in handy when we consider the modulus operation $|x|$ on a real number.

In passing, we mention the following: an ordered field that does not satisfy the Archimedean property is said to be **non-Archimedean ordered field.** An element ϵ such that $0 < \epsilon$ and $\epsilon < r$ for every positive $r \in \mathbb{R}$ is called an infinitesimal. By definition, an infinitesimal number is not a real number but belongs to an extension of $\mathbb{R}$, ${}^*\mathbb{R}$, called the hyperreal numbers.

Definition 94 *Let $\mathbb{F}$ be a field. A subset $\mathbb{K}$ that is itself a field under the operations of $\mathbb{F}$ is called a subfield of $\mathbb{F}$. The field $\mathbb{F}$ is called an **extension field** of $\mathbb{K}$. If $\mathbb{K} \neq \mathbb{F}$, $\mathbb{K}$ is called a proper subfield of $\mathbb{F}$.*

To simplify notation and terminology, one says that $\mathbb{F}/\mathbb{K}$ is a field extension to signify that $\mathbb{F}$ is an extension field of $\mathbb{K}$. Field extensions are the main object of study in field theory. The general idea is to start with a base field and construct in some manner a larger field that contains the base field and satisfies additional properties. For instance, the set $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$ is the smallest extension of $\mathbb{Q}$ that includes every real solution to the equation $x^2 = 2$. Such an extension is called a quadratic extension. Given a field extension $\mathbb{F}/\mathbb{K}$, the larger field $\mathbb{F}$ can be considered as a vector space over $\mathbb{K}$. The elements of $\mathbb{F}$ are the "vectors" and the elements of $\mathbb{K}$ are the "scalars", with vector addition and scalar multiplication obtained from the corresponding field operations. The dimension of this vector space is called the degree of the extension and is denoted by $[\mathbb{F} : \mathbb{K}]$. In the above example, the dimension of the extension is 2. It is common to construct an extension field of a given field $\mathbb{K}$ as a quotient ring of the polynomial ring $\mathbb{K}[X]$ in order to "create" a root for a given polynomial $f(X)$. Suppose for instance that $\mathbb{K}$ does not contain any element x with $x^2 = -1$. Then the polynomial $X^2 + 1$ is irreducible over $\mathbb{K}$ or in $\mathbb{K}[X]$ i.e. the coefficients of $X^2 + 1$ in $\mathbb{K}[X]$ cannot be factored into the product of two non-constant polynomials with coefficients in $\mathbb{K}$. Consequently the ideal $(X^2 + 1)$ generated by this polynomial is maximal, and $\mathbb{F} = \mathbb{K}[X]/(X^2 + 1)$ is an extension field of $\mathbb{K}$ which contains an element whose square is -1 (namely the residue class of X).

Given a field extension $\mathbb{K} \subset \mathbb{F}$, an element $\alpha \in \mathbb{F}$ is said to be **algebraic over** $\mathbb{K}$ if α is the root of a polynomial with coefficients in $\mathbb{K}$. So $\sqrt{2}$ is algebraic over $\mathbb{Q}$ since it is a root of $x^2 - 2$, which is a polynomial over $\mathbb{Q}$, but π isn't algebraic over $\mathbb{Q}$. If a number is not algebraic, it is called transcendental.

$\mathbb{R}(\alpha) = \{a + b\alpha \mid a, b \in \mathbb{R}\}$ will denote the structure generated by the real numbers $\mathbb{R}$ and the element α. This is isomorphic to the Complex Numbers.

Over general fields $\mathbb{F}(\alpha)$ can look very different from the complex numbers. For example, for $\alpha = \sqrt[n]{2}$ we have

$$Q(\alpha) = \{a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_{n-1}\alpha^{n-1} \mid a_i \in \mathbb{Q}\}$$

In the extension $\mathbb{Q} \subset \mathbb{R}$, the extension $\mathbb{Q}(\pi)$ is actually isomorphic to $\mathbb{Q}[X]$. The right picture to carry is that there exists n such that they have a nontrivial linear combination that equals zero. If there exists no such n, then the set of powers of α looks like the set of powers of x, and then we can see why $\mathbb{F}(\alpha)$ would be isomorphic to $\mathbb{F}[X]$.

Definition 95 *The **conjugate** elements of an algebraic element α, over a field extension $\mathbb{F}/\mathbb{K}$, are the (other) roots of the minimal polynomial. A **minimal polynomial** is defined relative to a field extension $\mathbb{K}/\mathbb{F}$ and is an element of the extension field $\mathbb{K}$. The minimal polynomial of an element, if it exists, is a member of $\mathbb{F}[x]$, the ring of polynomials in the variable x with coefficients in $\mathbb{F}$. Two elements α, β of a field $\mathbb{K}$, which is an extension field of a field $\mathbb{F}$, are called **conjugate** (over $\mathbb{F}$) if they are both algebraic over $\mathbb{F}$ and have the same minimal polynomial.*

Two complex conjugates $z = a + ib$ and $\overline{z} = a - ib$ (a,$b \in \mathbb{R}$,$b \neq 0$) are conjugate in this more abstract meaning, since they are the roots of the monic polynomial.

$$p(x) = x^2 - 2ax + a^2 + b^2$$

Moreover, the conjugate of a quaternion $a = a_1 + a_2 i + a_3 j + a_4 k$ is defined by

$$\overline{a} = a_1 - a_2 i - a_3 j - a_4 k$$

$S = \{\alpha \in \mathbb{K} \mid \overline{\alpha} = \alpha\}$ the set of all symmetric elements of $\mathbb{K}$. This is a **subring**.

Proof. Clearly, for $\alpha, \beta \in S$, we have $\alpha - \beta = \bar{\alpha} - \bar{\beta} = \overline{\alpha - \beta}$. Thus, $\alpha - \beta \in S$. Similarly, for $\alpha\beta = \bar{\alpha}\bar{\beta} = \overline{\alpha\beta}$ provided the elements are commutative. ■

1.6.1 Homomorphism of Fields

Definition 96 *A **ring homomorphism** is a map $f : R \longrightarrow S$ between two rings R and S such that*

1. *Addition is preserved: $f(r_1 + r_2) = f(r_1) + f(r_2)$,*
2. *The zero element is mapped to zero: $f(0_R) = 0_S$, and*
3. *Multiplication is preserved: $f(r_1 r_2) = f(r_1) f(r_2)$,*

where the operations on the left-hand side is in R and on the right-hand side in S.

Note that a homomorphism must preserve the additive inverse map because $f(g)+f(-g) = f(g-g) = f(0_R) = 0_S$ so $-f(g) = f(-g)$. A ring homomorphism for unit rings (i.e., rings with a multiplicative identity) satisfies the additional property that one multiplicative identity is mapped to the other, i.e., $f(1_R) = 1_S$

A ring homomorphism which is a bijection (one-one and onto) is called a **ring isomorphism**. If $f \longrightarrow S$ is such an isomorphism, we call the rings R and S isomorphic and write $R \cong S$. $f : K \longrightarrow K$ is a ring **antiautomorphism**, if $f(a+b) = f(a) + f(b)$ and $f(ab) = f(b) f(a)$ for all $a, b \in K$. In the commutative case, this is unnecessary. The antiautomorphism is **involutary** if $f^2(a) = a$.

Definition 97 *Let $(\mathbb{F}, +, \cdot)$ and $(\mathbb{K}, +', \cdot')$ be two fields. A **field homomorphism** is a function $\psi : \mathbb{F} \longrightarrow \mathbb{K}$ such that:*

1. $\psi(a+b) = \psi(a) +' \psi(b)$ *for all* $a, b \in F$

2. $\psi(a \cdot b) = \psi(a) \cdot' \psi(b)$

3. $\psi(1) = 1', \psi(0) = 0'$

If ψ is injective and surjective, then we say that ψ is a field **isomorphism**. If $\mathbb{F} = \mathbb{K}$, and ψ bijective, then ψ is a **field automorphism**. For example, complex conjugation is a field automorphism of $\mathbb{C}$, the complex numbers, because

$$\overline{0} = 0$$

$$\overline{1} = 1$$

$$\overline{a+b} = \overline{a} + \overline{b}$$

$$\overline{ab} = \overline{a}\overline{b}$$

1.7 Exercise

1. Prove that the set 2×2 of matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ forms a group under multiplication for $ad \neq bc$. (see appendix for a review of matrices).
2. Prove that $\sigma(X)$, the set of permutations on any set X, is a group.
3. Decide whether or not the following is a field. If it is, state a proof. If it isn't, state a counter-example:
   (a) $\mathbb{Q}$ under usual multiplication and addition.
   (b) $\mathbb{Z}$ under usual multiplication and addition.
   (c) $\mathbb{R} - \{0\}$ under usual multiplication and addition.
   (d) $\mathbb{C}$ under usual multiplication and addition.
   (e) Set of continuous functions with point-wise addition and multiplication.
4. Let $(G, +)$ be an Abelian group, A a nonempty set and $M(A, G)$ the set of all functions $f : A \longrightarrow G$. Define $+ : M(S, G) \times M(S, G) \longrightarrow M(S, G)$ as $(f + g) : A \longrightarrow G$ given by $f(a) + g(a) \in G$ for $a \in A$. Prove that $M(A, G)$ is an Abelian group.
5. In a group $(G, *)$ and for any $a, b, c \in G$, prove that following
   (a) the identity e and the inverse a^{-1} for any element a are both unique.
   (b) $a^2 = a \Rightarrow a = e$
   (c) $a^2 = e \Rightarrow G$ is Abelian
   (d) $ab = ac \Rightarrow b = c$ and $ba = ca \Rightarrow b = c$ (this is called the left and right cancellation law)
   (e) $\left(a^{-1}\right)^{-1} = a$
   (f) $(ab)^{-1} = b^{-1}a^{-1}$
   (g) for unknown $x, y \in G$, the solutions to the equations $ax = b$ and $ya = b$ exists and is unique
6. Show that the intersection of countably finite groups is a group but that the union of countably finite groups is not necessarily a group. (Hint: start with two groups).
7. Prove that the existence of left inverse and left identity, as has been defined for the definition of a group, is implied by the existence of a right inverse and right identity and conversely.
8. Let $\sim$ be an equivalence relation on a group G such that $g_1 \sim g_2$ and $h_1 \sim h_2$ imply $g_1 h_1 \sim g_2 h_2$ for all $g_i, h_i \in G$ where $i = 1, 2$. Then, prove that the set $G/\sim$ of all equivalence classes of G under $\sim$ is a group under the binary operation defined by $[g]\,[h] = [gh]$.

9. Prove that if G is an Abelian group, then so is $G/\sim$.

10. Let $(G, *)$ be a group. Prove that if G is Abelian, then for $g, h \in G$ and $\forall n, m$

    (a) $(gh)^n = g^n h^n$.
    (b) $g^{n+m} = g^n g^m$
    (c) $(g^m)^n = g^{mn}$

Spaces

Basically, linear algebra has to do with the algebra of matrices, vectors and the spaces formed by the collection of either. The idea of a three-dimensional vector can be viewed as a member of $\mathbb{R}^3$. This is called the Euclidean space. The addition, subtraction, scalar and cross multiplication of two vectors is generally well-known and so is the algebra of matrices (a recollection is added in the appendix). However, a mathematical treatment such as the one at our disposal is far more general and has in store some great surprises. Vectors are not vectors anymore in the usual sense yet sound very familiar. Their algebra, too, looks downright distinct but a closer analysis reveals surprising commonalities. In mathematics, the general idea is look for a common structure, formulate some rules such structures obey and then see where else the structure lies. Of course, intuition may play a role in the converse.

Anyway, vector spaces are exactly such spaces. It is to be remarked that mathematically, a space is any set endowed with a particular structure. Hence, vector spaces are just that – a set of vectors with an additional structure.

1.8 Vector Spaces

A group is a mathematical structure – a set that obeys certain axioms. If you want to prove that a particular collection of elements, say integers, forms a group under a certain binary operation, you prove that each and every element of the set satisfies the axioms. Similarly, a vector space satisfies certain axioms and to prove that a collection is a vector space, one follows the same route. The axioms for a vector space concern with addition and scalar multiplication of vectors, which can be intuitively understood. For a set of vectors V, by addition we mean a rule for associating with each pair of objects $\mathbf{u}, \mathbf{v} \in V$ an object $\mathbf{u} + \mathbf{v}$, called the sum of $\mathbf{u}$ and $\mathbf{v}$; by scalar multiplication, we mean a rule for associating with each scalar α and each object $\mathbf{u} \in V$ an object $\alpha\mathbf{u}$, called the scalar multiple of $\mathbf{u}$ by α. A scalar is an object that is a part of a field $\mathbb{F}$. We say that the vectors are scaled over the field $\mathbb{F}$.

More rigorously,

Definition 98 *Let V be a non-empty set and $\mathbb{F}$ be a field. Then, V is a* ***vector space*** *over $\mathbb{F}$ if for $+ : V \times V \longrightarrow V$ and $\cdot : \mathbb{F} \times V \longrightarrow V$*

1. $\mathbf{u}+\mathbf{v} \in V$, as implied by the definition of the function $+$.
2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$
3. $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}$
4. There is an object $\mathbf{0}$ in V, called a zero vector for V, such that $\mathbf{0}+\mathbf{u}=\mathbf{u}+\mathbf{0}=\mathbf{u}$ for all $\mathbf{u}$ in V.
5. For each $\mathbf{u}$ in V, there is an object $-\mathbf{u}$ in V, called a negative of $\mathbf{u}$, such that $\mathbf{u}+(-\mathbf{u})=(-\mathbf{u})+\mathbf{u}=\mathbf{0}$.
6. If α is any scalar and $\mathbf{u}$ is any object in V, then $\alpha\mathbf{u}$ is in V, as implied by the scalar multiplication function.
7. $\alpha(\mathbf{u}+\mathbf{v})=\alpha\mathbf{u}+\alpha\mathbf{v}$
8. $(\alpha+\beta)\mathbf{u}=\alpha\mathbf{u}+\beta\mathbf{u}$
9. $\alpha(\beta\mathbf{u})=(\alpha\beta)(\mathbf{u})$
10. $1\mathbf{u}=\mathbf{u}$

for $\mathbf{u},\mathbf{v},\mathbf{w} \in V$ and $\alpha,\beta,1 \in \mathbb{F}$.
If $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$, then we have a **linear vector space**.

The notation for $+(\mathbf{u},\mathbf{v})=\mathbf{u}+\mathbf{v}$ is adopted to neglect annoying rigour but has been mentioned only to emphasise that there's no magic going on. Also, note that the bold dot ("$\cdot$") has been dropped in favour of the juxtaposition.

The first five axioms for the vector space may be compressed to say that V is an Abelian group under vector addition. The fifth axiom may be derived from the fact that $\alpha=-1$ is also a scalar in $\mathbb{F}$, following from the tenth axiom. One can imagine that such concepts can readily be interpreted physically to be the usual convention of "arrows" in physics. However, the above axiomatisation is a vast generalisation of other spaces, as well. Functions and even sequences can be interpreted as vectors in the above sense, as can be seen from the examples.

Exercise 99 *Prove the following:*

1. $0\mathbf{x}=\mathbf{0}$
2. $\alpha\mathbf{0}=\mathbf{0}$
3. $(-1)\mathbf{x}=-\mathbf{x}$

Example 100 *It is easy to see that $\mathbb{R}$ satisfies the above 10 axioms with scalars belonging to $\mathbb{R}$ and vectors belonging to $\mathbb{R}$ as well. A vector over here is a "ray" from the origin on the x-axis. We say that $\mathbb{R}$ is a vector space over itself. In general, any field is a vector space over itself.*

Example 101 *The set $\mathbb{R}^n$ is a vector space over the set of reals $\mathbb{R}$ under the "usual" vector addition and scalar multiplication. For $\mathbf{x} = (x_1, x_2, ..., x_n)$ and $\mathbf{y} = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$ and for $\alpha \in \mathbb{R}$, we can define vector addition and scalar multiplication as follows: $+ : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that*

$$\begin{aligned} +(\mathbf{x}, \mathbf{y}) &= \mathbf{x} + \mathbf{y} \\ &= (x_1, x_2, ..., x_n) + (y_1, y_2, ..., y_n) \\ &= (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) \end{aligned}$$

and for $. : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that

$$. (\alpha, \mathbf{x}) = \alpha\mathbf{x} = \alpha (x_1, x_2, ..., x_n) = (\alpha x_1, \alpha x_2, ..., \alpha x_n)$$

To begin proving that these operations defined in the manner above do indeed form a vector space, the definitions will have to be applied directly. To see this, clearly,

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) \in \mathbb{R}^n$$

since $x_i + y_i \in \mathbb{R}\ \forall i \in I_n$. Next, $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ because addition is commutative in the set of reals. Again, inverses will exist for any $\mathbf{x} = (x_1, x_2, ..., x_n)$ since we can construct $-\mathbf{x} = (-x_1, -x_2, ..., -x_n)$ by resorting to the fact that every tuple, being real, has an inverse. Furthermore, since $0 \in \mathbb{R}$, then $\mathbf{0} = (0, 0, ..., 0) \in \mathbb{R}^n$ and is an additive identity. Therefore, $\mathbb{R}^n$ is an abelian group. Axiom 6 is satisfied by definition. For Axiom 7,

$$\alpha (\mathbf{x} + \mathbf{y}) = \alpha (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$$

using the rule of addition. Since the element $(x_i + y_i)$ is a member of the set of reals, therefore $(\alpha (x_1 + y_1), \alpha (x_2 + y_2), ..., \alpha (x_n + y_n))$ is justified. Furthermore, $\mathbb{R}$ is a field, hence $\alpha (x_i + y_i) = \alpha x_i + \alpha y_i$. Thus,

$$\begin{aligned} & \alpha (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) \\ = \ & (\alpha (x_1 + y_1), \alpha (x_2 + y_2), ..., \alpha (x_n + y_n)) \\ = \ & (\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2, ..., \alpha x_n + \alpha y_n) \end{aligned}$$

Now, since αx_i and αy_i are real numbers, we can split the tuple

$$(\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2, ..., \alpha x_n + \alpha y_n)$$

to get

$$(\alpha x_1, \alpha x_2, ..., \alpha x_n) + (\alpha y_1, \alpha y_2, ..., \alpha y_n)$$

which is, according to our definition of the scalar multiplication,

$$= \alpha (x_1, x_2, ..., x_n) + \alpha (y_1, y_2, ..., y_n)$$

Thus, we have justified $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$. For the 8th Axiom, we follow a similar series of steps:

$$\begin{aligned}
& (\alpha+\beta)\mathbf{x} \\
= \; & (\alpha+\beta)\,(x_1, x_2, ..., x_n) \\
= \; & ((\alpha+\beta)\,x_1, (\alpha+\beta)\,x_2, ..., (\alpha+\beta)\,x_n) \\
= \; & (\alpha x_1 + \beta x_1, \alpha x_2 + \beta x_2, ..., \alpha x_n + \beta x_n) \\
= \; & (\alpha x_1, \alpha x_2, ..., \alpha x_n) + (\beta x_1, \beta x_2, ..., \beta x_n) \\
= \; & \alpha\,(x_1, x_2, ..., x_n) + \beta\,(x_1, x_2, ..., x_n) \\
= \; & \alpha\mathbf{x} + \beta\mathbf{x}
\end{aligned}$$

The last two axioms can be similarly proved.

Notice how everything reduces to the pre-established axioms and the fact that the base is a field and that the tuples form from the same field. This painstaking mode of argument was only meant to serve as a motivation for how a vector space ought to be proved: the axioms must hold!

To prove that $\mathbb{R}^n$ is a vector space over the field $\mathbb{Q}$, we can similarly resort to the fact that each tuple, that is, every real number when added or multiplied by a rational number will yield another real number.

Equivalently, we can take n-tuple of the Complex plane $\mathbb{C}$ and generate for ourselves a vector space. Try and prove it first for $n = 1$ and then proceed via induction.

Example 102 *The space $C[a, b]$, the set of all continuous real valued functions defined on the interval $I = [a, b]$, forms a real vector space with the algebraic operations defined in the usual way:*
$(x + y)(t) = x(t) + y(t)$
$(\alpha x)(t) = \alpha\,(x\,(t))$ *for* $k \in \mathbb{R}$

The use of x and y as function is rather intentional. We will prove that the sum of continuous numbers is continuous. Since we're using the real line, we'll use the $\epsilon - \delta$ definition of continuity, defined in real analysis. To recall

Example 103 ***Definition 104*** *For $A, B \subset \mathbb{R}$, a function $x : A \longrightarrow B$ is* ***continuous*** *at a point t_0 if for every $\epsilon > 0$, there exists a $\delta > 0$ such that*

$$|x(t) - x(t_0)| < \epsilon \text{ whenever } |t - t_0| < \delta$$

A function is continuous if it is continuous at every point of its domain.

Now, if the function x is continuous, then for any point t_0, we have

$$|x(t) - x(t_0)| < \epsilon/2$$

whenever $|t - t_0| < \delta_1$ and if y is continuous at the same point, then we have

$$|y(t) - y(t_0)| < \epsilon/2$$

whenever $|t-t_0|<\delta_2$. *Take* $\delta=\min(\delta_1,\delta_2)$. *Then,*

$$\begin{aligned}|(x+y)(t)-(x+y)(t_0)| &= |x(t)+y(t)-x(t_0)-y(t_0)|\\ &= |[x(t)-x(t_0)]+[y(t)-y(t_0)]|\\ &\leq |x(t)-x(t_0)|+|y(t)-y(t_0)|\\ &< \epsilon/2+\epsilon/2=\epsilon.\end{aligned}$$

The same can be said for the continuity of $(\alpha x)(t)$ *if we let* $|x(t)-x(t_0)|<\epsilon/|\alpha|$. *Hence the definitions for the addition and scalar multiplication are justified.*

It is now rather routine to prove that the space of continuous functions is a vector space by showing that under the defined "vector" addition and scalar multiplication, the remaining 8 axioms hold. In particular, one should prove that the zero function is continuous; any additive inverse of a continuous function is also continuous.

We will pause over here to recall a fact from real analysis: there is a fine distinction between continuity and uniform continuity. In the above definition, given that the function is continuous, the ϵ can depend on x so that for each open interval in the range, we might be able to find an open interval in the domain with differing lengths (the length of an interval is defined as the difference between its endpoints so that $|(a,b)|=|[a,b]|=b-a$). However, if the function under consideration is uniformly continuous, then this ϵ will not depend upon the domain so that the lenghts of each pre-image of open sets is the same. Thus, continuity of a function is a local property since it depends upon x whereas uniform continuity of a function is a global property which does not depend upon x. Needless to say, if a function is uniformly continuous, then it is continuous but the converse is not true in geneal.

Hopefully, you have just proved that this collection of functions forms a vector space. In other words, functions can be viewed as vectors in a sense – a single point in space. Such a space is also called the function space.

Wherever possible, vectors will be made bold. However, this will not be strictly followed, especially when it comes to functions and sequences, since this may be a cause of confusion with some pre-established notations. This should in no way mean that functions are not vectors.

Example 105 *The set of polynomials* $P[a,b]$ *of order at most* n *over the set* $[a,b]$ *is also a vector space. Try to prove this yourself. Take the element*

$$P(x)=\sum_{i=0}^{n}\alpha_i x^i$$

and prove that this forms a vector space under the "ordinary" addition and scalar multiplication.

Before moving on to another vey important example, let us pause to consider the following: any vector space can be "constructed" from given vector spaces

over a similar field. For instance, in the first example, we see that $\mathbb{R}$ is a vector space over field $\mathbb{R}$ and $\mathbb{R}^n$, too, is a vector space over $\mathbb{R}$. Here, $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times ... \times \mathbb{R}$. Generally, if we have vector spaces $X_1, X_2, ..., X_n$ over a similar field $\mathbb{K}$, then

Lemma 106 *$X := \{(x_1, x_2, ..., x_n) \mid x_i \in X_i\}$ is a vector space*

Proof sketch. Define $(x_1, x_2, ..., x_n) + (y_1, y_2, ..., y_n) := (x_1 + y_1, ..., x_n + y_n)$ and $\lambda (x_1, x_2, ..., x_n) := (\lambda x_1, \lambda x_2, ..., \lambda x_n)$. This is natural because for each i-th tuple, the addition makes sense because the underlying vector space X_i is closed under addition and mulitplication.

It is now a routine matter to verify that the resulting space is indeed a vector space. ■

Such an X is a vector space and is called the Cartesian product (see Set Theory in preliminaries for details). Usually, when the context is clear, the word Cartesian is dropped and just the words "product of spaces" is used. However, this is not standard since there are other ways in which a product for a vector space may be defined. The product space is denoted by $X = X_1 \times X_2 \times ... \times X_n$

Before we move on to consider our next example, a definition is in order:

Definition 107 *The **modulus** of a real number x is a function defined such that* $|x| = \begin{cases} x \text{ if } x \geq 0 \\ -x \text{ if } x < 0 \end{cases}$

This modulus simply converts negative numbers to positive numbers and lets positive numbers be positive numbers. In reality, it actually measures the distance of a number from the zero point. In other words, the magnitude of the one-dimensional ray formed on the real line is measured by this definition.

The above definition was meant to serve as a gentle introduction what the modulus (actually, norm) is about. But enough with gentle introductions! Using this definition, the reader is invited to flex his/her muscles by proving the following:

Exercise 108 *For any real $a \geq 0$, then* $|x| \leq a \iff -a \leq x \leq a$

Exercise 109 $|x| = 0 \iff x = 0$

Exercise 110 $|xy| = |x|\,|y|$

Exercise 111 $|-x| = |x|$

Exercise 112 $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$ *for* $|y| \neq 0$

Exercise 113 $|x - y| = |y - x|$

Exercise 114 $\left|x^2\right| = x^2$

Exercise 115 $|x| = \sqrt{x^2}$

Exercise 116 $|x+y| \leq |x| + |y|$

Exercise 117 $|x-y| \leq |x| + |y|$

Exercise 118 $|x| - |y| \leq |x-y|$

Exercise 119 $||x| - |y|| \leq |x-y|$

We will only prove the triangle inequality, since it will be made fair use of. We start with $-|s| \leq s \leq |s|$ and $-|t| \leq t \leq |t|$ for $s, t \in \mathbb{R}$. Adding these two together, we have $-(|s| + |t|) \leq s + t \leq |s| + |t|$ and this is exactly the first statement for $s + t = x$ and $|s| + |t| = a$.

Considering that we've proved $\mathbb{R}^n$ is a vector space, it is now time to move ahead. $\mathbb{R}^n$ consists of n-tuples. To make these tuples infinite essentially means that we're considering seqeunces – which are tuples, by the way, but in a more magnified form. Now, a sequence may converge or diverge. This idea is made more formal in metric which you can read ahead for yourself for a refresher. You have probably passed a Calculus, Real Analysis and Topology course so I will just go ahead with myself: convergence in $\mathbb{R}^n$ requires the usual Euclidean metric. We can collect for ourself bounded sequences, convergent sequences or even convergent series. This is where the natural generalisation of l^p steps in for $p \in [1, \infty]$

Example 120 *The space l^2 (the space of two summable convergent and hence bounded sequences) with the algebraic operations defined similar to those for n-tuples in connection with sequences, that is, for $\xi, \varsigma \in l^2$ such that*

$$\sum_{i=1}^{\infty} |\xi_i|^2 < \infty$$

we have

$$\boldsymbol{\xi} + \boldsymbol{\varsigma} = (\xi_1, \xi_2, ..) + (\varsigma_1, \varsigma_2, ...) = (\xi_1 + \varsigma_1, \xi_2 + \varsigma_2, ..)$$

and

$$\alpha\boldsymbol{\xi} = (\alpha\xi_1, \alpha\xi_2, ..)$$

forming a vector space.

Proof. It is easy to show that these operations are well-defined because for any two equal sequences $\boldsymbol{\xi} = \boldsymbol{\varsigma}$, we have $\alpha\boldsymbol{\xi} = \alpha\boldsymbol{\varsigma}$ and its companion addition. However, we still won't be done as being well-defined is no guarantee that $\boldsymbol{\xi} + \boldsymbol{\varsigma}$ and $\alpha\boldsymbol{\xi} \in l^2$ so we prove this first. Let $\sum_{i=1}^{\infty} |\xi_i|^2 < \infty$ and let α be a finite scalar. Then, if $\alpha\boldsymbol{\xi} \notin l^2$, we must have $\sum_{i=1}^{\infty} |\alpha\xi_i|^2 \longrightarrow \infty$ or $\sum_{i=1}^{\infty} |\alpha|^2 |\xi_i|^2 \longrightarrow \infty$ or $|\alpha|^2$

$\sum_{i=1}^{\infty} |\xi_i|^2 \longrightarrow \infty$. Now, clearly, $\sqrt{\sum_{i=1}^{\infty} |\xi_i|^2} < \infty$ so that $|\alpha|^2 \longrightarrow \infty$ which is a contradiction. For the first axiom, first we note that $\sum_{i=1}^{\infty} |\xi_i|^2, \sum_{i=1}^{\infty} |\varsigma_i|^2 < \infty$ implies $\sqrt{\sum_{i=1}^{\infty} |\xi_i + \varsigma_i|^2} \leq \sqrt{\sum_{i=1}^{\infty} |\xi_i|^2} + \sqrt{\sum_{i=1}^{\infty} |\varsigma_i|^2} < \infty$

The inequality used over here is the Minkowski inequality (proved later) for $p = 2$.

Now we can move ahead with the real axioms! Clearly, we can have $\boldsymbol{\xi} + (\boldsymbol{\varsigma} + \boldsymbol{\eta}) = (\boldsymbol{\xi} + \boldsymbol{\varsigma}) + \boldsymbol{\eta}$, by resorting to the associativity of each ith element. Next, the sequence $\mathbf{0} = (0, 0, 0, ...)$ converges and hence belongs to l^2 so that

$$\begin{aligned} \boldsymbol{\xi} + \mathbf{0} &= (\xi_1, \xi_2, ..) + (0, 0, ...) \\ &= (\xi_1 + 0, \xi_2 + 0, ..) \\ &= (\xi_1, \xi_2, ..) \\ &= \xi \end{aligned}$$

Furthermore,

$$\sum_{i=1}^{\infty} |\xi_i|^2 = \sum_{i=1}^{\infty} |-\xi_i|^2$$

so that for every $\boldsymbol{\xi}$, we have $-\boldsymbol{\xi}$. For commutativity, again, we resort to the fact the base field is commutative so that

$$\begin{aligned} \boldsymbol{\xi} + \boldsymbol{\varsigma} &= (\xi_1 + \varsigma_1, \xi_2 + \varsigma_2, ..) \\ &= (\varsigma_1 + \xi_1, \varsigma_2 + \xi_2, ..) \\ &= \boldsymbol{\varsigma} + \boldsymbol{\xi} \end{aligned}$$

Now that additive operation forms an Abelian group, we can similarly prove that $\alpha(\boldsymbol{\xi} + \boldsymbol{\varsigma}) = \alpha\boldsymbol{\xi} + \alpha\boldsymbol{\varsigma}$, $(\alpha + \beta)\boldsymbol{\xi} = \alpha\boldsymbol{\xi} + \beta\boldsymbol{\xi}$, $\alpha(\beta\boldsymbol{\xi}) = (\alpha\beta)\boldsymbol{\xi}$ and $1\boldsymbol{\xi} = \boldsymbol{\xi}$ ■

These are known as Hilbert sequence spaces. This should not to be confused with the Hilbert space, which will be studied in the coming chapters. A remark, however, has to be in order; this was the first example of a Hilbert space presented in history. From this, one can get a sense that Hilbert spaces are not just about vectors in the usual sense and that their development had rather broad applications in mind.

In order to avoid confusion, we will refer to the space in the example as the sequence space. In this example, we have l^p for $p = 2$. This p can range between $1 \leq p < \infty$ and the power 2 in $\Sigma |\xi_i|^2 < \infty$ gets replaced accordingly. For $p = \infty$, the only requirement is that we have bounded sequences i.e. we have $|\xi_i| \leq c_x$ where c_x is a real number which may depend on x but does not depend on i. This is natural since the the bound will depend on the sequence x but has to be valid for all i. In fact, we might even have a divergent series if this constant depended on each i, giving us an unbounded sequence.

The keen eye should have noted the idea of boundedness is entirely dependent on how distance is defined. This topic can be put aside until normed spaces are defined.

If the sequences for $p = 2$ are real-valued, then the space is a collection of all convergent sequences on the real line, which forms a vector space, as mentioned. Instead of real numbers ξ_i, we can also have complex numbers as the base elements of the sequences. Note that this should not be confused with the Lebesgue Space L^p, which deals with functions with a p-norm. The l^p space is a special case of the L^p space, which is not covered in MTH327. However, because of the knowledge of measure that is required, such spaces are studied in MTH427. L^p spaces are defined using natural generalisations of p-norms for finite-dimensional vector spaces. They are named after the French mathematician Henri Lebesgue (June 28, 1875 – July 26, 1941).

We will prove Hölder's inequality on the real numbers. The inequality is valid for $p > 1$ and a q such that $\frac{1}{p} + \frac{1}{q} = 1$. In a sense, p and q are inverses or conjugates of each other. These are called conjugate exponents. For instance, $p = q = 2$. 1 and ∞ are also regarded as conjugate exponents.

Theorem 121 *For convergent (why?) real sequences* (x_i) *and* (y_i)

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} \left(\sum_{i=1}^{\infty} |y_i|^q \right)^{1/q}$$

Proof. From $\frac{1}{p} + \frac{1}{q} = 1$, we have $\frac{p+q}{pq} = 1$
$pq - p - q - 1 = 1$
$p(q-1) - (q-1) = 0$ or $(p-1)(q-1) = 1$
Thus,

$$\frac{1}{p-1} = q - 1$$

so that from a function $u = t^{p-1}$, we can have $u^{1/(p-1)} = t$ or $t = u^{q-1}$. Now let $a, b \in \mathbb{R}$ for $a, b > 0$. Then, we can think of ab as an area of a rectangle with sides a and b.

Let $f(t) = t^{p-1}$. Then,

$$\begin{aligned} ab &\leq \int_0^a f(t)\,dt + \int_0^b f^{-1}(t)\,dt \\ &= \int_0^a t^{p-1} dt + \int_0^b t^{q-1} dt \\ &= \frac{a^p}{p} + \frac{b^q}{q} \end{aligned}$$

Assume that we have the sequence (ξ_i) and (η_i) such that

$$\sum_{i=1}^{\infty} |\xi_i|^p = \sum_{i=1}^{\infty} |\eta_i|^q = 1$$

Set $a = |\xi_i|$ and $b = |\eta_i|$. Then, we have the inequality

$$|\xi_i|\,|\eta_i| \leq \frac{|\xi_i|^p}{p} + \frac{|\eta_i|^q}{q}$$

Summing such i objects, we get the inequality

$$\sum_{i=1}^{\infty} |\xi_i|\,|\eta_i| \leq \sum_{i=1}^{\infty} \frac{|\xi_i|^p}{p} + \sum_{i=1}^{\infty} \frac{|\eta_i|^q}{q}$$

from which we have

$$\sum_{i=1}^{\infty} |\xi_i|\,|\eta_i| = \sum_{i=1}^{\infty} |\xi_i \eta_i| \leq \frac{1}{p} + \frac{1}{q} = 1 \tag{1.2}$$

Now, let

$$\xi_i = \frac{x_i}{\left(\Sigma\,|x_i|^p\right)^{1/p}}$$

and

$$\eta_i = \frac{y_i}{\left(\Sigma\,|y_i|^q\right)^{1/q}}$$

Then,

$$\xi_i^p = \frac{x_i^p}{\left(\Sigma\,|x_i|^p\right)}$$

and

$$\eta_i^q = \frac{y_i^q}{\left(\Sigma\,|y_i|^q\right)}$$

which implies

$$|\xi_i|^p = \frac{|x_i|^p}{\left(\Sigma\,|x_i|^p\right)}$$

and

$$|\eta_i|^q = \frac{|y_i|^q}{\left(\Sigma\,|y_i|^q\right)}$$

which, on summing the i index, will yield

$$\sum_{i=1}^{\infty} |\xi_i|^p = \frac{\Sigma\,|x_i|^p}{\left(\Sigma\,|x_i|^p\right)} \text{ and } \sum_{i=1}^{\infty} |\eta_i|^q = \frac{\Sigma\,|y_i|^q}{\left(\Sigma\,|y_i|^q\right)}$$

both of which equal 1. Hence,

$$\xi_i = \frac{x_i}{\left(\Sigma\,|x_i|^p\right)^{1/p}} \text{ and } \eta_i = \frac{y_i}{\left(\Sigma\,|y_i|^q\right)^{1/q}}$$

is a valid substitution. Multiplying these two and placing in the equality (1.2), we get

$$\sum_{i=1}^{\infty}\left|\frac{x_i}{\left(\Sigma\left|x_i\right|^p\right)^{1/p}}\frac{y_i}{\left(\Sigma\left|y_i\right|^q\right)^{1/q}}\right|\leq 1$$

$$\sum_{i=1}^{\infty}\left|x_iy_i\right|\leq\left(\sum_{i=1}^{\infty}\left|x_i\right|^p\right)^{1/p}\left(\sum_{i=1}^{\infty}\left|y_i\right|^q\right)^{1/q}$$

This proof is valid only for l^p spaces and not norm spaces in general. ■

Exercise 122 *Show that the geometric mean of two positive numbers does not exceed their arithmetic mean.*

Continuing with the real sequences, for $p=2$, we have a special name for the **Hölder inequality**: the **Cauchy-Schwarz inequality**

$$\sum_{i=1}^{\infty}\left|x_iy_i\right|\leq\left(\sum_{i=1}^{\infty}\left|x_i\right|^2\right)^{1/2}\left(\sum_{i=1}^{\infty}\left|y_i\right|^2\right)^{1/2}$$

Equality holds if x and y are scalars of each other.

Idea. Trivial for zero vectors. Take non-zero vectors x,y such that

$$x=(x_1,x_2,...,x_n,...)=\lambda\left(y_1,y_2,...,y_n,...\right)=\lambda y$$

Then,

$$\begin{aligned}
&\Longrightarrow\left(\sum_{i=1}^{\infty}\left|x_iy_i\right|\right)^2\Longrightarrow\left(\sum_{i=1}^{\infty}\left|\lambda\right|\left|y_iy_i\right|\right)^2=\left(\sum_{i=1}^{\infty}\left|\lambda\right|\left|y_i\right|^2\right)^2\\
&=\lambda^2\left(\sum_{i=1}^{\infty}\left|y_i\right|^2\right)^2=\lambda^2\left(\sum_{i=1}^{\infty}\left|y_i\right|^2\right)\left(\sum_{j=1}^{\infty}\left|y_j\right|^2\right)\\
&=\left(\sum_{i=1}^{\infty}\left|\lambda y_i\right|^2\right)\left(\sum_{j=1}^{\infty}\left|y_j\right|^2\right)\\
&=\left(\sum_{i=1}^{\infty}\left|x_i\right|^2\right)\left(\sum_{j=1}^{\infty}\left|y_j\right|^2\right)\ \blacksquare
\end{aligned}$$

We will now prove the Minkowski Inequality for the same conditions as the Hölder inequality. Specifically, for $p\geq 1$

Theorem 123 $\left(\sum_{i=1}^{\infty}\left|x_i+y_i\right|^p\right)^{1/p}\leq\left(\sum_{i=1}^{\infty}\left|x_i\right|^p\right)^{1/p}+\left(\sum_{i=1}^{\infty}\left|y_i\right|^q\right)^{1/q}$

Proof. For $p=1$, we can directly apply the triangle inequality. For $p>1$

$\left|x_i+y_i\right|^p=\left|x_i+y_i\right|\left|x_i+y_i\right|^{p-1}\leq\left(\left|x_i\right|+\left|y_i\right|\right)\left|x_i+y_i\right|^{p-1}$

Summing over i,

$$\sum_{i=1}^{\infty} |x_i + y_i|^p \leq \sum_{i=1}^{\infty} |x_i| \, |x_i + y_i|^{p-1} + \sum_{i=1}^{\infty} |y_i| \, |x_i + y_i|^{p-1}$$

From the Hölder inequality, we have

$$\sum_{i=1}^{\infty} |x_i| \, |x_i + y_i|^{p-1} \leq \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} \left(\sum_{i=1}^{\infty} |x_i + y_i|^{(p-1)q} \right)^{1/q}$$

$$= \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{p-1} \left(\sum_{i=1}^{\infty} |x_i + y_i|^p \right)^{1/q}$$

and

$$\sum_{i=1}^{\infty} |y_i| \, |x_i + y_i|^{p-1} \leq \left(\sum_{i=1}^{\infty} |y_i|^p \right)^{1/p} \left(\sum_{i=1}^{\infty} |x_i + y_i|^p \right)^{1/q}$$

because $pq = p + q$

Hence we have

$$\sum_{i=1}^{\infty} |x_i + y_i|^p \leq \left(\left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p \right)^{1/p} \right) \left(\sum_{i=1}^{\infty} |x_i + y_i|^p \right)^{1/q}$$

$$\implies \left(\sum_{i=1}^{\infty} |x_i + y_i|^p \right)^{1-1/q} \leq \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p \right)^{1/p}$$

$$\implies \left(\sum_{i=1}^{\infty} |x_i + y_i|^p \right)^{1/p} \leq \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p \right)^{1/p} \quad \blacksquare$$

Now that you have these inequalities at your disposal, prove that l^p is a metric space

1.9 Normed Spaces

The modulus operation for a real number basically converts a negative number into a positive one and lets the positive ones be. If we think of the real number line as one-dimensional arrows eminating from zero, we can interpret the modulus as a function that tells how far the real number is from zero. Of course there is no reason to have a magnitude function (the modulus) for real numbers only. We can move ahead and generalise it for other vectors as well (strictly, elements of a vector space).

Definition 124 *Let N be a linear space over a field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A **norm** on N is a real-valued function $\|.\| : N \longrightarrow [0, \infty)$ such that*

N1: $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in N$ and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$

N2: $\|\alpha\mathbf{x}\| = |\alpha| \, \|\mathbf{x}\|$ for all $\alpha \in \mathbb{F}$, $\mathbf{x} \in N$

N3: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for arbitrary $\mathbf{x}, \mathbf{y} \in N$

Imagine that these definitions apply to two dimensional vectors. Then, justify to yourself that they apply to three dimensional vectors as well. There's no reason to stop there and we can continue with n-dimensional spaces, as well (don't worry, the example done below covers just that). Moving on, as can be sensed, this is a generalisation of the concept of "measure" of a vector or its length. A linear vector space together with a norm defined on it is called a **normed space**, denoted by $(N, \|.\|)$. Since the use of the extra brackets is tedious, we will often shorten it to just the name of the set. Bear in mind that a different norm on the same set does not make the same normed space.

There is absolutely no reason why we should restrict ourselves to two fields but is only because it makes matter a little easier to handle. We also have the added advantage of completeness in both and that of order in the former. We could have very well made use of the field of rational numbers, which may prove to be easy but because of the absence of completeness, we will not be covering this more general case. The imagination dries when the quaternions or even the field of functions are used as scalars.

The first part of N1 is the positivity property or positive definiteness, N2 is called the Absolute Homogenity Axiom whereas condition N3 is called the Triangle Inequality or the Subadditive property. The second part of N1 ensures that any two given vectors are two separate points.

The condition on N1 can be shortened to $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$

Proof. $\|x\| \leq 2\|x\| \implies \|x\| \geq 0$ ■

It is only included for emphasis.

Exercise 125 *Show that* $|\|\mathbf{y}\| - \|\mathbf{x}\|| < \|\mathbf{y} - \mathbf{x}\|$ *for any vectors* $\mathbf{x}, \mathbf{y}$ *in any norm space* N.

The first intuitive example of the magnitude of two vectors will serve as a motivation for the definition.

Example 126 *Let* $\mathbb{R}^2$ *be the Euclidean space. The length of a vector, understood to be derived from the Pythagorean theorem, is* $\|\mathbf{x}\| = \sqrt{(x_1)^2 + (x_2)^2}$. *Prove that this is indeed a normed space. We will prove the more general norm*

$$\|\mathbf{x}\| = \sqrt{(x_1)^2 + (x_2)^2 + ... + (x_n)^2}$$

for the space $\mathbb{R}^n$. *Clearly, this norm is positive since we're dealing with square roots. The only way* $\sqrt{(x_1)^2 + (x_2)^2 + ... + (x_n)^2}$ *or* $\|\mathbf{x}\|$ *could be zero is when* $x_1 = x_2 = ... = x_n = 0$. *This in turn implies that* $\mathbf{x} = (0, 0, ..., 0)$ *or* $\mathbf{x} = \mathbf{0}$. *To prove N2, we have*

$$\|\alpha\mathbf{x}\| = \sqrt{(\alpha x_1)^2 + (\alpha x_2)^2 + ... + (\alpha x_n)^2}$$

since $\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, ..., \alpha x_n)$. *We then have*

$$\begin{aligned}
\sqrt{(\alpha x_1)^2 + (\alpha x_2)^2 + ... (\alpha x_n)^2} &= \sqrt{\alpha^2 \left((x_1)^2 + (x_2)^2 + ... (x_n)^2 \right)} \\
&= \sqrt{\alpha^2} \sqrt{\left((x_1)^2 + (x_2)^2 + ... (x_n)^2 \right)} \\
&= |\alpha| \sqrt{\left((x_1)^2 + (x_2)^2 + ... (x_n)^2 \right)} \\
&= |\alpha| \, \|\mathbf{x}\|
\end{aligned}$$

N3 follows from Minkowski's inequality.

The discussion of $n = 1$ in $\mathbb{R}^n$ as a definition of the modulus $|x| = \sqrt{x^2}$ can now tell us that the real numbers are vectors in one dimension on the real line. For $n = 2$, do it yourself. For $x = (x_1, x_2, ..., x_n)$, $\|x\| = \left(x_1^2 + x_2^2 + ... + x_n^2 \right)^{1/2}$ is the usual Euclidean norm satisfying the above axioms.

This is further generalised to the p-norm $\|x\|_p = \left(|x_1|^p + |x_2|^p + ... + |x_n|^p \right)^{1/p}$. For $p = 1$, the norm is called Manhattan distance, or taxicab distance because distance is defined in terms of "blocks". For $p = \infty$, the norm changes to $\|x\|_\infty = max\{|x_1|, |x_2|, ..., |x_n|\}$. This is also known as the uniform norm.

Note that the norm defined as above, the Euclidean norm, is not the only norm on $\mathbb{R}^n$. For instance,

Example 127 *We can define* $\|.\| : \mathbb{R}^n \longrightarrow [0, \infty)$ *such that for* $\mathbf{x} \in \mathbb{R}^n$,

$$\|\mathbf{x}\| = \sum_{i=1}^{n} |x_i|$$

Then, clearly $x_i \geq 0$ *and*

$$\begin{aligned}
\|\mathbf{x}\| &= 0 \\
&\iff \sum_{i=1}^{n} |x_i| = 0 \\
&\iff x_i = 0 \forall i \\
&\iff \mathbf{x} = \mathbf{0}
\end{aligned}$$

Furthermore,

$$\begin{aligned}\|\alpha\mathbf{x}\| &= \sum_{i=1}^{n}|\alpha x_i| \\ &= \sum_{i=1}^{n}|\alpha|\,|x_i| \\ &= |\alpha|\sum_{i=1}^{n}|x_i| \\ &= |\alpha|\,\|\mathbf{x}\|\end{aligned}$$

N3 follows by inductively applying the triangle inequality for real numbers. That is,

$$\begin{aligned}& |x_1 + x_2 + ... + x_n| \\ \leq\ & |x_1| + |x_2 + ... + x_n| \\ & \vdots \\ \leq\ & |x_1| + |x_2| + ... + |x_n|\end{aligned}$$

Hence, we have

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\| &= \sum_{i=1}^{n}|x_i + y_i| \\ &\leq \sum_{i=1}^{n}|x_i| + \sum_{i=1}^{n}|y_i| \\ &= \|\mathbf{x}\| + \|\mathbf{y}\|\end{aligned}$$

Other norms on the same set yield different norm spaces, as discussed in the lectures. Rephrasing the lectures, two different norms on the same set do not necessarily generate the same norm space. The norms are defined according to different uses and priorities.

Exercise 128 *Show that*

$$\|z\| = \sqrt{|z_1| + |z_2| + ... + |z_n|}$$

for $(z_1, z_2, ..., z_n) \in \mathbb{C}^n$ *forms a norm and therefore,* $(\mathbb{C}^n, \|.\|)$ *a normed space.*

Exercise 129 *The norm of* l^p *was hinted at in the example for vector spaces; for any sequence* $x \in l^p$*, we have*

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

for $1 \leq p < \infty$. The proof that this definition satisfies the axioms for a norm, making it a bona fide normed space, follows a similar pattern as that for $\mathbb{R}^n$. Here, one can use the fact that $|\alpha\xi_i|^p = |\alpha|^p |\xi_i|^p$, thanks to the definition of the modulus, given that α is a real number. The proof of N3 is the proof of the Minkowski inequality.

Example 130 *The definition for $p = \infty$ can be recollected from the former material and can be proved by the reader.*

Example 131 *$C[a, b]$, the space of continuous functions in the interval $[a, b]$ for $a < b$ is a normed space under the norm such that for any $x(t) \in C[a, b]$, we have*

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

N1 holds clearly since (a) the modulus is always positive and (b) if the maximum value of a non-negative quantity is 0, then that quantity is itself 0. That is,

$$\max_{t\in[a,b]} |x(t)| = 0 \iff x(t) = 0(t)$$

where $0(t) = 0\ \forall t$. Next, the scalar α does not range over t, hence

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

proving N2. N3 follows by the use of the triangle inequality of real numbers.

We can also define a norm on $C[a, b]$ such that

$$\|\mathbf{x}\| = \|x(t)\| = \int_a^b |x(t)|\, dt$$

The reader is invited to prove that this is normed space.

The concept of a norm is important because length of a vector, distance between two vectors and hence the idea of convergence can be defined accordingly. Needless to say, the idea of convergence is essential in almost all areas of mathematics. Also, based on the definition of the norm and the vectors, convergence, magnitude and distance can then be decided for, as we shall see.

Set Topology

1.10 Metric Spaces

Definition 132 *Let X be any non-empty subset. A **metric** defined on X is a function $d : X \times X \longrightarrow [0, \infty)$ such that*

- *D1* $d(x, y) \geq 0$
- *D2* $d(x, y) = 0 \iff x = y$
- *D3* $d(x, y) = d(y, x)$
- *D4* $d(x, y) \leq d(x, z) + d(z, y)$

for $x, y, z \in X$. A **metric space** is a pair (X, d) where X is a set and d is a metric on X. Metric generalises the concept of distance between two points. A natural question that should arise is the following: how is the metric and a norm related? In a sense, the distance between two vectors can also be decided using the metric function. This is can be observed as follows: the difference between two vectors yields a third vector, which connects the tips of the two. The length of this vector is the distance between the tips of the vectors, which is what vectors are. Mathematically, $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. But we are getting ahead of ourselves! For now, here are a few examples of metric spaces.

Example 133 *A trivial metric which can be defined on any set is the discrete metric $d(x, y) = 1$ for $x \neq y$ and 0 otherwise. $d(x, y) \geq 0$,*

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

and $d(x, y) = d(y, x)$ are satisfied by definition. $d(x, y) \leq d(x, z) + d(z, y)$ can be verified exhaustively by considering cases $x \neq y$, $x = y$, $x \neq z$ and $x = z$.

Example 134 *On the real line $\mathbb{R}$, we can define the usual metric $d(x, y) = |x - y|$ (aha!). This can be generalised for the Euclidean plane $\mathbb{R}^n$ with metric*

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2}$$

For $n = 1$, we can get $\sqrt{(x_1 - y_1)^2} = |x - y|$, which should've been proved by the student in the exercise in the previous topic. For $n = 2$, we have the familiar Pythagorean Theorem in plane. Note that the metric d is also the Pythagorean Theorem in n-dimension. Dimensionality will be rigorously defined later. For now, we'll make do with the normal understanding of the word.

Example 135 *Another metric definable on the same set is*

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

This is called the taxicab or Manhattan metric.

This example illustrates the important fact that from a given set with more than one element, we can obtain various metric spaces by choosing different metrics, just like the normed space. The reader is invited to prove that both are certified metrics by satisfying the axioms first and then by induction.

Example 136 *A metric similar to the usual metric can also be defined for the complex plane $\mathbb{C}$. Try to recall the definition of the modulus $|z| = \sqrt{x^2 + y^2}$ for $z = x + iy$ and then derive a metric which can be generalised for $\mathbb{C}^n$*

Example 137 *The sequence space l^p for $1 \leq p < \infty$ forms a metric space under the metric*

$$d(x, y) = \left(\sum_{i=1}^{\infty} |\xi_i - \varsigma_i|^p \right)^{1/p}$$

Example 138 *As a set l^∞ we take the set of all bounded sequences of complex (or real) numbers; that is, every element x of l^∞ is a complex (resp. real) sequence $x = (\xi_1, \xi_2, ...)$, briefly $x = (\xi_i)$. If we have $x = (\xi_i)$ and $y = (\varsigma_i)$, we can have the metric defined by $d(x, y) = \sup A$ where $A = \{\alpha_i \mid \alpha_i = |\xi_i - \varsigma_i|\}$. This can be compactly written as*

$$\sup_{i \in \mathbb{N}} |\xi_i - \varsigma_i|$$

The supremum exists since the set is bounded (why?) and is unique. To prove that this is a metric is easy: D1, D2 and D3 can be easily satisfied. For D4, we can construct the sets $\{\alpha_i \mid \alpha_i = |\xi_i - \varsigma_i|\}$, $\{\beta_i \mid \beta_i = |\xi_i - \eta_i|\}$ and $\{\gamma_i \mid \gamma_i = |\eta_i - \varsigma_i|\}$ for $x = (\xi_i)$, $y = (\varsigma_i)$ and $z = (\eta_i)$ but before that, we need $|\xi_i - \varsigma_i| = |\xi_i - \eta_i + \eta_i - \varsigma_i| \leq |\xi_i - \eta_i| + |\eta_i - \varsigma_i|$ $\forall i$ which we can denote with α_i, β_i and γ_i respectively. Since $\alpha_i \leq \beta_i + \gamma_i$ is valid $\forall i$, we can have $\sup\{\alpha_i\} \leq \sup\{\beta_i\} + \sup\{\gamma_i\}$ or $d(x, y) \leq d(y, z) + d(z, x)$.

Example 139 *What about sequences that are unbounded? We can also have a metric on the space s of all bounded and unbounded sequences defined as*

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|\xi_i - \varsigma_i|}{1 + |\xi_i - \varsigma_i|}$$

for $x = (\xi_i)$ and $y = (\varsigma_i)$. Proving D1 to D3 is easy. For D4, let $z = (\eta_i)$ such that

$$\begin{aligned} |\xi_i - \varsigma_i| &\leq |\xi_i - \eta_i| + |\eta_i - \varsigma_i| \\ \implies & \frac{1}{1+|\xi_i - \eta_i|} + \frac{1}{1+|\eta_i - \varsigma_i|}. \\ &\leq \frac{1}{2+|\xi_i - \varsigma_i|} \\ &\leq \frac{1}{1+|\xi_i - \varsigma_i|} \end{aligned}$$

Note that

$$\begin{aligned} d(x,y) &= \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|\xi_i - \varsigma_i|}{1+|\xi_i - \varsigma_i|} \\ &= \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{-1+1+|\xi_i - \varsigma_i|}{1+|\xi_i - \varsigma_i|} \\ &= \sum_{i=1}^{\infty} \frac{1}{2^i} \left(1 - \frac{1}{1+|\xi_i - \varsigma_i|}\right) \\ &= \sum_{i=1}^{\infty} \frac{1}{2^i} - \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{1}{1+|\xi_i - \varsigma_i|} \\ &= 1 - \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{1}{1+|\xi_i - \varsigma_i|} \end{aligned}$$

Since we have $\frac{1}{2^i}\frac{1}{1+|\xi_i-\eta_i|} + \frac{1}{2^i}\frac{1}{1+|\eta_i-\varsigma_i|.} \leq \frac{1}{2^i}\frac{1}{1+|\xi_i-\varsigma_i|}$, we can equivalently have

$$\begin{aligned} & 1 - \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{1}{1+|\xi_i - \eta_i|} + 1 - \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{1}{1+|\eta_i - \varsigma_i|}. \\ \geq\ & 1 - \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{1}{1+|\xi_i - \varsigma_i|} \end{aligned}$$

which is what is required.

Example 140 *For $C[a,b]$, we have a bona fide metric space under the metric $d(x,y) = \max\limits_{t\in[a,b]} |x(t) - y(t)|$, which the reader is required to prove.*

In real analysis, you might have proved that every bounded function has a supremum using the fact that every function which is bounded above will have a least upper bound. You should have also noticed, while trying to prove the above example, that we have a maximum over here, instead of the supremum. This is because not every function has a supremum but it will have a maximum. This leads us to

Example 141 *The space $B[a,b]$ of real, bounded functions. By definition, each element $x(t) \in B[a,b]$ is a function defined and bounded on a given set $[a,b]$ and the metric is defined by*

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

Now,

$$\begin{aligned} d(x,y) &= 0 \\ &\iff \sup_{t\in[a,b]} |x(t) - y(t)| = 0 \\ &\iff |x(t) - y(t)| = 0 \forall t \in [a,b] \end{aligned}$$

since if the supremum of non-negative numbers zero, then all the numbers are themselves zero. Then, we have $x(t) = y(t)\ \forall t \in [a,b]$. Hence, $x = y$. Second, the supremum of non-negative numbers is non-negative, which the reader is required to prove rigorously. For D3,

$$\begin{aligned} d(x,y) &= \sup_{t\in[a,b]} |x(t) - y(t)| \\ &= \sup_{t\in[a,b]} |-(x(t) - y(t))| \\ &= \sup_{t\in[a,b]} |-x(t) + y(t)| \\ &= \sup_{t\in[a,b]} |y(t) - x(t)| \\ &= d(y,x) \end{aligned}$$

Finally,

$$\sup_{t\in[a,b]} |x(t) - y(t)| \leq \sup_{t\in[a,b]} |x(t) - z(t)| + \sup_{t\in[a,b]} |z(t) - y(t)|$$

This can be made rigorous by resorting to the definition of order and applying the triangle inequality.

Exercise 142 *Show that another metric can be obtained on s by replacing $1/2^i$ with $\mu_i > 0$ such that $\Sigma\mu_i < \infty$*

1.10.1 Balls and Spheres

To make use of sequences in a space, we need machinery which can tell us whether or not the sequence converges or not. Other than that, the space itself can give us good hints about the behaviour of sequences and series. This is possible when we know what open and closed sets are. In a metric space (X,d), we have the following

Definition 143 *Given a point $x_0 \in X$ and a real number $r > 0$, we define three types of sets:*

1. $B(x_0; r) = \{x \in X \mid d(x, x_0) < r\}$ (Open ball)
2. $\bar{B}(x_0; r) = \{x \in X \mid d(x, x_0) \leq r\}$ (Closed ball)
3. $S(x_0; r) = \{x \in X \mid d(x, x_0) = r\}$ (Sphere)

Intuitively, it is clear that in all three cases, x_0 is the centre and r the radius. Mathematically, the open ball of radius r is the set of all points in X whose distance from the centre of the ball is than r. Try to prove that $S(x_0; r) = \bar{B}(x_0; r) - B(x_0; r)$. Remember, these are two sets and you will have to employ set-theoretic arguments.

In Analysis, an open set is one in which every element has an open neighbourhood contained in that set, using the $\epsilon - \delta$ definition. In complete analogy (generalisation, rather), we have

Definition 144 *$M \subset X$ is said to be **open** if it contains an open ball about each of its points. $K \subset X$ is said to be **closed** $K^{\complement} = X - K$ is open*

According to Real or Complex Analysis, the open set should contain ϵ-neighbourhood. We can have this when we replace r with ϵ in an open ball. Thus, $B(x_0; \epsilon)$ is an ϵ-neighbourhood of x_0 where $\epsilon > 0$. We will agree to call a simple neighbourhood N of x_0 as a subset of X which contains an ϵ-neighbourhood of x_0. This ϵ can be arbitrary (but positive!). The letter N will be reserved for such neighbourhoods. It is unfortunate that our standard choice for a symbol of a normed space is the same but the context will clear any confusions.

Trivially, every neighbourhood of x_0 contains x_0. This will be called an **interior point**. We can collect all interior points of an open set M and give it a special name – the **interior** of M. This will be denoted by M^o. You might find that $\text{Int}(M)$ is reserved for such a set in some literature. This set has some interesting properties, as we shall see.

Clearly, if any neighbourhood N of x_0 is contained in an open set M, then M is also a neighborhood of x_0 (proof?). An alternative definition of open sets is as follows, with a trivial proof:

Exercise 145 *A set M is open in a metric space (X, d) if and only if every point is an interior point.*

Theorem 146 *M^o is the largest open set contained in M*

Proof. Clearly, $M^o \subseteq M$ by definition. To prove that this is the largest such set, assume that there exists another open set O such that $M^o \subseteq O \subseteq M$. Then, let $x \in A = O - M^o \implies x$ is not an interior point of X. In particular, it is not an interior point of O. Since x is arbitrary, therefore O is not an open set, establishing the required theorem. ■

It is not difficult to show that the collection τ of all open subsets of X has the following properties:

(τ_l) $\varnothing \in \tau$, $X \in \tau$.

(τ_2) The union of any members of τ is a member of τ.

(τ_3) The intersection of finitely many members of τ is a member of τ.

Proof. Vacuously, every point of $\varnothing$ is an interior point. Hence, it is open. Second, a ball of any radius, when and if constructed around any point of X will naturally be contained in X. Hence, X is open, too. This settles τ_l. For τ_l, let $B_1(x_0; r_1)$ and $B_2(y_0; r_2)$ be two open sets. Assign the value $\max(r_1, r_2, d(x_0, y_0))$ to r. Then, let A be a collection of x such that $d(x, y) < r$ and for a $y \in B_1(x_0; r_1) \cup B_2(y_0; r_2)$. This set is clearly open for the interior point y. The argument can be applied to arbitrary sets. Moving on to τ_3, if $B_1(x_0; r_1) \cap B_2(y_0; r_2) = \varnothing$, then we are done. Assume

$$B_1(x_0; r_1) \cap B_2(y_0; r_2) \neq \varnothing$$

Fix $y \in B_1(x_0; r_1) \cap B_2(y_0; r_2)$, let $r = \min(r_1, r_2)$ and take

$$\forall x \in B_1(x_0; r_1) \cap B_2(y_0; r_2)$$

such that $d(x, y) < r$ and we have our open set with interior point y. This cannot be extended indefinitely since $\min(r_1, r_2, ...)$ may be zero, giving us a singleton as the intersection, or even the empty set. As an example, consider the interval $\left(-\frac{1}{n}, \frac{1}{n}\right) = A_n$. Then, $\bigcap_n A_n = \{0\}$ ■

In analogy to interior points, we define the boundary point of a closed set: a boundary point or a limit point (synonymous with point of accumulation) $x_0 \in X$ of a set $A \subset X$ is such that $\forall \epsilon > 0$, an open ball $B(x_0; \epsilon)$ contains points of A other than x_0. Notice that x_0 need not be a member of A. The set of all limit points of a set A is denoted by A^d. The closure of a set A, written as $Cl(A)$ or even $\bar{A}$ is $A \cup A^d$. If a point is not a limit point, then it is an isolated point.

Exercise 147 *For any A, $Cl(A)$ is closed. Moreover, $Cl(A)$ is the smallest closed set containing A.*

Proposition 148 *If a set contains isolated points only, then that set is finite*

Proof. Let (X, d) be a metric space. Assume that we have a set $M \subset X$ which is infinite. Since we have infinite elements, we can collect elements x such that $d(x, x_0) < \epsilon$ for all ϵ. Such a collection implies that x_0 is a limit point, contradicting our hypothesis. Therefore, M is finite. ■

The crucial point of the proof relies on the fact that this collection is valid *for all* ϵ

Corollary 149 *Every neighbourhood of a limit point contains infinitely many points.*

Exercise 150 *A set is closed if and only if it contains all its boundary points. That is, A is closed if and only if $A = Cl(A)$*

Exercise 151 *A set S is closed if and only if every sequence converges in S.*

Definition 152 *Let (X,d) be a metric space and $A \subset X$. The **diameter** of a set A is $\delta(A) = diam(A) = \sup\{d(x,y) \mid x,y \in A\}$*

Example 153 *For the usual (Euclidean) metric space, $diam(\mathbb{Z}) = diam(\mathbb{P}) = diam(\mathbb{Q}) = diam(\mathbb{R}) = \infty$ whereas for $A = \{y \mid y = x^2, 0 \leq x \leq 5\} = [0,25]$ the diameter for this set is 25.*

That is, the greatest distance between any two points. If the set were circular in shape, then this definition would make sense. A set is called **bounded** if its diameter is finite.

Proposition 154 *In the usual metric space $(\mathbb{R},d)$, a subset A is bounded if and only if it is bounded above and bounded below*

Proof. $diam(A) = \sup\{d(x,y) \mid x,y \in A\} = c < \infty \iff d(x,y) \leq c$ for all $x,y \in A \iff |x-y| \leq c$ for all $x,y \in A \iff -c \leq x-y \leq c \iff -b \leq x \leq b$ for all x. ■

Definition 155 *A subset M of a metric space X is said to be **dense** in X if $\bar{M} = X$ or X is said to be **separable** if it has a countable subset which is dense in X.*

That is, we can have limit points of A within it and the result can equal to the parent set. One good example is the set of rationals and the real set. We all know that the set of rationals is not complete. In Analysis, real numbers can be constructed using Dedekind cuts or the addition of limits to every Cauchy sequence. That is, $\bar{\mathbb{Q}} = \mathbb{R}$ or that the set of rationals are dense in the set of reals. The complex plane, too, can be separated from the irrational real and imaginary parts against the rational ones. This has importance in the theory of operators, which will be glimpsed upon.

More technically, if M is dense in X, then $\forall x_0 \in X$ and $\forall \epsilon > 0$, $B(x_0;\epsilon)$ will contain points of M; or, in other words, in this case there is no point $x_0 \in X$ which has a neighbourhood that does not contain points of M. This is the direct consequence of the definition of a limit point.

Exercise 156 *The following conditions are equivalent:*

1. *M is dense in X*
2. *For every $x \in X$, there exists a sequence in M which converges in X.*
3. *Every nonempty open subset of X contains an element of M.*

To hit the point home, we will prove that the space l^p is separable for $1 \leq p < \infty$. We will proceed as follows: we will first construct a countable subset then basing our argument on the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$, construct limits for every sequence of elements of l^p which will be limit points of sequences in the constructed subset.

Proof. Let M be the set of all sequences x of the form $x = (\xi_1, \xi_2, \ldots, \xi_n, 0, 0, \ldots)$ where n is any integer. Now, we can assume that $\xi_i \in \mathbb{Q}\ \forall i$ since we're only discussing real or complex sequences. Since $\mathbb{Q}$ is countable, $\mathbb{Q}^n$ is therefore countable, leading us to a countable M. That justifies one part of the definition. To prove that $\bar{M} = l^p$. Take $y = (\eta_i) \in l^p$. Since this is convergent, we have

$$\sum_{k=n+1}^{\infty} |\eta_k|^p < \frac{\epsilon^p}{2}$$

Since $\bar{\mathbb{Q}} = \mathbb{R}$, for each η_k there is a rational ξ_k close to it. Hence, we can find an $x \in M$ such that

$$\sum_{k=1}^{n} |\xi_k - \eta_k|^p < \frac{\epsilon^p}{2}$$

Since

$$d(x, y) = \left(\sum_{k=1}^{\infty} |\xi_k - \eta_k|^p \right)^{1/p}$$

We therefore have

$$[d(x, y)]^p = \sum_{k=n+1}^{\infty} |\eta_k - 0|^p + \sum_{k=1}^{n} |\xi_k - \eta_k|^p < \frac{\epsilon^p}{2} + \frac{\epsilon^p}{2} = \epsilon^p$$

or that $d(x, y) < \epsilon$. That is, every sequence x will have a limit point y. Hence, $\bar{M} = l^p$ ■

This should in no way mean that every collection of convergent sequences forms a separable set.

Proposition 157 *The space l^∞ is not separable.*

As a reminder, it is mentioned that the proof that l^∞ forms a metric space was left to the reader.

Proof. Let $x = (\xi_1, \xi_2, \ldots)$ be a sequence of zeros and ones. Then x converges and hence $x \in l^\infty$. With x we associate $y = \left(\frac{\xi_i}{2^i} \right)$. Clearly, $y \in [0, 1]$. Since

$$d(x, y) = \sup_i \left| \xi_i - \frac{\xi_i}{2^i} \right| = \sup_i \left| \frac{\xi_i}{2^i} \right| = \frac{1}{2^i}$$

If these sequences are the centre of an open ball of diameter $\frac{1}{2^{i+1}}$ then these balls do not intersect and we have uncountably many of them, since $[0, 1]$ is uncountable (this is because the set of reals is uncountable and any interval is isomorphic to the real line). If $M \subset l^\infty$ such that $\bar{M} = l^\infty$, then each of these nonintersecting balls must contain an element of M implying that M is uncountable, contradicting the hypothesis that M is countable. ■

Proposition 158 *A discrete metric space (X, d) is separable if and only if X is countable.*

Proof. Let $X = \{x_1, x_2, ..., x_n\}$ be a countable set and let $M = \{x_i, ..., x_j\}$ be a subset for $1 \leq i, j \leq n$. Then, for $d(x_i, x_j) < \epsilon$, we have $x_i = x_j$ if ϵ is close to zero. Hence, any open ball $B(x_k; \epsilon)$ for $1 \leq k \leq n$ will contain only the element x_k and no point is a limit point. Hence, no proper subset of X can have a limit point. Therefore, any M will not be dense in X. Since there are no limit points in X, $X^d = \varnothing$. We therefore have $\bar{X} = X \cup X^d = X$. Hence, X is dense in X and, therefore, separable.

Conversely, assume that X is separable, that is, $\exists M \subset X$ such that $\bar{M} = X$ but M^d is empty for the same reason as above. Therefore, $\bar{M} = M \cup M^d = X$ implies $M = X$ is the only possible subset. Since M (or X) does not have any limit points, every point is an isolated point. Hence, the set is countable. ■

Simply put, no proper subset of separable X can have limit points if the metric is discrete. It is surprising that we can have limit points in a specific metric space whenever we can count the elements of that space, and conversely.

1.10.2 Sequences

Definition 159 *A sequence (x_n) in a metric space (X, d) is said to* ***converge*** *or* ***to be convergent*** *if there is an $x \in X$ such that $\lim_{n\to\infty} d(x_n, x) = 0$. Alternatively, a sequence is called convergent if $\forall \epsilon > 0$, $\exists N$ such that $d(x_n, x) < \epsilon$ whenever $n > N$.*

In such a case, x is called the limit of (x_n) and x_n is said to converge to x. This is denoted by $\lim_{n\to\infty} x_n = x$ or $x_n \longrightarrow x$ as $n \longrightarrow \infty$. If we cannot find an N for any given ϵ, or that if the sequence fails to be convergent, we say that this sequence diverges. A sequence (x_n) is bounded if its range (x_n) is bounded.

Exercise 160 *In the case of norm spaces, we have $\lim_{n\to\infty} \|x_n - x\| = 0$. Equivalently, $\forall \epsilon > 0$, $\exists N$ such that $\|x_n - x\| < \epsilon$ whenever $n > N$. Show that these definitions are equivalent.*

If a sequence converges, its limit is unique

Proof. Let $\lim_{n\to\infty} x_n = l_1$ and $\lim_{n\to\infty} x_n = l_2$ be two limits. Then, $\forall \epsilon > 0$, we can find N_1 and N_2 such that $d(x_n, l_1) < \epsilon/2$ and $d(x_n, l_2) < \epsilon/2$ for $n > N_1, N_2$. Let $N = \max(N_1, N_2)$. Then, $d(l_1, l_2) < d(x_n, l_1) + d(x_n, l_2) < \epsilon/2 + \epsilon/2 = \epsilon$ whenever $n > N$, $\forall \epsilon > 0$. The condition $d(l_1, l_2) < \epsilon$ implies $l_1 = l_2$ ■

Exercise 161 *Let α, α_n be scalars. If $x_n \longrightarrow x$ and $\alpha_n \longrightarrow \alpha$, then*

$$\alpha_n x_n \longrightarrow \alpha x$$

Sticking with our convention with analysis, a function is continuous at x if $\lim_{n\to\infty} f(x_n) = f(x)$ provided that $\lim_{n\to\infty} x_n = x$. This can be extended to mulitple variables, of course. Therefore, by the above exercise, scalar multiplication is a continuous function.

Exercise 162 *If $x_n \longrightarrow x$ and $y_n \longrightarrow y$, then $x_n + y_n \longrightarrow x + y$*

Thus, addition is a continuous function. You can take hints from the introduction of the space $C[a, b]$ in this monologue.

Other than a sequence, we can also take out some members of the sequence to form a subsequence. That "order" of the sequence has to be maintained so that in a subsequence formed by x_n and, say, x_{n+5}, we can have $x_n = a_1$ and $x_{n+5} = a_2$. In this case, we have $x_{n_k} = a_k$. Thus, for x_i, a certified subsequence is $x_{n_1}, x_{n_2}, ...$ where n_1 is some i, n_2 is another natural number greater than i and so on.

Exercise 163 *If a sequence converges to a point, then any subsequence will converge to that point.*

Proposition 164 *If a sequence converges, then it is bounded.*

Proof. Let $x_n \longrightarrow x$. Then, we can be assured that we will definitely have a (very large) natural number N such that $d(x_n, x) < \epsilon \ \forall n > N$ and $\forall \epsilon > 0$. Let $m = \max(d(x_1, x), d(x_2, x), ..., d(x_N, x), \epsilon)$. Then, $d(x_n, x) < m \ \forall n$ which really means that every element of the sequence is bounded. ■

The converse, however, is usually false. Consider the series $\frac{(-1)^n}{2}$. This series is bounded by ± 1 yet does not converge as it keeps on alternating between $-1/2$ and $1/2$.

Now, try to prove that the sum of two bounded series and a scalar multiple of a bounded sequence is bounded, to complete the proof that a collection of such sequences forms a vector space.

Corollary 165 *If a sequence is unbounded, then it is divergent.*

Proposition 166 *If x is a limit point of a subset A of a metric space (X, d), then there exists a sequence such that $x_n \longrightarrow x$.*

Proof. What we need to do is construct a sequence that converges to this limit point. Since x is a limit point, then we can rest assured that we have an open ball centred at x of ϵ radius contained in X and containing points other than x, by definition. Hence, we can collect such points and call them x_n. Therefore, $d(x_n, x) < \epsilon$. What we need now is to prove that we have an N such that this is valid for $n > N$. Since epsilon was arbitrary, we can let it depend on the index n. So, $\epsilon = 1/n$, say. From a collection of the natural numbers, we will always have an N such that $N\epsilon < 1$. This can be seeing by applying the Archimedean property of real numbers. Now, we have $1 > N\epsilon$ or $1 > N/n$ or $n > N$, which establishes the proof. ■

Proposition 167 *Every real sequence has a monotone subsequence*

Proof. A x_n sequence is monotonic if $x_k \leq x_{k+1}$ or if $x_k \geq x_{k+1}$ for all k. Hence we can collect such points and form the required subseuqence. ■

Proposition 168 (Monotone convergence theorem) *Every bounded above monotonically increasing sequence converges to its supremum*

For supermum to make sense, we must have order on the set. We, therefore, restrict this proof to that of the reals.

Proof. Let $x_k \leq x = \sup_k x_k$ for all k. From $x_k \leq x_{k+1}$, for every $\epsilon > 0$, we have an integer N such that $x - \epsilon < x_N \leq x$ for otherwise $x - \epsilon$ would be an upper bound of x_k. Since x_k increases, $n \geq N$ implies $x - \epsilon < x_n \leq x < x + \epsilon$ or $|x_n - x| < \epsilon$ for all $n \geq N$ ■

A similar proof is left as an exercise:

Exercise 169 *Every bounded below monotonically decreasing sequence converges to its infimum*

Theorem 170 *Let M be a nonempty subset of a metric space (X, d). Then*

1. $x \in \bar{M} \iff \exists x_n \in M$ such that $x_n \longrightarrow x$
2. M is closed $\iff$ $x_n \in M$ such that $x_n \longrightarrow x$ implies that $x \in M$.

Proof. For bullet 1, we've proven that any limit point will have a sequence convergent to it. The converse is a trivial result of the definition of convergence and limit points. Bullet two follows by observing that if M is closed, then $M = \bar{M}$ ■

Theorem 171 *Every real, bounded sequence has at least one convergent subsequence*

Proof. If we call the bound x, then we can let a monotonic sequence converge to that point, establishing the theorem. Details are left to the reader. ■

This is the famous Bolzano-Weistrass theorem and will be made use of extensively.

Proposition 172 *If $x_n \longrightarrow x$ and $y_n \longrightarrow y$ in X, then $d(x_n, y_n) \longrightarrow d(x, y)$.*

Proof. $\forall \epsilon > 0$, we can find N_1 and N_2 such that

$$d(x_n, x) < \epsilon/2$$

and

$$d(y_n, y) < \epsilon/2$$

for $n > N_1, N_2$. Let $N = \max(N_1, N_2)$. Then,

$$\begin{aligned} d(x_n, y_n) &\leq d(x_n, x) + d(x, y) + d(y_n, y) \\ &\Longrightarrow d(x_n, y_n) - d(x, y) \leq d(x_n, x) + d(y_n, y) \end{aligned}$$

Also,

$$\begin{aligned} d(x, y) &\leq d(x, x_n) + d(x_n, y_n) + d(y_n, y) \\ &\Longrightarrow -[d(x, x_n) + d(y_n, y)] \leq d(x_n, y_n) - d(x, y) \end{aligned}$$

Using the two inequalities, we have

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y) < \epsilon/2 + \epsilon/2 = \epsilon$$

i.e. $|d(x_n, y_n) - d(x, y)| < \epsilon \ \forall n > N$. ■

Notice that in this proof, we've treated $d(x_n, y_n)$ as a sequence with the index n. Thus, the metric function is continuous.

Now it's time to move on to another very useful type of sequence.

Definition 173 *A sequence (x_n) in a metric space (X, d) is said to **Cauchy** if $\forall \epsilon > 0$, $\exists N$ such that $d(x_n, x_m) < \epsilon$ whenever $n, m > N$.*

This can understood to mean that after a certain number of elements, the elements of the sequence become arbitrarily close together. However, this in no way means that every Cauchy sequence converges. Find a counter example. First, try to prove that every convergent sequence is Cauchy i.e. if a sequence approaches to a point, then after a certain number of elements of the sequence, the elements of the sequence themselves will come arbitrarily close together.

Example 174 *The sequence $a_n = \sum_{k=1}^{n} 1/k^2$ is Cauchy since. Assuming that $n \geq m$ with out loss of generality*

$$\begin{aligned}
|a_n - a_m| &= |a_m - a_n| \\
&= \left| \sum_{k=1}^{n} 1/k^2 - \sum_{k=1}^{m} 1/k^2 \right| \\
&= \left| \sum_{k=m+1}^{n} 1/k^2 \right| \\
&= \sum_{k=m+1}^{n} 1/k^2 \leq \sum_{k=N}^{\infty} 1/k^2 \\
&\leq \sum_{k=N}^{\infty} 1/k\,(k-1) \\
&\leq \sum_{k=N}^{\infty} 1/(k-1) - \sum_{k=N}^{\infty} 1/k \\
&= 1/(N-1)
\end{aligned}$$

Proposition 175 *Every Cauchy sequence is bounded*

Proof. Let (x_n) be Cauchy. Then, we can be assured that we will definitely have a (very large) natural number N such that $d(x_n, x_m) < \epsilon \ \forall n, m > N$ and $\forall \epsilon > 0$. Let

$$\max_{1 \leq i,j \leq N} (d(x_i, x_j), \epsilon) = c$$

Then, $d(x_n, x_m) < c \ \forall n, m$. ■

Exercise 176 *Let (x_n) be a sequence in a metric space (X, d). Let*

$$A_k = \{x_{k+1}, x_{k+2}, ...\}$$

Show that $\{x_n\}$ is a Cauchy sequence if and only if $\lim_{k\to\infty} \delta(A_k) = 0$ where $\delta(A_k) = \sup\{d(x,y) \mid x, y \in A_k\}$ is the diameter of the set A_k.

Solution 177 ($\Longrightarrow$)

$\{x_n\}$ *is Cauchy implies* $\forall \epsilon > 0$, $\exists N < m, n$ *such that* $d(x_n, x_m) < \epsilon$. *Now,* $\delta(A_N) = \sup\{d(x_n, x_m) \mid x, y \in A_N\} < \epsilon$

$\Longrightarrow \lim_{N\to\infty} \delta(A_N) = 0$

($\Longleftarrow$)

$\lim_{k\to\infty} \delta(A_k) = 0$

$\Longrightarrow \lim_{k\to\infty} \sup\{d(x_n, x_m) \mid x, y \in A_k\} = 0$ *for* $n, m > k$

$\Longrightarrow d(x_n, x_m) \longrightarrow 0$ *for* $n, m > k$

$\Longrightarrow \{x_n\}$ *is Cauchy*

Some spaces in which every Cauchy sequence does converge have a special place in functional analysis, playing a crucial role in most theorems. Any such space is said to be complete if every Cauchy sequence in it converges (that is, has a limit which is an element in that set). If any space has a Cauchy sequence which does not converge, then that space is incomplete. Otherwise, the space is set to be complete. For instance, the set of rationals, as mentioned that the completion of the set of rationals will be assumed via Dedekind cuts, forming the reals. We can prove that every Cauchy, real sequence converges:

Proof. If we have a Cauchy sequence of real numbers (x_n), we know that it is bounded and if the Cauchy sequence is bounded, then it will have a convergent subsequence (x_{n_k}), as was proved earlier. To make things rigorous, let $\epsilon > 0$. We know that we have a K so that $d(x_{n_k}, l) < \frac{\epsilon}{2}$ whenever $k > K$ and also N so that $d(x_n, x_m) < \frac{\epsilon}{2}$ whenever $n, m > N$. Notice that n_k is a sequence which increases without bound, so that we don't have $n_k > K$ but instead have $k > K$. Now, we can pick a $k > K$ so that $n_k > N$. Then, for every $n > N$

$$d(x_n, l) \leq d(x_n, x_{n_k}) + d(x_{n_k}, l) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus, $d(x_n, l) < \epsilon$ ■

This leads us to the complex numbers and the generalised Euclidean space (now on, we will often use the relation $d(x, y) = \|x - y\|$ frequently without any warning).

Proof. Let $\mathbf{x}_k \in \mathbb{R}^n$ be a Cauchy sequence in the generalised Euclidean space. Here, the subscript k indicates the index, instead of n, which is reserved for dimensionality. Then, we have $\|\mathbf{x}_k - \mathbf{x}_j\| < \frac{\epsilon}{n}$ $\forall j, k > N$ for $N \in \mathbb{N}$ This norm is the usual norm

$$\|\mathbf{x}\| = \sqrt{(x_1)^2 + (x_2)^2 + ... + (x_n)^2}$$

which, as already mentioned, is a generalised version of the Pythagorean theorem. Hence, from $\|\mathbf{x}_k - \mathbf{x}_j\| < \frac{\epsilon}{n}$, we have

$$\sqrt{\left(x_1^{(k)} - x_1^{(j)}\right)^2 + \left(x_2^{(k)} - x_2^{(j)}\right)^2 + ... + \left(x_n^{(k)} - x_n^{(j)}\right)^2} < \frac{\epsilon}{n}$$

$$\implies \left(x_1^{(k)} - x_1^{(j)}\right)^2 + \left(x_2^{(k)} - x_2^{(j)}\right)^2 + ... + \left(x_n^{(k)} - x_n^{(j)}\right)^2 < \epsilon^2$$

$$\implies \left(x_i^{(k)} - x_i^{(j)}\right)^2 < \frac{\epsilon^2}{n^2}$$

$$\implies \sqrt{\left(x_i^{(k)} - x_i^{(j)}\right)^2} < \frac{\epsilon}{n}$$

$$\implies \left|x_i^{(k)} - x_i^{(j)}\right| < \frac{\epsilon}{n}$$

for all the i tuples. Since every i tuple is a real number, we're actually talking about the set of reals, which is complete. That is, every Cauchy sequence converges. Hence, $\left|x_i^{(k)} - x_i^{(j)}\right| < \epsilon$ will converge to, say, x_i. We can thus construct a number $\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$. Now,

$$\begin{aligned}
& \|\mathbf{x}_k - \mathbf{x}\| \\
= \; & \sqrt{\left(x_1^{(k)} - x_1\right)^2 + \left(x_2^{(k)} - x_2\right)^2 + ... + \left(x_n^{(k)} - x_n\right)^2} \\
\leq \; & \sqrt{\left(x_1^{(k)} - x_1\right)^2} + \sqrt{\left(x_2^{(k)} - x_2\right)^2} + ... + \sqrt{\left(x_n^{(k)} - x_n\right)^2} \\
= \; & \left|x_1^{(k)} - x_1\right| + \left|x_2^{(k)} - x_2\right| + ... + \left|x_n^{(k)} - x_n\right| \\
< \; & \frac{\epsilon}{n} + \frac{\epsilon}{n} + ... + \frac{\epsilon}{n} \\
= \; & \epsilon
\end{aligned}$$

That is, $\|\mathbf{x}_k - \mathbf{x}\| < \epsilon$. ■

Exercise 178 $\mathbb{C}^n$ *is complete*

This is first established by the fact that $\mathbb{R}^2 \cong \mathbb{C}$ and then using the above construction.

What we've done over here is that we've taken an arbitrary Cauchy sequence, constructed a limit and then proved that this sequence converges to this limit. The $\cong$ means that the two sets are isomorphic or similar in structure to each other. Loosely speaking, their operations, their behaviour, their multiplication and addition rules behave the same in each individual space.

Theorem 179 *A subspace M of a complete metric space (X, d) is itself complete if and only if $M = \bar{M}$.*

Proof. If M is complete, we can have a limit point onto which a sequence converges belonging to M, establishing the existence of limit points in M. Thus, $M = \bar{M}$.

Conversely, if M is closed, then it contains all its limits points. For a Cauchy sequence x_n in M, $x_n \longrightarrow x \in X$ since X is complete. Since x is a limit point of x_n, it is a limit point of M, establishing that any Cauchy sequence in M converges in M. ■

Theorem 180 *If a Cauchy sequence x_n has a convergent subsequence $x_{n_k} \longrightarrow x$, then*

$$x_n \longrightarrow x$$

Proof. Since we have a convergent subsequence, we therefore have $d(x, x_{n_k}) < \epsilon/2$ for all ϵ, where we are promised the existence of a natural number N_1 such that $n_k > N$. We can also enumerate these x_{n_k}'s such that $d(x_{n_k}, x_n) < \epsilon/2$ for $n > N$ because we have a Cauchy sequence. It remains trivial then to see that $d(x, x_n) \leq d(x, x_{n_k}) + d(x_{n_k}, x_n) < \epsilon/2 + \epsilon/2 = \epsilon$ ■

Exercise 181 *A Cauchy sequence of real or complex numbers is convergent if and only if it has a convergent subsequence.*

1.10.3 Continuity

Open sets also play a role in connection with continuous mappings, where continuity is a natural generalisation of the continuity known from complex and real analysis and is defined as follows:

Definition 182 *Let (X, d_1) and (Y, d_2) be metric spaces. A mapping*

$$T : X \longrightarrow Y$$

*is said to be **continuous** at a point $x_0 \in X$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that*

$$d_2(T(x), T(x_0)) < \epsilon$$

whenever $d_1(x, x_0) < \delta$.

T is said to be continuous if it is continuous at every point of X. Alternatively, this definition could be phrased as follows:

Theorem 183 *A mapping T of a metric space X into a metric space Y is continuous if and only if the inverse image of any open subset of Y is an open subset of X.*

Proof. Let $B \subset Y$ be an open set and let $T^{-1}(B) = \{x \mid T(x) \in B\}$. We need to prove that $T^{-1}(B)$ is open. Let $T(x_0) \in B$. Since $T(x_0)$ is an interior point, we have $d_2(T(x), T(x_0)) < \epsilon \ \forall \epsilon > 0$. Since T is continuous, this ensures the existence of a δ such that $d_1(x, x_0) < \delta$. Hence for any ϵ or for any open set, we

can find a δ or an open set $A(x_0;\delta) = \{x \mid d_1(x, x_0) < \delta\}$. Hence the inverse image of every open set is open.

The converse of the proof is trivial. We start with $B(T(x_0);\epsilon)$, an open set, such that $T^{-1}(B) = A$ is open by suggesting that this set satisfies $d(x, x_0) < \delta$ for every $x \in A$, guaranteeing the existence of the required δ. ■

Theorem 184 *Let (X, d_1) and (Y, d_2) be metric spaces. A mapping*

$$T : X \longrightarrow Y$$

is continuous at a point $x \in X \iff$

$$x_n \longrightarrow x \implies T(x_n) \longrightarrow T(x)$$

Proof. If T is continuous at x, then for every $\epsilon > 0$ there exists a $\delta > 0$ such that $d_2(T(y), T(x)) < \epsilon$ whenever $d_1(y, x) < \delta$. If $x_n \longrightarrow x$, then we can label the points y when we have an $n > N$ so that $d_1(x_n, x) < \delta$, which is possible when $d_2(T(x_n), T(x)) < \epsilon$, for every $\epsilon > 0$ and $n > N$.

The converse of the proof is trivial. ■

Thus, the metric function is continuous. We therefore see that we have continuity regarding functions themselves but what about sequences?

Of particular interest is the convergence of sequence of functions f_n. However, in this case, we also have to consider the domain of the functions, as well. In the ordinary notion of continuity, this convergence will depend on each point of the domain, giving the name "point-wise convergence". Apart from this notion of continuity, we also have the notion of uniform continuity, in which the elements of the domain do not matter. Thus, in uniform continuity, we have

Definition 185 *A sequence of functions $f_n(x)$ **converges uniformly** if $\forall \epsilon > 0$, $\exists N$ such that $d(f_n(x), f(x)) < \epsilon$ $\forall x$ whenever $n > N$*

In uniform convergence, we have convergence of functions for every element of the domain. This type of convergence is important when dealing with spaces involving continuous functions. In fact, if a sequence of function converges to a function, that is $\lim_{n\to\infty} f_n = f$, then this is valid for all x. That is, $\lim_{n\to\infty} f_n(x) = f(x)$. That ϵ in the definition will depend upon x if the convergence is point-wise and will not if the convergence is uniform.

There is another notion of convergence in the space of functions:

Definition 186 *A sequence of functions $f_n(x)$ **converges pointwise** if $\forall \epsilon > 0$ $\forall x$, $\exists N$ such that $d(f_n(x), f(x)) < \epsilon$ whenever $n > N$.*

The difference is subtle: here N depends both on x and ϵ whereas in the former, for each ϵ you need to be able to find an N for all x in the domain of the function. In other words, N can depend on ϵ but not on x. Like uniform and ordinary continuity, the former definition is global in nature whereas the other talks about convergence depending on the domain.

Exercise 187 *Show the uniform convergence implies pointwise convergence but not conversely.*

Theorem 188 *If a series of functions converges uniformly, then the limit is continuous*

Proof. Let $f_n(x) \longrightarrow f(x)$ uniformly. Then, we have an N such that $\forall \epsilon > 0 \, \forall x$, $d(f_n(x), f(x)) < \frac{\epsilon}{3}$ whenever $n > N$. We also have continuous $f_n(x)$ so that $\forall \epsilon > 0$, $\exists \delta$ such that $d(f_n(x), f_n(y)) < \frac{\epsilon}{3}$ whenever $d(x,y) < \delta$. The uniform continuity is valid for all $x \in \mathcal{D}(f)$ and, in particular, whenever $d(x,y) < \delta$. Hence, whenever $d(x,y) < \delta$, we have

$$\begin{aligned} d(f(x), f(y)) &\leq d(f_n(x), f(x)) + d(f_n(x), f_n(y)) + d(f_n(y), f(y)) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon \end{aligned}$$

so that $f(x)$ is continuous. ■

Note that the above proof has to be valid for all x. Hence, convergence in $C[a,b]$ is always uniform and never point-wise.

Now that we have some machinery regarding Cauchy sequence, let's see some other examples of complete metric spaces, which include l^p, l^∞, c and $C[a,b]$.

Proof. The completeness of l^p will follow a similar pattern; we will take an arbitrary Cauchy sequence, construct a limit in l^p and then show that this sequence converges to that limit. So we have

$$d(x_n x_m) = \left(\Sigma \left| \xi_i^{(n)} - \xi_i^{(m)} \right|^p \right)^{1/p} < \epsilon$$

from which we have

$$\left| \xi_i^{(n)} - \xi_i^{(m)} \right| < \epsilon$$

Whether real or complex, each $\xi_i^{(n)} \longrightarrow \xi_i$ because each ith sequence is a member of a space of convergent Cauchy sequences. Constructing $x = (\xi_i)$, we need to show that this sequence of sequences is bounded under the p-norm for it to belong to this space. Since we're talking about a sequence of sequences, we can have $\left(\Sigma \left| \xi_i^{(n)} - \xi_i^{(m)} \right|^p \right)^{1/p} < \frac{\epsilon}{\sqrt[p]{2}}$. Using this, we have

$$\begin{aligned} & \left(\Sigma \left| \xi_i - \xi_i^{(m)} \right|^p \right) \\ = & \left(\sum_{i=1}^{k} \left| \xi_i - \xi_i^{(m)} + \xi_i^{(m)} - \xi_i^{(n)} \right|^p \right) \\ \leq & \left(\sum_{i=1}^{k} \left| \xi_i - \xi_i^{(m)} \right|^p + \sum_{i=1}^{k} \left| \xi_i^{(m)} - \xi_i^{(n)} \right|^p \right) \\ < & \left(\frac{\epsilon}{\sqrt[p]{2}} \right)^p + \left(\frac{\epsilon}{\sqrt[p]{2}} \right)^p \\ = & \epsilon^p \end{aligned}$$

We will then have $d(x_m, x) = \left(\Sigma \left|\xi_i^{(m)} - \xi_i\right|^p\right)^{1/p} < \epsilon$, implying that the Cauchy sequence converges. In particular, we have

$$d(x_m, x) = \left(\Sigma \left|\xi_i^{(m)} - \xi_i\right|^p\right) < \epsilon^p = c$$

implying that the sequence $\left(\xi_i^{(m)} - \xi_i\right)$ belongs to the space. Since x_m was a member and $x_m - x$ is a member, we can use the fact the vector addition is closed to say that $x = x_m - (x_m - x)$ belongs to the space. ■

In a similar way, we can prove that l^∞ is complete.

Proof. Taking an arbitrary Cauchy sequence $x_m = \left(\xi_i^{(m)}\right)$, we have

$$d(x_m, x_n) = \sup_i \left|\xi_i^{(m)} - \xi_i^{(n)}\right| < \epsilon$$

for all $n, m > N$. The definition of supremum implies that $\left|\xi_i^{(m)} - \xi_i^{(n)}\right| < \epsilon$ for all i. Again, the ith Cauchy sequence $\xi_i^{(n)}$ converges to, say, ith limit ξ_i. Now, we construct (ξ_i). To prove that $(\xi_i) \in l^\infty$, observe that

$$\left|\xi_i^{(m)} - \xi_i^{(n)}\right| < \epsilon \implies \left|\xi_i^{(m)} - \xi_i\right| \leq \epsilon$$

This is valid whenever $m > N$. For any sequence $y = (\eta_i) \in l^\infty$, we have $|\eta_i| \leq c_y$. Similarly, $\xi_i^{(m)} \in l^\infty$ implies that $\left|\xi_i^{(m)}\right| \leq c_m$ where this constant depends on the mth term of x_m. Then,

$$\begin{array}{rl} & |\xi_i| \\ \leq & \left|\xi_i^{(m)} - \xi_i\right| + \left|\xi_i^{(m)}\right| \\ \leq & \epsilon + c_m = k \\ \implies & |\xi_i| \text{ is bounded} \\ \implies & \sup|\xi_i| \text{ exists} \\ \implies & \sup\left|\xi_i^{(m)} - \xi_i\right| \text{ exists} \end{array}$$

Since $\left|\xi_i^{(m)} - \xi_i\right| \leq \epsilon$, we have $\sup\left|\xi_i^{(m)} - \xi_i\right| \leq \epsilon$ whenever $m > N$ and $\epsilon > 0$ ■

An analogous proof holds for n-tuples $\mathbf{x} = (\xi_1, \xi_2, ..., \xi_n)$ and $\mathbf{y} = (\eta_1, \eta_2, ..., \eta_n)$ under the metric

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

The space c is a subspace of l^∞ and involves not only bounded sequences but also sequences which are convergent. Hence, the collection of convergent sequences also forms a vector space. Try to prove first that the sum of two convergent sequences is a convergent sequence and that a scalar multiple of a convergent sequence is convergent. This space is complete, as well.

Proof. We can pursue this proof by the normal way or take a detour and apply the fact that subspace of a complete metric space is itself complete if and only if the space is closed. Hence, if we can prove that c is closed in l^{∞}, we're done. We take a limit point x of c. This $x = (\xi_i)$. Then, we clearly have a sequence x_n in c which converges to this point. But c is a collection of convergent sequences. Hence, this x belongs to c. Therefore, c is closed. ■

The space $C[a,b]$ is also complete under its usual norm

Proof. For any Cauchy sequence (x_n) we have

$$d(x_n, x_m) = \max_{t\in[a,b]} |x_n(t) - x_m(t)| < \epsilon$$

which says that we can have a $t_0 \in [a,b]$ such that $|x_n(t_0) - x_m(t_0)| < \epsilon$. These are, then, not continuous functions but points in a complete (complex or real) space. Hence, we can have a $x_n(t_0) \longrightarrow x(t_0)$ but this is a real or complex number, not a function. What we need is an arbitrary t, which we can still get by remembering that we had a maximum function. So, we then associate for each t, $x_n(t) \longrightarrow x(t)$. Notice that the convergence will depend on t. So, in this case, we have point-wise convergence. To show that the limit $x(t)$ belongs to the space, we let $n \longrightarrow \infty$ in $d(x_n, x)$ to get $d(x_m, x) \leq d(x_n, x_m) + d(x_n, x) < \epsilon$. Hence, we have

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

This convergence is now uniform since the limit is valid for any $t \in [a,b]$. Since if continuous functions converge uniformly, then the limit will also be continuous, as has been proved earlier. Therefore, $x(t)$ is continuous or that $x(t) \in C[a,b]$. That is, any Cauchy sequence of continuous functions has a limit in the same space, establishing completeness. ■

The above proof indicates why the metric is also referred to as the uniform metric. If we change the norm to the integral norm, the space $C[a,b]$ becomes incomplete. We will see the case for $a = 0$ and $b = 1$.

Theorem 189 *In space $C[a,b]$, convergence is always uniform*

Proof. From $x_n(t) \longrightarrow x(t)$, we have

$$\begin{aligned} d(x_n, x) &= \max_{t\in[a,b]} |x_n(t) - x(t)| < \epsilon \\ &\Longrightarrow \quad |x_n(t) - x(t)| < \epsilon \ \forall t \end{aligned}$$

i.e. the epsilon is independent of the points of the domain. ■

Apart from these well-structured spaces, the discrete metric space is always complete

Proof. Let x_n be a Cauchy sequence. Then, $d(x_n, x_m) < \epsilon$. If this $0 < \epsilon < 1$, then $x_n = x_m$ for $n, m > N$. Thus, with $x = x_N$, we have $d(x_n, x) = 0 < \epsilon$ ■

For incomplete spaces, we start off with the usual set of rationals.

Proof. To show that the set of rationals $\mathbb{Q}$ is incomplete, we take the Cauchy sequence $x_n = \left(1 + \frac{1}{n}\right)^n$. The limit of this sequence is the exponential e. We

show that this sequence is indeed Cauchy. Let $m = n + q$ for an integer q. For sufficiently large n, m under the metric $d(x, y) = |x - y|$, we have

$$\begin{aligned}
d(x_n, x_m) &= \left| \left(1 + \frac{1}{n}\right)^n - \left(1 + \frac{1}{n+q}\right)^{n+q} \right| \\
&= \left| \left(1 + \frac{1}{n}\right)^n - \left(1 + \frac{1}{n+q}\right)^{n+q} \right| \\
&= \left| \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{n}\right)^k - \sum_{k=0}^{n+q} \binom{n+q}{k} \left(\frac{1}{n+q}\right)^k \right| \\
&= \left| \begin{array}{c} \left(1 + 1 + \frac{n-1}{2!n} + \frac{(n-1)(n-2)}{3!n^2} + ... + \frac{1}{n^n}\right) - \\ \left(1 + 1 + \frac{n+q-1}{2!(n+q)} + \frac{(n+q-1)(n+q-2)}{3!(n+q)^2} + ... + \frac{1}{(n+q)^{n+q}}\right) \end{array} \right| \\
&= \left| \begin{array}{c} \frac{n-1}{2!n} + \frac{(n-1)(n-2)}{3!n^2} + ... + \frac{1}{n^n} \\ -\frac{n+q-1}{2!(n+q)} - \frac{(n+q-1)(n+q-2)}{3!(n+q)^2} - ... - \frac{1}{(n+q)^{n+q}} \end{array} \right|
\end{aligned}$$

To this last experssion, we can apply the triangle inequality to each term individually. Hence, for a large n, we have smaller and smaller distances between n and m. In other words, for any ϵ, we can always find a value for n so that the Cauchy condition is certified.

From real analysis, it is well known that $e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + ...$ Hence, we'll show that $\lim_{n\to\infty} x_n = e = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + ...$ is irrational. We'll proceed by contradiction. Assume $e = \frac{p}{q}$. Then,

$$q!e = q! + \frac{q!}{1!} + \frac{q!}{2!} + ... + \frac{q!}{q!} + R$$

where $R = \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + ...$ is the remainder of the terms. Since

$$q!e = (q-1)!p$$

is an integer and

$$q! + \frac{q!}{1!} + \frac{q!}{2!} + ... + \frac{q!}{q!}$$

is also the finite sum of integers (implying that it, itself, is an integer), hence R is an integer but

$$\frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + ... < \frac{1}{q+1} + \frac{1}{(q+1)^2} + ... = \frac{1}{q}$$

so that $R < \frac{1}{q}$ means that R is not an integer. This is our contradiction so we can conclude that $e \neq \frac{p}{q}$

In summary, we have $\lim_{n\to\infty} x_n \notin \mathbb{Q}$ where x_n is Cauchy, implying that $\mathbb{Q}$ is incomplete. ■

Other than the rationals, the set $\mathbb{Z}$ is also incomplete under the metric

$$d(m,n) = |m-n|$$

For, from a Cauchy sequence (x_n), we have

$$d(x_n, x_m) = |m-n| < \epsilon$$

This sequence does not converge for $\epsilon = \frac{1}{N}$ for $n > N$.

The completion for this metric space is the set of reals, as already mentioned.

The space of polynomials $P[a,b]$ is also not complete.

Proof. The Cauchy sequence

$$x_n = \sum_{k=0}^{n} \frac{x}{k!}$$

converges uniformly to the function e^x which is not a polynomial. For, if $n, m > N$, we have

$$\begin{aligned} |x_n - x_m| &= \left| \sum_{k=0}^{n} \frac{x}{k!} - \sum_{k=0}^{m} \frac{x}{k!} \right| \\ &= \left| \sum_{k=\min(n,m)}^{\max(n,m)} \frac{x}{(k+1)!} \right| \end{aligned}$$

which can be made as small as we like. Notice, also that the the epsilon will not depend on x. ■

The space $C[0,1]$ under the norm

$$\|x(t)\| = \int_0^1 |x(t)|\, dt$$

is incomplete.

Proof. If we have a sequence

$$x_n = \begin{cases} 0 \text{ if } 0 \leq t \leq \frac{1}{2} \\ 1 \text{ if } \frac{1}{m} + \frac{1}{2} \leq t \leq 1 \end{cases}$$

of functions, we get a Cauchy sequence if $m, n > 1/\epsilon$. This is because

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

Now, for any $x \in C$, from $\|x_n(t) - x(t)\| < \epsilon$ and $n > N$, we should have

$$\begin{aligned} & \int_0^1 |x_n(t) - x| \, dt \\ = & \int_0^{1/2} |x_n(t) - x| \, dt + \int_{1/2}^{1/2+1/n} |x_n(t) - x| \, dt + \int_{1/2+1/n}^1 |x_n(t) - x| \, dt \\ = & \int_0^{1/2} |x(t)| \, dt + \int_{1/2}^{1/2+1/n} |x_n(t) - x| \, dt + \int_{1/2+1/n}^1 |x(t) - 1| \, dt < \epsilon \end{aligned}$$

In other words, each integral is less than ϵ. Now, recalling that the choice of ϵ is arbitrary, we have $x(t) = 0$ for $0 \leq t < 1/2$ from $\int_0^{1/2} |x(t)| \, dt < \epsilon$. $t = 1/2$ fails because of the middle integral in the range $(1/2, 1/2 + 1n)$ where n is very large. Finally, if the last integral has to be arbitrarily small, we can safely say that $x(t) = 1$ in the range $(1/2 + 1n, 1]$. From this, we can observe that $\lim_{t \to 1/2} x(t)$ fails to exist because $\lim_{t \to 1/2^-} x(t) \neq \lim_{t \to 1/2^+} x(t)$, implying that the function is not continuous, so that we cannot have any limit for the particular Cauchy sequence $x_n(t)$.

This special case of the interval $[0, 1]$ has no particular importance. For instance, the interval $[a, b]$ can be mapped from $[0, 1]$ using the transformation $x = (b - a)t + a$ for $t \in [0, 1]$ so that $C[0, 1]$ is ismorphic to $C[a, b]$. ■

Other than this machinery, we can also complete a metric space. Intuitively, this is done by "adding" limit points so that every Cauchy sequence converges. Of course we can't include elements in the set on our own accord but what we can do is make the set "equal" to another complete subset of another space. This "equality" is not the true equality which we are wired to think of and is based on the definition of a type of isomorphism which follows this paragraph. This strange "equality" of sets makes the two sets indiscernable with respect to their structure and properties but the substance itself differs. For instance, as is known, the set $\mathbb{Q}$ is constructed from $\mathbb{Z}\times(\mathbb{Z}-\{0\})$. The removal of 0 ensures that zero is excluded in the denominator. The resulting set $\mathbb{Q}$ is not TRULY equal to an extension of the set of integers so that it is unfair to state that $\mathbb{Z}\subset\mathbb{Q}$, strictly speaking. $\mathbb{Z}$ happens to a single set whereas $\mathbb{Q}$ is the Cartesian product of both. What can actually be said is that $\mathbb{Z}\times\{1\}\subset\mathbb{Q}$. However, there exists an isomorphism between $\mathbb{Z}\times\{1\}$ and $\mathbb{Z}$, so that the sets are equal in their properties and structure but not substance, as you can clearly see. For all practical purposes, people usually don't beat about the bush and simply state $\mathbb{Z}\subset\mathbb{Q}$, which is safe to say because of the concept of isomorphism. To make the point relatable to completion, it is not true, strictly speaking, that $\bar{\mathbb{Q}}=\mathbb{R}$ because $\mathbb{Q}$ has holes in it and we have absolutely no authority to add points to complete the set of rational numbers but what we can do is make the set $\mathbb{Q}$ isomorphic to subset of complete set. Thus, while we're not really adding points, we're still making the set complete by accounting for the missing holes, which is why we have the colloquial "addition of points".

Here is the promised definition:

Definition 190 *A mapping T from a metric space (X,d) into $\left(\hat{X},\hat{d}\right)$ is said to be an **isometry** if T $\forall x,y\in X$, $\hat{d}(T(x),T(y))=d(x,y)$. Two metric spaces are said to be **isometric** or **isomorphic** as metric spaces if there is a bijective isometry between them.*

If such a bijective mapping $T:X\to\hat{X}$ can be found, then X is said to be isometric with $\hat{X}$. This is the isomorphism for two metric spaces. Intuitively, an isometry preserves distances so nearby points in one space and equivalently near in another metric space. Remember, in metric spaces, one is concerned with distance between two points so that if two metric spaces have the same structure and properties i.e. must be isometric, then the distance between two points must be conserved and nothing else matters – not the names of the points, at least.

In order to complete any metric space, we can show that it is isomorphic to a dense subspace of a complete metric space and that this complete metric space must necessarily exist. Furthermore, this metric space is unique (up to isomorphism).

Theorem 191 *For a metric space (X,d), there exists a complete metric space $\left(\hat{X},\hat{d}\right)$ which has a subspace W that is isometric with $\hat{X}$ such that $\bar{W}=\hat{X}$. Furthermore, this space is unique except for isometries.*

Proof. First, we focus on the construction of $\left(\hat{X}, \hat{d}\right)$. Let x_n and $\acute{x}_n$ be Cauchy sequences in X. We will call two Cauchy sequences equivalent if they have the same limit i.e.

$$\lim_{n\to\infty} d\left(x_n, \acute{x}_n\right) = 0$$

This will be written as $(x_n) \sim (\acute{x}_n)$. We can then gather all such equivalent sequences and form an equivalent class. Indeed, $(x_n) \sim (x_n)$ is trivial, so this relation is reflexive. Also, since the arguments of a metric function are symmetric, the relation $\sim$ is symmetric. Finally, if $(x_n) \sim (y_n)$ and $(y_n) \sim (z_n)$, we have

$$d\left(x_n, z_n\right) \leq d\left(x_n, y_n\right) + d\left(x_n, y_n\right)$$

Taking limits on both sides and using the fact that the metric function is always positive, we have

$$\lim_{n\to\infty} d\left(x_n, z_n\right) = 0$$

so that $(x_n) \sim (z_n)$, implying transitivity. Thus, we can have for ourselves an equivalence class $\hat{x} = \{\bar{x}_n\}$ of Cauchy sequences. We can collect all such equivalence classes $\hat{x}, \hat{y}, ...$ and form the set $\hat{X}$. For this set, we can have the metric function

$$\hat{d}\left(\hat{x}, \hat{y}\right) = \lim_{n\to\infty} d\left(x_n, y_n\right)$$

where $x_n \in \hat{x}$ and $y_n \in \hat{y}$. Note that this is not equal to zero since x_n and y_n are members of a different equivalence class. To show that this limit is well-defined or that this definition is sensible and not ambiguous with different results for the same choice of inputs, we will first show that this limit exists and then show that it is independent of the choice of representatives. First, we have

$$d\left(x_n, y_n\right) \leq d\left(x_n, x_m\right) + d\left(x_m, y_m\right) + d\left(y_m, y_n\right)$$

$$\implies$$

$$d\left(x_n, y_n\right) - d\left(x_m, y_m\right) \leq d\left(x_n, x_m\right) + d\left(y_m, y_n\right)$$

Similarly,

$$d\left(x_m, y_m\right) \leq d\left(x_m, x_n\right) + d\left(x_n, y_n\right) + d\left(y_n, y_m\right)$$

$$\implies$$

$$d\left(x_m, y_m\right) - d\left(x_n, y_n\right) \leq d\left(x_m, x_n\right) + d\left(y_n, y_m\right)$$

$$\implies$$

$$-\left(d\left(x_m, x_n\right) + d\left(y_n, y_m\right)\right) \geq d\left(x_n, y_n\right) - d\left(x_m, y_m\right)$$

this is basically $b \geq a$ and $-b \geq a$ so that we have $|a| \leq b$. Hence,

$$\left|d\left(x_n, y_n\right) - d\left(x_m, y_m\right)\right| \leq d\left(x_n, x_m\right) + d\left(y_m, y_n\right)$$

Now, since x_n is Cauchy, we have $d\left(x_n, x_m\right) < \epsilon/2$ and similarly $d\left(y_m, y_n\right) < \epsilon/2$. This in turn implies that for $n, m > N$

$$\left|d\left(x_n, y_n\right) - d\left(x_m, y_m\right)\right| < \epsilon$$

so that

$$\lim_{n\to\infty} d(x_n, y_n) = \lim_{m\to\infty} d(x_m, y_m)$$

Hence, $\hat{d}(\hat{x}, \hat{y})$ is just as valid for any Cauchy sequence. Now, we prove that $\left(\hat{X}, \hat{d}\right)$ is a metric space. $\hat{d}(\hat{x}, \hat{y}) = 0 \iff \lim_{n\to\infty} d(x_n, y_n) = 0 \iff \lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n$ so that $(x_n) \sim (y_n)$, making them members of the same equivalence class. Since members of an equivalence class are either disjoint or the same, therefore $\hat{x} = \hat{y}$. Next, since $d(x_n, y_n) \geq 0$, we have $\hat{d}(\hat{x}, \hat{y}) \geq 0$. Furthermore, $d(x_n, y_n) = d(y_n, x_n)$ so that $\hat{d}(\hat{x}, \hat{y}) = \hat{d}(\hat{y}, \hat{x})$. Finally,

$$d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n)$$

so that $\hat{d}$ obeys the triangle inequality.

We have just proved that for any metric space (X, d), we will have another metric space $(\hat{X}, \hat{d})$ by accounting for the limits of the Cauchy sequences, made possible by clumping all Cauchy sequences with common limits. Let $W \subset \hat{X}$ and let $T : X \longrightarrow W$ be a mapping such that $T(a) = \hat{a}$ where $\hat{a}$ is an equivalence class of constant Cauchy sequences. Here, W is a subclass of constant Cauchy sequences. Since if two sequences are both constant and converge to the same limit, then the two sequences are equal. Thus, every equivalence class of constant Cauchy sequences will be a singleton so that $\hat{b}$ will only contain the Cauchy sequence $(b, b, ...)$. We will now prove that this is an isometry.

First, notice that the mapping is clearly onto. This can be understood by recalling how we arrived at $\hat{X}$ and hence W. Next, for $T(b_1) = T(b_2)$, we have $\hat{b}_1 = \hat{b}_2$ so that the mapping is one-to-one. Hence T is bijective. Finally, T is an isometry since

$$\hat{d}(T(a), T(b)) = \hat{d}\left(\hat{a}, \hat{b}\right) = \lim_{n\to\infty} d(a_n, b_n) = d(a, b)$$

To show that this W is dense in $\hat{X}$. For that, we need to show that the limit points of W are in $\hat{X}$. That is, if $\hat{x} \in \hat{X}$, we should have $\hat{d}(\hat{x}, x) < \epsilon \ \forall \epsilon > 0$ contained in W for $x \in W$. For $\hat{x} \in X$ and $(x_n) \in \hat{x}$. Now, for any Cauchy sequence x_n the inequality

$$d(x_n, x_N) < \epsilon/2$$

will be valid $\forall \epsilon > 0$ whenever $n > N$. For the constant sequence

$$(x_N, x_N, ...) = \hat{x}_N \in W$$

we have

$$\hat{d}(\hat{x}, x) = \lim_{n\to\infty} d(x_n, x_N) \leq \epsilon/2 < \epsilon$$

so that every neighbourhood of $\hat{x}$ will contain a point of W.

To show that $\hat{X}$ is complete, let $(\hat{x}_n)$ be any Cauchy sequence in $\hat{X}$. Now, since W is dense in $\hat{X}$, every point $\hat{x}_n \in \hat{X}$ and $\forall \epsilon > 0$, we can find a point

$\hat{z}_n \in W$ so that $\hat{d}(\hat{x}_n, \hat{z}_n) < \epsilon$. We can choose this $\epsilon = 1/n$ so that the sequence $(\hat{z}_n)$ becomes Cauchy. This can be observed as follows:

$$\begin{aligned} \hat{d}(\hat{z}_m, \hat{z}_n) &\leq \hat{d}(\hat{z}_m, \hat{x}_m) + \hat{d}(\hat{x}_m, \hat{x}_n) + \hat{d}(\hat{x}_n, \hat{z}_n) \\ &< 1/m + \hat{d}(\hat{x}_m, \hat{x}_n) + 1/n \end{aligned}$$

Since the element of W, $\hat{z}_n$, is Cauchy, $(z_m) = T^{-1}(\hat{z}_m)$ is also Cauchy in X. If (z_m) is contained in the class $\hat{x}$, then

$$\begin{aligned} \hat{d}(\hat{x}_n, \hat{x}) &\leq \hat{d}(\hat{x}_n, \hat{z}_n) + \hat{d}(\hat{z}_n, \hat{x}) \\ &< 1/n + \hat{d}(\hat{z}_n, \hat{x}) \\ &= 1/n + \lim_{m \to \infty} d(z_n, z_m) \end{aligned}$$

Since the sequence (z_m) is an element of the equivalence class of Cauchy sequences $\hat{x}$ and $\hat{z}_n$ is an equivalence class of Cauchy sequences and is contained in W, we have $(z_n, z_n, z_n, ...) \in \hat{z}_n$ and thus the inequality can be made as small as we like, implying that the limit of $\hat{x}_n$ is $\hat{x}$

If $\left(\tilde{X}, \tilde{d}\right)$ is another complete space with a subspace $\tilde{W}$ which is isometric with X such that $\tilde{W}$ is dense in $\tilde{X}$. Then, for any $\tilde{x}, \tilde{y} \in \tilde{X}$, we apply the same method as above to get

$$\left|\tilde{d}(\tilde{x}, \tilde{y}) - \tilde{d}(\tilde{x}_n, \tilde{y}_n)\right| \leq \tilde{d}(\tilde{x}, \tilde{x}_n) + \tilde{d}(\tilde{y}, \tilde{y}_n)$$

so that $\tilde{d}(\tilde{x}_n, \tilde{y}_n) \longrightarrow \tilde{d}(\tilde{x}, \tilde{y})$, implying that $\hat{X}$ and $\tilde{X}$ are isometric. ■

Example 192 *If (X, d) is complete, then $\left(X, \tilde{d}\right)$ is complete for $\tilde{d} = d/(1+d)$ since for any Cauchy sequence x_n,*

$$\begin{aligned} \tilde{d}(x_n, x) &= \frac{d(x_n, x)}{1 + d(x_n, x)} \\ &= 1 - \frac{1}{1 + d(x_n, x)} \end{aligned}$$

For a large n, $d(x_n, x) < \frac{1}{1-\epsilon} - 1$ is valid since the distance between the points of the Cauchy sequence and its limit point can be made arbitrarily small. Now

$$\begin{aligned} 1 + d(x_n, x) &< \frac{1}{1-\epsilon} \\ 1 - \epsilon &< \frac{1}{1 + d(x_n, x)} \\ \epsilon - 1 &> -\frac{1}{1 + d(x_n, x)} \\ \epsilon &> 1 - \frac{1}{1 + d(x_n, x)} \end{aligned}$$

so that $\tilde{d}(x_n, x) < \epsilon$ for any Cauchy sequence.

Just for a sneak peak:

Definition 193 *A **series** associated with a sequence x_n is the sum of all the terms of the sequence.*

More compactly, a series can be written using the summation symbol Σx_n. A series converges if this sum is finite.

Example 194 *If a rabbit hops infinite hopes such that every hop is half the distance of its previous hop, then the total distance will be represented by $a\Sigma\frac{1}{2^n}$ where a is the distance covered by the first hop.*

Problem 195 *Show that the open set of reals (a, b) is incomplete whereas $[a, b]$ is complete.*

Solution 196 *The trick over here is to construct a particular Cauchy sequence which converges to a or b. This does mean that the sequence does not converge in the particular open subset. For this, $x_n = b - 1/n$ or even $x_n = a + 1/n$ does fairly well. The fact that the latter interval is closed automatically qualifies it for completeness. For a particular construction, see the related theorem.*

Problem 197 *Show if the infinite sum $\sum x_n$ converges, then $x_n \longrightarrow 0$*

More spaces

There is a way in which we can form new vector spaces out of old ones

Definition 198 *The **direct sum** of two vector spaces V and W is the set $V \oplus W$ of pairs of vectors (v, w) in V and W, with the operations $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$ and $c(v, w) = (cv, cw)$ where c is a scalar.*

Thanks to the axiom of choice, we have the following result

Theorem 199 *Let V be a vector space over a field $\mathbb{F}$. Then, V is isomorphic to $\bigoplus \mathbb{F}$*

This direct sum is isomorphic to the Cartesian product if the dimension of I is finite.

However, we will focus our attention on which spaces can be taken out from existing ones.

1.11 Subspaces

A subspace of a vector space is so that the subset inherits the structure. So, for a vector space V, if we have a $A \subset V$ and if this satisfies all the axioms of a vector space using the addition and scalar multiplication defined for V, we have a subspace. Needless to say, the improper subspaces are the trivial subspace $\{0\}$ or the space V itself. Any other subset, if it satisfies the ten axioms using the induced vector addition and scalar multiplication, is a subspace. However, this method of verification is too lengthy. We can make use of the fact that vector addition is a binary operation. As already established, we can have a subgroup H of a G if and only if $\forall a, b \in H$, $ab^{-1} \in H$. This will take care of the fact that we have an Abelian group for vector multiplication. But what about scalar multiplication? For that we have a theorem:

Theorem 200 *A subspace of a vector space V is a nonempty subset A of V $\iff$ for all $\mathbf{u}, \mathbf{v} \in A$ and all scalars α, β we have $\alpha\mathbf{u} - \beta\mathbf{v} \in A$.*

The proof for groups has been worked out. Try to prove this on your own.

Just like the intersection of two subgroup is again a subgroup, we have following:

Exercise 201 *Show that the intersection of two vector subspaces A and B of a vector space X is a vector space.*

Function spaces are not restricted to one-dimensional intervals. Recall that functions can be defined for mutliple variables. Also remember that we do have complex valued functions. Note that differentiability is usually defined on an open interval (a, b). In our case above, we have the closed interval $[a, b]$ but the derivative at the end points can be easily defined as the right and left derivative for the left and right end point, respectively. This is a weaker version of differentiability but it does not harm in one-dimension. The situation is troublesome in more than one-dimension so that we are forced to consider open sets in the mutiple variable case. Thus, if we let Ω be an open subset of $\mathbb{R}^n$, then we can have the following subspaces all functions from Ω into $\mathbb{R}$, which the reader is required to prove:

- $C(\Omega)$, the space of all continuous complex valued functions defined on Ω.
- $C^k(\Omega)$, the space of all complex valued functions defined on Ω with continuous partial derivatives of order k.
- $C^\infty(\Omega)$, the space of infinitely differentiable functions defined on Ω.
- $P(\Omega) =$ the space of all polynomials of n variables (considered as functions on Ω).

In case we have complex-valued functions, then recall that a function is analytic (derivative of every order exists) if and only if the derivative of the first order exists. Thus, the difference (except for the space of polynomials) disappears.

Thus, what we have covered so far for $C[a, b]$ applies equally well to these more general cases.

Example 202 *If we have a vector space of the usual plane $\mathbb{R}^2$, a subspace would be $\mathbb{R}$ – that's the x-axis. Similary, the y-axis is a subspace, too*

The union of two vector spaces is not a vector space in general. For instance, in the example above, $(a, 0)$ belongs to the subspace corresponding to the x-axis. Similary, $(0, b)$ is a vector belonging to the subspace corresponding to the y-axis. However, in the union of the x-axis and y-axis, the element $(a, 0)+(0, b) = (a, b)$ does not belong to the union of y-axis and x-axis (it belongs to the span – covered later)

But what do we mean by dimension? And is the dimension of c, the space of all convergent sequences the same, less or greater than the dimensionality of l^∞, of which it is a subspace? For that, a few more definitions.

Definition 203 *A **linear combination** of vectors $x_1, x_2, ..., x_n$ of a vector space X is an expression of the form $\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n$ where $\alpha_i \in \mathbb{F}$.*

We will side step the definition, admittedly, at times: for us, linear combination will not refer only to the expression but also to the value of $\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n$. This will be clear from context.

Example 204 *For scalars* $2, 3$ *and vectors* $(1, 2)$ *and* $(2, 4)$*, we have the linear combination* $2(1, 2) + 3(2, 4) = (2, 4) + (6, 12) = (8, 16)$.

The word "linear" indicates that the vectors are not non-linear i.e. we do not have any square, square roots, or cube-roots. For vectors, this consideration is meaningless, all the same, since we don't know how to multiply vectors yet, let alone take their roots. Of course, this does not mean we're discussing trivial matters and just beating about the bush – such ideas to come in advance functional anslysis.

From linear combinations, we can move forward to define dimensions but for now, another definition.

Definition 205 *For any nonempty subset* A *of a vector space* V*, the set of all linear combinations of vectors of* A *called the* ***span*** *of* A*, written span*(A)*.* A *set* $T \subset V$ *is called* ***total*** *in* V *if* $\overline{spanT} = V$*.*

Example 206 *The span of the vector* $(1, 2, 3)$ *is* $\alpha(1, 2, 3)$ *i.e. all vectors that are parallel to it. Note that in this case, anti-parallel vectors are also included (anti-parallel means that the vectors are parallel but with an opposite direction)*

Example 207 $span\{(2, 4), (1, 3)\} = \alpha(2, 4) + \beta(1, 3)$

For any such A, span$A$ is a subspace of the vector space V. In fact, it is the smallest vector subspace of V containing A

Proof. For any $\mathbf{x}, \mathbf{y} \in A$, $\alpha\mathbf{x} - \beta\mathbf{y} \in$span$A$. ■

We're just a step away from defining dimensions.

Example 208

- $A = \{(x, x, 0) \mid x \in \mathbb{R}\}$

  This is a subspace of $\mathbb{R}^2$ *because for any scalars* α *and* β*,*

$$\begin{aligned} \alpha\mathbf{x} - \beta\mathbf{y} &= \alpha(x, x, 0) - \beta(y, y, 0) \\ &= (\alpha x - \beta y, \alpha x - \beta y, 0) \in A \end{aligned}$$

- $A = \{(x - 1, x, z) \mid x \in \mathbb{R}\}$

  For any scalars α *and* β*,* $\alpha\mathbf{x} - \beta\mathbf{y} = \alpha(x - 1, x, z) - \beta(y - 1, y, z) = (\alpha(x - 1) - \beta(y - 1), \alpha x - \beta y, \alpha z - \beta z) \notin A$ *since the first element is not* $\alpha x - \beta y - 1$*. Hence* A *is not a subspace of* $\mathbb{R}^2$

- $A = \{(x, y, z) \mid x, y, z \in \mathbb{R}^+\}$

  This set will not have additive inverses so is not a subspace of $\mathbb{R}^2$

- $A = \{(x, y, z) \mid x, y, z \in \mathbb{R}$ *and* $x - y + z = constant\}$

  The additive identity does not belong to this set hence A *is not a subspace of* $\mathbb{R}^2$

Definition 209 *A linear combination of vectors $x_1, x_2, ..., x_n$ is **linearly independent** in a vector space V if $\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n = 0 \iff \alpha_i = 0 \; \forall i$.*

Such a combination is not linearly independent or is linearly dependent if $\exists \alpha_i \neq 0$. This also means that for any j, we cannot have (proof?)

$$x_j = \alpha_1 x_1 + \alpha_2 x_2 + ... \alpha_{j-1} x_{j-1} + \alpha_{j+1} x_{j+1} + ... + \alpha_n x_n$$

Thus, no such vector can be written as a linear combination of other vectors which, together, form a linear independent set.

To show that this definition does not just apply to ordinary physical vectors we're used to, consider $x_i(t) = t^i \in P[a, b]$. That is, the set of finite degree polynmials Then, if

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

This is only possible $\iff \alpha_i = 0 \; \forall i$ i.e. if a polynomial equals zero, then all the scalars have to be zero. Notice that there is no scalar in this polynomial. What we're looking for is the roots of this polynomial. One root, clearly, is $t = 0$. We can then factor the remaining polynomial and find its roots, in which case we'll have non-zero scalars which can form a linearly dependent combination! The trick is to know that this cannot be repeated forever or that the field may not be algebraically closed. This will have to evoke some ring and field theory, which I, sadly, have to skip. Why bring this up in the first place? This is a very important fact for practical purposes. Consider a differential equation which must have two linearly independent solutions. You can also think of this basis as the basis for different transcendental functions: recall the polynomial expansion of e^x, $\sin x$ and $\cos x$.

An important use of this is in the theory of Ordinary Differential Equations. An n-th order homogenous ODE will have n linearly independent solutions. These form the basis of the "solution space". Thus, a linear combination of any of the solution of the given ODE are also a solution, provided the ODE is homogenous.

The name derives from the fact that a linear combination of such vectors can never be produced by any single other vector i.e. if the vectors are zero and some scalar is non-zero, then the non-zero scalar can be written as the linear combination of the remaining vectors. For instance, the vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ are linearly independent i.e. the usual $\mathbf{i}, \mathbf{j}, \mathbf{k}$ axis cannot create any one from each other whereas the vectors $(1, 2)$ and $(2, 4)$ are not. Also, if any third vector were written out, it would depend on these three linearly independent vectors – not more, not less.

Lemma 210 *If there are n such linearly independent vectors, then $n+1$ vectors are linearly dependent*

A subset B of V is called linearly independent if every finite linear combination of vectors in B is linearly independent. This is equivalent to the condition

that every $x \in V$ can be written in precisely one way as $\sum_{i=1}^{n} c_i x_i$, $c_i \in \mathbb{F}$ and $x_i \in B$ for some finite i. That is, B is linearly independent and every vector $\mathbf{v} \in V$ can be obtained as a finite linear combination of vectors from B.

The use of the word finite is very important here. We know how to add two vectors and how to scale one, thanks to the axioms. However, from the axioms, all we can do is repeat this finite number of times. In particular, we cannot have an infinitely long linear combination. As discussed, there are two types of infinities – countable and uncountable. For the uncountable case, it can be proven that the sum of an uncountable number of elements always diverges, which makes the summation notation of uncountable terms useless in our case. The trouble is with the countable case. This gives us a series. We are in no position to talk about the convergence of a series in a vector space because we don't have a norm.

Thus, $\varnothing \neq B \subset V$ is called a **Hamel basis** in V if $\text{span}(B) = V$. By span, we mean that every vector in V can be represented as a finite linear combination of elements of B. Note that we cannot have a Hamel basis for the space l^2. The elements e_i with the tuple 1 in the i-th position and zero otherwise does not form a Hamel basis for l^2. For instance, the element $\left(1, \frac{1}{2}, \frac{1}{4}, ...\right) \in l^2$ cannot be written as a finite linear combination of e_i's

This leads to our long-awaited notion of dimension.

Definition 211 *A vector space V is said to have **dimension** n, written $\dim V = n$, if every vector $x \in V$ can be written out as a linear combination of a linearly independent set of n vectors*

For any choice of basis, we will need n such vectors in order to span the whole space. It can be proven that for any basis X and Y of a vector space V, $|X| = |Y|$, the proof of which requires significant set theory. We shall mention another idea in passing: the axiom of choice roughly states that we can have a function for any set which can "pick" or "choose" elements from the set. This has been shown to be equivalent to the fact that every vector space has a basis. For finite dimenionsal spaces, the use of the word "finite dimensional" itself guarantees that there must exist such a Hamel basis. In the infinite dimensional case, this can be done by evoking another very fundamental theorem known as Zorn's lemma. This will not be covered in the current course.Hence, when $\dim V = n$ is mentioned, this n is well-defined.

One clear insight for this is that if B is the set of n basis for vector space V, we have span$B = V$. Note that if we happen to remove one element of the basis, then we'll be left with $n-1$ elements which will not span the entire space.

Exercise 212 *Show that if $|B| = n$ and $a \in B$ for span$B = V$, then $A = B \setminus \{a\}$ is not a basis*

By default, this means that any set of $n+1$ or more vectors of V is linearly dependent. By definition, $V = \{0\}$ is finite dimensional and $\dim V = 0$. If n is

not finite, then V is said to be infinite dimensional. Again, the $\mathbf{i}, \mathbf{j}, \mathbf{k}$ axis set the intuition for the dimension and clearly satisfy this definition, as highlighted above. As mentioned, such linearly independent vectors are called basis if they are used as a spanning set. For $\mathbb{R}^n$, we can have the basis

$$\begin{aligned} e_1 &= (1, 0, 0, ..., 0) \\ e_2 &= (0, 1, 0, ..., 0) \\ &\vdots \\ e_2 &= (0, 0, 0, ..., 1) \end{aligned}$$

which are also called the canonical basis or the standard basis.

The basis of a finite dimensional vector space are not unique. In fact, we can make do with any linearly independent set so as long as both the sets span the vector space. What is unique is the representation of any vector in the specific basis.

Proof. Consider α_i and β_i. Now, if

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

are two different representations of the same vector in the same basis, then clearly $\alpha_i = \beta_i$ because we can compare the e_i's ■

Theorem 213 *Let V be an n-dimensional vector space. Then, for any proper subspace A, $\dim A < \dim V$.*

Proof. If $n = 0$, $V = \{0\}$ so that there is no proper subspace and there is nothing left to prove. If V is not a trivial vector space, then let $\dim V = n$ and $A \neq \{0\}$ be a proper subspace. Clearly, the subspace of a finite vector space is finite so that we can have $\dim A = m$. By the law of trichotomy, we have three options viz. $m < n$, $m = n$ and $m > n$. The last options says that A is not a subspace because it has more basis than its parent space. If $m = n$, then both spaces have the same dimension and hence the span, making them equal. So, we are left with $m < n$. ■

We've mentioned that l^2 is a vector space. We've also mentioned that every space has a basis. But we've also mentioned that the basis in the usual sense (Hamel basis) are non-existent for l^2. What do we do? Make a new definition: take a sequence of basis (e_n) such that the series

$$\sum_{i=1}^{\infty} \alpha_i e_i$$

converges for a choice of a sequence of scalars (α_n). Since the limit of any series is unique (proof?), therefore we can safely write out $v = \Sigma \alpha_i e_i$. That is, the expansion of any vector v is determined by the choice of a sequence of scalars.

Such a basis is called the **Schauder Basis**. Needless to say, the idea of convergence is not defined in a vector space since it is not necessarily a norm space. In fact, even if it were a normed space, we might not have a limit (not every normed space is a Banach space). Thus, Schauder basis usually exist in a Banach space because for any element of a Banach Space, we can find a sequence that converges to that point and this sequence can be the partial sum of the series mentioned above.

Just like vector spaces, we have subspaces in norm spaces. Formally, for a norm space N and a subset A of N, A is a norm space if it satisfies the properties of norm. The norm of the subspace is said to be induced from the parent space. Just like studied earlier, this subspace can be closed or complete even if the parent space is not.

An important result for Banach spaces, seen further, is that the closure of a subspace is again a subspace

Proof. Since the only difference between a subspace and its closure is that of limit points, we will focus on the fact that limit points obey the structure of the norm space.

Let A be the subspace of a norm space N. If x and y are limit points, then we will have sequences $x_n \longrightarrow x$ and $y_n \longrightarrow y$. Since $x_n + y_n \in A$, and $x, y \in \bar{A}$, we have $x_n + y_n \longrightarrow x + y \in \bar{A}$ proving that the subspace is closed under addition. As already asked to prove, $x_n \longrightarrow x$ implies $\alpha x_n \longrightarrow \alpha x$ (for norm spaces, see **Proposition 218)**. But $\alpha x \in \bar{A}$ so that scalar multiplication is well-defined. Thus, for $\alpha x_n + \beta y_n$, $\alpha x + \beta y \in \bar{A}$ ■

Needless to point out, there are subspaces of norm spaces which are open and not closed.

Example 214 *The subspace Q comprising of convergent sequences of rational numbers is not closed. One example is the sequence $S_n = \sum^{n} 1/k^2$ which is convergent to $\pi^2/6$ in l^2 but this point does not belong to $\mathbb{Q}$. See Basel problem in the start for further details. Since this limit point does not belong to $\mathbb{Q}$, therefore $\mathbb{Q}$ is not closed.*

Example 215 *For l^∞ the subspace comprising of bounded sequences but not convergent ones is not closed.*

1.12 Metric Spaces and Norm Spaces

As already explained why, a norm space can be converted into a metric space by defining $d(x, y) = \|x - y\|$. This is called the metric induced by the norm. The norm of any single vector, then, can be viewed as its distance from the zero vector. In $\mathbb{R}^n$, this is the origin. However, there is still a distinction between the two in terms of structure. For instance, objects of a norm space are necessarily vectors whereas in a metric space, one can have ordinary points. In fact, one does not need any other relation between the points in a metric space. This should be easy to understand because the discrete metric can convert any ordinary set

into a metric space, even the set $\{Alina,\ Ajmal,\ Junaid,\ Haroon\}$. Metric spaces are "coarse" whereas normed spaces have a richer structure. In short, every norm space is a metric space but not conversely.

Since we now have a way of "equating" any norm with a metric, convergence of sequences and related concepts in normed spaces follow readily from the corresponding definitions. For instance, a sequence (x_n) in a normed space $(N, \|.\|)$ is convergent if $\iff \exists \mathbf{x}$ such that $\lim_{n\longrightarrow\infty} \|x_n - x\| = 0 \iff \forall \epsilon > 0$, $\exists N$ such that $\|x_n - x\| < \epsilon$ whenever $n > N$.

Example 216 *To show that the sequence* $(a_n) = \left(1/n^2\right)$ *converges to* 0 *under any norm, we can let* $\epsilon > 1/N$ *so that* $n > N$ *implies*

$$\begin{aligned} & \left\| \frac{1}{n^2} - 0 \right\| \\ = & \left\| \frac{1}{n^2} \right\| < \left\| 1/N^2 \right\| < 1/N < \epsilon \end{aligned}$$

Thus, if we have any given epsilon, we can find an N *so that the criterion of convergence is satisfied.*

Just like in ordinary calculus or Real/Complex analysis, the following proposition carries over to this more generalised space

Proposition 217 *The sum of two convergent sequences is convergent*

Proof. Let $x_n \longrightarrow x$ and $y_n \longrightarrow y$. Then, for any $\epsilon > 0$, we can have N_1 such that $\|x_n - x\| < \epsilon/2$ for $n > N_1$. Similarly, we have an N_2 such that $\|y_n - y\| < \epsilon/2$ for $n > N_2$. Now,

$$\begin{aligned} & \|x_n + y_n - (x + y)\| \\ \leq & \|x_n - x\| + \|y_n - y\| \\ < & \epsilon \end{aligned}$$

for $n > N = \max(N_1, N_2)$ ■

Proposition 218 *The scalar multiple of a convergent sequence is convergent.*

Proof. For $x_n \longrightarrow x$, we can have $\|x_n - x\| < \epsilon/|\alpha|$ so that for $\|\alpha x_n - \alpha x\| = |\alpha| \|x_n - x\| < \epsilon$ ■

Try to prove that the norm on $C[a, b]$ is uniform.

Lemma 219 *A metric* d *induced by the norm on a normed space* $(N, \|.\|)$ *is invariant to translation and scales appropriately. That is,*

1. $d(x + a, y + a) = d(x, y)$
2. $d(\alpha x, \alpha y) = |\alpha| d(x, y)$

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

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

Proposition 220 $|\|y\| - \|x\|| \leq \|x-y\|$

The reader will recall this proposition as an exercise for the modulus (norm on $\mathbb{R}$)

Proof. $d(y,0) \leq d(y,x) + d(x,0)$ so that $d(y,0) - d(x,0) \leq d(y,x) = d(x,y)$

Next, $d(x,0) \leq d(x,y) + d(y,0)$ so that $d(x,y) \leq -(d(y,0) - d(x,0))$. Combining the two, we have

$$|d(x,0) - d(y,0)| \leq d(x,y)$$

from which we get the required result. ■

Using this, we can prove that the norm is continuous

Proof. Since for continuity, we must have a $\delta > 0$ for every $\epsilon > 0$ in

$$|\|y\| - \|x\|| < \epsilon$$

whenever $\|x-y\| < \delta$. It is clear that from $\|x-y\| < \delta$, we have

$$|\|y\| - \|x\|| < \|x-y\| < \delta = \epsilon$$

■

Alternatively, this can be proved by evoking the sequential definition of continuity: let $\|x - x_n\| < \delta$. Then, $|\|x\| - \|x_n\|| < \delta = \epsilon$. In other words, if $x_n \longrightarrow x$, then $\|x_n\| \longrightarrow \|x\|$.

1.13 Convex Spaces

Informally, we can have convex spaces if any line originating from one point and ending at another is contained within the set.

Definition 221 *A subset A of a vector space V is said to be convex if $x, y \in A$ implies $M = \{z \in X \mid z = \alpha x + (1-\alpha) y,\ 0 \leq \alpha \leq 1\}$ is contained in A.*

Such a subset has boundary points x, y and any other point is an interior point.

Example 222 *Any closed interval in the set of real numbers is convex*

Example 223 *The closed unit ball is convex*

We introduce over here what is commonly called the Parallelogram Equality: $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$. The equality has a geometrical description: Recall that the vector from y to x is given by $x-y$. Thus, $x-y$ happens to be a diagonal of a parallelogram with one side x and the other side y. On the other hand, $x+y$ is also the other diagonal.

Exercise 224 *Construct a normed space which does not satisfy the parallelogram equality.*

1.14 Complete Norm Spaces

We have just proved instances of norm spaces which are complete and which can be completed. Moving on to a more general notion, we have

Definition 225 *A **Banach Space** is a complete normed space (complete in the metric defined by the norm).*

Needless to say, a sequence $(\mathbf{x}_n)$ in a normed space $(N, \|.\|)$ is Cauchy if $\Longleftrightarrow \lim_{n\longrightarrow\infty} \|\mathbf{x}_n - \mathbf{x}_m\| = 0 \iff \forall\epsilon > 0\,,\ \exists N$ such that $\|\mathbf{x}_n - \mathbf{x}_m\| < \epsilon$ whenever $n, m > N$.

We have already seen examples of complete metric spaces. The examples we've covered so far (except for the discrete metric space) are also norm spaces and, therefore, Banach Spaces.

Theorem 226 *A subspace A of a Banach space N is complete if and only if the set A is closed in N*

That is, A subspace A of a Banach space N is complete if and only if $\bar{A} = A$. The proof of this theorem is similar to that proved for metric spaces. Only the metric has to be replaced by the norm. We state it here for emphasis
Proof. If (x_n) is Cauchy in A if and only if there exists a limit x in N. A is closed if and only if $x \in A$. Hence A is Banach. ■

Theorem 227 *Let $(N, \|.\|)$ be a normed space. Then there is a Banach space B and an isometry T from N onto a subspace A of B such that $\bar{A} = B$.The space B is unique, except for isometries.*

Again, the proof for this theorem follows similar lines of reasoning as that for the completion of metric spaces. Again, try to do this yourself.

Apart from the notion of sequences carrying over to norm spaces courtesy of the induced norm by the metric, we can make use of the additional structure of a norm space and define series.

Definition 228 *The **partial sum** S_n of a sequence (x_n) is defined as*

$$S_n = x_1 + x_2 + ... + x_n$$

This can be seen as a sequence in itself. As $n \longrightarrow \infty$, the whole sequence is being added. That is, $S_n \longrightarrow S = x_1 + x_2 + ...$ Making use of the norm, we have $\|S - S_n\| < \epsilon$ for some N such that $n > N$. If this S is finite, then the series converges. Otherwise, it diverges. If $\|x_1\| + \|x_2\| + ...$ converges, then the series is said to be **absolutely convergent**. For real numbers, it is true that absolute convergence implies convergence but for norm spaces in general, this is not valid. Consider the sequence with elements

$$\begin{aligned} y_1 &= \left(1/1^2, 0, 0, ...\right) \\ y_2 &= \left(0, 1/2^2, 0, 0, ..\right) \\ y_3 &= \left(0, 0, 1/3^2, 0, ..\right) \end{aligned}$$

The series formed by the addition of elements of this sequence is not convergent. For l^2 norm, we have

$$\begin{aligned} & \|y_1\| + \|y_2\| + ... \\ = \; & 1 + 1/2^2 + 1/3^2 + ... \\ = \; & \pi^2/6 \end{aligned}$$

so that $S_n = \left(1, 1/2^2, 1/3^2, ..., 1/n^2, 0, 0, ...\right)$ and

$$\|S_n - S\| \geq 1/N^2$$

for $n > N$ where $S = \lim_{n\to\infty} S_n$ so that we cannot make this difference arbitrarily small.

The argument is invalid for Banach spaces. In fact, N is Banach if and only if every absolutely convergent series is convergent.

Proof. If B is a Banach space, then let $\sum \|x_k\|$ be convergent for a sequence x_k. What this means is that we can have an integer N such that

$$\sum_{k=N}^{\infty} \|x_k\| < \epsilon$$

If we have an N such that for $n, m > N$, the partial sums S_n and S_m for

$$S_n = \sum_{k=1}^{n} x_k$$

can give us

$$\|S_n - S_m\| = \left\| \sum_{k=n+1}^{m} x_k \right\|$$

for $n < m$. Then,

$$\left\| \sum_{k=n+1}^{m} x_k \right\| \leq \sum_{k=n+1}^{m} \|x_k\| < \epsilon$$

i.e. $\|S_n - S_m\| < \epsilon$ is Cauchy. Since this is a Cauchy sequence, we must have a limit point S. Thus, $S = \lim_{n\to\infty} S_n$ exists and is finite so that the series converges.

Conversely, let every absolutely convergent series be convergent and let x_n be a Cauchy sequence. Since $\|x_n - x_m\| < \epsilon$, we can have an integer n_1 so that $\|x_n - x_m\| < 2^{-1}$ for $n, m \geq n_1$. Again, we can find a n_2 such that

$$\|x_n - x_m\| < 2^{-2}$$

for $n, m \geq n_2$. Moving on, we can find n_k such that

$$\|x_n - x_m\| < 2^{-k}$$

for $n, m \geq n_k$. In particular, since

$$n_{k+1} > n_k \geq n_k$$

we can have $n = n_k$ and $m = n_{k+1}$ so that we have

$$\left\|x_{n_{k+1}} - x_{n_k}\right\| < 2^{-k}$$

Thus, we can have a subsequence n_k such that

$$\left\|x_{n_{k+1}} - x_{n_k}\right\| < 2^{-k}$$

If we substitute y_k for $x_{n_{k+1}} - x_{n_k}$, then

$$\sum \|y_k\| < \epsilon$$

implying that we have an absolutely convergent series. By our hypothesis, it should converge. Thus, $\sum y_k \longrightarrow S$ and the sequence of partial sums of y_k converges and this is a subsequence of x_n. Since x_n has a convergent and is Cauchy, it will also converge to the same limit as its subsequence (see previous proofs on sequences). Thus, this particular Cauchy sequence converges, which implies our space is Banach. ■

As already brought up, the basis for transcendental functions can be thought of as the linearly independent polynomials. Other than that, we have a power series expansion of any continuous function about a point or 0, called the Taylor or McLaurin series, respectively. This asks us to consider the relationship between norm spaces, series and their convergence when they are used to represent the basis of a space; in short motivating the following definition:

Definition 229 *If there exists a sequence e_n in a norm space N such that*

$$\|x - (\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n)\| \longrightarrow 0$$

*for every $x \in N$, then e_n is called a **Schauder basis**.*

The series $x = \sum \alpha_n e_n$ represents x and is called the expansion of x with respect to (e_n). This basis could be powers of x in the case of, say, the cosine function. Here's another example to make this a bit more abstract:

Example 230 *l^p has a Schauder basis (e_n) where*

$$\begin{aligned} e_1 &= (1, 0, 0, 0, ...) \\ e_2 &= (0, 1, 0, 0, ...) \\ e_3 &= (0, 0, 1, 0, ...) \\ &... \end{aligned}$$

This is just like the ordinary basis for $\mathbb{R}^n$ except that n is infinite.

This can be shortly written as $e_n = (\delta_{nj})$. Here,

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

is the Kronecker delta "function".

Theorem 231 *If a normed space has a Schauder basis, then the space is separable.*

We digress a little to talk about norm spaces. Remember that from given vector spaces, we can form a new vector space by simply taking their Cartesian product and defining addition and scalar mutliplication component-wise. We can do something similar with norm spaces. Given two norm spaces, $(N_1, \|.\|_1)$ and $(N_2, \|.\|_2)$ over the same field $\mathbb{F}$, we can form a new norm space by taking the Cartesian product of the given norm spaces. We can define addition and scalar multiplication for

$$N_1 \times N_2 = \{(x, y) \mid x \in N_1, y \in N_2\}$$

such that $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and $\alpha\,(x, y) = (\alpha x, \alpha y)$ for a scalar α and $x, x_1, y_1 \in N_1$ and $y, x_2, y_2 \in N_2$.

Exercise 232 *Show that the norm $\|.\| : N_1 \times N_2 \longrightarrow \mathbb{R}$ such that $\|(x, y)\| = \sqrt{\|x\|_1^2 + \|y\|_2^2}$ defines a norm, effectively showing that the product of norm spaces is a norm spaces. Furthermore, show that if N_1 and N_2 are Banach spaces, then so is $N_1 \times N_2$ under the same norm $\|.\|$*

1.15 Finite Dimensional Spaces

In this section, we'll focus on finite dimensional vector spaces in general and finite dimensional norm spaces in particular. Some of the properties are fairly generic. One intuitively understandable property is that of completeness. Let us start with a lemma.

Lemma 233 *Let $\{x_1, x_2, ..., x_n\}$ be a set of linearly independent vectors in a normed space N. Then, there is a number $c > 0$ such that for any choice of scalars $\alpha_1, \alpha_2, ..., \alpha_n$*

$$\|\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n\| \geq c\,(|\alpha_1| + |\alpha_2| + ... + |\alpha_n|)$$

Proof. The lemma holds trivially if all the scalars are zero. Let $|\alpha_1| + |\alpha_2| + ... + |\alpha_n| > 0$. Then $\|\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n\| \geq c\,(|\alpha_1| + |\alpha_2| + ... + |\alpha_n|)$

$$
\begin{aligned}
&\Rightarrow \quad \frac{\|\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n\|}{(|\alpha_1| + |\alpha_2| + ... + |\alpha_n|)} \geq c \\
&\Rightarrow \quad \frac{\|\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n\|}{\||\alpha_1| + |\alpha_2| + ... + |\alpha_n|\|} \geq c \\
&\Rightarrow \quad \left\| \begin{array}{c} \frac{\alpha_1}{|\alpha_1|+|\alpha_2|+...+|\alpha_n|} x_1 + \frac{\alpha_2}{|\alpha_1|+|\alpha_2|+...+|\alpha_n|} x_2 + \\ .... + \frac{\alpha_n}{|\alpha_1|+|\alpha_2|+...+|\alpha_n|} x_n \end{array} \right\| \geq c \\
&\Rightarrow \quad \|\beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n\| \geq c \text{ (say)}
\end{aligned}
$$

It should be remarked that $\sum_{j=1}^{n} \left|\beta_j\right| = 1$

Suppose $\|\beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n\| < c$ with $\sum_{j=1}^{n} \left|\beta_j\right| = 1$. Then, $\exists$ a sequence $y_m = \beta_1^{(m)} x_1 + \beta_2^{(m)} x_2 + ... + \beta_n^{(m)} x_n$ such that $y_m \longrightarrow 0$ with $\sum_{j=1}^{n} \left|\beta_j^{(m)}\right| = 1$ from which we have $\left|\beta_j^{(m)}\right| \leq 1$

Since $\beta_j^{(m)}$ is a bounded sequence for every j, then there must exist a corresponding convergent subsequence for each j, according to the Bolzano Weierstrass theorem. Let $y_{1,m}$ denote the corresponding subsequence for y_m. Since y_m is bounded, $y_{1,m}$ is also bounded and has a convergent subsequence $y_{2,m}$ (say). Continuing this way, we can obtain n such subsequnces. β_j is a limit for each j subsequence and, consequently, $y_{n,m} \longrightarrow y = \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n$ as m increases without bound. It is clear that $\beta_j \neq 0$ for each j otherwise $y = 0$ but $y_m \longrightarrow 0 \Longrightarrow y_{nm} \longrightarrow 0 \Longrightarrow y = 0$ which is a contradiction. ■

Example 234 *For $\mathbb{R}^2$, this c can be found as follows:*

$\|\alpha_1 e_1 + \alpha_2 e_2\| \geq c\,(|\alpha_1| + |\alpha_2|)$ *where* $e_1 = (1, 0)$ *and* $e_2 = (0, 1)$ *are the natural orthogonal basis. Thus, utilising the norm generalised by the Pythagorean theorem, we have*

$$c \leq \sqrt{\alpha_1^2 + \alpha_2^2} / \left(|\alpha_1| + |\alpha_2|\right)$$

In general, for $\mathbb{R}^n$ we have, $c \leq \sum_{j=1}^{n} \sqrt{|\alpha_j|^2} / \sum_{j=1}^{n} |\alpha_j|$

Theorem 235 *Every finite dimensional normed space N is complete.*

Proof. Consider the Cauchy sequence x_k and a set of linearly independent basis $\{e_1, e_2, ..., e_n\}$. We can represent the k-th term of this sequence as $x_k = \alpha_1^{(k)} e_1 + \alpha_2^{(k)} e_2 + ... + \alpha_n^{(k)} e_n$ where the superscript is not a power but rather serves as a reminder that the scalars α_i will depend on k. Now, For $n, m > N$

$$
\begin{aligned}
\|x_n - x_m\| \quad &< \quad c\epsilon \\
&\Longrightarrow \quad \left\| \sum_{j=1}^{n} \left(\alpha_j^{(n)} - \alpha_j^{(m)}\right) e_j \right\| < \epsilon
\end{aligned}
$$

$\Longrightarrow \exists c$ such that $c\left|\sum_{j=1}^{n}\alpha_j^{(n)}-\alpha_j^{(m)}\right| \leq \left\|\sum_{j=1}^{n}\left(\alpha_j^{(n)}-\alpha_j^{(m)}\right)e_j\right\| < c\epsilon$

i.e. $\left|\sum_{j=1}^{n}\left(\alpha_j^{(n)}-\alpha_j^{(m)}\right)\right| < \epsilon$ which is a Cauchy sequence of scalars belonging to a complete field. Needless to say, we have convergence so that we can use n such limits of the form α_i to construct $x=\alpha_1 e_1+\alpha_2 e_2+...+\alpha_n e_n$ so that $x\in N$. Now,

$$\begin{aligned}\|x_k-x\| &= \left\|\sum_{j=1}^{n}\left(\alpha_j^{(k)}-\alpha_j\right)e_j\right\| \\ &\leq \left|\sum_{j=1}^{n}\left(\alpha_j^{(k)}-\alpha_j\right)\right|\|e_j\| \\ &\leq b\left|\sum_{j=1}^{n}\left(\alpha_j^{(k)}-\alpha_j\right)\right|\end{aligned}$$

where $b=\max_j e_j$

$$\begin{aligned}\|x_k-x\| &\leq b\left|\sum_{j=1}^{n}\left(\alpha_j^{(k)}-\alpha_j\right)\right| \\ &< b\epsilon\end{aligned}$$

so that the Cauchy sequence converges, implying convergence. ■

Corollary 236 *Every finite dimensional subspace A of a normal space N is closed in N.*

Proof. If A is finite dimensional, then it is complete and hence closed. Details left to the reader. ■

The argument is invalid for infinite dimensional spaces. Here's a counter-example:

Example 237 *The infinite dimensional space $C[0,1]$ with basis $\left(t,t^2,t^3,...\right)$ for $x(t)\in C[0,1]$ does not have a limit. Consider $e=\lim_{n\to\infty}a_n$ where $a_n=\left(1+\frac{1}{n}\right)^n$. We know that*

$$x(t)=e^t=t^0/0!+t^1/1!+t^2/2!+...$$

in the given basis. We have already proved that a_n is Cauchy in the proof for the incompletness of rational numbers. Since a_n is Cauchy, we can prove that $x_n(t)=t^0/0!+t^1/1!+t^2/2!+...+t^n/n!$ is also Cauchy without a limit.

Definition 238 *The norm $\|.\|_1$ on any normed space N is said to be **equivalent** to another norm $\|.\|_2$ if $\exists\ a,b\in\mathbb{R}^+$ such that $\forall\mathbf{x}\in N$*

$$a\|\mathbf{x}\|_1\leq\|\mathbf{x}\|_2\leq b\|\mathbf{x}\|_1$$

This condition can be shown to be equivalent to the following: equivalent norms have similar limits for Cauchy sequences, which is what the definition actually means. This should be related to the example of the sequence $a_n = 1/n^2$.

Proof sketch. If x_n is a Cauchy sequence under $\|.\|_1$, then $\|x_n - x_m\|_1 < a\epsilon$ for $n, m > N$ but if this norm is equivalent to $\|.\|_2$, then from $a\|x\|_2 \leq \|x\|_1$ implies $\|x_n - x_m\|_2 < \epsilon$ which shows that the Cauchy sequences are the same if the reverse argument $\|x\|_1 \leq b\|x\|_2$ is applied. Using similar reasoning, we can show that the limits are the same. ■

We can move a step ahead and find for ourselves an equivalence class of similar Cauchy sequences.

Other than the equivalence of Cauchy sequence, this concept is motivated by the following fact: equivalent norms on N define the same topology for N.

Proof sketch. Suppose that a norm $\|.\|_1$ generates a topology τ_1 and $\|.\|_2$ generates the topology τ_2. What this means is that open sets (balls) in either topology will be formed from their respective norms. To prove that the topologies are equivalent, we need to prove that they are both subsets of each other. This can be done by showing that every open set in one topology is contained in another because of the radius of the ball is less than or equal to than a constant times the radius of the other. Details are left to the reader. ■

Theorem 239 *Any two norms on a finite dimensional space are equivalent*

Proof. Let X be an n-dimensional vector space with norms $\|.\|_1$ and $\|.\|_2$

$\forall \mathbf{x} \in X$, $\mathbf{x} = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n$ where $e_1, e_2, ..., e_n$ are the basis for X and $\alpha_1, \alpha_2, ..., \alpha_n$ are scalars

Then,

$$\begin{aligned} \|\mathbf{x}\|_1 &= \|\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n\|_1 \\ &\leq \|\alpha_1 e_1\|_1 + \|\alpha_2 e_2\|_1 + ... + \|\alpha_n e_n\|_1 \\ &= |\alpha_1| \|e_1\|_1 + |\alpha_2| \|e_2\|_1 + ... + |\alpha_n| \|e_n\|_1 \end{aligned}$$

Let $k = \max_{1 \leq i \leq n} \|e_i\|_1$

Then,

$$\begin{aligned} & |\alpha_1| \|e_1\|_1 + |\alpha_2| \|e_2\|_1 + ... + |\alpha_n| \|e_n\|_1 \\ \leq\ & |\alpha_1| k + |\alpha_2| k + ... + |\alpha_n| k \\ =\ & (|\alpha_1| + |\alpha_2| + ... + |\alpha_n|) k \end{aligned}$$

Also,

$$\begin{aligned} \frac{\|\mathbf{x}\|_2}{c} &= \frac{\|\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n\|_2}{c} \\ &\geq (|\alpha_1| + |\alpha_2| + ... + |\alpha_n|) \end{aligned}$$

Combining the two inequalities, we get

$$\|\mathbf{x}\|_1 \leq (|\alpha_1| + |\alpha_2| + ... + |\alpha_n|)\, k \leq \frac{k\,\|\mathbf{x}\|_2}{c}$$

or

$$\frac{c}{k}\,\|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2$$

which is the first half of the inequality for $a = c/k$. The second inequality $\|\mathbf{x}\|_2 \leq b\,\|\mathbf{x}\|_1$ can be obtained in a similar fashion. ■

In particular, the norms on $\mathbb{R}^2$ are equivalent:
$\|(x,y)\|_1 = |x| + |y|$ (the Manhattan norm)
$\|(x,y)\|_2 = \sqrt[2]{x^2+y^2}$ (the Pythagorean norm)
$\|(x,y)\|_\infty = \max(|x|, |y|)$ (the infinity norm)

1.16 Compact Spaces

Definition 240 *A metric space (X,d) is said to be **compact** if every sequence in X has a convergent subsequence.*

Example 241 *All complete metric spaces are compact.*

Example 242 *Any continuous and bounded curve in $\mathbb{R}^n$ with a finite length is compact.*

A general property of compact sets is expressed in:

Lemma 243 *A compact subset A of a metric space (X,d) is closed and bounded.*

Proof. If A is compact, then there is a sequence x_n with a convergent subsequence in A. If x is a limit point of A, then we can have this sequence x_n with a convergent subsequence so that $x_n \longrightarrow x \in A$. This implies that A is closed since this limit point was arbitrary. If A is not bounded, then we would have an unbounded and hence divergent sequence in A, which cannot be true since A must necessarily have convergent subsequences. ■

$\mathbb{R}^n$ and thus $\mathbb{C}^n$ are not compact because these spaces are neither closed nor bounded. In particular, the sequences $x_m = (m-n, m-n+1, ..., m-1, m)$ do not have a convergent subsequence. However, closed subsets of these spaces are, as will be proved later.

The infinite discrete metric space is also not compact. Consider any sequence x_n. In fact, a discrete metric space is compact if and only if it is finite.

The converse of this lemma is in general false. Consider a subspace of l^2 with sequence (e_n). This sequence is bounded since $\|e_n\| = 1$ and for every element of the sequence, $\|e_n\| < 1+\epsilon$ for $\epsilon > 0$. However, (e_n) does not converge so that the set containing (e_n) does not have a convergent subsequence.

For a finite dimensional normed space we have:

Theorem 244 *Let N be a finite dimensional normed space. $A \subseteq N$ is compact if and only if A is closed and bounded.*

Proof. We have already proved the sufficient conditions for this theorem in the lemma. For the necessary condition, we assume that A is $n-$dimensional, closed and bounded. Take the sequence x_m so that

$$x_m = \alpha_1^{(m)} e_1 + \alpha_2^{(m)} e_2 + ... + \alpha_n^{(m)} e_n$$

This sequence is bounded because the set A is itself bounded. Say $\|x_m\| \leq k$. Then, $\exists c > 0$ such that

$$k \geq c \sum_{j=1}^{n} \left| \alpha_j^{(m)} \right|$$

The $j-$sequence of numbers $\left(\alpha_j^{(m)}\right)$ are bounded so that by the Bolzano, Weistrass theorem, it will have a convergent subsequence. ■

Theorem 245 *If a subset A of a compact metric space X is closed, then A is compact.*

Proof. Since A is closed, for a limit point x of A, there exists a sequence x_n such that $x_n \longrightarrow x$. Clearly, a convergent sequence will have a convergent subsequence, implying that A is compact. ■

A good example would be closed and bounded subsets of $\mathbb{R}^n$ but for infinite dimensional spaces, of course the above stated property is not valid. A counter example has already been presented. We now move ahead with another result useful for the analysis of compact and even convex sets.

Theorem 246 (Riesz's lemma) *Let $A = \bar{A} \subset B$ be subspaces of a normed space N. Then for every real number $\theta \in (0,1)$ there is a $b \in B$ such that $\|b\| = 1$, $\|b - a\| \geq \theta \ \forall a \in A$.*

Proof. Take an element v belonging to B but not A. The distance from v to the set A is

$$d = \inf_{a \in A} \|v - a\|$$

Since A is closed, therefore no element not contained in A has to be at a distance of at least $\epsilon > 0$. Thus, $d > 0$. By the definition of infimum, we will an $a_0 \in A$ such that $d \leq \|v - a_0\| \leq \frac{d}{\theta}$ for $\theta \in (0,1)$. Let $b = c\,(v - a)$ where $c = \frac{1}{\|v - a_0\|}$. Then, $\|b\| = 1$. Also, $\|b - a\| = \|c\,(v - a_0) - a\| = |c| \left\|v - a_0 - c^{-1}a\right\| = c\,\|v - a_1\|$ where $a_1 = a_0 + c^{-1}a$ where this a_1 is variable. Now, $\|v - a_1\| \geq d$ so that we now have $\|b - a\| = c\,\|v - a_1\| \geq cd = d/\,\|v - a_0\| \geq d/\,(d/\theta) = \theta$ so that $\|b - a\| \geq \theta$ ■

This lemma is useful in proving the following, thus offering a characterisation of compactness.

Theorem 247 *If a normed space N has the property that the closed unit ball $A = \{x \mid \|x\| \leq 1\}$ is compact, then N is finite dimensional.*

The converse obviously holds, given our development so far.

Continuity has to do with open sets, as already seen – it maps inverse images of open sets to open sets. Similarly, it maps inverse images of closed sets to closed sets. For bounded operators, bounded sets are mapped to bounded sets. Since boundedness and continuity are equivalent at a point, we can say the same for closed and bounded sets; in other words, compact sets:

Theorem 248 *Let (X, d_1) and (Y, d_2) be metric spaces and $T : X \longrightarrow Y$ a continuous mapping. Then, the image of a compact subset A of X under T is compact.*

Proof. Let x_n be a sequence with a convergent subsequence in A. That is, $x_{n_k} \longrightarrow x$. Then, $T(x_{n_k}) \longrightarrow T(x)$ so that the subsequence $T(x_{n_k})$ of $T(x_n)$ has a point of convergent, giving us a compact set $T(A)$. ■

Corollary 249 *A continuous mapping T of a compact subset A of a metric space (X, d_1) into $\mathbb{R}$ assumes a maximum and a minimum at some points of A.*

Proof. Since A is compact, it is closed and bounded so that there will exist points which correspond to maximum and minimum. ■

Operators

In real and complex analysis, real and complex valued functions are used with a specified domain and range. In a similar analogy, we will develop a theory of a more general class of functiosn called **operators** which will act from general spaces to general spaces.

Definition 250 *Let X and Y be vector spaces over the same field. Then, an operator $T : X \longrightarrow Y$ is **linear** if for all $x, y \in X$ and scalars α, $T(x+y) = T(x) + T(y)$ and $T(\alpha x) = \alpha T(x)$.*

Example 251 *The identity operator $\hat{1}(x) = x$ is clearly linear.*

Example 252 *The zero operator $\hat{0}(x) = \mathbf{0}$ is also linear by default.*

Example 253 *The differential operator*

$$\frac{d}{dx} : P[a,b] \longrightarrow P[a,b]$$

is linear. For, let $f(x) = \sum_{i=0}^{n} \alpha_i x^i$ by any polynomial. Then,

$$\frac{d}{dx} f(x) = \sum_{i=1}^{n} \beta_i x^{i-1} \in P[a,b]$$

where $\beta_i = i\alpha_i$. Then,

$$\frac{d}{dx}\left[f(x) + g(x)\right] = \frac{d}{dx} f(x) + \frac{d}{dx} g(x)$$

and

$$\frac{d}{dx} \alpha f(x) = \alpha \frac{d}{dx} f(x)$$

Example 254 *Another operator T from $C[a,b]$ into itself can be defined by*

$$T(x(t)) = \int x(\tau) d\tau$$

is linear. See the example on bounded operator for details.

Example 255 *An operator T from $C[a,b]$ into itself defined by*

$$T(x(t)) = tx(t)$$

This operator is linear. For $T(\alpha x(t)) = t(\alpha x)(t) = \alpha(tx(t)) = \alpha T(x(t))$. Furthermore,

$$\begin{aligned} & T(x(t) + y(t)) \\ = \; & t(x(t) + y(t)) \\ = \; & tx(t) + ty(t) \\ = \; & T(x(t)) + T(y(t)) \end{aligned}$$

Lemma 256 *An operator $T : X \longrightarrow Y$ is linear if and only if $\forall x, y \in X$ and $\forall \alpha, \beta \in F$, $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$*

Proof. $T(\alpha x + \beta y) = T(\alpha x) + T(\beta y) = \alpha T(x) + \beta T(y)$

Conversely, by definition, $T(\alpha x) = \alpha(T(x)) = \alpha T(x)$

For the first property of linearity, we have

$$\begin{aligned} & T(\alpha x + \beta y) \\ = \; & \alpha T(x) + \beta T(y) \\ = \; & T(\alpha x) + T(\beta y) \end{aligned}$$

If $\alpha x = u$ and $\beta y = v$, then from $T(\alpha x + \beta y) = T(\alpha x) + T(\beta y)$, we have $T(v + u) = T(u) + T(v)$ ■

Exercise 257 *The graph of a linear operator is a vector space.*

For the definition of a graph of a function, which is a general case of an operator, see Chapter 1, Set Theory.

Example 258 *The operator $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ defined by $T(\mathbf{w}) = \mathbf{v} \times \mathbf{w}$ is linear where $\mathbf{v} = (v_1, v_2, v_3)$ is a fixed vector in $\mathbb{R}^3$ and $''\times''$ is the usual cross product for vectors. This can asily be seen as follows: $T(\alpha\mathbf{x} + \beta\mathbf{y}) = \mathbf{v} \times (\alpha\mathbf{x} + \beta\mathbf{y})$*

$= (\mathbf{v} \times \alpha\mathbf{x}) + (\mathbf{v} \times \beta\mathbf{y}) = \alpha(\mathbf{v} \times \mathbf{x}) + \beta(\mathbf{v} \times \mathbf{y})$

$= \alpha T(\mathbf{x}) + \beta T(\mathbf{y})$, where we have resorted to the well-established identities for cross multiplication, which may easily be verified by resorting to the determinant notation for cross product. Since this operator is linear, it will and does, indeed, preserve linear dependence.

Example 259 *$M_2(\mathbb{F})$ is called the collection of all 2x2 matrices with elements from a field $\mathbb{F}$. This is a ring with multiplicative identity*

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and additive identity

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

This collection forms a group and a ring with identity. If we restrict this collection to have matrices with non-zero determinants, then we have for ourselves inverses for each element. If scalars are taken from the field of complex or real numbers, then we have for ourselves a vector space. Note that multiplication is not commutative. We can also define for ourselves a norm and get a normed space. At any rate, if we take a fixed matrix M in this space and let $T(x) = Mx$ be an operator. This operator is clearly linear. Also, T^{-1} will exist if M^{-1} exists or if M has a non-zero determinant.

Thus the check for linearity may be shortened to the statement of this lemma instead of refering to the original definition. As an immediate consequence, linear operators preserve linear dependence. Of special importance is the fact that for any linear operator T and $\mathbf{0}$ vector, $T(\mathbf{0}) = \mathbf{0}$

Proof. $T(\mathbf{0}) = T(x - x)$
$= T(x + (-1x))$
$= T(x) + T(-1x)$
$= T(x) - T(x)$
$=\mathbf{0}$ ■

This linearity can be looked upon as a structure preserving operator between vector spaces. To hit the point, two vector spaces are said to be isomorphic if there exists a bijective linear operator between them. The theorems that follow might make this clearer.

Theorem 260 *Linear operators preserve linear dependence*

Proof. Let $\alpha_1 e_1 + ... + \alpha_n e_n = 0 \iff \alpha_i = 0$. Then,

$$T(0) = \alpha_1 T(e_1) + ... + \alpha_n T(e_n) = 0$$

$\iff \alpha_i = 0$ ■

Theorem 261 *$\mathcal{R}(T)$ is a vector space if T is linear*

Proof. Each axiom for the vector space can be checked individually for verification here. ■

Theorem 262 *Let T be a linear operator. Then $\dim \mathcal{D}(T) = n < \infty$ implies $\dim \mathcal{R}(T) \leq n$.*

Proof. $\dim \mathcal{D}(T) = n < \infty \implies$ for $x \in \mathcal{D}(T)$, we have $x = \alpha_1 e_1 + ... + \alpha_n e_n$ so that $T(x) = \alpha_1 T(e_1) + ... + \alpha_n T(e_n)$ clearly implying that $\dim \mathcal{R}(T) \leq n$ since if $\alpha_1 e_1 + ... + \alpha_n e_n = \iff \alpha_i = 0$ so that $\alpha_1 T(e_1) + ... + \alpha_n T(e_n) = 0$ $\iff \alpha_i = 0$ ■

Definition 263 *Let Let X and Y be vector spaces and $T : X \longrightarrow Y$ be an operator. Then, the **null space** $\mathcal{N}(T)$ or kernel of T, denoted by $\ker T$ is the set $\{x \mid T(x) = 0\}$.*

This is the complement of the supp$T$

Example 264 *The null space for* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *defined by* $T(\mathbf{w}) = \mathbf{v} \times \mathbf{w}$ *is* $\{w \mid w = \alpha \mathbf{v}\}$.

Theorem 265 *Let* T *be a linear operator. Then* $\ker T$ *is a vector space.*

Proof. We will prove that $\ker T \subseteq \mathcal{D}(T)$ is a subspace. For $x, y \in \ker T$, $T(\alpha x - \beta y) = \alpha T(x) - \beta T(y) = 0$ so that $\alpha x - \beta y \in \ker T$ ■

After this, we consider inverses of operators. As it turns out, inverses exist only when the function is into, just as in calculus. For a review on functions, see the introductory chapter. We can shorten the criteria for bijectiveness and consider only the following:

Theorem 266 *Let* X, Y *be vector spaces, both real or complex. Let*

$$T : X \longrightarrow Y$$

be a linear operator with range $\mathcal{R}(T) \subset Y$. *Then, the inverse*

$$T^{-1} : \mathcal{R}(T) \longrightarrow X$$

exists if and only if $T(x) = 0$ *implies* $x = 0$. *This* T^{-1} *is a linear operator. Furthermore,* $\dim \mathcal{D}(T) = n < \infty$ *and* T^{-1} *exists* $\implies \dim \mathcal{R}(T) = n$.

Proof. We replace the notion as follows: T^{-1} exists if and only if $\ker T = \{0\}$ .Suppose that T^{-1} exists. Then, $T(x_1) = T(x_2)$ implies $x_1 = x_2$ so that if $x_2 = 0$, $T(x_1) = T(0) = 0$ or $x_1 = 0$. Thus, $x_1 = 0$ is the only element in $\ker T$.

Conversely, Suppose that $\ker T = \{0\}$. Then, $T(x) = 0$ implies $x = 0$. Let $T(x_1) = T(x_2)$. Then, $T(x_1) - T(x_2) = 0$

or $T(x_1 - x_2) = 0$ or $x_1 - x_2 = 0$ or $x_1 = x_2$. Thus, T is injective. Thus, T^{-1} exists.

To show that T^{-1} is linear, we proceed as follows: since T^{-1} exists, we can expect T to be into. Then, for $y_1 = T(x_1)$ and $y_2 = T(x_2)$ so that $T^{-1}(y_1) = x_1$ and $T^{-1}(y_2) = x_2$. Now, if T is linear,

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2) = \alpha y_1 + \beta y_2$$

Thus, $T^{-1}(\alpha y_1 + \beta y_2)$
$= T^{-1}T(\alpha x_1 + \beta x_2)$
$= \alpha x_1 + \beta x_2$
$= \alpha T^{-1}(y_1) + \beta T^{-1}(y_2)$
i.e. $T^{-1}(\alpha y_1 + \beta y_2) = \alpha T^{-1}(y_1) + \beta T^{-1}(y_2)$ ■

Corollary 267 *Let* $T : X \longrightarrow Y$ *be linear and* $\dim X = \dim Y = n < \infty$. *Then,* $\mathcal{R}(T) = Y \iff T^{-1}$ *exists.*

Proof. If we can prove that $T(\mathbf{x}) = \mathbf{0}$ for only $\mathbf{x} = \mathbf{0}$, then we're done. Let $e_1, e_2, ..., e_n$ be the basis for X. Then, $\{T(e_1), T(e_2), ..., T(e_n)\}$ is linearly independent so that if $T(\mathbf{x}) = \mathbf{0}$, then we have

$$T\left(\sum_{i=1}^{n} \alpha_i e_i\right) = \mathbf{0}$$

or

$$\left(\sum_{i=1}^{n} \alpha_i T(e_i)\right) = \mathbf{0}$$

This is possible only when $\alpha_i = 0$ because $\{T(e_1), T(e_2), ..., T(e_n)\}$ is linearly independent so that $\mathbf{x} = \mathbf{0}$.

Conversely, if T^{-1} and $\dim \mathcal{D}(T) = \dim X = n$ implies $\dim \mathcal{R}(T) = n$. Also, $\dim Y = n$. So, we have $\dim Y = \dim \mathcal{R}(T) = n$. Now, if T was not surjective, we must have had $\mathbf{y} \in Y$ such that $T(\mathbf{x}) \neq \mathbf{y}$ for any $\mathbf{x}$ so that $\mathcal{R}(T) \subset Y$, implying $\dim Y > \dim \mathcal{R}(T)$, which is the required contradiction. ■

Corollary 268 *Let $T : X \longrightarrow Y$ and $S : Y \longrightarrow Z$ be bijective linear operators, where X, Y and Z are vector spaces. The inverse $(ST)^{-1} : Z \longrightarrow X$ of the composite ST exists and $(ST)^{-1} = T^{-1}S^{-1}$*

Proof. Since S and T are bijective, ST is bijective. Thus, $(ST)^{-1}$ exists. We also know that $(ST)^{-1} ST = \hat{1}$ so that $(ST)^{-1} S = T^{-1}$ and further

$$(ST)^{-1} = T^{-1}S^{-1}$$

■

This combination of operators is said to **commute** if $S(T(x)) = T(S(x))$ $\forall x$. This is shortened to $ST = TS$.

In light of this exploration, another definition is in order:

Definition 269 *Let $(N, \|.\|_1)$ and $(M, \|.\|_2)$ be normed spaces and $T : N \longrightarrow M$ a linear operator. N and M are said to be isomorphic as normed spaces if $\|T(x)\|_2 = \|x\|_1$ $\forall x \in N$*

This links the idea of isometry between normed spaces as metric spaces. We've already explored the concept of isomorphism between two metric spaces (isometry) and we know that every norm space is a metric space. In this light, you have the following exercise:

Proposition 270 *The linear combination of bijective operators is bijective*

Proof. It is straightforward to show that a combination of linear operators is linear. Let T, S be bijective on the same domain. Then, $T(y) = 0$ and $S(y) = 0$ implies $y = 0$ and we need $(\alpha T + \beta S)(x) = 0$ implies $x = 0$. If $x \neq y$, then $T(x) = 0$ so that $\ker T \neq \{0\}$ ■

Exercise 271 *Show that if two spaces are isomorphic as norm spaces, then they are isomorphic as metric spaces.*

Let us turn our attention to the norm of operators. Before that, we need to set some definitions straight.

Definition 272 *Let* $(N, \|.\|_1)$ *and* $(M, \|.\|_2)$ *be normed spaces and* $T : N \longrightarrow M$ *a linear operator. The operator* T *is said to be* ***bounded*** *if there is a real number* $c > 0$ *such that* $\forall x \in N \ \|T(x)\|_2 \leq c \|x\|_1$

Note that the norms do not have to be necessarily the same. However, since both norms obey particular properties viz. homogenity and the triangular inequality, we can safely perform an algebra on both sides. From now on, we will drop this convention of invoking subscripts to remind us where the norm comes from since the situation will normally make this clear.

An alternative way to frame this definition is as follows:

Theorem 273 *A linear operator* $T : X \longrightarrow Y$ *is bounded if and only if* T *maps bounded sets* X *into bounded sets* Y*.*

Proof. The definition of boundedness in a metric space is as follows:

$$diam(A) = \sup \{d(x, y) \mid x, y \in A\} = b < \infty$$

$\iff d(x, y) \leq b$ for all $x, y \in A$ which can be rephrased as

$$diam(A) = \sup \|x - y\| = b \forall x, y \in A$$

where $A \subseteq X$ or, equivalently, $-b \leq \|x - y\| \leq b$. So, if we have a bounded set A, then using

$$\|T(z)\|_Y \leq c \|z\|_X$$

for all $z \in A$ and $z = x - y$, we get

$$\|T(z)\|_Y \leq a = b/c$$

Trivially, $\|T(z)\|_Y \geq 0$ so that we can have $\|T(z)\|_Y \geq -a$ for $a > 0$

Conversely, assume bounded sets are mapped to bounded sets. Assume that T is not bounded. Then,

$$\|T(z)\| \geq c \|z\|$$

for some $c > 0$. This implies

$$\frac{\|T(z)\|}{b} \geq \frac{\|T(z)\|}{\|z\|} \geq c$$

or $\|T(z)\| \geq cb$, which is our required contradiction. ■

This goes out to say that the range of a bounded operator need not be closed. This enables us to differentiate between bounded operators and compact operators. Since the operator is continuous, the inverse images of open (resp.

closed) sets must be open (resp. closed). If the image of a subset of the range is not closed, then the domain must necessarily not be closed.

Recall the definition of supremum. It is an upper bound and the lowest of the upper bounds of a set. Thus, we can collect all x such that $\frac{\|T(x)\|}{\|x\|} \leq c$ and define a supremum out of it. If we can find a smallest such c, then we have

Definition 274 *The **norm** of a bounded linear operator T, denoted by $\|T\|$, is defined as $\|T\| = \sup_x \frac{\|T(x)\|}{\|x\|}$.*

Needless to say, this is valid when $\|x\| \neq 0$. Also, the norm of $\|T\|$ can be taken over normalised vectors so that $\|T\| = \sup_{\|x\|=1} \|T(x)\|$.

Note the requirement for a norm to exist: the operator must be bounded.

Since $\|T\| = \sup_x \frac{\|T(x)\|}{\|x\|}$, we can safely say that $\|T\| \geq \frac{\|T(x)\|}{\|x\|}$ for all x so that we have for ourselves the inequality

$$\|T(x)\| \leq \|T\| \|x\|$$

Thus, the following definitions are equivalent:

$\|T\| := \sup_{x\neq 0} \frac{\|T(x)\|}{\|x\|} = \sup_{\|x\|=1} \|T(x)\| = \sup_{0<\|x\|\leq 1} \|T(x)\| = \sup_{0<\|x\|<1} \|T(x)\| = \inf\{k : \|T(x)\| \leq k\|x\|, \forall x\}$

Proof. Let $A = \left\{ \frac{\|T(x)\|}{\|x\|} : x \in X \setminus \{0\} \right\}$

$B = \{\|T(x)\| : x \in X\setminus\{0\} \text{ and } \|x\| = 1\}$

$C = \{\|T(x)\| : x \in X\setminus\{0\} \text{ and } \|x\| \leq 1\}$

$D = \{\|T(x)\| : x \in X\setminus\{0\} \text{ and } \|x\| < 1\}$

Since equal sets have the same supremum, we will show that $A = B = C = D$

Clearly, A contains B, C and D.

Let $a \in A$

$\iff a = \frac{\|T(x)\|}{\|x\|}$ for some $x \in X\setminus\{0\}$

Since X is a norm space and closed under scalar multiplication, we can let $\|y\| x = y$

$\iff y \neq 0$ and $\|y\| = 1$ so that $a = \|T(y)\|$ for some $y \in X\setminus\{0\}$

$\iff a \in B$

$\iff A = B$

It is clear that $D \subseteq C$ and that $B \cup D = C$ so that $B \subseteq C$ as well.

Further, $B = A \subseteq C$ so that we have $B = A = C$

To show that $B \subseteq D$

$a \in B$

$\implies a = \|T(x)\|$ for some $x \in X\setminus\{0\}$ and $\|x\| = 1$

Assume that $\exists (x_n) \in B$ such that $x_n \longrightarrow x$.

Let $y_n = \frac{n-1}{n} x_n$. Then, $y_n \longrightarrow y$ since $\|x_n\| \longrightarrow \|x\|$ and $\|y\| < 1$

Then, $a_n = \|T(y_n)\| = \left\|\frac{n-1}{n}\right\| \|T(x_n)\| \longrightarrow \|T(y)\| = a$

Finally, we show that $\sup_{x\neq 0} \frac{\|T(x)\|}{\|x\|} = \inf\{k : \|T(x)\| \leq k\|x\|, \forall x\}$

Assume that $\sup_{x\neq 0}\frac{\|T(x)\|}{\|x\|}=\alpha$
Then, $\|T(x)\|\leq\alpha\|x\|$
$\Longrightarrow \sup_{x\neq 0}\frac{\|T(x)\|}{\|x\|}\geq\inf\{k:\|T(x)\|\leq k\|x\|,\forall x\}$
Next, $\inf\{k:\|T(x)\|\leq k\|x\|,\forall x\}\geq\frac{\|T(x)\|}{\|x\|}\geq\alpha-\frac{1}{n}$ for all n
So that $\inf\{k:\|T(x)\|\leq k\|x\|,\forall x\}=\alpha$ ■

This norm satisfies all the conditions of a norm space:

Proof. For N1, $\|T\|\geq\frac{\|T(x)\|}{\|x\|}\geq 0$. Next, $\|T\|=0$ if and only if $\sup\frac{\|T(x)\|}{\|x\|}=0$ which implies $\sup\|T(x)\|=0$. Since the supremum of non-negative numbers is zero, therefore $\|T\|=0$ if and only if $\|T(x)\|=0$ for all x. This is only possible when T is the zero operator.

For N2, $\|\alpha T\|=\sup_x\frac{\|\alpha T(x)\|}{\|x\|}=\sup_x\frac{|\alpha|\|T(x)\|}{\|x\|}=|\alpha|\sup_x\frac{\|T(x)\|}{\|x\|}=|\alpha|\|T\|$. In the second step, the homogenity property is applied because of the norm of $\mathcal{R}(T)$. In the third step, the scalar can be factored out because it has no role in the supremum since it does not depend on x.

For N3, $\sup\|(T_1+T_2)(x)\|=\sup\|T_1(x)+T_2(x)\|$. Since $\|T_1(x)+T_2(x)\|\leq\|T_1(x)\|+\|T_2(x)\|$ and so also their supremum, thus $\sup\|T_1(x)+T_2(x)\|\leq\sup\|T_1(x)\|+\sup\|T_2(x)\|$

From this, we can have $\|T_1+T_2\|\leq\|T_1\|+\|T_2\|$ ■

Here's an important conclusion from this jibber-jacky: we can collect all bounded operators $T:X\longrightarrow Y$ and get for ourselves a norm space! This space is called space of bounded linear operators from X to Y and is denoted by $B(X,Y)$. Some important theorems result from this idea. For instance, we can determine the completeness of the domain, given the completeness of the range, amongst other beautiful ideas but they will have to wait, for now.

Another useful formula goes as follows:

Proposition 275 $\|T^n\|\leq\|T\|^n$

Proof.

$$\begin{aligned}&\|TT\|\\ =\;&\sup\|T(T(x))\|\\ \leq\;&\sup\|T\|\,\|(T(x))\|\\ =\;&\|T\|\sup\|(T(x))\|\\ =\;&\|T\|^2\end{aligned}$$

The proof then follows in a similar method for n by induction. ■

Example 276 *The identity operator $\hat{1}:N\longrightarrow N$ on a normed space $N\neq\varnothing$ is bounded and has norm $\left\|\hat{1}\right\|=1$.*

Example 277 *The zero operator $\hat{0}:N\longrightarrow M$ on a normed space N to norm space M is bounded and has norm $\left\|\hat{0}\right\|=0$.*

Example 278 *Let N be the normed space of all polynomials on $J = [0,1]$ with norm given $x = \max_{t \in J} \|x(t)\|$. A differentiation operator $T = \frac{d}{dt}$ is defined on N by $T(x(t)) = x\prime(t)$ where the prime denotes differentiation with respect to t. This operator is linear but not bounded. Indeed, let $x_n(t) = t^n$ where $n \in \mathbb{N}$. Then $\|x_n\| = 1$ and $T(x(t)) = x\prime(t) = nt^{n-1}$ so that $\|T(x_n)\| = n$ and $\frac{\|T(x_n)\|}{\|x_n\|} = n$. Since $n \in \mathbb{N}$ is arbitrary, this shows that there is no fixed number c such that $\frac{\|T(x_n)\|}{\|x_n\|} \leq c$. From this, we conclude that T is not bounded.*

Example 279 *Let $T(x_n) = nx$ for any sequence $x_n \longrightarrow x$. Then, this operator, too, is not bounded.*

Example 280 *For $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ defined by $T(\mathbf{w}) = \mathbf{v} \times \mathbf{w}$, we have $T(\mathbf{w}) = \|\mathbf{v}\| \|\mathbf{w}\| \sin\theta \leq c\|\mathbf{w}\|$ where $c = \|\mathbf{v}\|$ so that T is bounded.*

Example 281 *The integral operator $T : C[0,1] \longrightarrow C[0,1]$ such that*

$$T(x(t)) = \int_0^1 k(t,\tau)\, x(\tau)\, d\tau$$

is bounded where $k(t,\tau)$ is continuous on the closed interval $[0,1] \times [0,1]$ and thus bounded itself.

$T(\alpha x(t) + \beta y(t)) =$

$$\begin{aligned} T(\alpha x(t) + \beta y(t)) &= \int_0^1 k(t,\tau)\,(\alpha x(\tau) + \beta y(\tau))\, d\tau \\ &= \int_0^1 k(t,\tau)\,\alpha x(\tau)\, d\tau + \int_0^1 k(t,\tau)\,\beta y(\tau)\, d\tau \\ &= \alpha \int_0^1 k(t,\tau)\, x(\tau)\, d\tau + \beta \int_0^1 k(t,\tau)\, y(\tau)\, d\tau \\ &= \alpha T(x(t)) + \beta(Ty(t)) \end{aligned}$$

To show that this operator is bounded, we first let $|k(t,\tau)| \leq c$. This assumption is justified since k is bounded. Next, since $x \in C[0,1]$, we can have $\|x\| = |x(t)| = \max_t x(t)$ so that

$$\begin{aligned} \|T(x)\| &= \max_t \|k(t,\tau)\, x(\tau)\, d\tau\| \\ &\leq \max_t \int_0^1 |k(t,\tau)|\, |x(\tau)|\, d\tau \\ &\leq c\|x\| \end{aligned}$$

Thus, T is bounded. From $\|T(x)\| / \|x\| \leq c$ we can define $\|T\| = c$.

We can even say more about the boundedness of the inverse:

Theorem 282 *Let $T : X \longrightarrow Y$ be bounded, onto and linear. If there exists a b such that $\|T(x)\| \geq b\|x\|$ for all x, then T^{-1} exists and is bounded.*

Note that the condition does not voilate the definition of boundedness because it has little to do with the c except that $b \leq c$

Proof. If $T(x) = 0$, then $\|T(x)\| \geq b\|x\|$ implies $0 \geq b\|x\|$ but this is only possible when $\|x\| = 0$ if and only if $x = 0$. Thus $T(x) = 0$ implies $x = 0$ so that T is into, proving that T^{-1} exists. From $T(x) = y$ and $T^{-1}(y) = x$ and $\|T(x)\| \geq b\|x\|$, we have $\|y\| \geq b\left\|T^{-1}(y)\right\|$ so that T^{-1} is bounded. ■

Boundedness is typical; it is an essential simplification which we always have in the finite dimensional case, as follows.

Theorem 283 *If a normed space N is finite dimensional, then every linear operator on N is bounded.*

Proof. For finite dimensional spaces,

$$\begin{aligned} \|T(x)\| &= \left\|\sum \alpha_i T(e_i)\right\| \\ &\leq \sum \|\alpha_i T(e_i)\| \\ &= \sum |\alpha_i| \|T(e_i)\| \\ &\leq c_1 \sum |\alpha_i| \end{aligned}$$

where $c_1 = \max_i \|T(e_i)\|$. Next, $\|x\| \geq c_2 \sum |\alpha_i|$. Thus, we have $\|T(x)\| \leq c_1 \sum |\alpha_i|$ and $\|x\| \geq c_2 \sum |\alpha_i|$ from which we have

$$c_2 \|T(x)\| \leq c_1 c_2 \sum |\alpha_i| \leq c_1 \|x\|$$

or $\|T(x)\| \leq c\|x\|$ where $c = c_1/c_2$. ■

Corollary 284 *In a finite dimensional space, every linear operator is continuous.*

Try to prove this corollary without using the next theorem.

Theorem 285 *Let T be a linear operator. Then, T is continuous if and only if it is bounded.*

Proof. Let $T : X \longrightarrow Y$ be continuous. Then, $\|T(\mathbf{x}) - T(\mathbf{x}_0)\|_Y < \varepsilon$ whenever $\|\mathbf{x} - \mathbf{x}_0\|_X < \delta$

or $\|T(\mathbf{x} - \mathbf{x}_0)\|_Y < \varepsilon$ whenever $\|\mathbf{x} - \mathbf{x}_0\|_X < \delta$

Let $\mathbf{x} - \mathbf{x}_0 = \frac{\varepsilon \mathbf{y}}{a\|\mathbf{y}\|_X}$ for $a > 0$. This is justified since the denominator is bounded and not equal to zero. Then,

$$\left\| T(\frac{\varepsilon \mathbf{y}}{c \left\| \mathbf{y} \right\|_X}) \right\|_Y < \varepsilon \Rightarrow \left\| T(\mathbf{y}) \right\|_Y < a \left\| \mathbf{y} \right\|_X$$

or $\left\| T(\mathbf{y}) \right\|_Y \leq c \left\| \mathbf{y} \right\|_X$ for some $0 < c < a$
Conversely, $\left\| T(\mathbf{y}) \right\|_Y \leq c \left\| \mathbf{y} \right\|_X$
Let $\mathbf{y} = \mathbf{x} - \mathbf{x}_0$ for $\left\| \mathbf{y} \right\|_X = \left\| \mathbf{x} - \mathbf{x}_0 \right\|_X < \delta$
Then, $\left\| T(\mathbf{x}) - T(\mathbf{x}_0) \right\|_Y < c\delta = \varepsilon$ whenever $\left\| \mathbf{x} - \mathbf{x}_0 \right\|_X < \delta$ ■

Corollary 286 *Let $T : N \longrightarrow M$ be a linear operator and N, M are normed spaces. Then, if T is continuous at a single point, then it is continuous.*

Corollary 287 *$x_n \longrightarrow x$ implies $T(x_n) \longrightarrow T(x)$*

Proof. Let $\left\| (x_n - x) \right\| < \epsilon / \left\| T \right\|$ for $n > N$. Then,

$$\begin{aligned} & \left\| T(x_n) - T(x) \right\| \\ = & \left\| T(x_n - x) \right\| \\ \leq & \left\| T \right\| \left\| (x_n - x) \right\| \\ < & \epsilon \end{aligned}$$

■

Theorem 288 *$\ker T$ is closed for linear, bounded T.*

Proof. For a limit point x of $\ker T$, there exists a sequence $x_n \longrightarrow x$. From this, $T(x_n) \longrightarrow T(x)$. Since $T(x_n) = 0$, then $T(x) = 0$ so that $x \in \ker T$ ■

Note that the definition of continuity implies that the inverse images of open sets is open. This is no way implies that open sets are mapped to open sets. Similarly, it in no way implies that closed sets are mapped to closed sets so that for a linear bounded T, we cannot say more about $\mathcal{R}(T)$.

Recall from calculus that two functions f_1 and f_2 are equal if $f_1(x) = f_2(x)$ $\forall x \in \mathcal{D}(f_1) = \mathcal{D}(f_2)$. In a similar vein, we say that two operators T_1 and T_2 are equal if $T_1(x) = T_2(x)$ $\forall x \in \mathcal{D}(T_1) = \mathcal{D}(T_2)$

We can restrict an operator by restricting the operator to a subset A of the domain $\mathcal{D}(T)$. Only a few elements are eliminated from the domain. Conversely, from this subset, we can go back to the domain by applying the operator to this superset – this is the extension of the operator. At times, some elements are added to give a new domain $M \supseteq \mathcal{D}(T)$, which will leave the mapping of already existing elements unchanged. This can give us many extensions but for practical purposes, the extension must preserve boundedness or linearity and norm. This is the case when $\mathcal{D}(T)$ is dense in M. Needless to say, this M must be complete. Thus, we have the following theorem:

Theorem 289 *Let X and Y be Banach spaces and let $T : \mathcal{D}(T) \longrightarrow Y$ be a linear, bounded operator. Then, there exists an extension $\tilde{T} : \overline{\mathcal{D}(T)} \longrightarrow Y$ such that T is linear, bounded and $\left\| T \right\| = \left\| \tilde{T} \right\|$*

Proof. Let x_n be a sequence in $\mathcal{D}(T)$ convergent in X to x. Since T is linear and bounded and Y is complete, $T(x_n) \longrightarrow T(x) = y$. Define $\tilde{T}(x) = y$. This definition is independent of the choice of x_n. Suppose $z_n \longrightarrow x$. Then, the sequence $v_m = (x_1, z_1, x_2, z_2, ...)$ converges to x so that the subsequences $T(x_n)$ and $T(z_n)$ have the same limit. Thus, the choice of sequence does not affect the uniqueness of $\tilde{T}$.

To show that $\tilde{T}$ is linear, $\tilde{T}(\alpha s + \beta t)$. Then,

$$\begin{aligned} & \tilde{T}(\alpha s + \beta t) \\ = & \lim_{n\to\infty} T(\alpha s_n + \beta t_n) \\ = & \alpha \lim_{n\to\infty} T(s_n) + \lim_{n\to\infty} \beta T(t_n) \\ = & \alpha\tilde{T}(s) + \beta\tilde{T}(t) \end{aligned}$$

Clearly, $\tilde{T}(x) = T(x)$ for $x \in \mathcal{D}(T)$ so that $\tilde{T}$ is a certified extension of T. To show that $\tilde{T}$ is bounded,

$$\begin{aligned} & \left\|\tilde{T}(x)\right\| \\ = & \left\|\lim_{n\to\infty} T(x_n)\right\| \\ \leq & \lim_{n\to\infty} \|T\| \|x_n\| \\ = & \|T\| \|x\| \end{aligned}$$

Thus $\tilde{T}$ is bounded.

We have also obtained $\left\|\tilde{T}\right\| \leq \|T\|$ if we divide both sides of the above obtained inequality by $\|x\|$ and take supermum over x. Trivially, $\left\|\tilde{T}\right\| \geq \|T\|$ because the norm cannot decrease in an extension. Combining, we have $\left\|\tilde{T}\right\| = \|T\|$ ■

What if X or Y is not complete? Let $T \in B(X,Y)$ and let $E \subseteq X$ be a normed subspace. If we know what $T|_E$ is and don't have any knowledge of T, we can recover part of T. Intuitively, this is done by adding limit points to domain of the restriction. Since a linear operator is continuous, the added limit points can, in some sense, recover the operator T.

Proposition 290 *$T|_E$ has a unique extension to a continuous linear mapping defined on $\bar{E}$.*

Proof. In order to avoid the cumbersome $T|_E$, we let $T|_E = L$. Also, this should serve as a reminder that we're dealing with a different operator L and we're looking for an extension of it, T, whose mapping we're not sure of right now.

If $x \in \bar{E}$, then there is a Cauchy sequence $x_n \in E$ such that $x_n \longrightarrow x$. That is, for every $\epsilon > 0$, there exists N such that $\|x_n - x_m\| < \epsilon / \|L\|$ for $m, n > N$

Then, $\|L(x_n) - L(x_m)\| = \|L(x_n - x_m)\| \leq \|L\| \|x_n - x_m\| < \epsilon$, implying that $L(x_n)$ is Cauchy. Since Y is complete, we have $z \in Y$ such that

$\lim_{n\to\infty} L(x_n) = z$. Now, in order for the extension to be sensible, we should have $\lim_{n\to\infty} L(x_n) = T(x)$. By this justification, we have $T(x) = L(x)$ for $x \in E$, making T linear.

This definition will be valid if limit z is unique for all sequences in E convergent to x. If $y_n \longrightarrow x$, then

$L(y_n) = L(y_n) - L(x_n) + L(x_n) \longrightarrow z$ hence $L(y_n) - L(x_n) \longrightarrow 0$ so that $L(y_n) = L(x_n)$

Next, $\|T(x)\| = \left\| \lim_{n\to\infty} L(x_n) \right\| = \lim_{n\to\infty} \|L(x_n)\| \leq \sup_{\|x_n\|=1} \|L(x_n)\| = \|L\|$ hence T is bounded, making it continuous. ■

In particular, if E is dense in X, that is, $\bar{E} = X$, then we can completely determine the extension and use the previous theorem.

Problem 291 *Let X and Y be normed spaces and T be a bounded, linear surjective operator. Suppose that there exists a constant $b > 0$ such that $\|T(x)\| \geq b\|x\|$ for all $x \in X$. Show that T^{-1} exists and is bounded*

Proof. First, we will show that T^{-1} exists. This will need two parts. One, that T^{-1} is well-defined. That is, it maps similar elements to similar images. That is, $x = y$ implies$T^{-1}(x) = T^{-1}(y)$. Surjectively will then imply that $\mathcal{R}(T) = \mathcal{D}\left(T^{-1}\right)$ so that there are no elements for which the mapping T^{-1} is undefined. This T^{-1} will have to be constructed, provided we can prove that T is bijective (surjectivity is given). Now, let We can never get injectivity from surjectivity so we resort to the norm. $T(x) = 0$. Then, $\|T(x)\| = 0$ and $\|T(x)\| \geq b\|x\|$ implies that $b\|x\| \leq 0$

$\implies \|x\| = 0$

$\implies x = 0$

Hence, T is bijective so that we can have $T^{-1}(y) = x$ for $y = T(x)$ so that $\|T(x)\| \geq b\|x\|$

$\implies \|y\| \geq b\left\|T^{-1}(y)\right\|$

$\implies \left\|T^{-1}(y)\right\| \leq c\|y\|$ where $c = 1/b$ ■

Problem 292 *Let $T : C[0,1] \longrightarrow C[0,1]$ be defined by*

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

Find $\mathcal{R}(T)$ and $T^{-1} : \mathcal{R}(T) \longrightarrow C[0,1]$. Is T^{-1} linear? Bounded?

Here's a motivating exercise for the next section:

Exercise 293 *Let N be a Banach space and M a norm space. Let $\{T_k\}$ be a sqeuence of bounded linear operators from N to M such that for each $x \in N$, the set $\{T_k(x)\}$ is bounded subset of M for each k. Then, the sequence $\{\|T_k\|\}$ of norms of T_k is also bounded*

1.17 Normed Space of Operators

As already proved, the operator of a norm satisfies all the axioms of a norm space. Thus, the space of bounded, linear operators $T : X \longrightarrow Y$ between vector spaces X and Y, denoted by $B(X,Y)$ is a norm space, without a shadow of doubt.. but we haven't proved that this is vector space, yet!

Exercise 294 *$B(X,Y)$ is a vector space*

In what case is this space complete i.e. Banach?

Theorem 295 *If Y is a Banach space, then so is $B(X,Y)$*

Proof. Now, remember, elements of $B(X,Y)$ are linear operators T so that if we want to show that an arbitrary Cauchy sequence in $B(X,Y)$ converges, we must take a sequence of operators and show that it converges. Let (T_n) be a Cauchy sequence of operators in $B(X,Y)$. Thus, by definition, for all $\epsilon > 0$, $\exists N$ such that $\|T_n - T_m\| < \epsilon$ whenever $n, m > N$. For all $x \in X$ and $n, m > N$, we have $\|T_n(x) - T_m(x)\| = \|(T_n - T_m)(x)\|$ (point-wise addition) $\leq \|T_n - T_m\| \|x\|$ (T_i's are bounded)

Therefore, $\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\| \|x\| < \epsilon \|x\|$. Now, for any fixed x and given ϵ', we may choose $\epsilon = \epsilon_x$ such that $\epsilon_x \|x\| < \epsilon'$. Then, $\|T_n(x) - T_m(x)\| < \epsilon'$, $\epsilon' > 0$ and $n, m > N$ implies $T_n(x)$ is Cauchy in Y. Since Y is complete, therefore there exists an element y such that the Cauchy sequence $T_n(x) \longrightarrow y \in Y$. Now, the limit y depends upon the choice of x because $\|T_n(x) - y\| \longrightarrow 0$. We can call this $y = T(x)$. Thus, we have $T_n(x) \longrightarrow T(x)$. To prove that $T(x) \in B(X,Y)$, we need to show that $T(x)$ is linear and bounded.

Linear:

$$
\begin{aligned}
&T(\alpha x + \beta y)\\
&= \lim_{n\to\infty} T_n(\alpha x + \beta y)\\
&= \lim_{n\to\infty} [\alpha T_n(x) + \beta T_n(y)]\\
&= \lim_{n\to\infty} \alpha T_n(x) + \lim_{n\to\infty} \beta T_n(y)\\
&= \alpha \lim_{n\to\infty} T_n(x) + \beta \lim_{n\to\infty} T_n(y)\\
&= \alpha T(x) + \beta T(y)
\end{aligned}
$$

Bounded:

$$
\begin{aligned}
&\|T_n(x) - T(x)\|\\
&= \left\| T_n(x) - \lim_{m\to\infty} T_m(x) \right\|\\
&= \lim_{m\to\infty} \|T_n(x) - T_m(x)\|\\
&\leq \lim_{m\to\infty} \|T_n - T_m\| \|x\|\\
&= \|T_n(x) - T(x)\| \|x\|\\
&< \epsilon \|x\|
\end{aligned}
$$

That is, $\|T_n(x) - T(x)\| \leq \epsilon \|x\|$. Hence the operator $(T_n - T)$ is bounded and $T_n - T \in B(X,Y)$. Since $T_n \in B(X,Y)$ and $B(X,Y)$ is closed under

addition, which you have hopefully proved above, therefore $T_n - (T_n - T) = T \in B(X, Y)$. ■

We mention an important tool in passing: a corresponding theorem for functionals (see next chapter) is known as the famous Hahn-Banach Theorem, without which the study of functional analysis is complete.

1.18 Operators on Finite Dimensional Spaces

We have already seen that finite dimensional spaces are much simpler than infinite dimensional ones in certain aspects. Of particular note is the role of operators and functionals on such spaces. We will show this by incorporating matrices into our discussions. For a review of matrices, see the appendix.

Recall that for an n-dimensional vector, an $r \times n$ matrix acts on it to give a r-dimensional vector. Thus, linear operators on finite dimensional spaces can be viewed as matrices. Matrix operation is associative, linear and in some cases, bounded and invertible, making it a perfect candidate for our present discussion. We also have the added advantage of going computational.

Exercise 296 *Determine the null space of the operator $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ if T is represented by the matrix*

$$\begin{bmatrix} 1 & 3 & 2 \\ -2 & 1 & 0 \end{bmatrix}$$

Here's how the equivalence can be made:

Let $T : X \longrightarrow Y$ be a linear operator with X, Y normed spaces:

Let $\dim X = n < \infty$ and $\dim Y = r < \infty$ and let basis of X be $e_1, e_2, ..., e_n$. Then, every vector x in the domain can be represented using scalars α_i's such that

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

Applying the linear operator, we get

$$y = T(x) = \sum_{i=1}^{n} \alpha_i T(e_i)$$

If $\bar{e}_1, \bar{e}_2, ..., \bar{e}_r$ are the basis of the range, then every vector y can be represented as

$$y = \sum_{i=1}^{r} \beta_i \bar{e}_i$$

Now, every $T(e_k)$ is a vector in the range. Hence, this, too can be represented as

$$T(e_k) = \sum_{i=1}^{r} \gamma_{ik} \bar{e}_i$$

where β_i, γ_i are scalars in the field of the codomain. The scalar γ will vary, depending on the vector $T(e_k)$, which justifies the subscript. Now, the two representations of y should agree. That is,

$$y = \sum_{k=1}^{r} \beta_k \bar{e}_k = \sum_{i=1}^{n} \alpha_k T(e_k)$$

This equation implicitly assumes that we can know the (unique) images of each member of the basis of the domain. The representation of the vector $T(e_k)$ is placed into this equation to give

$$\begin{aligned} y &= \sum_{i=1}^{n} \alpha_k T(e_k) = \sum_{i=1}^{n} \alpha_k \sum_{i=1}^{r} \gamma_{ik} \bar{e}_i \\ &= \sum_{i=1}^{r} \sum_{i=1}^{n} (\alpha_k \gamma_{ik}) \bar{e}_i \end{aligned}$$

Now, y cannot have two different representations. We must have $\beta_i = \sum_{i=1}^{n} (\alpha_k \gamma_{ik})$ for each i. This should look familiar: it is a tuple of a vector b if you perform the matrix multiplication $Ax = b = (\beta_i)$. Now, if we can determine these β_i's, we know the value of $T(x)$. By now, it should be clear that in order to determine the matrix equivalent A of T, we can safely say that $A = (a_{ik}) = (\gamma_{ik})$. Notice that this depends on the choice of basis for the domain so that there can be many different matrices by changing the choice of basis of the domain.

Note that this is valid only for finite dimensional spaces!

Let us do an example to hit the point home.

Example 297 *Let's say we have an operator that skews a vector and reduces a dimension. That is,* $T(x, y, z) = (3x, 2y)$. *To make our lives simple, we will assume* $e_1 = (1, 0, 0), e_2 = (0, 1, 0)$ *and* $e_3 = (0, 0, 1)$ *as our basis in both spaces. Now,* $T(e_1) = (3, 0)$, $T(e_2) = (0, 2)$ *and* $T(e_3) = (0, 0)$. *Therefore, the corresponding matrix is*

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}$$

so that

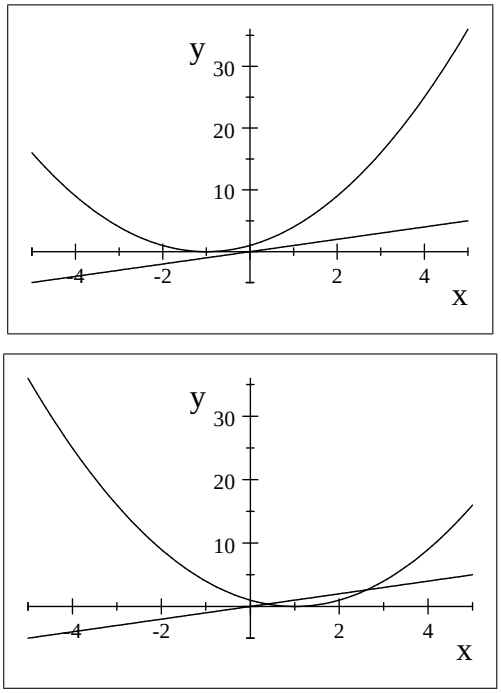
$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Exercise 298 *Let* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *be defined as follows:* $T(x, y, z) = (x, y, -x - y)$. *Find* $\mathcal{N}(T)$, $\mathcal{R}(T)$ *and the matrix that forms this operator.*

1.19 Application: Fixed Point Theory

For a real valued function $f : \mathbb{R} \longrightarrow \mathbb{R}$, a **fixed point** is a point $x \in \mathbb{R}$ such that $f(x) = x$. This can be seen as the intersection of the graph of a function

with the line $y = x$. As another example, the point $x = -1$ for $f(x) = 3x + 2$ is a fixed point. There may be cases were the graph of $f(x)$ does not intersect with $y = x$ or may even have more than one points of intersection. A parabola, for instance, will usually have two. In particular, the function $f(x) = x$ has infinite fixed points. Thus, a fixed point may not exist, may exist but may not be unique. One way this idea is important is the use in finding roots of an equation as follows: write out $f(x) = 0$ in the form $g(x) = x$. Thus, a fixed point of g will be the same as the root of f.



An interesting problem, therefore, is the determination of the existence (and uniqueness) of a fixed point. Of course we have no reason to limit ourselves to the real numbers. Let $(X, \|.\|_X)$, $(Y \|.\|_Y)$ be two normed spaces and let $T : X \longrightarrow Y$ be an operator. The subscript is a reminder of the norm being defined for a particular set.

Definition 299 *T is **Lipschitzian** if $\exists$ Lipschitzian constant α such that $\|T(x) - T(y)\|_Y \leq \alpha \|x - y\|_X$. T is **non-expansive** if $\alpha = 1$. T is a **contractive** map if $\|T(x) - T(y)\|_Y < \|x - y\|_X$. T is a **contraction** if $\alpha \in (0, 1)$*

Note that every contraction map is contractive but the converse is not true. $T(x) = 2x$ is Lipschitzian for $\alpha = 2$ under the usual Euclidean norm but not non-expansive. $T(x) = x$ is nonexpansive but not contractive since $\|T(x) - T(y)\|_Y = \|x - y\|_X$. For $X = (1, \infty)$, $f(x) = x + 1/x$ is contractive but not a contraction since $|f(x) - f(y)| = |(x - y) + (1/x - 1/y)| = |x - y|\,|1 - 1/xy| < |x - y|$

We will now drop the vexing subscript notation.

Theorem 300 *Every Lipschitz mapping is uniformly continuous.*

Proof. Let $\epsilon > 0$. Choose $\delta = \epsilon/\alpha$. Then, $\|x - y\| < \delta \implies \|T(x) - T(y)\| \leq \alpha \|x - y\| < \epsilon$ ■

Theorem 301 *For $X = \mathbb{R}$ and usual metric d, $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a contraction $\iff$ $|g'(x)| \leq \alpha < 1$ for continuous g.*

Proof. $(\Longleftarrow)$ $\frac{g(x)-g(y)}{x-y} = g'(t)$ for $t \in (x - y - \delta, x - y + \delta)$. From Mean Value Theorem, we have $|g(x) - g(y)| \leq \alpha |x - y|$

$(\Longrightarrow)$ $|g(x + h) - g(x)| \leq \alpha |h|$ ■

We now have a look at the solution to the problem posed at the beginning of this section.

Theorem 302 (Banach's Fixed Point Theorem) *Every contraction map $T : X \longrightarrow X$ on a complete norm space X has a unique fixed point x.*

Proof. Let $x_0 \in X$. We need to have repeated images of T. Define $T(x_0) = x_1$, $x_2 = T(x_1) = T^2(x_0), ...$ so that $x_n = T^n(x_0)$ and $x_{n+1} = T(x_n)$. A routine verification will show that $\|x_{m+1} - x_m\| \leq \alpha^m \|x_0 - x_1\|$. Using the triangle inequality and the formula for the sum of a geometric progression, assuming $n > m$, we can have $\|x_m - x_n\| \leq \frac{\alpha^m}{1-\alpha} \|x_0 - x_1\|$. Now, $\alpha < 1$ implies $\alpha^m \longrightarrow 0$ so that x_n is a Cauchy series. Hence there exists x such that $x_n \longrightarrow x$ because of completeness.

Next, $\|x - T(x)\| \leq \|x - x_m\| + \|x_m - T(x)\|$

$\leq \|x - x_m\| + \alpha^m \|x_{m-1} - x\| \longrightarrow 0$ so that $\|x - T(x)\| \leq 0$ in the limiting case, implying $T(x) = x$. This is our fixed point.

If there were another fixed point $y = T(y)$, then $\|x - y\| = \|T(x) - T(y)\| \leq \alpha \|x - y\|$, implying the contradiction $\alpha \geq 1$ ■

The proof covered in the lectures is for a closed space of a complete metric space. This theorem is rather more general (every closed space of a complete space is complete). The idea is to form a Cauchy sequence by repeatedly applying T to X. This has the effect of "reducing the domain" because of the scalar < 1 forcing T to converge to "singleton" – the fixed point. This exists since the domain is complete. Hence, the proof is invalid if completeness is taken away.

Exercise 303 *Come up with an example. Hint: use the rational numbers.*

In analysis, a usual sufficient conidtion for the convergence of an iteration $x_n = g(x_{n-1})$ is that g be continuously differentiable and $|g'(x)| \leq \alpha < 1$. Verify this by Banach's fixed point theorem.

We have to show that the mapping g is contractive. Then it will have a fixed point and hence the iteration will converge. This is a valid mode of reasoning, as we can see from the proof of Banach's fixed point theorem covered in the lecture notes.

We know that $g'(x)$ exists and is continuous. Thus, we have $g'(x) = \lim_{y\to x} \frac{g(y)-g(x)}{y-x} = \lim_{y\to x} \frac{g(x)-g(y)}{x-y}$ on which we apply the mean value theorem. This says that if f is a real continuous function on $[a,b]$ which is differentiable on (a,b), then there is a point $c \in (a,b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Hence for any interval, we apply MVT to get $g'(t) = \frac{g(x)-g(y)}{x-y}$ for some t. Now, by hypothesis, $1 > \alpha \geq |g'(t)| = \left|\frac{g(x)-g(y)}{x-y}\right|$

which is $|g(x) - g(y)| \leq \alpha |x-y|$. This is the familiar $d(g(x), g(y)) \leq \alpha d(x,y)$ for $\alpha < 1$

Exercise 304 *Show that* $\|x_m - x\| \leq \frac{\alpha^m}{1-\alpha} \|x_1 - x_0\|$ *and* $\|x_m, x\| \leq \frac{\alpha}{1-\alpha} \|x_m - x_{m-1}\|$.

For obvious reasons, the former is called prior estimate and the latter is called the posterior estimate. The problems below will show the significance of the prior and posterior estimate.

Problem 305 *Let f be a real-valued twice differntiable function on the interval $[a,b]$. Let $f(\hat{x}) = 0$ for some $\hat{x} \in (a,b)$. Newton's method defined as $x_{n+1} = g(x_n)$ and $g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$. Show that g a contraction in some neighbourhood of $\hat{x}$*

Solution 306 $g(x) = x + \frac{f(x)}{f'(x)} \implies g'(x) = 1 = f(x) f''(x) / [f'(x)]^2 \implies \lim_{x\to\hat{x}} g'(x) = 0 \implies |g'(x)| < \epsilon$

What if we need to find $x = \sqrt[r]{c}$? That is, the solution to $x^r - c = 0$. Let $f(x) = x^r - c$. Aply Newton's formula to get $x_{n+1} = x_n + \frac{f(x_n)}{f'(x_n)} = g(x_n) = \frac{1}{r}\left(x_n + \frac{c}{x_n}\right)$. This map is contraction with $\alpha = |1/r\,(1 - c/xy)|$

Exercise 307 *Let $X = [1,\infty)$ and let $T : X \longrightarrow X$ be such that $T(x) = \frac{x}{2} + \frac{1}{x}$. Show that T is a contraction and find the smallest α*

Solution 308 $d(T(x), T(y)) = \left|\frac{x}{2} + \frac{1}{x} - \frac{y}{2} - \frac{1}{y}\right| = \left|\frac{x-y}{2} - \frac{x-y}{xy}\right| = |x-y|\left|\frac{1}{2} - \frac{1}{xy}\right| = \alpha d(x,y)$ *where* $\alpha = \left|\frac{1}{2} - \frac{1}{xy}\right|$. *Now, what is the smallest such α? What is the largest value of $\frac{1}{xy}$? 1. What is the smallest? 0. Hence $\alpha \leq 1/2$ and therefore T is a contraction*

Exercise 309 *Consider an iteration process for solving $f(x) = x^3 + x - 1$. Form $x_n = g(x_{n-1})$. One way is by considering $g(x) = 1 - x^3$. Is $|g'(x)| < 1$?*

We can also have other forms as well. For each form, try $x_0 = 1$. Which converges faster? The real root is 0.682328.

Solution 310 *We have to find a root for* $f(x) = x^3 + x - 1$.

That is, a value of x *such that* $x^3 + x - 1 = 0$. *Now, we can have* $x = 1 - x^3$, *which we call* $g_1(x)$. *We, therefore have to find a fixed point of* $g_1(x)$. *One other form* $g_2(x)$ *comes from* $x = \sqrt[3]{1-x}$. *Thus,* $g_2(x) = \sqrt[3]{1-x}$. *Yet another form comes from* $x\left(1+x^2\right) - 1 = 0$ *so that* $x = \frac{1}{1+x^2} := g_3(x)$. *Yet another way is given by* $x^2 = \frac{x}{1+x^2}$ *which implies* $g_4(x) := x = \sqrt{\frac{x}{1+x^2}}$. *In each case, we can set up an iterative procedure as follows:*

$$x_n = g(x_{n-1})$$

A deeper look at the proof of Banach's fixed point theorem will tell you that this is the correct way of looking at it. Let's start with $x_0 = 1$. *Then,* $g_3(1) = 0.500 = x_1$

Now, $x_2 = g_3(x_1) = 0.800 = x_3$

$g_3(x_2) = 0.610$

In this case, $|g_3'(x)| < 1$ *because*

$$g_3'(x) = -\frac{2x}{\left(1+x^2\right)^2}$$

and $\left(1+x^2\right)^2 > 1$ *if* $x \neq 0$ *which implies* $\frac{1}{(1+x^2)^2} < 1$ *which implies* $\frac{2x}{(1+x^2)^2} < 1$ *if* $x \in (0,1)$

To apply Banach's fixed point theorem, first we need to consider a domain. Let's take $[0,1]$. *This is closed subspace of a complete space and hence complete. You may argue that this was created out of thin air but you can apply Newton's bijection method to get this domain. Since the function does not have a derivative at the end points, therefore the open interval is justified.*

In the case of $g_1'(x) = -3x^2$, *the absolute value* $3x^2 \geq 1$ *for* $x \in (0,1)$. *Thus, the answer to the first part is NO.*

Now, by the discussion on $|g'(x)| < 1$, *a fixed point will exist by Banach's fixed point theorem if the domain is complete and if* $|g'(x)| \leq \alpha < 1$ *and that* $g(x)$ *is continuously differentiable, which, being a polynomial, it is. The convergence will depend upon the different values of* α. *Think of a shorter value for* α *as a value that compresses the domain much faster. I will use my knowledge of calculus to find* α *for* g_3 *only and you will do the rest. Look at*

$$g_3'(x) = -\frac{2x}{\left(1+x^2\right)^2}$$

as a function, say $f(x)$. *How do we find it's maximum value? Take its derivative and set it to zero.*

$\frac{d}{dx}\left(-\frac{2x}{(1+x^2)^2}\right) = \frac{2\left(3x^2-1\right)}{(x^2+1)^3} = 0$ *implies* $3x^2 - 1 = 0$ *and, therefore,* $x = \frac{\sqrt{3}}{3}$. *Put this in* $g_3'(x)$ *to get* $\frac{2\frac{\sqrt{3}}{3}}{\left(1+\left(\frac{\sqrt{3}}{3}\right)^2\right)^2} = 0.649\,52$. *This is our* α.

Part II

Hilbert Spaces

This part of the handouts will focus on Hilbert spaces. We have set some terminology and concepts straight using our pre-established knowledge and shedding light on it as well. To begin with a study of functional analysis, we will first need to go through functionals and then move on to study Hilbert spaces

1.20 Functionals

Just like operators are mappings between spaces, we have functionals operating on vector and norm spaces to their fields.

Definition 311 *Let X be a norm space over $\mathbb{F}$. Then, a **functional** is a mapping $f : X \longrightarrow \mathbb{F}$.*

Since we are chiefly concerned with real and complex fields, functionals with only such ranges will be considered. From now on, f, g, h,... will be used to denote functionals. $f(x)$ will be an element of underlying (real or complex) field. The same notation $\mathcal{D}(f)$ and $\mathcal{R}(f)$ will be used to indicate domain and range of the functional f.

Essentially, functionals are "operators" so that all the previous theorems and definitions or linearity, boundedness and norm apply. We will revisit the definitions. Note that the only changes involved are those related to the difference between fields and norm spaces.

Definition 312 *A **linear** functional is a functional such that*

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

where α and β are any scalars and x, y are any vectors.

Needless to say, $f(x)$ and $f(y)$ are not vectors anymore but rather elements of a field.

Exercise 313 *Show that f is linear if and only if $f(\alpha x) = \alpha f(x)$ and $f(x + \beta) = f(x) + f(y)$*

Definition 314 *A linear functional is **bounded** if there exists a number $c > 0$ such that $|f(x)| \leq c \|x\|$.*

Note the norm on $f(x)$.

This can again give us the definition of the bound for a functional:

$$\|f\| = \sup_x \frac{|f(x)|}{\|x\|}$$

Clearly, $|f(x)| \leq \|f\| \|x\|$.

Example 315 *The norm $\|.\| : N \longrightarrow \mathbb{R}$ is a functional on a normed space N. This operator is not linear. To see why, it is sufficient to note that $\|\alpha x + \beta y\| \leq |\alpha| \|x\| + |\beta| \|y\|$ but rarely equal to $\alpha \|x\| + \beta \|y\|$. Can you tell when the norm becomes linear?*

Example 316 *The dot product*

$$f(x) = \sum^{n} x_i v_i$$

with a fixed vector $v = (v_1, v_2, ..., v_n)$ *is a functional on* $\mathbb{R}^n$. *This operator is linear and bounded. Linearity is clear:*

$$\begin{aligned} & f(\alpha x + \beta y) \\ = & \sum^{n} (\alpha x_i + \beta y_i) v_i \\ = & \alpha \sum^{n} x_i v_i + \beta \sum^{n} y_i v_i \\ = & \alpha f(x) + \beta f(y) \end{aligned}$$

To see why this functional is bounded, $|f(x)| = \|\sum x_i v_i\| \leq \|x\| \|v\|$. *Here,* $\|v\|$ *is the required* c *and so* $\|f\| = \|v\|$

Example 317 *This can be extended on* l^2 *where such a dot-product is simply extended for infinite tuples. Linearity can be seen from the above example whereas boundedness can be seen from the Cauchy-Schwarz inequality.*

Example 318 *We've considered the integral operator. Notice that the range of the opeartor consisted of functions since the integral sign was without limits. Even if there were limits, they were of an independent variable. What happens when we have constants as limits? The integral opertor then gives us the area under a given function. This can be rephrased as follows:*

$$f(x(t)) := \int_a^b x(t)\, dt$$

is a functional that gives us the area of the region under $x(t)$. *Thus, it is easy to see why* $f : C[a,b] \longrightarrow \mathbb{R}$ *acts on a norm space to an underlying field. We have already proved that this operator is linear. To prove that it is bounded, recall that* $\|x\| = \max_t x(t)$. *Now,*

$$\begin{aligned} |f(x(t))| &= \left| \int_a^b x(t)\, dt \right| \\ &\leq (b-a) \max_t x(t) \\ &= c \|x\| \end{aligned}$$

where we have used the geometric argument for area under a curve and area of a rectangle. Here, $\|f\| = c = b - a$

Example 319 *Another functional* $f : C[a,b] \longrightarrow \mathbb{R}$ *defined as*

$$f(x) = \int_a^b x(t)\, y_o(t)\, dt$$

is linear and bounded for any $y_o(t) \in C[a,b]$*. The proof follows the same pattern as above.*

Example 320 *Yet another functional on the same space can be defined as* $f_{t_0}(x(t)) = x(t_0)$*, the rationale being that a polynomial can be evaluated for a certain* $t_0 \in [0,1]$*. However, we move a step ahead and consider another functional* $f : C[a,b] \longrightarrow \mathbb{R}$ *defined as*

$$f(x) = k_1 x(a) + k_2 x(b)$$

for some fixed $k_1, k_2 \in \mathbb{R}$ *and* $x(t) \in C[a,b]$*. We will show that this is linear. For* $f(\alpha x + \beta y)$*, we have*

$$\begin{aligned} & f(\alpha x + \beta y) \\ = \;& k_1(\alpha x(a) + \beta y(a)) + k_2(\alpha x(b) + \beta y(b)) \\ = \;& \alpha k_1 x(a) + k_1 \beta y(a) + k_2 \alpha x(b) + k_2 \beta y(b) \\ = \;& \alpha(k_1 x(a) + k_2 x(b)) + \beta(k_1 y(a) + k_2 y(b)) \\ = \;& \alpha f(x) + \beta f(y) \end{aligned}$$

Furthermore,

$$\begin{aligned} |f(x)| &= |k_1 x(a) + k_2 x(b)| \\ &\leq |k_1|\,|x(a)| + k_2\,|x(b)| \\ &\leq (k/2)(|x(a)| + |x(b)|) \\ &\leq k\,|x(t)| \\ &\leq k \max_{t\in[a,b]} |x(t)| \\ &= k\,\|x(t)\| \end{aligned}$$

where $k = \max(|k_1|, |k_2|)$ *and* $x(t) = \max(x(a), x(b))$ *so that* $\|f\|$ *exists and can be found by varying* x *and thus* t *in* $\|f\| = \sup_x \frac{|f(x)|}{\|x(t)\|}$*. On the other hand, for the former case,* $\|f_{t_0}\| = \sup_{\|x\|=1} |x(t_0)| = 1$

Example 321 *We've already proved that the linear operator* $T(x_n) = nx$ *is not bounded. A some-what related functional is* $f(x_n) = x_i$ *where this* i *is a fixed integer. Clearly, this functional is linear. To show that it is bounded, we show that it is continuous, instead. Recall that this is valid for operators. An equivalent formulation is available for functionals, which you will have to prove.*

Now, take a sequence of sequences (this might get messy) X_m which converge to a sequence x_n that is, $\lim_{m\to\infty} X_m = x_n$. Now, $f(X_m) = x_i$ (element, not sequence) for each m, that is for each sequence X_k, we will have an x_i. Apply limits on both sides to get $\lim_{m\to\infty} f(X_m) = x_i = f(x_n)$ by the limit definition of continuity.

Example 322 *It is easy to see that the functional $f : C[-1,1] \longrightarrow \mathbb{F}$ defined by*

$$f(x) = \int_{-1}^{0} x(t)\,dt - \int_{0}^{1} x(t)\,dt$$

is linear. To find its norm, $\|f\| = \sup_{\|x\|=1} \left| \int_{-1}^{0} x(t)\,dt - \int_{0}^{1} x(t)\,dt \right| = 1 + 1 = 2$ by geometric arguments.

Example 323 *The functional $f(x) = \max_{t\in J} x(t)$ is not linear since*

$$f(x+y) \leq \max_{t\in J} x(t) + \max_{t\in J} y(t)$$

However, the norm for this functional is clearly 1. Similarly,

$$g(x+y) = \min_{t\in J} [x(t) + y(t)] \geq g(x) + g(y)$$

is not linear but bounded, as well.

We will have more to say on the following but we mention this in passing: to every space, we can define an operator on that space which takes its elements to its subspace. That is, $T : V \longrightarrow W$ where $W \subset V$. If this operator is linear, then we have for ourselves an endomorphism. Such operators project the space to a subspace, in a loose sense. For instance, a three dimensional vector may be reduced to a two dimensional one or even a single dimensional vector. This is a scalar in a practical sense of the word. To this end, we have the following example: let $f : l^p \longrightarrow \mathbb{R}$ be a functional such that $f(x) = f((\xi_i)) = \xi_n$ for some fixed n. This functional is linear and bounded

Proof. $f(\alpha x + \beta y)$
$= f((\alpha\xi_i) + (\beta\eta_i))$
$= f((\alpha\xi_i + \beta\eta_i))$
$= \alpha\xi_n + \beta\eta_n$
$= \alpha f(x) + \beta f(y)$
Also, $\|f\| = \sup_{\|x\|=1} |\xi_n| \leq 1$ ■

If we were to define a new functional from this such that $g = \overline{f}$, then g would no longer be linear: indeed, $g(\alpha x) = \overline{f}(\alpha x) = \overline{\alpha}\overline{f}(x) = \overline{\alpha} g(x)$. However, since $\|z\| = \|\overline{z}\|$, we therefore have that $\|f\| = \left\|\overline{f}\right\| = \|g\|$ so that g is bounded.

Definition 324 (Recollection) *A linear functional is called* ***continuous*** *at* $x \in \mathcal{D}(f)$ *if* $\forall \epsilon > 0$*, there exists* $\delta > 0$ *such that*

$$\|x - y\| < \delta \implies |f(x) - f(y)| < \epsilon$$

This is the same definiton for function continuity except that we have taken the liberty to invoke the norm defined on the given vector space and field.

Try to prove the following:

Exercise 325 $\|f^n\| \leq \|f\|^n$

Exercise 326 *f is continuous if and only if f is bounded*

Exercise 327 *$\|f\|$ obeys the properties of a norm.*

This exercise is specially useful for what follows in the next chapter.

Exercise 328 *If f is continuous at a single point, then it is continuous.*

Exercise 329 *If f is bounded and $x_n \longrightarrow x$, then $f(x_n) \longrightarrow f(x)$*

Exercise 330 *If f is linear, onto and bounded and there exists $b > 0$ such that $|f(x)| \geq b\|x\|$, then f^{-1} exists and is bounded.*

Theorem 331 *Let V be a vector space over field $\mathbb{F}$ and let $f : V \longrightarrow \mathbb{F}$. Then, f is either trivial (equal to 0 everywhere) or surjective.*

Proof sketch. This follows since just as the image of a vector subspace under a linear transformation is a subspace, so is the image of V under f. ■

Theorem 332 *A linear functional is continuous if and only if its kernel is closed*

1.20.1 Dual Spaces

We now move on to another fundamental study of the subject matter. Just like we can have for ourselves a norm space of bounded linear operators, we can have for ourselves a space of functionals. All we do is collect bounded functionals and have for ourselves a norm space. Hopefully, you will have proved this in the exercise of the previous chapter.

Definition 333 *The collection of all functionals on a vector space V over $\mathbb{F}$ is called the* ***algebraic dual space*** *V^* of V.*

Since these are functionals, we can use our previous knowledge of functions to give us our addition and scalar multiplication binary operators. That is, for $f_1 + f_2 \in V^*$, then

$$+ : V^* \times V^* \longrightarrow V^*$$

and

$$\cdot : \mathbb{F} \times V^* \longrightarrow V^*$$

Then

$$+ (f_1, f_2) = (f_1 + f_2)(x) = f_1(x) + f_2(x)$$

and

$$\cdot (\alpha, f) = (\alpha f)(x) = \alpha (f(x))$$

In this way, the additive identity is the zero function $\hat{O}(x) = 0$ to give us a vector space V^*.

We can go a step further ahead and consider the algebraic dual space $(V^*)^*$ of the dual space V^*, called the second algebraic dual V^{**}. This is the space of functionals on the dual space itself. We can move on and on but for now, second algebraic dual spaces will suffice.

Here is one purpose of considering the second algebraic dual space: we can define functionals of V^* as follows: $g(f)$. Remember, $g \in V^{**}$ and f acts as an input variable, much like $f \in V^*$ and acts on elements $x \in V$. Just like we can vary x to find different values for $f(x)$, likewise we can vary f to find different values of g. If we fix an $x \in V$, then one way of defining $g(f)$ is as follows: $g(f) = g_x(f) = f(x)$, with the subscript reminding us what to do with f.

This g is linear, keeping the x fixed.

Proof. $g(\alpha f_1 + \beta f_2)$
$= (\alpha f_1 + \beta f_2)(x)$
$= (\alpha f_1)(x) + (\beta f_2)(x)$
$= \alpha (f_1(x)) + \beta (f_2(x))$
$= \alpha g(f_1) + \beta g(f_2)$ ■

Since V^{**} is the collection of linear and bounded functionals on V^*, g_x really is an element of V^{**}. Just as we have kernels or null spaces of a specific mapping, that is, $\ker g = N(g) = \{x \mid g(x) = 0, x \in \mathcal{D}(g)\}$, we can also have a null space or a kernel of the entire vector space itself. In this case,

$$N(V) = \{x \mid f(x) = 0 \ , \forall f \in V^*\}$$

We can of course do the same with the algebraic dual space in which every such functional is considered. That is,

$$N(V^*) = \{x \mid g_x(f) = f(x) = 0 \ , \forall g \in V^{**}\} = N(V)$$

This is indeed a vector subspace of V.

Proof. Let $x, y \in N(V^*)$ such that $f(x) = f(y) = 0$ for any $f \in V^*$. Then, $f(\alpha x + \beta y) = 0$ hence $\alpha x + \beta y \in N(V^*)$ ■

In either case, $\dim N(g)$ and $\dim N(V^*) \leq n$ if $\dim V = n$. These facts follow from the fact that both are subspaces and from the fact that $\dim V = \dim V^*$ (proved later)

Notice the similarities and differences between the null space of an operator and the null space the vector space itself.

We now move to a justification of that subscripted x to consider a relationship between V and V^{**}. Let us define a mapping as follows and call it the

canonical mapping: $C : V \longrightarrow V^{**}$ such that $C(x) = g_x$. This mapping is linear

Proof. $C(\alpha x + \beta y)$
$= g_{\alpha x+\beta y}$
$= f(\alpha x + \beta y) \ \forall f$
$= \alpha f(x) + \beta f(y)$
$= \alpha g_x + \beta g_y$
$= \alpha C(x) + \beta C(y)$ ■

In mathematical literature, this mapping is also called the canonical embedding of V into V^{**}. Since this operator C is linear and takes elements from a vector space to another vector space, we have for ourselves a vector space homomorphism! Provided that this mapping is bijective, we then have for ourselves an isomorphism. The choice of the word "embedding" should be clear from the choice of domain and range and the fact that we have a isomorphism to a subset of the codomain. This is also stated as follows: V is embeddable into V^{**}. The mapping C is one-to-one provided that the functionals f are injective. i.e. if we have two functionals g and h, then they are influenced because of different elements from the domain.

Proof. Let $C(x) = C(y)$. Then, $g_x = g_y$ and $f(x) = f(y) \ \forall f \implies x = y$ ■

Thus, if we limit the codomain to the range and assume that every functional f on V is injective, then we have for ourselves a bijective C and thus an isomorphism. If the codomain and the range are already the same, then we have for ourselves an isomorphism without limiting the codomain.

This requirement is equivalent to the following: if $f(x) = 0$ for all f, then $x = 0$ and will be proved when we have the proper machinery for it.

We have, therefore, justified the following definition:

Definition 334 *A vector space V is said to be **algebraically reflexive** if it is isomorphic to its second algebraic dual space V^{**}.*

All finite dimensional spaces are algebraically reflexive. To prove this, you need to prove that the canonical mapping is bijective. The proof for this fact can wait for now, as we move on to look at some more properties for finite dimensional spaces.

By now, you should have had a fair idea that algebra, analysis and geometry go hand-in-hand and that this three-way traffic often helps to bring about surprises in each field. In a similar vein, we have the following definition:

Definition 335 *Let $f : N \longrightarrow \mathbb{R}$ be a non-zero functional on a real normed space N. Then, for any scalar c, we have a **half space** $H_c = \{x \in N \mid f(x) = c\}$.*

The half space is a subset of those vectors in a norm space that have a specific functional value and is closed and convex. Try to prove this. The use of the word indicates that this collection seperates N into two spaces by the law of trichotomy: $\{x \in N \mid f(x) \leq c\}$ and $\{x \in N \mid f(x) \geq c\}$. Of course this is possible only if the underlying field is real (which in our case it is). If $c = 1$, then we will call such a space a **hyper space**. Usually, the dimension of such

hyperspaces is one less than its ambient space i.e. the space surrounding the object. In our case, this space is N.

Now, just as we can have a matrix for an operator, we can have a matrix for a functional.

Definition 336 *Let $E = \{e_1, ..., e_n\}$ be a basis of X. $E^* = \{e_1^*, ..., e_n^*\}$ is an (algebraic) **dual basis** for the algebraic dual space X^* of X.*

Now this definition may not be exactly enlightening but was only mentioned to set some record straight. We will justify this definition in the theorem that follows but for now, let's look at it from a computational point of view (note that index i varies finitely). This is important because we want to be able to find the elements of a dual space. The computation follows the manner for operators. Thus, if $x = \sum_{i=1}^{n} \alpha_i e_i$, then $f(x) = \sum_{i=1}^{n} \alpha_i f(e_i) = \sum_{i=1}^{n} \alpha_i e_i^*$. For now, $f(e_i) = e_i^*$ is just a notation but we are trying to go in accord with the definition given above, as will hopefully be made clear. Notice that

$$\left[\begin{array}{ccccc} e_1^* & e_2^* & ... & e_{n-1}^* & e_n^* \end{array} \right] \left[\begin{array}{c} \alpha_1 \\ \vdots \\ \alpha_n \end{array} \right] = \sum_{i=1}^{n} \alpha_i e_i^*$$

Clearly, our required $1 \times n$ matrix A is, therefore, $A = (e_{1i}^*) = (f(e_i))$. By the construction principle for linear maps (above), there exists a linear functional $f_i = e_i^* \in E^*$ which maps e_i to 1 and the other basis vectors to 0. That is, for each basis e_k^*, $e_k^*(e_j) = f_k(e_j) = \delta_{kj}$ where δ_{ij} is the Kronecker delta function. Thus, if we have a vector $v = \sum_{i=1}^{n} \alpha_i e_i$ we must have

$$e_k^*(v) = f_k(v) = f_k\left(\sum_{i=1}^{n} \alpha_i e_i\right) = \sum_{i=1}^{n} \alpha_i f_k(e_i) = \alpha_k$$

which shows that the linear functional e_i^* maps every vector of X to its i-th coordinate with respect to the basis B. In order to be able to thus say that for all $x \in X$, $x = \sum_{i=1}^{n} f_i(x) e_i$, we need to show that the E^* is a linearly independent set and this will be done in the theorem below but before that, notice how the space and its dual are connected and that this construction works for any basis $E = \{e_1, ..., e_n\}$.

Example 337 *The dual basis for the basis $(1,0,0), (0,1,0)$ and $(0,0,1)$ in $\mathbb{R}^3$ can be found as follows: $(1,0,0)^T, (0,1,0)^T$ and $(0,0,1)^T$ where the superscript T indicates the transpose of this vector. Justify it to yourself that these transposed vectors do indeed form functionals and the basis for the algebraic dual of $\mathbb{R}^3$*

Exercise 338 *Let* $E=\{e_1,e_2,e_3\}$ *be a basis of* $\mathbb{R}^3$ *where*

$$\begin{aligned} e_1 &= (1,1,1) \\ e_2 &= (1,1,-1) \\ e_3 &= (1,-1,-1) \end{aligned}$$

Find $E^*=\{e_1^*,e_2^*,e_3^*\}=\{f_1,f_2,f_3\}$ *and determine* $f_1(x),f_2(x),f_3(x)$ *where* $x=(1,0,0)$

Theorem 339 *Let* X *be a vector space and* $E=\{e_1,...,e_n\}$ *be a basis of* X. *Then,* $E^*=\{e_1^*,...,e_n^*\}=\{f_1,f_2,...,f_n\}$ *is the basis for the agebraic dual* X^* *of* X *and* $\dim X=\dim X^*=n$

Proof. Take any linear combination of an element of X^* as

$$\sum_{i=1}^{n}\alpha_i f_i(x)=0$$

for any $x\in X$. Set $x=e_j$ to get

$$\sum_{i=1}^{n}\alpha_i f_i(e_j)=\sum_{i=1}^{n}\alpha_i\delta_{ij}=\alpha_j=0$$

so that the chosen linear combination is linearly independent. To show that every element $f\in X^*$ can be written in the linear combination above, observe that

$$f(x)=\sum_{i=1}^{n}\alpha_i f(e_i)=\sum_{i=1}^{n}\alpha_i e_i^*$$

where $f(e_i)=e_i^*$ for any $x\in X$. On the other hand, we also have

$$e_k^*(v)=f_k(v)=f_k\left(\sum_{i=1}^{n}\alpha_i e_i\right)=\sum_{i=1}^{n}\alpha_i f_k(e_i)=\alpha_k$$

Placing this value of α_k in the chosen linear combination, we get

$$f(x)=\sum_{i=1}^{n}\alpha_i f_i(x)$$

Since this is valid for $x\in X$, we must have

$$f=\sum_{i=1}^{n}\alpha_i f_i$$

■

We have just fermented an algebraic form of a functional in terms of functionals f_i's! We can now refine $\dim N(f)$ from $\leq n$ to $=n$

To prepare an interesting application of this basis, we present the following lemma

Lemma 340 *Let X be a finite dimensional vector space. If $x_0 \in X$ has the property that $f(x_0) = 0$ for all $f \in X^*$, then $x_0 = 0$*

Proof. Take $x_0 = \sum_{i=1}^{n} \alpha_i e_i$. For all $f \in X^*$, we have $0 = f(x_0)$
$= f\left(\sum_{i=1}^{n} \alpha_i e_i\right) = \sum_{i=1}^{n} \alpha_i f(e_i) = \sum_{i=1}^{n} \alpha_i e_i^* = 0$
Since $e_i^{*\prime}$'s are linearly independent, we must have $\alpha_i = 0 \forall i$. Hence $x_0 = \sum_{i=1}^{n} \alpha_i e_i \implies x_0 = 0$ ■

Exercise 341 *The canonical mapping is always injective.*

Solution 342 $C_x(f) = C_x(g) \implies f(x) = g(x) \implies (f-g)(x) = 0 \forall x \implies f - g = 0 \implies f = g$

Thus, $N(X^*) = \{0\}$ for a finite dimensional X. Note that this does not say that $N(f) = \{0\}$ for some $f \in X^*$. In fact, the dimension of the null space of a functional is always less than or equal to the dimension of the space X on which the functional is defined (why?).

Theorem 343 *A finite dimensional space is algebraically reflexive*

Proof. We need to prove that a canonical mapping $C : X \longrightarrow X^{**}$ is bijective. This mapping is linear, as has been proved. Let $C(x) = 0$. Then, $g_x = f(x) = 0$ for all $f \in X^*$ implies $x = 0$ so that C is one-to-one. Recall that

$$\dim X = \dim X^* = n$$

From this, we also have $\dim X^{**} = \dim X^* = n$. Thus, $\mathcal{R}(C) = X^{**}$ hence C is surjective. ■

Corollary 344 *The second algebraic dual space of $\mathbb{R}^n$ is $\mathbb{R}^n$*

In addition to the algebraic dual space for vector spaces, we have an equivalent concept, called simply dual space, for norm spaces.

Definition 345 *The collection of all functionals on a norm space N over $\mathbb{F}$ is called the **dual space** N' of N.*

Exercise 346 *Hopefully, you will have proved that for any functional*

$$f : N \longrightarrow \mathbb{F}$$

$\|f\|$ satisfies the axioms of a norm space. Therefore, the dual space is also a norm space, which you are required to prove.

Let us abuse some terminology to hit a point. Let us call $N' = C(X, \mathbb{F})$ where the $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ so that $\mathbb{F}$ is complete. Just like $B(X, Y)$ is Banach if Y is Banach, the dual space (not algebraic) is Banach whether or not the norm space N is Banach. We give a more general proof.

Proposition 347 *If X is a norm space and Y is a complete, then $C(X,Y)$, with the norm $\|f\| = \sup_x |f(x)|$, is complete.*

Proof. Suppose (f_n) is a Cauchy sequence in $C(X,Y)$, so, as $n \longrightarrow \infty$, $\|f_n - f_m\| \longrightarrow 0$. In particular $(f_n(x))$ is a Cauchy sequence in Y for each $x \in X$ since $\|f_n(x) - f_m(x)\| = \|(f_n - f_m)(x)\| \leq \|f_n - f_m\| \|x\| \longrightarrow 0$ so it converges, say to $f(x) \in Y$. It remains to show that $f \in C(X,Y)$ and that $f_n \longrightarrow f$. We have that $\|f_n(x) - f(x)\| \leq \|f_n - f\| \|x\| \longrightarrow 0 \ \forall x$ i.e., $f_n \longrightarrow f$ uniformly. It remains only to show that f is continuous. For this, let $x_k \longrightarrow x$ in X and let $\epsilon > 0$. Pick N so that $\epsilon_N < \epsilon$. Since f_N is continuous, there exists $K \in \mathbb{N}$ such that $k \geq K$

$\implies \|f_N(x_k) - f_N(x)\| < \epsilon$. Hence $k \geq K$

$\implies \|f(x_k) - f(x)\|$

$= \|f(x_k) - f_N(x) + f_N(x) - f(x) + f(x_k) - f(x_k)\|$

$\leq \|f(x_k) - f_N(x_k)\| + \|f_N(x_k) - f_N(x)\| + \|f(x) - f_N(x)\| \longrightarrow 0$ ■

Theorem 348 *The dual space of $\mathbb{R}^n$ is $\mathbb{R}^n$*

The words of the theorem might be slightly misleading: in one case, we have vectors whereas in the other we have functionals. How can they be equivalent? In particular, if we interpret $\mathbb{R}^n$ as the space of columns of n real numbers, its dual space is typically written as the space of **rows** of n real numbers. The situation works vice versa, as well. In such a case, this transposed vector acts as a functional, basing itself on matrix multiplication (see construction of dual basis). We have already seen that any functional can itself be given a basis representation and, therefore, is a vector in its own right. Thus, we need to be looking for an isomorphism – in particular an isometric isomorphism – that respects the structure of both spaces so that they are, in fact, equivalent.

Proof. Since $\mathbb{R}^n$ is both a vector and a norm space, $\mathbb{R}^{n\prime} = \mathbb{R}^{n*}$. If we take the standard basis, then any vector $x \in \mathbb{R}^n$ has the representation $x = \sum_{i=1}^{n} \alpha_i e_i$. Applying $f \in \mathbb{R}^{n*}$, we get $f(x) = \sum_{i=1}^{n} \alpha_i f(e_i) = \sum_{i=1}^{n} \alpha_i \gamma_i$. Then, $|f(x)| = \left|\sum_{i=1}^{n} \alpha_i \gamma_i\right|$

$\leq \left(\sum_{i=1}^{n} |\alpha_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |\gamma_i|^2\right)^{1/2}$ by Cauchy-Schwarz inequality

$= \|x\| \left(\sum_{i=1}^{n} |\gamma_i|^2\right)^{1/2}$

That is, $\frac{|f(x)|}{\|x\|} \leq \left(\sum_{i=1}^{n} |\gamma_i|^2\right)^{1/2}$ for $\|x\| \neq 0$

$\implies \sup_{\|x\|=1} \frac{|f(x)|}{\|x\|} \leq \sup_{\|x\|=1} \left(\sum_{i=1}^{n} |\gamma_i|^2\right)^{1/2}$

$\implies \|f\| \leq \left(\sum_{i=1}^{n} |\gamma_i|^2\right)^{1/2}$

If we choose $x = (\gamma_1, \gamma_2, ..., \gamma_n)$, then we must have equality instead of inequality so that $\|f\| = \left(\sum_{i=1}^{n} |\gamma_i|^2 \right)^{1/2} = \|x\|$

Since f is linear and $\|f\| = \|x\|$, that is, elements in the norm space $\mathbb{R}^{n\prime}$ preserve norm under f, we therefore have an isometric isomorphism between $\mathbb{R}^{n\prime}$ and $\mathbb{R}^n$ ■

Theorem 349 *The dual space of l^1 is l^∞*

Similar remarks of the previous theorem apply except for one detail: in the infinite dimensional case, we cannot contrsuct a corresponding basis for the dual space as our previous construction for a technical reason involving the Axiom of Choice, which we will not mention. The following proof, nevertheless, is still easy to understand and does not violate anything we have learned so far.

Proof. A Schauder basis for l^1 is (e_k) where $e_k = (\delta_{kj})_{j \in \mathbb{N}^+}$ so that $\|e_k\| = 1$. Hence we can have

$$x = \sum_{k=1}^{\infty} \xi_k e_k$$

Let $l^{1'}$ be the dual space of l^1. We have to prove that $l^{1'}$ is the same as l^∞. Applying $f \in l^{1'}$ on x, we get

$$f(x) = \sum_{k=1}^{\infty} \xi_k f(e_k)$$

We have such a representation because f is linear. The series converges since f is bounded (remember that $f \in l^{1'}$). Now, what can we say about the sequence $(f(e_1), f(e_2), f(e_3), ...)$? It, too, is bounded above since

$$|f(e_k)| \leq \|f\| \|e_k\| = \|f\|$$

from which we get

$$\|(f(e_k))\| = \sup_k |f(e_k)| \leq \|f\|$$

hence $(f(e_k)) \in l^\infty$. In a sense, we can therefore say that the mapping $(e_k) \mapsto (f(e_k))$ embeds l^1 to l^∞. We now do the converse.

Let $(\beta_k) \in l^\infty$ such that $g(x) = \sum_{k=1}^{\infty} \xi_k \beta_k$ for a functional g on l^1 for $x = \sum_{k=1}^{\infty} \xi_k e_k \in l^1$. This construction can be justified by appeal to the Axiom of Choice but we will let that pass for now.

In order to prove that $g \in l^{1'}$, we need to prove that g is linear and bounded. Linearity following by comparing

$$\sum_{k=1}^{\infty} \xi_k \beta_k = g\left(\sum_{k=1}^{\infty} \xi_k e_k \right)$$

which is only possible if $g(e_k) = \beta_k$. Next,

$$|g(x)| = \left|\sum_{k=1}^{\infty} \xi_k \beta_k\right| \leq \sum_{k=1}^{\infty} |\xi_k \beta_k|$$

by the triangle inequality on $\mathbb{R}$ and $\leq \left(\sum_{k=1}^{\infty} |\xi_k|^p\right)^{1/p} \left(\sum_{k=1}^{\infty} |\beta_k|^q\right)^{1/q}$ by the Hölder inequality for $\frac{1}{p} + \frac{1}{q} = 1$. Here, $p = 1$ and $q = \infty$ (don't take this notation too seriously. It has only been mentioned for emphasis). Thus,

$$|g(x)| \leq \sum_{k=1}^{\infty} |\xi_k| \sup_k \beta_k = \|x\| c$$

where $c = \sup_k \beta_k < \infty$ since $(\beta_k) \in l^{\infty}$. Thus, $g(x)$ is bounded and hence $g \in l^{1'}$, proving the required converse. We're not done yet, though. We haven't validated isometry, yet, so here it is:

$|f(x)| = \left|\sum_{k=1}^{\infty} \xi_k f(e_k)\right| \leq \sup_k |f(e_k)| \sum_{k=1}^{\infty} |\xi_k|$ again by the Hölder inequality so that $|f(x)| \leq \|x\| \sup_k |f(e_k)|$

$\Longrightarrow \frac{|f(x)|}{\|x\|} \leq \sup_k |f(e_k)|$

$\Longrightarrow \sup_x \frac{|f(x)|}{\|x\|} \leq \sup_x \sup_k |f(e_k)|$

We cannot vary x on the right hand side. Thus, $\|f\| \leq \sup_k |f(e_k)|$. We have also seen that $\sup_k |f(e_k)| \leq \|f\|$ so that $\|f\| = \sup_k |f(e_k)|$ establishing the required isometry. ■

As might be guessed from the proof above, dual space of l^p is l^q where $\frac{1}{p} + \frac{1}{q} = 1$. The proof of this fact can be seen from Lecture 18, MTH327.

Theorem 350 (Hahn-Banach Theorem) *Let $(X, \|.\|)$ be a normed space and let $Y \subseteq X$ be a subspace. For any $f \in X'$, there exists $\tilde{f} \in X^*$ such that $\tilde{f}$ is an extension of f ($\tilde{f}(y) = f(y)$ for any $y \in Y$) and $\left\|\tilde{f}\right\| = \|f\|$*

Corollary 351 *If $f(x) = 0$ for all $f \in X'$, then $x = 0$*

Corollary 352 *Let X be a normed space and let $x_0 \neq 0$ be any element of X. Then, there exists a linear bounded functional $\tilde{f}$ on Xsuch that $\left\|\tilde{f}\right\| = 1$ and $\tilde{f}(x_0) = \|x_0\|$*

Proof. Consider the subspace Y consisting of $x = \alpha x_0$. Define f on Y by $f(x) = \alpha \|x_0\|$. f is bounded has norm $\|f\| = 1$ because $|f(x)| = |f(\alpha x_0)| = |\alpha| \|x_0\| = \|x\|$

From the Hahn-Banach theorem, $\|f\| = \left\|\tilde{f}\right\| = 1$. Further, $\tilde{f}(x_0) = f(x_0) = \|x_0\|$ ■

Corollary 353 *For every $x \in X$, $\|x\| = \sup_{f \in X'} \frac{|f(x)|}{\|f\|}$*

Proof. $\sup_{f \in X'} \frac{|f(x)|}{\|f\|} \geq \frac{\left|\tilde{f}(x)\right|}{\left\|\tilde{f}\right\|} = \frac{\|x\|}{1} = \|x\|$

Conversely, $|f(x)| \leq \|f\| \|x\|$ implies $\sup_{f \in X'} \frac{|f(x)|}{\|f\|} \leq \|x\|$ ■

Corollary 354 *X" and X are isometric. That is, $\|g_x\| = \|x\|$*

Proof. $\|g_x\| = \sup_{f \in X'} \frac{|g_x(f)|}{\|f\|} = \sup_{f \in X'} \frac{|f(x)|}{\|f\|} = \|x\|$ ■

Theorem 355 (Principle of uniform boundedness (Banach-Steinhaus)) *Let $(X, \|.\|_X)$ be a Banach space, $(Y, \|.\|_Y)$ a normed space and $T_n : X \longrightarrow Y$ a bounded operator for each $n \in N$. Suppose that for any $x \in X$ there exists $C_x > 0$ such that $\|T_n x\|_Y \leq C_x$ for all n. Then there exists $C > 0$ such that $\|T_n\| \leq C$ for all n.*

Theorem 356 *If (x_n) is a sequence in a Banach space and $(f(x_n))$ is bounded for all $f \in X'$, show that $\|x_n\|$ is bounded.*

Proof. We will apply the uniform boundedness principle to the dual space X^*. This is complete, whether or not X is. The role of T_n will be played by $\hat{x}_n \in X^{**}$. Recall that $\hat{x}_n$ is defined as the bounded linear functional on X^* for which $\hat{x}_n(f) = f(x_n)$ $(f \in X^*)$. The assumption that $(f(x_n))$ is bounded means that for any vector f in our space X^* the sequence $\hat{x}_n(f)$ is bounded. Using the uniform boundedness principle we get that there exists C such that $\|\hat{x}_n\| \leq C$ for all n. From the corollary of Hahn-Banach theorem, $\|x_n\| = \|\hat{x}_n\|$.
■

Pre-Hilbert Space

A Pre-Hilbert space or an inner product space is so called because it is a vector space with an additional structure – that of the inner product. Why do we need inner product spaces in the first place? We have considered generalisations of the vector and the length of a vector but what we don't have is a generalisation of the ordinary dot product. Such spaces are thus richer are more important from a geometric point of view.

Recall that the dot product is originally $(a,b,c).(x,y,z) = ax + by + cz$ from which we have $(x,y,z).(x,y,z) = x^2 + y^2 + z^2$. This dot product can be generalised and be called the inner product, as the example will show but first its axioms.

Definition 357 *A vector space I over $\mathbb{F}$ endowed with the operation*

$$\langle .,. \rangle : I \times I \longrightarrow \mathbb{F}$$

called the inner product, is called an ***inner product space*** *or pre-Hilbert space if $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in I$ and $\alpha \in \mathbb{F}$ it satisfies the following axioms:-*

- $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
- $\langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$
- $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$
- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$
- $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}$

An inner product space I is compactly written as $(I, \langle .,. \rangle)$.

Example 358 *For $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$,*

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle (x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n) \rangle = \sum_{i=1}^{n} \bar{x}_i y_i$$

satisfies the above conditions and hence is an inner product space. The first axiom of the inner product space is satisfied since

$$\begin{aligned}
& \langle \mathbf{x}+\mathbf{y}, \mathbf{z}\rangle \\
= & \langle (x_1+y_1, x_2+y_2, ..., x_n+y_n), (z_1, z_2, ..., z_n)\rangle \\
= & \overline{(x_1+y_1)}z_1 + \overline{(x_2+y_2)}z_2 + ... + \overline{(x_n+y_n)}z_n \\
= & \overline{x_1}z_1 + \overline{y_1}z_1 + \overline{x_2}z_2 + \overline{y_2}z_2 + ... + \overline{x_n}z_n + \overline{y_n}z_n \\
= & (\overline{x_1}z_1 + \overline{x_2}z_2 + ... + \overline{x_n}z_n) + (\overline{y_1}z_1 + \overline{y_2}z_2 + ... + \overline{y_n}z_n) \\
= & \langle \mathbf{x}, \mathbf{z}\rangle + \langle \mathbf{y}, \mathbf{z}\rangle
\end{aligned}$$

Next,

$$\begin{aligned}
& \langle \alpha\mathbf{x}, \mathbf{y}\rangle \\
= & \langle \alpha(x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n)\rangle \\
= & \sum_{i=1}^{n} \alpha \bar{x}_i y_i = \alpha \sum_{i=1}^{n} \bar{x}_i y_i \\
= & \alpha \langle \mathbf{x}, \mathbf{y}\rangle
\end{aligned}$$

Also,

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y}\rangle &= \sum_{i=1}^{n} \bar{x}_i y_i \\
&= \sum_{i=1}^{n} \bar{x}_i \overline{\overline{y_i}} = \sum_{i=1}^{n} \overline{\overline{y_i}} \bar{x}_i \\
&= \sum_{i=1}^{n} \overline{\overline{y_i} x_i} = \overline{\sum_{i=1}^{n} \overline{y_i} x_i} \\
&= \overline{\langle \mathbf{y}, \mathbf{x}\rangle}
\end{aligned}$$

$\langle \mathbf{x}, \mathbf{x}\rangle \geq 0$ *and* $\langle \mathbf{x}, \mathbf{x}\rangle = 0 \iff \mathbf{x} = \mathbf{0}$ *are easy to see. Notice that in the field of real numbers and for* $n = 3$*, this is the dot product of vectors, as made extensive use of in Physics. If we let* $n \rightarrow \infty$*, then we have the space* l^2*.*

Example 359 *Let* Ω *be an open set.*

$$L^p(\Omega) = \{f : f \text{ is measurable on } \Omega \text{ and } \|f\|_p < \infty\}$$

where for $1 \leq p < \infty$ *,* $\|f\|_p = \left(\int_\Omega \|f\|^p \, dx\right)^{1/p}$ *and* $\|f\|_\infty = \sup_x \frac{|f(x)|}{\|x\|}$*.* L^p *spaces. It is from these spaces that Quantum Mechanics adopts its mechinary.* $L^2(-\infty, \infty)$ *is a particular example of a space of square-integrable function. Let's take a close look at the special case* $p = 2$ *where our ordinary rules of integration will be of aid withour resorting to concepts of Measure Theory. We can define a dot product of two real-valued functions as*

$$\langle x, y\rangle = \int x(t)y(t)dt$$

If we have a domain, then the integral becomes definite. If we have complex-valued functions, then the dot product becomes

$$\langle x, y\rangle = \int_a^b x(t)\overline{y(t)}dt$$

To prove that this is an inner product space, see that

$$\begin{aligned}
\langle \mathbf{x}+\mathbf{y}, \mathbf{z}\rangle &= \int_a^b (x+y)(t)\overline{z(t)}dt \\
&= \int_a^b (x(t)+y(t))\,\overline{z(t)}dt \\
&= \int_a^b \left[x(t)\overline{z(t)}+y(t)\overline{z(t)}\right]dt \\
&= \int_a^b x(t)\overline{z(t)}dt + \int_a^b y(t)\overline{z(t)}dt \\
&= \langle \mathbf{x}, \mathbf{z}\rangle + \langle \mathbf{y}, \mathbf{z}\rangle
\end{aligned}$$

Next,

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y}\rangle &= \int_a^b x(t)\overline{y(t)}dt \\
&= \int_a^b \overline{y(t)}x(t)dt \\
&= \int_a^b \overline{y(t)\overline{x(t)}}dt \\
&= \overline{\int_a^b y(t)\overline{x(t)}dt} \\
&\quad \overline{\langle \mathbf{y}, \mathbf{x}\rangle}
\end{aligned}$$

For the scaling axiom, we have

$$\begin{aligned}\langle \alpha\mathbf{x}, \mathbf{y}\rangle &= \int_a^b \alpha x(t)\overline{y(t)}dt \\ &= \alpha\int_a^b x(t)\overline{y(t)}dt \\ &= \alpha\langle \mathbf{x}, \mathbf{y}\rangle\end{aligned}$$

Postivity axioms $\langle \mathbf{x}, \mathbf{x}\rangle \geq 0$ *and* $\langle \mathbf{x}, \mathbf{x}\rangle = 0 \iff \mathbf{x} = \mathbf{0}$ *can similarly be proved.*

By now you must have probably studied random variables in STA365. Here's a supplementary example:

Example 360 *For random variables* X *and* Y*, the expected value of their product is an inner product. In this case,* $\langle X, X\rangle = 0$ *if and only if* $Pr(X = 0) = 1$*(i.e.,*$X = 0$ *almost surely). This definition of expectation as inner product can be extended to random vectors as well.*

Recall the definition of a linear operator. The inner product looks linear in the first argument. However, because of the axiom $\langle \mathbf{x}, \mathbf{y}\rangle = \overline{\langle \mathbf{y}, \mathbf{x}\rangle}$, we don't have linearity in the second argument with regards to scalar multiplication. In fact, we have $1\frac{1}{2}$ linearity, so to speak. This is called **sesquilinear**:

Proposition 361 *Let* $(X,\langle .,.\rangle)$ *be an inner product space. Then, for* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in X$ *over* $\mathbb{F}$ *and* $\alpha, \beta \in \mathbb{F}$*, the following properties hold:-*

- $\langle \alpha\mathbf{v}_1 + \beta\mathbf{v}_2, \mathbf{v}_3\rangle = \alpha\langle \mathbf{v}_1, \mathbf{v}_3\rangle + \beta\langle \mathbf{v}_2, \mathbf{v}_3\rangle$
- $\langle \mathbf{v}_1, \alpha\mathbf{v}_2 + \beta\mathbf{v}_3\rangle = \bar{\alpha}\langle \mathbf{v}_1, \mathbf{v}_2\rangle + \bar{\beta}\langle \mathbf{v}_1, \mathbf{v}_3\rangle$

Proof. For i)

$$\begin{aligned}\langle \alpha\mathbf{v}_1 + \beta\mathbf{v}_2, \mathbf{v}_3\rangle &= \langle \alpha\mathbf{v}_1, \mathbf{v}_3\rangle + \langle \beta\mathbf{v}_2, \mathbf{v}_3\rangle \\ &= \alpha\langle \mathbf{v}_1, \mathbf{v}_3\rangle + \beta\langle \mathbf{v}_2, \mathbf{v}_3\rangle\end{aligned}$$

and for ii) we have

$$\begin{aligned}\langle \mathbf{v}_1, \alpha\mathbf{v}_2 + \beta\mathbf{v}_3\rangle &= \overline{\langle \alpha\mathbf{v}_2 + \beta\mathbf{v}_3, \mathbf{v}_1\rangle} \\ &= \overline{\alpha\langle \mathbf{v}_2, \mathbf{v}_1\rangle} + \overline{\beta\langle \mathbf{v}_3, \mathbf{v}_1\rangle} \\ &= \bar{\alpha}\overline{\langle \mathbf{v}_2, \mathbf{v}_1\rangle} + \bar{\beta}\overline{\langle \mathbf{v}_3, \mathbf{v}_1\rangle} \\ &= \bar{\alpha}\langle \mathbf{v}_1, \mathbf{v}_2\rangle + \bar{\beta}\langle \mathbf{v}_1, \mathbf{v}_3\rangle\end{aligned}$$

■

Recall that the dot product is originally

$$(a,b,c)\,.\,(x,y,z) = ax + by + cz$$

from which we have $(x,y,z)\,.\,(x,y,z) = x^2+y^2+z^2$, which is basically $\|(x,y,z)\|^2$. In this way, we can call an inner product space I a norm space.

Example 362 *For the same space $\mathbb{F}^n$, we have the norm*

$$\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^{n} \overline{x_i} x_i = \sum_{i=1}^{n} |x_i|^2 = \|\mathbf{x}\|^2$$

from the dot product of a vector with itself. Now we also see why we need the conjugate and why the order makes sense, notwithstanding the conjugate. If these were real numbers, of course $|x_i|^2 = x_i$, which brings us back to the Euclidean norm and, of course, the usual metric.

Example 363 *For inner product space $L^2[a,b]$, we also have a norm induced from the dot product:*

$$\begin{aligned} & \langle \mathbf{x}, \mathbf{x} \rangle \\ = & \int_a^b x(t)\overline{x(t)}dt \\ = & \int_a^b |x(t)|^2\, dt \\ = & \|\mathbf{x}\|^2 \end{aligned}$$

It has already been proved that this norm makes $L^2[a,b]$ a normed space.

And now for the theorem itself.

Theorem 364 *Every inner product space is a normed space*

Proof. For $\|.\| : V \times V \longrightarrow \mathbb{F}$, define $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. This is justified since $\langle .,. \rangle : V \times V \longrightarrow \mathbb{F}$ carries the same "structure" as $\langle .,. \rangle : V \times V \longrightarrow \mathbb{F}$

Let V be an inner product space over $\mathbb{F}$.

$\forall\ \alpha \in \mathbb{F}$ and $\mathbf{x}, \mathbf{y} \in V$

By default, $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0} \implies \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \geq 0$ and

$$\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = 0 \iff \mathbf{x} = \mathbf{0}$$

Next,

$$\begin{aligned}
\|\alpha\mathbf{x}\| &= \sqrt{\langle\alpha\mathbf{x},\alpha\mathbf{x}\rangle} \\
&= \sqrt{\alpha\langle\mathbf{x},\alpha\mathbf{x}\rangle} \\
&= \sqrt{\alpha\bar{\alpha}\langle\mathbf{x},\mathbf{x}\rangle} \\
&= \sqrt{|\alpha|^2\langle\mathbf{x},\mathbf{x}\rangle} \\
&= |\alpha|\sqrt{\langle\mathbf{x},\mathbf{x}\rangle} \\
&= |\alpha|\,\|\mathbf{x}\|
\end{aligned}$$

Finally,

$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 &= \langle\mathbf{x}+\mathbf{y},\mathbf{x}+y\rangle \\
&= \langle\mathbf{x},\mathbf{x}+\mathbf{y}\rangle + \langle\mathbf{y},\mathbf{x}+\mathbf{y}\rangle \\
&= \langle\mathbf{x},\mathbf{x}\rangle + \langle\mathbf{x},\mathbf{y}\rangle + \langle\mathbf{y},\mathbf{x}\rangle + \langle\mathbf{y},\mathbf{y}\rangle \\
&= \|\mathbf{x}\|^2 + \langle\mathbf{x},\mathbf{y}\rangle + \langle\mathbf{y},\mathbf{x}\rangle + \|\mathbf{y}\|^2 \\
&= \left|\|\mathbf{x}\|^2 + \langle\mathbf{x},\mathbf{y}\rangle + \langle\mathbf{y},\mathbf{x}\rangle + \|\mathbf{y}\|^2\right| \\
&\leq \left|\|x\|^2\right| + |\langle\mathbf{x},\mathbf{y}\rangle| + |\langle\mathbf{y},\mathbf{x}\rangle| + \left|\|\mathbf{y}\|^2\right| \\
&= \|\mathbf{x}\|^2 + |\langle\mathbf{x},\mathbf{y}\rangle| + |\langle\mathbf{y},\mathbf{x}\rangle| + \|\mathbf{y}\|^2 \\
&= \|\mathbf{x}\|^2 + |\langle\mathbf{x},\mathbf{y}\rangle| + \sqrt{\overline{\langle\mathbf{y},\mathbf{x}\rangle}\langle\mathbf{y},\mathbf{x}\rangle} + \|\mathbf{y}\|^2 \\
&= \|\mathbf{x}\|^2 + |\langle\mathbf{x},\mathbf{y}\rangle| + \sqrt{\overline{\overline{\langle\mathbf{x},\mathbf{y}\rangle}\langle\mathbf{x},\mathbf{y}\rangle}} + \|\mathbf{y}\|^2 \\
&= \|\mathbf{x}\|^2 + |\langle\mathbf{x},\mathbf{y}\rangle| + \sqrt{\langle\mathbf{x},\mathbf{y}\rangle\overline{\langle x,y\rangle}} + \|\mathbf{y}\|^2 \\
&= \|\mathbf{x}\|^2 + |\langle\mathbf{x},\mathbf{y}\rangle| + |\langle\mathbf{x},\mathbf{y}\rangle| + \|\mathbf{y}\|^2 \\
&= \|\mathbf{x}\|^2 + 2\,|\langle\mathbf{x},\mathbf{y}\rangle| + \|\mathbf{y}\|^2 \\
&\leq \|\mathbf{x}\|^2 + 2\,\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2 \\
&= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2
\end{aligned}$$

i.e.

$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 &\leq (\|\mathbf{x}\| + \|\mathbf{y}\|)^2 \\
&\Longrightarrow \|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|
\end{aligned}$$

■

The converse is not true in general since an additional structure is needed. A counter example is the space $C[0,\pi/2]$ for a norm

$$\|f\| = \underset{x\in[0,\pi/2]}{Sup}\,|f(x)|$$

However, what we can do is we can make use of the following to "retrieve" an inner product corresponding to a norm on the complex (and hence real) numbers.

Theorem 365 (Polarization identity) *In a Complex Inner Product Space, that is, an inner product space with the underlying field $\mathbb{F} = \mathbb{C}$, we have*

$$\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) + \frac{i}{4}\left(\|x+iy\|^2 - \|x-iy\|^2\right)$$

Proof.

$$\begin{aligned} & \|x+y\|^2 - \|x-y\|^2 \\ = \; & \langle x+y, x+y\rangle - \langle x-y, x-y\rangle \\ = \; & 2\langle x, y\rangle + 2\langle y, x\rangle \end{aligned}$$

From which we have $\|x+iy\|^2 - \|x-iy\|^2 = 2\langle x, iy\rangle + 2\langle iy, x\rangle$

From

$$\begin{aligned} 4\langle x, y\rangle &= 2\langle x, y\rangle + 2\langle x, y\rangle \\ &= 2\langle x, y\rangle + 2\langle x, y\rangle - 2\langle y, x\rangle + 2\langle y, x\rangle \\ &= 2\langle x, y\rangle + 2\langle y, x\rangle + 2\langle x, y\rangle - 2\langle y, x\rangle \\ &= 2\langle x, y\rangle + 2\langle y, x\rangle - i^2 2\langle x, y\rangle + 2i^2\langle y, x\rangle \\ &= 2\langle x, y\rangle + 2\langle y, x\rangle + 2i\langle x, iy\rangle + 2i\langle iy, x\rangle \\ &= \|x+y\|^2 - \|x-y\|^2 + i\left(\|x+iy\|^2 - \|x-iy\|^2\right) \end{aligned}$$

■

If we have a real space, then the imaginary part will equal to zero and $2\langle x, y\rangle + 2\langle y, x\rangle = 4\langle x, y\rangle$ because in a real space, conjugate of a real number is equal to the real number. That disposes of the $\mathbb{C}$ and $\mathbb{R}$ but more generally, we have the following:

Theorem 366 *The norm on an inner product space satisfies the parallelogram equality* $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$

Proof.

$$\begin{aligned} & \|x+y\|^2 + \|x-y\|^2 \\ = \; & \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\ = \; & \langle x, x+y\rangle + \langle y, x+y\rangle + \langle x, x-y\rangle - \langle y, x-y\rangle \\ = \; & \langle x, x\rangle + \overline{\langle x, y\rangle} + \overline{\langle y, x\rangle} + \overline{\langle y, y\rangle} + \overline{\langle x, x\rangle} - \overline{\langle x, y\rangle} - \overline{\langle y, x\rangle} + \overline{\langle y, y\rangle} \\ = \; & \langle x, x\rangle + 0 + 0 + \langle y, y\rangle + \langle x, x\rangle + \langle y, y\rangle \\ = \; & 2\left(\|x\|^2 + \|y\|^2\right) \end{aligned}$$

■

We repeat: not every norm space satisfies the parallelogram equality. We, therefore, have the following exercises, the first part of which has already been done:

Exercise 367 *A normed space X is an inner product space if and only if it satisfies the parallelogram law for every pair of vector.*

Exercise 368 *The polarization identity is valid for any norm space if the norm space satisfies the parallelogram law.*

The basic idea is to get an inner product from the given norm space, provided that $\|x+y\|^2+\|x-y\|^2 = 2\left(\|x\|^2+\|y\|^2\right)$. Let $4\langle x,y\rangle = \|x+y\|^2-\|x-y\|^2$. Since the norm is well-defined function of one variable, we can first view this as a function of x keeping y fixed. This works if the underlying field is that of the real numbers since the norm is a function from a vector space to the set of non-negative reals. Try to accomplish this first so that you might get an idea of where to tweak this definition to work for the complex field. The theorems/exercises that follow might be of help but good luck trying to do this exercise without what follows.

Lemma 369 *Let I be an inner product with a corresponding norm. For $\mathbf{x},\mathbf{y}\in I$, the space satisfies the Cauchy-Schwartz inequality $|\langle \mathbf{x},\mathbf{y}\rangle|^2 \leq \langle \mathbf{x},\mathbf{x}\rangle\langle \mathbf{y},\mathbf{y}\rangle$*

Proof. For $\mathbf{y}=\mathbf{0}, \langle \mathbf{x},\mathbf{0}\rangle = \langle \mathbf{x},\mathbf{x}-\mathbf{x}\rangle = \langle \mathbf{x},\mathbf{x}\rangle - \langle \mathbf{x},\mathbf{x}\rangle = 0$

Let $\alpha\in\mathbb{F}$. For $\mathbf{y}\neq 0$,

$$\begin{aligned}
\langle \mathbf{x}-\alpha\mathbf{y},\mathbf{x}-\alpha\mathbf{y}\rangle &\geq 0\\
&\Rightarrow \langle \mathbf{x},\mathbf{x}-\alpha\mathbf{y}\rangle - \alpha\langle \mathbf{y},\mathbf{x}-\alpha\mathbf{y}\rangle \geq 0\\
&\Rightarrow \langle \mathbf{x},\mathbf{x}\rangle - \bar{\alpha}\langle \mathbf{x},\mathbf{y}\rangle - \alpha\left[\langle \mathbf{y},\mathbf{x}\rangle - \bar{\alpha}\langle \mathbf{y},\mathbf{y}\rangle\right] \geq 0
\end{aligned}$$

For $\bar{\alpha} = \frac{\langle \mathbf{y},\mathbf{x}\rangle}{\langle \mathbf{y},\mathbf{y}\rangle}$, we have

$$\begin{aligned}
&\langle \mathbf{x},\mathbf{x}\rangle - \bar{\alpha}\langle \mathbf{x},\mathbf{y}\rangle - \alpha\left[\langle \mathbf{y},\mathbf{x}\rangle - \langle \mathbf{y},\mathbf{x}\rangle\right]\\
= \;& \langle \mathbf{x},\mathbf{x}\rangle - \frac{\langle \mathbf{y},\mathbf{x}\rangle}{\langle \mathbf{y},\mathbf{y}\rangle}\langle \mathbf{x},\mathbf{y}\rangle \geq 0\\
\Rightarrow\; & \langle \mathbf{x},\mathbf{x}\rangle - \frac{\overline{\langle \mathbf{x},\mathbf{y}\rangle}}{\langle \mathbf{y},\mathbf{y}\rangle}\langle \mathbf{x},\mathbf{y}\rangle\\
= \;& \langle \mathbf{x},\mathbf{x}\rangle - \frac{|\langle \mathbf{x},\mathbf{y}\rangle|^2}{\langle \mathbf{y},\mathbf{y}\rangle} \geq 0\\
\Rightarrow\; & \langle \mathbf{x},\mathbf{x}\rangle \geq \frac{|\langle \mathbf{x},\mathbf{y}\rangle|^2}{\langle \mathbf{y},\mathbf{y}\rangle}\\
\Rightarrow\; & \langle \mathbf{x},\mathbf{x}\rangle\langle \mathbf{y},\mathbf{y}\rangle \geq |\langle \mathbf{x},\mathbf{y}\rangle|^2\\
\Rightarrow\; & |\langle \mathbf{x},\mathbf{y}\rangle|^2 \leq \langle \mathbf{x},\mathbf{x}\rangle\langle \mathbf{y},\mathbf{y}\rangle
\end{aligned}$$

If $\mathbf{x}=\alpha\mathbf{y}$, then

$$\begin{aligned}
&\langle \mathbf{x}-\alpha\mathbf{y},\mathbf{x}-\alpha\mathbf{y}\rangle\\
= \;& \langle \alpha\mathbf{y}-\alpha\mathbf{y},\alpha\mathbf{y}-\alpha\mathbf{y}\rangle\\
= \;& \langle \mathbf{0},\mathbf{0}\rangle\\
= \;& 0
\end{aligned}$$

or

$$\begin{aligned}\langle \mathbf{x}-\alpha y, \mathbf{x}-\alpha \mathbf{y}\rangle &= 0\\ &\Longrightarrow \quad \langle \mathbf{x}, \mathbf{x}-\alpha \mathbf{y}\rangle - \alpha\langle \mathbf{y}, \mathbf{x}-\alpha \mathbf{y}\rangle = 0\\ &\Longrightarrow \quad \langle \mathbf{x}, \mathbf{x}\rangle - \bar{\alpha}\langle \mathbf{x}, \mathbf{y}\rangle - \alpha\left[\langle \mathbf{y}, \mathbf{x}\rangle - \bar{\alpha}\langle \mathbf{y}, \mathbf{x}\rangle\right] = 0\end{aligned}$$

Again, for $\bar{\alpha} = \frac{\langle \mathbf{y},\mathbf{x}\rangle}{\langle \mathbf{y},\mathbf{y}\rangle}$, we have

$$\begin{aligned}\langle \mathbf{x}, \mathbf{x}\rangle - \frac{\langle \mathbf{y}, \mathbf{x}\rangle}{\langle \mathbf{y}, \mathbf{y}\rangle}\langle \mathbf{x}, \mathbf{y}\rangle &= 0\\ &\Longrightarrow \quad \langle \mathbf{x}, \mathbf{x}\rangle - \frac{\overline{\langle \mathbf{x}, \mathbf{y}\rangle}}{\langle \mathbf{y}, \mathbf{y}\rangle}\langle \mathbf{x}, \mathbf{y}\rangle\\ &= \quad \langle \mathbf{x}, \mathbf{x}\rangle - \frac{|\langle \mathbf{x}, \mathbf{y}\rangle|^2}{\langle \mathbf{y}, \mathbf{y}\rangle} = 0\\ &\Longrightarrow \quad \langle \mathbf{x}, \mathbf{x}\rangle = \frac{|\langle \mathbf{x}, \mathbf{y}\rangle|^2}{\langle \mathbf{y}, \mathbf{y}\rangle}\\ &\Longrightarrow \quad |\langle \mathbf{x}, \mathbf{y}\rangle|^2 = \langle \mathbf{x}, \mathbf{x}\rangle\langle \mathbf{y}, \mathbf{y}\rangle\end{aligned}$$

i.e. equality will hold if the vectors are multiples of each other. ■

Proposition 370 (Appolonius's Identity) $\|z-x\|^2+\|z-y\|^2 = \frac{1}{2}\|x-y\|^2 + 2\left\|z-\frac{1}{2}(x+y)\right\|^2$

Proof.

$$\begin{aligned}&\langle z-x, z-x\rangle + \langle z-y, z-y\rangle\\ = \quad &\frac{1}{2}\langle x-y, x-y\rangle + 2\left\langle z-\frac{1}{2}(x+y), z-\frac{1}{2}(x+y)\right\rangle\end{aligned}$$

$\Longrightarrow$

$$\begin{aligned}&\langle z, z-x\rangle - \langle x, z-x\rangle + \langle z, z-y\rangle - \langle y, z-y\rangle\\ = \quad &\frac{1}{2}\langle x, x-y\rangle - \frac{1}{2}\langle y, x-y\rangle + 2\left\langle z, z-\frac{1}{2}(x+y)\right\rangle - \frac{2}{2}\left\langle x+y, z-\frac{1}{2}(x+y)\right\rangle\end{aligned}$$

$\Longrightarrow$

$$\begin{aligned}&\overline{\langle z, z\rangle} - \overline{\langle z, x\rangle} - \overline{\langle x, z\rangle} + \overline{\langle x, x\rangle} + \overline{\langle z, z\rangle} - \overline{\langle z, y\rangle} - \overline{\langle y, z\rangle} + \overline{\langle y, y\rangle}\\ = \quad &\tfrac{1}{2}\overline{\langle x, x\rangle} - \tfrac{1}{2}\overline{\langle x, y\rangle} - \tfrac{1}{2}\overline{\langle y, x\rangle} + \tfrac{1}{2}\overline{\langle y, y\rangle} + 2\overline{\langle z, z\rangle} -\\ &\tfrac{2}{2}\langle z, x+y\rangle - \left\langle x, z-\tfrac{1}{2}(x+y)\right\rangle - \left\langle y, z-\tfrac{1}{2}(x+y)\right\rangle\end{aligned}$$

$\Longrightarrow$

$$\begin{aligned}&-\langle x, z\rangle - \overline{\langle x, z\rangle} + \frac{1}{2}\overline{\langle x, x\rangle} - \overline{\langle z, y\rangle} - \langle z, y\rangle + \frac{1}{2}\langle y, y\rangle\\ = \quad &-\frac{1}{2}\overline{\langle x, y\rangle} - \frac{1}{2}\langle x, y\rangle - \langle z, (x+y)\rangle - \left\langle x, z-\frac{1}{2}(x+y)\right\rangle - \left\langle y, z-\frac{1}{2}(x+y)\right\rangle\end{aligned}$$

$$\Longrightarrow$$

$$\begin{aligned} & -2\operatorname{Re}\langle x,z\rangle+\frac{1}{2}\langle x,x\rangle-2\operatorname{Re}\langle z,y\rangle+\frac{1}{2}\langle y,y\rangle \\ = \; & -\operatorname{Re}\langle x,y\rangle-\langle z,(x+y)\rangle-\overline{\langle x,z\rangle}+\frac{1}{2}\overline{\langle x,(x+y)\rangle}-\overline{\langle y,z\rangle}+\frac{1}{2}\overline{\langle y,x+y\rangle} \end{aligned}$$

$$\Longrightarrow$$

$$\begin{aligned} & -2\operatorname{Re}\langle x,z\rangle+\frac{1}{2}\langle x,x\rangle-2\operatorname{Re}\langle z,y\rangle+\frac{1}{2}\langle y,y\rangle \\ = \; & -\operatorname{Re}\langle x,y\rangle-\overline{\langle z,x\rangle}-\overline{\langle z,y\rangle}-\overline{\langle x,z\rangle}+\frac{1}{2}\langle x+y,x\rangle-\overline{\langle y,z\rangle}+\frac{1}{2}\langle x+y,y\rangle \end{aligned}$$

$$\Longrightarrow$$

$$\begin{aligned} & -2\operatorname{Re}\langle x,z\rangle+\frac{1}{2}\langle x,x\rangle-2\operatorname{Re}\langle z,y\rangle+\frac{1}{2}\langle y,y\rangle \\ = \; & -\operatorname{Re}\langle x,y\rangle-\overline{\langle z,x\rangle}-\overline{\langle z,y\rangle}-\overline{\langle x,z\rangle}+\frac{1}{2}\langle x,x\rangle+\frac{1}{2}\langle y,x\rangle-\overline{\langle y,z\rangle}+\frac{1}{2}\langle x,y\rangle+\frac{1}{2}\langle y,y\rangle \end{aligned}$$

$$\Longrightarrow$$

$$\begin{aligned} & -2\operatorname{Re}\langle x,z\rangle-2\operatorname{Re}\langle z,y\rangle \\ = \; & -\operatorname{Re}\langle x,y\rangle-2\operatorname{Re}\langle x,z\rangle-2\operatorname{Re}\langle y,z\rangle+\operatorname{Re}\langle x,y\rangle \end{aligned}$$

$$\Longrightarrow$$

$$0=0$$

■

Proposition 371 *In an inner product space, if* $\langle \mathbf{x},\mathbf{u}\rangle=\langle \mathbf{x},\mathbf{v}\rangle$ *for all* $\mathbf{x}$*, then* $\mathbf{u}=\mathbf{v}$

Proof.

$$\begin{aligned} \langle \mathbf{x},\mathbf{u}\rangle-\langle \mathbf{x},\mathbf{v}\rangle &= 0 \\ &\Longrightarrow \langle \mathbf{x},\mathbf{u}-\mathbf{v}\rangle=0 \;\forall \mathbf{x} \end{aligned}$$

if $\mathbf{x}=\mathbf{u}-\mathbf{v}$, then

$$\begin{aligned} \langle \mathbf{u}-\mathbf{v},\mathbf{u}-\mathbf{v}\rangle &= 0 \\ &\Longrightarrow \mathbf{u}-\mathbf{v}=0 \\ &\Longrightarrow \mathbf{u}=\mathbf{v} \end{aligned}$$

■

Proposition 372 *Under the inner product space* $\mathbb{R}^n$, $\|\mathbf{x}\|=\|\mathbf{y}\| \implies \langle \mathbf{x}+\mathbf{y},\mathbf{x}-\mathbf{y}\rangle=0$

Proof.

$$\begin{aligned}\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle &= \langle \mathbf{x}, \mathbf{x}-\mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}-\mathbf{y}\rangle \\ &= \langle \mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle - \langle \mathbf{y}, \mathbf{y}\rangle \\ &= (\langle \mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{y}, \mathbf{y}\rangle) + (-\langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle) \\ &= (\langle \mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{x}, \mathbf{x}\rangle) + (-\langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle) \\ &= 0\end{aligned}$$

■

Definition 373 *Two elements* $\mathbf{x}, \mathbf{y}$ *of an inner product space are* ***orthogonal*** *if* $\langle \mathbf{x}, \mathbf{y}\rangle = 0$.

This is written as $\mathbf{x} \perp \mathbf{y}$. If, furthermore, the norm of the two elements is 1, then the two are said to be orthonormal to each other. Two inner product spaces A and B are orthogonal if $\forall \mathbf{x} \in A$ and $\forall \mathbf{y} \in B$, $\mathbf{x} \perp \mathbf{y}$. This is written as $A \perp B$. Also, we define $M^{\perp} = \{x \in M : x \perp M\}$. Note that M itself does not have to be a vector space.

Exercise 374 *$M^{\perp}$ is a subspace*

If M is a subspace, $M \perp M^{\perp}$, that is, $\langle 0, 0\rangle = 0$ implying that $0 \in M^{\perp}$,M we have $M \cap M^{\perp} = \{0\}$

Theorem 375 *For any subset M of an inner product space I, the set $M^{\perp}$ is a closed subspace of I*

Proof. For any two scalars α,$\beta \in \mathbb{K}$, and two elements x, $y \in M^{\perp}$, we have $\langle \alpha x + \beta y, z\rangle = \alpha \langle x, z\rangle + \beta \langle y, z\rangle = 0$ for every $z \in M$. This means that $\alpha x + \beta y \in M^{\perp}$. We next prove a stronger statement than closedness viz. completeness. Let $\{x_n\}$ be a convergent Cauchy sequence in $M^{\perp}$. If I was not complete, then for any limit point x of $M^{\perp}$ and for any $x_n \longrightarrow x$ and from the continuity of the inner product, we have $\langle x, y\rangle = \left\langle \lim_{n\longrightarrow\infty} x_n, y \right\rangle = \lim_{n\longrightarrow\infty} \langle x_n, y\rangle = 0$ for every $y \in M$. This shows that $x \in M^{\perp}$, and thus $M^{\perp}$ is complete hence closed. ■

Lemma 376 *If X is a closed subspace of E, then $X^{\perp\perp} = X$*
Proof. *Clearly $X \subset X^{\perp\perp}$. If $X \neq X^{\perp\perp}$, then there exist a non-zero vector $u \in X^{\perp\perp}$ such that $u \in X^{\perp}$. It follows that $\langle u, u\rangle = 0$ or $u = 0$ which is a contradiction. Hence $X = X^{\perp\perp}$.* ■

A subspace X of E will, thus, be called **closed** when $X = X^{\perp\perp}$.

Theorem 377 *An orthonormal set is linearly independent*

Proof. Let $\{e_1, e_2, ..., e_n\}$ be orthonormal. Consider $\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n = 0$

Then, for any $j \in I_n$

$$\begin{aligned}\left\langle \sum_{i=1}^{n} \alpha_i e_i, e_j \right\rangle &= 0 \\ &\Longrightarrow \sum_{i=1}^{n} \alpha_i \langle e_i, e_j \rangle = 0 \\ &\Longrightarrow \alpha_j \langle e_j, e_j \rangle = 0 \\ &\Longrightarrow \alpha_j = 0\end{aligned}$$

Since j is arbitrary, therefore $\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n = 0$
$\Longrightarrow \alpha_1 = \alpha_2 = ... = \alpha_n = 0$ ■

Definition 378 *Let I be a Inner Product space and $B \subset I$. B is an **orthonormal system** of I if $\|\mathbf{x}\| = 1$ $\forall \mathbf{x} \in B$ and $\mathbf{x} \perp \mathbf{y}$ $\forall \mathbf{x}, \mathbf{y} \in B$*

A great advantage of orthonormal sequences over arbitrary linearly independent sequence is the following: if we know that a given x can be represented as a linear combination of some elements of an orthonormal sequence, then the orthonormality makes the actual determination of the coefficients very easily. In fact $(e_1, e_2, ...)$ is an orthonormal sequence in an inner product space X and we have $x \in$span$\{e_1$,$...e_n\}$ where n is fixed, then by the definition of the span. Thus,

$$x = \sum_{k=1}^{n} \alpha_k e_k \tag{1.3}$$

and if we take inner product by a fixed e_j, we obtain

$$\langle x, e_j \rangle = \left\langle \sum \alpha_k e_k, e_j \right\rangle = \sum \alpha_k \langle e_k, e_j \rangle = \alpha_j \tag{1.4}$$

With these coefficients, above equation becomes

$$x = \sum_{k=1}^{n} \langle x, e_k \rangle e_k$$

This is **Parvesal's identity**, an important application of which is the Fourier series expansion of specific functions. In fact, $\langle x, e_k \rangle$ are called the Fourier coefficients.

This shows that the determination of the unknown coefficients in 1.3 is simple. Another advantage of orthonormality becomes apparent if in 1.3 and 1.4, we want to add another term $\alpha_{n+1} e_{n+1}$, to take care of

$$\widetilde{x} = x + \alpha_{n+1} e_{n+1} \in \text{span}\{e_1, ..., e_{n+1}\}$$

then we need to calculate only one more coefficient since the other coefficients remain unchanged.

Theorem 379 *Nonzero pairwise orthogonal vectors are linearly independent*

Proof. Suppose that the non-zero vectors $v_1, v_2, ..., v_k$ are all orthogonal to each other. Set up a dependency relation

$0 = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_k e_k.$

$\implies \langle v_j, 0\rangle = \langle v_j, \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_k v_k\rangle$

$\implies 0 = \alpha_1 \langle v_j, v_1\rangle + \alpha_2 \langle v_j, v_2\rangle + ... + \alpha_j \langle v_j, v_j\rangle + ... + \alpha_n \langle v_j, v_k\rangle$

$\implies 0 = 0 + 0 + ... + \alpha_j \langle v_j, v_j\rangle + 0 + ... + 0$

$\implies 0 = \alpha_j \langle v_j, v_j\rangle$

Since no $\langle v_j, v_j\rangle$ can be 0 (no v_j is 0), therefore $\alpha_j = 0$ for each α_j. Since the α_i's were the coefficients of the dependency relation for $v_1, v_2, ..., v_k$, the vectors $v_1, v_2, ..., v_k$ are linearly independent. ■

We state the following as an exercise:

Exercise 380 *Let $\{x_n \mid n = 1, 2, 3, ...\}$ be an orthonormal sequence. The the following statements are equivalent.*

Theorem 381

1. *$\{x_n \mid n = 1, 2, 3, ...\}$ is maximal (that is, it is not a proper subset of any orthonormal set)*
2. *If $\alpha_n = \langle h, x_n\rangle = 0$, for all n then $h = 0$.*
3. *(Fourier expansion) For all $h \in H$ we have $h = \sum_{n=1}^{\infty} \alpha_n x_n$*
4. *(Parseval's relation) For all $h, g \in H$ we have $\langle h, g\rangle = \sum_{n=1}^{\infty} \alpha_n \overline{\beta}_n$*
5. *(Bessel's equality) For all $h \in H$ we have $\|h\|^2 = \sum_{n=1}^{\infty} |\alpha_n|^2$*

Here, $\alpha_n = \langle h, x_n\rangle$ and $\beta_n = \langle g, x_n\rangle$

Thus, orthogonal vectors may be taken as basis. It is important to see that for any orthogonal basis $(x_n)_{n\in\mathbb{N}}$, we can have an orthonormal series: $\left(\frac{x_n}{\|x_n\|}\right)_{n\in\mathbb{N}}$. In fact, a maximal orthonormal sequence is called an orthonormal basis.

Theorem 382 (Pythagorean formula) *If $x_1.....x_n$ are orthonormal vectors in an inner product space, then*

$$\left\|\sum_{k=1}^{n} x_k\right\|^2 = \sum_{k=1}^{n} \|x_k\|^2 \tag{1.5}$$

Proof. If $x_1 \perp x_2$, then

$$\begin{aligned} & \|\mathbf{x}_1 + \mathbf{x}_2\|^2 \\ = & \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_1 + \mathbf{x}_2 \rangle \\ = & \langle \mathbf{x}_1, \mathbf{x}_1 + \mathbf{x}_2 \rangle + \langle \mathbf{x}_2, \mathbf{x}_1 + \mathbf{x}_2 \rangle \\ = & \overline{\langle \mathbf{x}_1, \mathbf{x}_1 \rangle} + \overline{\langle \mathbf{x}_1, \mathbf{x}_2 \rangle} + \overline{\langle \mathbf{x}_2, \mathbf{x}_1 \rangle} + \langle \mathbf{x}_2, \mathbf{x}_2 \rangle \\ = & \|\mathbf{x}_1\|^2 + 0 + 0 + \|\mathbf{x}_2\|^2 \\ = & \|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 \end{aligned}$$

This is the famous Pythagorean Theorem in n-dimensions since the vector $\mathbf{x}$ is left as it is, without resorting to tuples. Also, if any two vectors (be they sequences, functions or ordinary lines) are orthogonal, then this identity will hold. The identity can be extended to m mutually orthogonal vectors:

$\langle \mathbf{x}_i, \mathbf{x}_j \rangle = k_{ij}\delta_{ij}$ for $1 \leq i, j \leq m$

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

is the Kronecker delta "function" and the constant k_{ij} depends on the vectors $\mathbf{x}_i$ and $\mathbf{x}_j$. Now,

$$\begin{aligned} & \|\mathbf{x}_1 + \mathbf{x}_2 + ... + \mathbf{x}_m\|^2 \\ = & \langle \mathbf{x}_1 + \mathbf{x}_2 + ... + \mathbf{x}_m, \mathbf{x}_1 + \mathbf{x}_2 + ... + \mathbf{x}_m \rangle \\ = & \langle \mathbf{x}_1, \mathbf{x}_1 + \mathbf{x}_2 + ... + \mathbf{x}_m \rangle + \langle \mathbf{x}_2, \mathbf{x}_1 + \mathbf{x}_2 + ... + \mathbf{x}_m \rangle + ... + \langle \mathbf{x}_m, \mathbf{x}_1 + \mathbf{x}_2 + ... + \mathbf{x}_m \rangle \\ = & \overline{\langle \mathbf{x}_1, \mathbf{x}_1 \rangle} + \overline{\langle \mathbf{x}_1, \mathbf{x}_2 \rangle} + ... + \overline{\langle \mathbf{x}_1, \mathbf{x}_m \rangle} + \overline{\langle \mathbf{x}_2, \mathbf{x}_1 \rangle} + ... + \overline{\langle \mathbf{x}_2, \mathbf{x}_m \rangle} + ... + \langle \mathbf{x}_m, \mathbf{x}_m \rangle \\ = & \langle \mathbf{x}_1, \mathbf{x}_1 \rangle + \langle \mathbf{x_2}, \mathbf{x_2} \rangle + ... + \langle \mathbf{x}_m, \mathbf{x}_m \rangle \\ = & \|\mathbf{x}_1\|^2 + ... + \|\mathbf{x}_m\|^2 \end{aligned}$$

■

Theorem 383 *The inner product is continuous*

Proof. Let $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$ be two convergent sequences in an inner product space I. Then, $\exists N_1$ such that $\|y_n - y\| < \frac{\epsilon}{2\|x_n\|}$ whenever $n > N_1$ and $\exists N_2$ such that $\|x_n - x\| < \frac{\epsilon}{2\|y\|}$ whenver $n > N_2$. Needless to say, these inequalities are valid $\forall \epsilon > 0$.

In order to establish continuity of the inner product, we need to prove that $\lim_{n\to\infty} \langle x_n, y_n \rangle = \langle x, y \rangle$ just like $\lim_{n\to\infty} f(x_n) = f(x)$. In other words, we need to

prove that the sequence $\langle x_n, y_n \rangle$ converges to $\langle x, y \rangle$.

$$\begin{aligned}
& |\langle x_n, y_n \rangle - \langle x, y \rangle| \\
= \; & |\langle x_n, y_n \rangle - \langle x_n, y \rangle + \langle x_n, y \rangle - \langle x, y \rangle| \\
\leq \; & |\langle x_n, y_n \rangle - \langle x_n, y \rangle| + |\langle x_n, y \rangle - \langle x, y \rangle| \\
= \; & |\langle x_n, y_n - y \rangle| + |\langle x_n - x, y \rangle| \\
\leq \; & \|x_n\| \|y_n - y\| + \|x_n - x\| \|y\| \\
< \; & \|x_n\| \frac{\epsilon}{2\|x_n\|} + \frac{\epsilon}{2\|y\|} \|y\| \\
= \; & \epsilon
\end{aligned}$$

i.e. $|\langle x_n, y_n \rangle - \langle x, y \rangle| < \epsilon$ whenever $n > N$ where $N = \max(N_1, N_2)$ ■

Corollary 384 *If* $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$ *in a inner product space and* $\mathbf{y} \perp \mathbf{x}_n$*, then* $\mathbf{y} \perp \mathbf{x}$

Proof. $\lim_{n\to\infty} \langle \mathbf{x}_n, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$
$\implies 0 = \langle \mathbf{x}, \mathbf{y} \rangle$
$\implies \mathbf{y} \perp \mathbf{x}$ ■

Just like subspaces for vectors, we have subspaces for inner product spaces where the inner product is restricted to the vectors of the subspace.

Theorem 385 *Every subset of a separable inner product space is separable.*

Not every sequence in an inner product space converges. In fact, there is a criterion under which any sequence x_n converges in an inner product space.

Exercise 386 *If* $\lim_{n\to\infty} \|x_n\| = \|x\|$ *and* $\lim_{n\to\infty} \langle x_n, x \rangle = \langle x, x \rangle = \|x\|^2$*, then* $\lim_{n\to\infty} x_n = x$

Hilbert Spaces

Definition 387 *A complete inner product space is known as a* ***Hilbert Space****.*

Needless to say, every Hilbert Space is a Banach space. One only needs to prove that every Cauchy sequence in a Hilbert space converges under the norm $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

Hilbert spaces are named after the German mathematician David Hilbert (January 23, 1862 – February 14, 1943).

Recalling the fact that every metric space can be completed, we can similarly complete an inner product space and make it into a Hilbert space. Similar to isometry, the isomorphism of inner product spaces is defined as follows:

Definition 388 *Two inner product spaces* $(X_1, \langle . \rangle_1)$ *and* $(X_2, \langle . \rangle_2)$ *are* ***isomorphic*** *if there exists a bijective linear operator* $T : X_1 \longrightarrow X_2$ *such that*

$$\langle T(x), T(y) \rangle_2 = \langle x, y \rangle_1$$

Such a mapping is called an ***isomorphism****.*

Being bijective and having a similar dot product guarantees that the Inner Product spaces are the same, except for the labelling of points. Since from inner product we can have norms and from norms we can have metrics, therefore, T is also an isometry of X_1 onto X_2. In particular, for isomorphism T, we have $\langle T(x), T(x) \rangle = \|T(x)\|^2 = \langle x, x \rangle = \|x\|^2$. Hence $\|T(x)\| = \|x\|$

Exercise 389 *Show that every isometry is one-to-one.*

A particular case of the completion of metric spaces, we have the following:

Exercise 390 *For every inner product space* I*, there exists a Hilbert space* H *and an isomorphism* $T : I \longrightarrow A$ *where* $\bar{A} = H$*. The space* H *is unique except for isomorphisms.*

Theorem 391 *Let* H *be a Hilbert space with an orthonormal system*

$$S = \{e_n : n \in \mathbb{N}\}$$

and let $x \in H$*. Then,*

$$\sum_{n=0}^{\infty} |\langle e_n, x \rangle|^2 \leq \|x\|^2$$

Proof. Let $x_k := x - \sum_{n=0}^{k} \langle e_n, x\rangle\, e_n \ \forall k \in \mathbb{N}$, then

$$\begin{aligned}
&\langle x_k, e_n\rangle \\
&= \left\langle x - \sum_{n=0}^{k} \langle e_n, x\rangle\, e_n, e_n \right\rangle \\
&= \langle x, e_n\rangle - \left\langle \sum_{n=0}^{k} \langle e_n, x\rangle\, e_n, e_n \right\rangle \\
&= \langle x, e_n\rangle - \left[\sum_{n=0}^{k} \langle e_n, x\rangle\right] \langle e_n, e_n\rangle \\
&= 0 \ \forall n = 0, ..., k
\end{aligned}$$

Hence $x_k \perp e_n$ and $x_k \perp \sum_{n=0}^{k} \langle e_n, x\rangle\, e_n$

By Pythagoras theorem,

$$\begin{aligned}
\|x\|^2 &= \|x_k\|^2 + \left\|\sum_{n=0}^{k} \langle e_n, x\rangle\, e_n\right\|^2 \\
&= \|x_k\|^2 + \sum_{n=0}^{k} |\langle e_n, x\rangle|^2 \|e_n\|^2 \\
&= \|x_k\|^2 + \sum_{n=0}^{k} |\langle e_n, x\rangle|^2 \\
&\geq \sum_{n=0}^{k} |\langle e_n, x\rangle|^2
\end{aligned}$$

i.e., $\sum_{n=0}^{\infty} |\langle e_n, x\rangle|^2 \leq \|x\|^2$. This is the famous Bessel's inequality. ■

Theorem 392 *Let H be a separable Hilbert space and $S = \{e_n : n \in \mathbb{N}\}$ be an orthonormal system. Then, the following are equivalent*

1. *S is an orthonormal basis*
2. *$\mathbf{x} \perp S$ implies $\mathbf{x} = \mathbf{0}\ \forall \mathbf{x} \in H$*
3. *$\mathbf{x} = \sum_{n=0}^{\infty} \langle e_n, \mathbf{x}\rangle e_n \ \forall \mathbf{x} \in H$*
4. *$\langle \mathbf{x}.\mathbf{y}\rangle = \sum_{n=0}^{\infty} \langle \mathbf{x}, e_n\rangle \langle e_n, \mathbf{y}\rangle \ \forall \mathbf{x}, \mathbf{y} \in H$*

5. $\|\mathbf{x}\|^2 = \sum_{n=0}^{\infty} |\langle e_n, \mathbf{x}\rangle|^2 \ \forall \mathbf{x} \in H$

Proof. "1.$\Longrightarrow$2."

Let $\mathbf{x} \perp S$ for any $\mathbf{x} \in H$. If $\mathbf{x} \neq \mathbf{0}$ or $\frac{\mathbf{x}}{\|\mathbf{x}\|} \neq \mathbf{0}$, then $\langle e_n, \frac{\mathbf{x}}{\|\mathbf{x}\|}\rangle = 0 \Longrightarrow \frac{\mathbf{x}}{\|\mathbf{x}\|} \in S$ i.e. S is not an orthonormal system rather $S \cup \{\frac{\mathbf{x}}{\|\mathbf{x}\|}\}$ is which is a contradiction

"2.$\Longrightarrow$3."

To prove that $\mathbf{x} = \sum_{n=0}^{\infty} \langle \mathbf{x}, e_n\rangle e_n \ \forall \mathbf{x} \in H$ converges, consider

$$\begin{aligned}
& \left\langle e_i, \mathbf{x} - \sum_{n=0}^{\infty} \langle \mathbf{x}, e_n\rangle \, e_n \right\rangle \forall i \\
= \ & \langle e_i, \mathbf{x}\rangle - \left\langle e_i, \sum_{n=0}^{\infty} \langle \mathbf{x}, e_n\rangle e_n, \right\rangle \\
= \ & \langle e_i, \mathbf{x}\rangle - \left[\sum_{n=0}^{\infty} \overline{\langle \mathbf{x}, e_n\rangle}\right] \langle e_i, e_n\rangle \\
= \ & \langle e_i, \mathbf{x}\rangle - \overline{\langle \mathbf{x}, e_i\rangle} \\
= \ & 0
\end{aligned}$$

Since $e_i \neq 0 \ \forall i$, therefore $\mathbf{x} - \sum_{n=0}^{\infty} \langle \mathbf{x}, e_n\rangle \, e_n = 0$

or $\mathbf{x} = \sum_{n=0}^{\infty} \langle \mathbf{x}, e_n\rangle \, e_n$

"3.$\Longrightarrow$4."

$$\begin{aligned}
\langle \mathbf{x}.\mathbf{y}\rangle &= \left\langle \sum_{n=0}^{\infty} \langle \mathbf{x}, e_n\rangle \, e_n, \sum_{n=0}^{\infty} \langle \mathbf{y}, e_n\rangle \, e_n \right\rangle \\
&= \sum_{n=0}^{\infty} \langle \mathbf{x}, e_n\rangle \, \overline{\langle \mathbf{y}, e_n\rangle} \, \langle e_n, e_n\rangle \\
&= \sum_{n=0}^{\infty} \langle \mathbf{x}, e_n\rangle \, \langle e_n, \mathbf{y}\rangle
\end{aligned}$$

"4.$\Longrightarrow$5."

Set $\mathbf{x} = \mathbf{y}$ in 4

"5.$\Longrightarrow$1."

Suppose S is not an orthonormal basis. Then, $\exists \ \mathbf{x} \in H$ such that $\|\mathbf{x}\| = 1$ and $S \cup \{\mathbf{x}\}$ is an orthonormal system.

But since $\|\mathbf{x}\|^2 = 1 = \sum_{n=0}^{\infty} |\langle e_n, \mathbf{x}\rangle|^2$ and $\mathbf{x}$ is orthonormal to $e_n \in S$, then $1 = 0$ which is absurd. ■

Just like subspaces for inner product spaces and vector spaces, we can have subspaces for Hilbert spaces. However, a subspace of a Hilbert space need not be complete. All the theorems of completeness and subspaces are applicable here. In particular,

Exercise 393 *Every subspace of a Hilbert space is complete if and only if the subspace is closed.*

Exercise 394 *Every finite dimensional subspace of a Hilbert space is complete.*

Lemma 395 (Minmising Vector) *Let M be a closed subspace in a Hilbert space I. For any point $x \in I \backslash M$ there is unique point $y \in M$ which is nearest point to x. The vector $x - y$ is orthogonal to M.*

Proof. Let d be the greatest lower bound for the distances $d(x, y)$ where $y \in M$. We can find $y_n \in M$ so that $d(x, y_n) < d + \frac{1}{n}$. Consider the parallelogram with vertices y_n, x, y_m, $y_n + y_m - x$. We have

$$2\|x - y_n\|^2 + 2\|x - y_m\|^2 = \|y_n - y_m\|^2 + 4\left\|x - \frac{y_n + y_m}{2}\right\|^2$$

Since the first two lengths are $< d + \frac{1}{n}$ and the last one is $\geq d$, we obtain

$$|y_n - y_m|^2 < 4(d + \frac{1}{n})^2 - 4d^2 = \frac{8d}{n} + \frac{4}{n^2}$$

We see, that $d(y_n, y_m) \longrightarrow 0$ when $n \longrightarrow \infty$. Therefore, $\{y_n\}$ is a Cauchy sequence. But I' is closed, hence complete, and the sequence $\{y_n\}$ has a limit y. For this y we have $d(x, y) = d$. ■

Let now w be any vector from M. We show that $\langle x - y, w\rangle = 0$ Assume the contrary. Multiplying w by the appropriate scalar, we can assume that $\langle x - y, w\rangle$ is real. Consider the function of the real variable t given by $f(t) = d(x, y + tw)^2$. By definition, this function has a minimum at $t = 0$, hence $f'(0) = 0$. On the other hand, we have

$$\begin{aligned} f(t) &= (x - y - tw)^2 = d^2 + 2t(x - y, w) + t^2\|w\|^2 \\ \text{and } f'(0) &= (x - y, w) = 0. \end{aligned}$$

Theorem 396 *Let I be an Inner product space and $M \subset I$ be a closed subspace. Then $I = M \oplus M^{\perp}$.*

Proof. By the above geometric lemma, consider any $x \in I$. If $x \notin M$, let x' be the nearest point to x in M. If $x \in M$, put $x' = x$. In both cases we have $x = x' + x''$ where $x \in I'$, $x'' \in M^{\perp}$. ■

Theorem 397 (Riesz's Theorem) *Every bounded linear functional on a Hilbert space H can be represented in terms of the inner product. That is, $f(x) = \langle x, z\rangle$ where z depends upong f and is uniquely determined by f and has norm $\|z\| = \|f\|$*

Proof. We prove that (a) f has the stated representation (b) z is unique and (c) $\|z\| = \|f\|$

(a) if $f = 0$, then we have the trivial relation for $z = 0$. Let $f \neq 0$. To motivate the idea of the proof, let us ask what properties z must have if the stated representation exists. First, $z \neq 0$ for otherwise $f = 0$. Second, $\langle x, z\rangle = 0$ for all x for which $x \in \ker f$. Hence $z \perp \ker f$. This suggests that we consider $\ker f$ and its orthogonal complement $\ker f^{\perp}$

We have already seen that $\ker f$ is a vector space and is closed. Furthermore, $f \neq 0$ implies $\ker f \neq H$ so that $\ker f^{\perp} \neq \{0\}$ by the projection theorem. Hence $\ker f^{\perp}$ contains $z_0 \neq 0$. We set $v = f(x) z_0 - f(z_0) x$ where $x \in H$ is arbitrary. It is easy to see that $v \in \ker f$. Since $z_0 \perp \ker f$, we have $0 = \langle v, z_0\rangle = \langle f(x) z_0 - f(z_0) x, z_0\rangle = f(x) \langle z_0, z_0\rangle - f(z_0) \langle x, z_0\rangle$. Noting that $\|z_0\|^2 = \langle z_0, z_0\rangle \neq 0$, we can solve for $f(x)$. The result is $f(x) = \frac{f(z_0)}{\langle z_0, z_0\rangle} \langle x, z_0\rangle$. That is, $z = \frac{\overline{f(z_0)}}{\langle z_0, z_0\rangle} z_0$. Since $x \in H$ was arbitrary, we have proved the representation.

(b) Suppose that $f(x) = \langle x, z_1\rangle = \langle x, z_2\rangle$. Then, $\langle x, z_1 - z_2\rangle = 0$ for all x. Hence $z_1 = z_2$.

(c) With $x = z$, we obtain $\|z\|^2 = \langle z, z\rangle = f(z) \leq \|f\| \|z\|$ so that $\|z\| \leq \|f\|$

Conversely, $|f(x)| = |\langle x, z\rangle| \leq \|x\| \|z\|$ by the Cauchy-Schwarz inequality so that $\|f\| = \sup_{\|x\|=1} |\langle x, z\rangle| \leq \|z\|$ ■

1.21 Classification of Hilbert Spaces

Hilbert spaces can be classified, up to isometric isomorphism, according to their dimension. Recall also that the dimension of a Hilbert space and hence a vector space is a well-defined concept, i.e. all orthonormal bases of an Hilbert space share the same cardinality. The classification theorem we describe here states that two Hilbert spaces H_1 and H_2 are isometrically isomorphic if and only if they have the same dimension, i.e. if and only if an orthonormal basis of H_1 has the same cardinality of an orthonormal basis of H_2. This will be achieved by proving that every Hilbert space is isometrically isomorphic to an $l^2(X)$ space, where X has the cardinality of any orthonormal basis of the Hilbert space in consideration.

Theorem 398 *Let H be a separable Hilbert space. If H is infinite dimensional then H is isomorphic to l^2.*

Proof. Let (x_n) be a complete orthonormal sequence in H. If H is infinite dimensional, then (x_n) is an infinite sequence. Let x be an element of H. Define $T(x) = (\alpha_n)$, where $\alpha_n = \langle x, x_n\rangle$, $n = 1, 2, \ldots$. T is one to one mapping from H onto l^2. It is clearly linear. Moreover, for $\alpha_n = \langle x, x_n\rangle$, x,$y \in H$, $n \in \mathbb{N}$, we have

$$\langle T(x), T(y)\rangle = \langle (\alpha_n), (\beta_n)\rangle$$

$$= \sum_{n=1}^{\infty} \alpha_n \overline{\beta_n} = \sum_{n=1}^{\infty} \langle x, x_n \rangle \overline{\langle y, y_n \rangle}$$

$$= \sum_{n=1}^{\infty} \langle x, \langle y, x_n \rangle x_n \rangle = \left\langle x, \sum_{n=1}^{\infty} \langle y, x_n \rangle x_n \right\rangle = \langle x, y \rangle$$

Thus, T is an isomorphism from H onto l^2. ■

Corollary 399 *Suppose H is an Hilbert space and let I be a set that indexes one (and hence, any) orthonormal basis of H. Then, H is isometrically isomorphic to $l^2(I)$.*

Theorem 400 (Classification of Hilbert Spaces) *Two Hilbert spaces H_1 and H_2 are isometrically isomorphic if and only if they have the same dimension.*

Proof. $\Longrightarrow$ If the Hilbert spaces H_1 and H_2 are isometrically isomorphic, with isometric isomorphism U, then if $\{e_i\}_{i \in I}$ is an orthonormal basis for H_1 than $\{Ue_i\}_{i \in I}$ is an orthonormal basis for H_2. Hence, H_1 and H_2 have the same dimension.

$\Longleftarrow$ If the Hilbert spaces H_1 and H_2 have the same dimension, then we can index any orthonormal basis of H_1 and any orthonormal basis of H_2 by the same set I. Using **Theorem 398** we see that H_1 and H_2 are both isometrically isomorphic to $l^2(I)$. Hence H_1 and H_2 are isometrically isomorphic. ■

1.22 Tensor Products of Hilbert Spaces

Just as we can form bigger vector spaces from smaller ones, we can form bigger Hilbert spaces from smaller ones.

Definition 401 *Let H_1 and H_2 be two Hilbert spaces of dimension n and k respectively. Given two vectors $(x_1, x_2, ..., x_n) \in H_1$ and $(y_1, y_2, ..., y_3) \in H_2$. The tensor product $\otimes$ of $\mathbf{x}$ and $\mathbf{y}$, written compactly as $\mathbf{x} \otimes \mathbf{y}$, or even $\mathbf{xy}$ is defined as*

$$\mathbf{x} \otimes \mathbf{y} := (x_1 y_1, x_1 y_2, ..., x_1 y_k, x_2 y_1, x_2 y_2, ..., x_2 y_k, ..., x_n y_1, x_n y_2, ..., x_n y_k)$$

One can even take the tensor product of two spaces altogether to form a "bigger" space by taking the tensor of each element of the former space with each element of the latter space i.e. $H_1 \otimes H_2 = \{x \otimes y, x \in H_1, y \in H_2\}$. This construction accounts for a system of particles..If, however, one wishes to break a Hilbert space into its constituent orthogonal spaces, then one considers the direct sum of two spaces. If finitely many Hilbert spaces $H_1, H_2, ..., H_n$ are given, one can construct their direct sum. Formally, this is done as follows. Let H_1 and H_2 be two Hilbert spaces over a Field $\mathbb{F}$. The Cartesian product $\mathrm{H}_1 \times H_2$ can be given a space structure by using the direct sum $H_1 \oplus H_2$ and then turn this into a Hilbert space by defining the inner product as

$$\langle (x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n) \rangle = \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle + ... + \langle x_n, y_n \rangle$$

and

$$\alpha(\mathbf{x},\mathbf{y}) = (\alpha\mathbf{x},\alpha\mathbf{y})$$

Proposition 402 *Let H_1, H_2, H_3 be Hilbert spaces of dimension n, k and m respectively, over a field $\mathbb{F}$. For $\alpha \in \mathbb{F}$, $\mathbf{y}, \mathbf{y}' \in H_1$, $\mathbf{x}, \mathbf{x}' \in H_2$ and $\mathbf{w} \in H_3$*

- $(\mathbf{x} \otimes \mathbf{y}) \otimes \mathbf{w} = \mathbf{x} \otimes (\mathbf{y} \otimes \mathbf{w})$
- $\alpha(\mathbf{x} \otimes \mathbf{y}) = (\alpha\mathbf{x}) \otimes \mathbf{y} = x \otimes (\alpha\mathbf{y})$
- $(\mathbf{x} + \mathbf{x}') \otimes y = (\mathbf{x} \otimes \mathbf{y}) + (\mathbf{x}' \otimes \mathbf{y})$
- $\mathbf{x} \otimes (\mathbf{y} + \mathbf{y}') = (\mathbf{x} \otimes \mathbf{y}) + (\mathbf{x} \otimes \mathbf{y}')$

Proof. $(\mathbf{x} \otimes \mathbf{y}) \otimes \mathbf{w}$

$= (x_1y_1, x_1y_2, ..., x_1y_k, x_2y_1, x_2y_2, ..., x_2y_k, ..., x_ny_1, x_ny_2, ..., x_ny_k) \otimes (w_1, w_2, ..., w_m)$

$$= \begin{pmatrix} x_1y_1w_1, x_1y_2w_1, ..., x_1y_kw_1, \\ x_2y_1w_1, x_2y_2w_1, ..., x_2y_kw_1, ..., \\ x_ny_1w_1, x_ny_2w_1, ..., x_ny_kw_1, \\ x_1y_1w_2, x_1y_2w_2, ..., x_1y_kw_2, \\ x_2y_1w_2, x_2y_2w_2, ..., \\ x_2y_kw_2, ..., \\ x_ny_1w_2, x_ny_2w_2, ..., \\ x_ny_kw_2, \\ \vdots \\ x_1y_1w_m, x_1y_2w_m, ..., x_1y_kw_m, \\ x_2y_1w_m, x_2y_2w_m, ..., x_2y_kw_m, ..., \\ x_ny_1w_m, x_ny_2w_m, ..., x_ny_kw_m \end{pmatrix}$$

$$= \begin{pmatrix} x_1(y_1w_1), ..., \\ x_1(y_kw_1), x_2(y_1w_1), ..., \\ x_2(y_kw_1), ..., x_n(y_1w_1), ... \\ x_n(y_kw_{1,}), x_1(y_1w_2), ..., \\ x_1(y_kw_2), x_2(y_1w_2), ..., \\ x_2(y_kw_2), ..., x_n(y_1w_2), ..., \\ x_n(y_kw_{2,}) \\ \vdots \\ x_1(y_1w_m), ..., \\ x_1(y_kw_m), x_2(y_1w_m), ..., \\ x_2(y_kw_m), ..., \\ x_n(y_1w_m), ..., \\ x_n(y_kw_m) \end{pmatrix}$$

$= \mathbf{x} \otimes (\mathbf{y} \otimes \mathbf{w})$

Next, $\alpha(\mathbf{x} \otimes \mathbf{y})$

$= \alpha(x_1y_1, x_1y_2, ..., x_1y_k, x_2y_1, x_2y_2, ..., x_2y_k, ..., x_ny_1, x_ny_2, ..., x_ny_k)$

$= (\alpha x_1y_1, \alpha x_1y_2, ..., \alpha x_1y_k, \alpha x_2y_1, \alpha x_2y_2, ..., \alpha x_2y_k, ..., \alpha x_ny_1, \alpha x_ny_2, ..., \alpha x_ny_k)$

Then, $((\alpha x_1)y_1, ..., (\alpha x_1)y_k, (\alpha x_2)y_1, ..., (\alpha x_2)y_k, ..., (\alpha x_{(n)})y_1, ..., (\alpha x_{(n)})y_k) = (\alpha\mathbf{x}) \otimes \mathbf{y}$

and $(x_1(\alpha y_1), ..., x_1(\alpha y_k), x_2(\alpha y_1), ..., x_2(\alpha y_k), ..., x_n(\alpha y_1), ..., x_n(\alpha y_k)) = \mathbf{x} \otimes (\alpha \mathbf{y})$

For the third proposition,

$$\begin{aligned} & (\mathbf{x} + \mathbf{x}') \otimes \mathbf{y} \\ = & \begin{array}{l} ((x_1 + x_1') y_1, (x_1 + x_1') y_2, ..., (x_1 + x_1') y_k, (x_2 + x_2') y_1, ..., \\ (x_2 + x_2') y_k, (x_3 + x_2') y_2, ..., (x_n + x_n') y_1, ..., (x_n + x_n') y_k) \end{array} \\ = & (x_1 y_1 + x_1' y_1, ..., x_1 y_k + x_1' y_k, ..., x_n y_k + x_n' y_k) \\ = & (x_1 y_1, ..., x_1 y_k, ..., x_n y_k) + (x_1' y_1, ..., x_1' y_k, ..., x_n' y_k) \\ = & (\mathbf{x} \otimes \mathbf{y}) + (\mathbf{x}' \otimes \mathbf{y}) \end{aligned}$$

The fourth proposition follows in a similar manner. ■

$H_1 \otimes H_2$ is a Hilbert space. Firstly, the set of the tensor product of elements from two different fields yields a field itself. Secondly, one can construct a vector space from two others since every vector space is free and finally, one can define the inner product $\langle ., . \rangle_{H_1 \otimes H_2} : H_1 \otimes H_2 \longrightarrow \mathbb{F}_1 \otimes \mathbb{F}_2$ by $\langle \mathbf{v}_1 \otimes \mathbf{v}_2, \mathbf{u}_1 \otimes \mathbf{u}_2 \rangle_{H_1 \otimes H_2} = \langle \mathbf{v}_1, \mathbf{u}_1 \rangle \langle v_2, \mathbf{u}_2 \rangle$ for $\mathbf{v}_1, \mathbf{u}_1 \in H_1$ and $\mathbf{v}_2, \mathbf{u}_2 \in H_2$. Completeness can be shown by using the same inner product.

If H_1 and H_2 have orthonormal bases $\{e_n\}$ and $\{e\prime_k\}$, respectively, then $\{e_n \otimes e\prime_k\}$ is an orthonormal basis for $H_1 \otimes H_2$. Furthermore, the dimension of $H_1 \otimes H_2$ is the product (as cardinal numbers) of the Hilbert dimensions i.e. the dimension of such a space is $n \times k$.

Proof. Let $e_1, e_2, ..., e_n$ and $e_1', e_2', ..., e_k'$ be linearly independent basis for H_1 and H_2 respectively. Then, From the basis $(e_1, e_2, ..., e_n) \otimes (e_1', e_2', ..., e_k')$ we can form the sum $c_1 c_1' e_1 e_1' + c_1 c_2' e_1 e + ... + c_1 c_k' e_1 e_k' + c_2 c_1' e_2 e_1' + c_2 c_2' e_2 e_2' + ... + c_2 c_k' e_2' e_k' + ... + c_n c_1' e_n e_1' + c_n c_2' e_n e_2' + ... + c_n c_k' e_n' e_k'$

If $(e_1, e_2, ..., e_n) \otimes (e_1', e_2', ..., e_k') = \mathbf{0}$

then $c_1 e_1 (c_1' e_1' + c_2' e_2' + ... + c_2' e_k') + c_2 e_2 (c_1' e_1' + c_2' e_2' + ... + c_2' e_k') + ... + c_n e_n (c_1' e_1' + c_2' e_2' + ... + c_2' e_k') = \mathbf{0}$

$\Rightarrow (c_1 e_1 + c_2 e_2 + ... + c_n e_n)(c_1' e_1' + c_2' e_2' + ... + c_2' e_k') = \mathbf{0}$

$\Rightarrow (c_1 e_1 + c_2 e_2 + ... + c_n e_n) = 0$ or $(c_1' e_1' + c_2' e_2' + ... + c_2' e_k') = \mathbf{0}$

$\Rightarrow c_1, c_2, ..., c_n = 0$ or $c_1', c_2', ..., c_n' = \mathbf{0}$

$\Rightarrow c_1 c_1', c_1 c_2', ..., c_1 c_k', ..., c_n c_k' = \mathbf{0}$

$\Rightarrow (e_1, e_2, ..., e_n) \otimes (e_1', e_2', ..., e_k')$ is linearly independent

This completes the proof. ■

1.23 Operators on Hilbert Spaces

Theorem 403 *The collection of all operators on a Hilbert space form a space themselvces*

This is known as the **dual space of a Hilbert Space**. Similar to the theorems we've done, we have the following:

Theorem 404 *Every Hilbert Space is reflexive*

Proof. Consider the operator $T : H \longrightarrow H^*$ such that $T(x) = f$ where $f \in H^*$. By Riesz's lemma, every functional can be represented as an inner product. That is, $\exists z$ such that $f(.) = \langle ., z\rangle$ and $\|f\| = \|z\|$. Together, this implies that T is isometric and bijective. ■

Lemma 405 *The hermitian conjugate of any vector in a Hilbert space H is an element of the dual space of H*

A dual space of any Hilbert space H is a collection of all functionals on H.

Proposition 406 *Let $T_1 : H \longrightarrow H$ and $T_2 : H \longrightarrow H$ be operators. Then, $\langle T_1(\mathbf{x}), \mathbf{x}\rangle = \langle T_2(\mathbf{x}), \mathbf{x}\rangle \ \forall \mathbf{x} \implies T_1 = T_2$*

Proof. $\langle T_1(\mathbf{x}), \mathbf{x}\rangle = \langle T_2(\mathbf{x}), \mathbf{x}\rangle \ \forall \mathbf{x}$
$\implies \langle T_1(\mathbf{x}), \mathbf{x}\rangle - \langle T_2(\mathbf{x}), \mathbf{x}\rangle = 0 \ \forall \mathbf{x}$
$\implies \langle T_1(\mathbf{x}) - T_2(\mathbf{x}), \mathbf{x}\rangle = 0 \ \forall \mathbf{x}$
$\implies T_1(\mathbf{x}) - T_2(\mathbf{x}) = 0 \ \forall \mathbf{x}$
$\implies T_1(\mathbf{x}) = T_2(\mathbf{x}) \ \forall \mathbf{x}$
$\implies T_1 = T_2$ ■

The Hermitian of any vector is seen by applying the Hermitian operator on the vector with the property that $\langle T(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, T^*(\mathbf{y})\rangle$

Definition 407 *Let H_1, H_2 be Hilbert spaces and $T : H_1 \longrightarrow H_2$ be a bounded, linear operator. Then, the Hilbert **adjoint** operator T^*: $H_2 \longrightarrow H_1$ of T is such that, $\forall \mathbf{x} \in H_1$ and $\mathbf{y} \in H_2$*

$$\langle T(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, T^*(\mathbf{y})\rangle$$

Theorem 408 *The Hilbert adjoint operator T^* of T is unique*

Proof. Let ${T^*}_1 : H_2 \longrightarrow H_1$ and ${T^*}_2 : H_2 \longrightarrow H_1$ be Hilbert adjoints of $T : H_1 \longrightarrow H_2$

Then, $\forall \mathbf{x} \in H_1$ and $\mathbf{y} \in H_2$
$\langle T(\mathbf{x}), \mathbf{y}\rangle$
$= \langle \mathbf{x}, {T^*}_1(\mathbf{y})\rangle$
$= \langle \mathbf{x}, {T^*}_2(\mathbf{y})\rangle$
$\implies {T^*}_1(\mathbf{y}) = {T^*}_2(\mathbf{y}) \ \forall \mathbf{y} \in H_2$
$\implies {T^*}_1 = {T^*}_2$ ■

Since operators can be viewed as matrices acting on a vector, we can view the adjoint of an operator as a conjugate transpose of a matrix.

Proposition 409 *Let H_1 and H_2 be Hilbert spaces and $T : H_1 \longrightarrow H_2$, $S : H_1 \longrightarrow H_2$ be bounded, linear operators and α be any scalar. Then,*

1. $\langle T^*(\mathbf{y}), \mathbf{x}\rangle = \langle \mathbf{y}, T(\mathbf{x})\rangle$
2. $(S + T)^* = T^* + S^*$
3. $(\alpha T)^* = \overline{\alpha} T^*$

4. $(T^*)^* = T$
5. $T^*T = \hat{0} \Leftrightarrow T = \hat{0}$
6. $(ST)^* = T^*S^*$

Proof. $\langle T^*(\mathbf{y}), \mathbf{x}\rangle = \overline{\langle \mathbf{x}, T^*(\mathbf{y})\rangle} = \overline{\langle T(\mathbf{x}), \mathbf{y}\rangle} = \langle \mathbf{y}, T(\mathbf{x})\rangle$
Next, $\forall \mathbf{x}, \mathbf{y}$, we have

$$\begin{aligned}
\langle (T+S)^*(\mathbf{x}), \mathbf{y}\rangle &= \langle \mathbf{x}, (T+S)(\mathbf{y})\rangle \\
&= \langle \mathbf{x}, T(\mathbf{y}) + S(\mathbf{y})\rangle \\
&= \langle \mathbf{x}, T(\mathbf{y})\rangle + \langle \mathbf{x}, S(\mathbf{y})\rangle \\
&= \langle T^*(\mathbf{x}), \mathbf{y}\rangle + \langle S^*(\mathbf{x}), \mathbf{y}\rangle \\
&= \langle (T^* + S^*)(\mathbf{x}), \mathbf{y}\rangle
\end{aligned}$$

Hence $(T+S)^* = T^* + S^*$
For 3.,

$$\begin{aligned}
\langle (\alpha T)^*(\mathbf{y}), \mathbf{x}\rangle &= \langle \mathbf{y}, (\alpha T)(\mathbf{x})\rangle \\
&= \langle \mathbf{y}, (\alpha T)(\mathbf{x})\rangle \\
&= \langle \mathbf{y}, \alpha T(\mathbf{x})\rangle \\
&= \bar{\alpha} \langle \mathbf{y}, T(\mathbf{x})\rangle \\
&= \bar{\alpha} \langle T^*(\mathbf{y}), \mathbf{x}\rangle \\
&= \langle \bar{\alpha} T^*(\mathbf{y}), \mathbf{x}\rangle
\end{aligned}$$

Apply 1. and then the definition of the adjoint to get 4.
For 5., $\langle T^*(T(\mathbf{y})), \mathbf{x}\rangle = \langle \hat{0}(\mathbf{y}), \mathbf{x}\rangle \Leftrightarrow \langle 0, \mathbf{x}\rangle = 0$
$\Leftrightarrow \langle T^*(T(\mathbf{y})), \mathbf{x}\rangle = \langle T(\mathbf{y}), T(\mathbf{x})\rangle = 0$
Since $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$ are arbitrary, therefore $T(\mathbf{x}) = 0$ $\forall \mathbf{x}$ hence $T = \hat{0}$
Lastly, $\langle (ST)^*(\mathbf{x}), \mathbf{y}\rangle = \langle (\mathbf{x}, S(T(\mathbf{y}))\rangle = \langle (S^*(\mathbf{x}), T(\mathbf{y})\rangle$
$= \langle (T^*(S^*(\mathbf{x})), \mathbf{y}\rangle$ $\forall \mathbf{x}, \mathbf{y}$ hene $(ST)^* = T^*S^*$ ■

Proposition 410 $\hat{0}^* = \hat{0}$

Proof. $\langle \hat{0}(\mathbf{x}), \mathbf{y}\rangle = \langle 0, \mathbf{y}\rangle = 0$ $\forall \mathbf{x}, \mathbf{y}$
Also, $\langle \hat{0}(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, \hat{0}^*(\mathbf{y})\rangle = 0$ $\forall \mathbf{x}, \mathbf{y}$
so that $\hat{0}^*(\mathbf{y}) = 0 = \hat{0}(\mathbf{x})$ $\forall \mathbf{x}, \mathbf{y}$
or $\hat{0}^*(\mathbf{x}) = \hat{0}(\mathbf{x})$ $\forall \mathbf{x}$
or $\hat{0}^* = \hat{0}$ ■

Proposition 411 $\hat{1}^* = \hat{1}$

Proof. $\langle \hat{1}(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle = 0$ $\forall \mathbf{x}, \mathbf{y}$
Also, $\langle \hat{1}(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, \hat{1}^*(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ $\forall \mathbf{x}, \mathbf{y}$
so that $\hat{1}^*(\mathbf{y}) = \mathbf{y}$
which implies $\hat{1}^* = \hat{1}$ ■

Exercise 412 *Show that T^{*-1} exists*

Proof. Since T^* exists for any T and $I = I^*$, we have $I^* = \left(T^{-1}T\right)^* = T^*T^{-1*}$. Thus, such an operator will exist. In fact, $T^*T^{-1*} = I$ so that $T^{*-1}\left(T^*T^{-1*}\right) = T^{*-1}I$ which implies $\left(T^{*-1}T^*\right)T^{-1*} = T^{*-1}$ or $IT^{-1*} = T^{-1*} = T^{*-1}$ ■

Theorem 413 *Let T_n be a bounded linear operator for every $n \in \mathbb{N}$. Show that if $T_n \longrightarrow T$, then $T^*_n \longrightarrow T^*$ where T^* is the adjoint of T*

Proof. $\|T^*_n - T^*\| = \|(T_n - T)^*\| = \|T_n - T\| \longrightarrow 0$ i.e. $\|T^*_n - T^*\| \longrightarrow 0$. Thus, $T^*_n \longrightarrow T^*$ ■

Theorem 414 *Let H_1 and H_2 be two Hilbert spaces and let $T : H_1 \longrightarrow H_2$ be a bounded linear operator. If $M_1 \subset H_1$ and $M_2 \subset H_2$ such that $M_1 \subset T^*(M_2)$. Then, $T(M_1) \subset M_2$*

Proof. Let $T(x) \in T(M_1)$ .Then, $\exists y \in M_2$ such that $T^*(y) = x$ ■

Theorem 415 *Let H_1 and H_2 be two Hilbert spaces and let $T : H_1 \longrightarrow H_2$ be a bounded linear operator. If $M_1 \subset H_1$ and $M_2 \subset H_2$ such that $T(M_1) \subset M_2$, then show that $T^*\left(M_2^{\perp}\right) \subset M_1^{\perp}$*

Proof. For the subspaces M_1 and M_2, we have $T : M_1 \longrightarrow M_2$ and T^*: $M_2 \longrightarrow M_1$ Since these are sets, so we'll employ set-theoretic arguments. Let $T^*(x) \in T^*\left(M_2^{\perp}\right)$ and let $y \in T(M_1) \subset M_2$. Then, $T^*(y) \in T^*(M_2)$ so that $\langle T^*(x), T^*(y)\rangle = 0$
or $\langle TT^*(x), y\rangle = 0$
or $\langle TT^*(x), T(y^*)\rangle = 0$ where $y^* \in M_1$
or $\langle T^*(x), T^*T(y^*)\rangle = 0$
$\implies T^*(x) \in M_1^{\perp}$ ■

Theorem 416 *Let H_1 and H_2 be two Hilbert spaces and let $T : H_1 \longrightarrow H_2$ be a bounded linear operator. If $M_1 \subset H_1$ and $M_2 \subset H_2$ be closed subspaces. Then, $T(M_1) \subset M_2$ if and only if $T^*\left(M_2^{\perp}\right) \subset M_1^{\perp}$*

Proof. The necessary condition holds without invoking the closedness of the subspaces, as was proved in the previous proof. To see the converse, from $T^*\left(M_2^{\perp}\right) \subset M_1^{\perp}$ we have

$$M_1^{\perp\perp} \subset T^*\left(M_2^{\perp}\right)^{\perp}$$

$$\implies M_1 \subset T^*\left(M_2^{\perp}\right)^{\perp}$$

Since T^* is bounded, it is continuous so that $T^*\left(M_2^{\perp}\right)$ will be closed because $M_2^{\perp}$ is closed from the closure of M_2. Hence, $T^*\left(M_2^{\perp}\right)^{\perp} = T^*(M_2)$ so that we have $M_1 \subset T^*(M_2)$

$$\implies T(M_1) \subset M_2$$

■

Definition 417 *A bounded linear operator $T : H \longrightarrow H$ on a Hilbert space H is said to be*

- ***self-adjoint*** *or Hermitian if $T^* = T$*
- ***unitary*** *if T is bijective and $T^* = T^{-1}$*
- ***normal*** *if $TT^* = T^*T$*

Thus, if T is self-adjoint or Hermitian, then $\langle T(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, T^*(\mathbf{y})\rangle = \langle \mathbf{x}, T(\mathbf{y})\rangle$. Hence, for a hermitian operator containing elements of the real field, the corresponding adjoint is equal to its transpose because for any real number, the complex conjugate of that real number is equal to that real number. If T is unitary, then $T^*T = T^{-1}T = I$, implying that T is a unitary matrix. From this, it is easy to see that the columns are orthogonal to rows. Also, $\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \left\langle (\mathbf{x}), T^{-1}T(\mathbf{y})\right\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ so that unitary operators preserve linear operators. This is how they derive their name.

Theorem 418 *For Hermitian operators T, $\langle T(\mathbf{x}), \mathbf{x}\rangle$ is real.*

Proof. Since we have $\langle T(\mathbf{x}), \mathbf{x}\rangle = \langle \mathbf{x}, T(\mathbf{x})\rangle$ from the property of Hermitian operators and $\langle T(\mathbf{x}), \mathbf{x}\rangle = \overline{\langle \mathbf{x}, T(\mathbf{x})\rangle}$ from the property of inner products, we can conclude that $\langle \mathbf{x}, T(\mathbf{x})\rangle = \overline{\langle \mathbf{x}, T(\mathbf{x})\rangle}$ so that $\langle T(\mathbf{x}), \mathbf{x}\rangle$ is real. ■

Theorem 419 *The eigenvalues of every Hermitian operator are real*

Proof. Let H be a Hilbert space over $\mathbb{F}$ and $T(\mathbf{x}) = \alpha\mathbf{x}$ for $\alpha \in \mathbb{F}$ and $\mathbf{x} \in H$
Then,

$$\begin{aligned} \langle T(\mathbf{x}), \mathbf{y}\rangle &= \langle T^*(\mathbf{x}), \mathbf{y}\rangle \\ &\Rightarrow \langle \alpha\mathbf{x}, \mathbf{y}\rangle = \langle \bar{\alpha}\mathbf{x}, \mathbf{y}\rangle \\ &\Rightarrow \alpha \langle \mathbf{x}, \mathbf{y}\rangle = \bar{\alpha} \langle \mathbf{x}, \mathbf{y}\rangle \\ &\Rightarrow \alpha = \bar{\alpha} \text{ if } \langle \mathbf{x}, \mathbf{y}\rangle \neq 0 \end{aligned}$$

■

Theorem 420 *The product of two Hermitian operators is Hermitian if and only if the operators commute.*

Proof. Let H_1 and H_2 be Hilbert spaces and $T : H_1 \longrightarrow H_2, S : H_1 \longrightarrow H_2$ be bounded, linear operators. Then, $(ST)^* = T^*S^* = TS$. Since we have assumed that the product of the operators is Hermitian, we have $ST = (ST)^*$ from which we conclude that $TS = ST$, implying commuativity.

Conversely, if $TS = ST$, then $(TS)^* = S^*T^* = ST = TS$ so that $(TS)^* = TS$, implying that the product is Hermitian. ■

Theorem 421 *For unitary operators U, V, the following holds:*

1. *U is isometric. Converse holds if the isometry is bijective*
2. *$\|U\| = 1$*
3. *$U^{-1} = U^*$ is unitary*
4. *UV is unitary*
5. *U is normal*

Proof. For 1, we have already proved that $\langle U(\mathbf{x}), U(\mathbf{y})\rangle = \left\langle \mathbf{x}, U^{-1}U(\mathbf{y})\right\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ so that the inner product as well as the metric defined from its norm are preserved. Hence the name "unitary". Conversely, from the property of isometries courtesy of the preservation of distance, we have $\langle U(\mathbf{x}), U(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ so that $\langle UU^*(\mathbf{x}), \mathbf{y}\rangle$. This is possible because every operator has to have an adjoint. From this, we have $UU^*(\mathbf{x}) = \mathbf{x}$ for all bijective isometric operators U, implying that $UU^* = \hat{1}$. Since U is a bijective isometry, then we can have U^{-1} and apply it to both sides of $UU^* = \hat{1}$ to get $U^* = U^{-1}$

For 2, $\|U\| = \sup\limits_{\mathbf{x}\neq\mathbf{0}} \frac{\|U(\mathbf{x})\|}{\|\mathbf{x}\|}$

$= \sup \frac{\sqrt{\langle U(\mathbf{x}), U(\mathbf{x})\rangle}}{\sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}}$

$= \sup \frac{\sqrt{\langle UU^*(\mathbf{x}), \mathbf{x}\rangle}}{\sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}}$

$= \sup \frac{\sqrt{\langle UU^{-1}(\mathbf{x}), \mathbf{x}\rangle}}{\sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}}$

$= \sup \frac{\sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}}{\sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}}$

$= \sup 1$

$= 1$

To prove that U^* is unitary, we need to prove that $(U^*)^* = U = (U^*)^{-1}$. $U = U^{**}$ and $U^* = U^{-1}$ imply that $U = (U^*)^{-1}$

Next, $(UV)^* = V^*U^* = V^{-1}U^{-1} = (UV)^{-1}$

Finally, $U^{-1}U = UU^{-1}$ so that $U^*U = UU^*$ ■

Definition 422 *Let H be a Hilbert space and $U \subset H$. An operator $\hat{P} : H \longrightarrow U$ is called the **projection operator** if $\hat{P}^\dagger = \hat{P}$ and $\hat{P}^2 = \hat{P}$.*

The product of two commuting projection operators is also a projection operator.

Proof. Let $\hat{P}_1$ and $\hat{P}_2$ be two projection operators. Then,

$$\begin{aligned}\left(\hat{P}_1\hat{P}_2\right)^\dagger &= \hat{P}_2^\dagger\hat{P}_1^\dagger \\ &= \hat{P}_2\hat{P}_1 \\ &= \hat{P}_1\hat{P}_2\end{aligned}$$

And

$$\begin{aligned}\left(\hat{P}_1\hat{P}_2\right)^2 &= \hat{P}_1\hat{P}_2\hat{P}_1\hat{P}_2\\ &= \hat{P}_1\hat{P}_1\hat{P}_2\hat{P}_2\\ &= \hat{P}_1^2\hat{P}_2^2\\ &= \hat{P}_1\hat{P}_2\end{aligned}$$

justifying the requirments for projectivity of the operator $\hat{P}_1\hat{P}_2$. ■

The sum of two projection operators is not necessarily a projection operator itself. Two projection operators are orthogonal if their product is zero. Thus, $\hat{P}_j\hat{P}_i = \delta_{ij}\hat{P}_i$. The sum of two projection operators is a projection operator if and only if the projection operators are mutually orthogonal

Proof.

$$\left(\hat{P}_1 + \hat{P}_2\right)^\dagger = \hat{P}_1^\dagger + \hat{P}_2^\dagger = \hat{P}_1 + \hat{P}_2$$

and

$$\begin{aligned}&\left(\hat{P}_1 + \hat{P}_2\right)^2\\ =\ &\left(\hat{P}_1 + \hat{P}_2\right)\left(\hat{P}_1 + \hat{P}_2\right)\\ =\ &\hat{P}_1^2 + \hat{P}_1\hat{P}_2 + \hat{P}_2\hat{P}_1 + \hat{P}_2^2\\ =\ &\hat{P}_1 + \hat{0} + \hat{0} + \hat{P}_2\end{aligned}$$

■

1.24 Strong and Weak Convergence

We know that in calculus one defines different types of convergnce. We've seen such types: ordinary convergence, absolute convergence and uniform convergence. We now move on to consider a weaker version of convergence but in order to justify the word "weak", we will call our usual understanding of convergence as strong convergence. More specifically,

Definition 423 *A sequence (x_n) in a normed space X is said to be **strongly convergent** if there is an $x \in X$ such that $\lim_{n\to\infty} \|x_n - x\| = 0$*

Again, this will be shortened to $x_n \longrightarrow x$ or $\lim_{n\to\infty} x_n = x$. x will be called a strong limit.

Weak converge provides a sense in which a sequence is convergent based on some particular support.

Definition 424 *A sequence (x_n) in a normed space X is said to be **strongly convergent** if there is an $x \in X$ such that for every $f \in X'$ $\lim_{n\to\infty} |f(x_n) - f(x)| = 0$.*

This will be written $x_n \xrightarrow{w} x$. In a sense, we are mapping each member of a sequence to a natural or real number, depending on the underlying field. That is, we have a sequence $(a_n) = (f(x_n))$. This allows us to resort to the familiar theorems specific for real and complex numbers.

By Riesz's lemma, every functional can be given a representation as an inner product. Hence this definition implies the definition studied in MTH427

Theorem 425 *Let $x_n \xrightarrow{w} x$. Then,*

1. *The weak limit x of (x_n) is unique*
2. *Every subsequence of (x_n) converges weakly to x*
3. *The sequence $\|x_n\|$ is bounded*

Proof. 1. Suppose $x_n \xrightarrow{w} x$ and $x_n \xrightarrow{w} y$. Then, $f(x_n) \longrightarrow f(x)$ and $f(x_n) \longrightarrow f(y)$. Since $f(x_n)$ is a sequence of real or complex numbers, its limit is unique. That is, $f(x) = f(y)$

$\implies f(x-y) = 0$ for all f

Hence $x = y$

2. This follows from the fact that if a real or complex sequence is convergent, then every subsequence converges to the same limit as the sequence

3. Since $(f(x_n))$ is convergent, it is bounded, say $|f(x_n)| \leq c_f$ for all n, where c_f depends on f but not on n. Define $g_{x_n}(f) = f(x_n)$. Then, $g_{x_n}(f)$ is bounded for every $f \in X'$. Since X' is complete regardless of the completion of X, we can apply the uniform boundedness theorem to X'' and get $\|g_{x_n}\|$ bounded. By another corollary, $\|x_n\| = \|g_{x_n}\|$ ■

Finite dimensional spaces make life easier; here's another reason why:

Theorem 426 *In a finite dimensional space, strong convergence and weak convergence are equivalent*

Proof. First we show that strong convergence implies weak convergence with the same limit. If $x_n \longrightarrow x$. Then, for any $f \in X'$ $|f(x_n) - f(x)| \leq \|f\| \|x_n - x\| \longrightarrow 0$ hence $x_n \xrightarrow{w} x$

Conversely, suppose $x_n \xrightarrow{w} x$ and $\dim X = k$. Then, $x_n = \alpha_1^{(n)} e_1 + ... + \alpha_k^{(n)} e_k$ and $x = \alpha_1 e_1 + ... + \alpha_k e_k$. By assumption, $f(x_n) \longrightarrow f(x)$ for any f. We take in particular $f_1, ..., f_k$ defined by $f_j(e_k) = \delta_{jk}$. Then, $f_j(x_n) = \alpha_j^{(n)}$ and $f_j(x) = \alpha_j$ hence $f_j(x_n) \longrightarrow f_j(x)$. From this, we readily obtain $\|x_n - x\| = \left\| \sum_{j=1}^{k} \left(\alpha_j^{(n)} - \alpha_j \right) e_j \right\| \leq \sum_{j=1}^{k} \left| \alpha_j^{(n)} - \alpha_j \right| \|e_j\| \longrightarrow 0$ hence $x_n \longrightarrow x$ ■

As might have been guessed, there are infinite dimensional spaces where a sequence may converge weakly but not strongly:

Take an orthonormal sequence (e_n) in a Hilbert Space H. Since every $f \in H'$ has a Riesz representation, $f(x) = \langle x, z \rangle$. Hence $f(e_n) = \langle e_n, z \rangle$. From the

Bessel inequality, $\sum_{j=1}^{\infty} |\langle e_n, z\rangle|^2 \leq \|z\|^2$ so that the series on the left converges to zero. That is, $\langle e_n, z\rangle = f(e_n) \longrightarrow 0$. Since f in arbitrary, we see that $e_n \longrightarrow 0$ but that is true since $\|e_n - e_m\|^2 = \langle e_m - e_n, e_m - e_n\rangle = 2$

Exercise 427 *If $x_n \in C[a,b]$ and $x_n \xrightarrow{w} x \in C[a,b]$, show that (x_n) is point-wise convergent on $[a,b]$*

Solution 428 *We have to show that $x_n(t) \longrightarrow x(t)$ for every $t \in [a,b]$. Functionals f_{t_0} of $C[a,b]$ are defined for vectors $x(t) \in C[a,b]$ such that $f_{t_0}(x(t)) = x(t_0)$ for $t_0 \in [a,b]$. Hence, for any sequence of functions (vectors) $x_n(t)$ in $C[a,b]$, $x_n(t) \xrightarrow{w} x(t)$*
$\Longrightarrow f_{t_0}(x_n(t)) \longrightarrow f_{t_0}(x(t))$
$\Longrightarrow x_n(t_0) \longrightarrow x(t_0)$ *for any $t_0 \in C[a,b]$. Hence weak convergence implies point-wise convergence in $C[a,b]$*

Exercise 429 *Let X and Y be normed spaces., $T \in B(X,Y)$ and (x_n) a sequence in X. If $x_n \xrightarrow{w} x_0$, show that $T(x_n) \xrightarrow{w} T(x_0)$*

Solution 430 *Let $x_n \xrightarrow{w} x_0$. Then, $|f(x_n) - f(x)| \longrightarrow 0$. From $\|f\| = \sup_{0\neq x\in X} \frac{|f(x)|}{\|x\|}$, we have $\|x\| = \sup_{0\neq f\in X'} \frac{|f(x)|}{\|f\|}$. Thus, $\|x_n - x_0\| = \sup_{0\neq f\in X'} \frac{|f(x-x_0)|}{\|f\|}$ so that for any $g \in Y'$ and for any $T \in B(X,Y)$, we have $|g(T(x_n)) - g(T(x))| = |g(T(x_n) - T(x))| = |g(T(x_n - x))|$*
$\leq \|g\| \|T(x_n - x)\|$
$\leq \|g\| \|T\| \|x_n - x\|$
$= \|g\| \|T\| \sup_{0\neq f\in X'} \frac{|f(x-x_0)|}{\|f\|} \longrightarrow 0$

Weak convergence covers scalar multiplication and vector addition.

Lemma 431 *If (x_n) and (y_n) are sequences in the same normed space X, show that $x_n \xrightarrow{w} x$ and $y_n \xrightarrow{w} y$ implies $x_n + y_n \xrightarrow{w} x + y$ as well as $\alpha x_n \xrightarrow{w} \alpha x$*

Proof. Let $x_n \xrightarrow{w} x$ and $y_n \xrightarrow{w} y$. Then, for all $\epsilon > 0$, we have N_1 such that $|f(x_n) - f(x)| < \epsilon/2 \ \forall n \geq N_1$ and N_2 such that $|g(y_n) - g(y)| < \epsilon/2 \ \forall n \geq N_2$ for all $g, f \in X'$. Let $N = \max\{N_1, N_2\}$ and choose the particular $f = g$. Then, $|f(x_n + y_n) - f(x - y)|$
$= |f(x_n) - f(x) + f(y_n) - f(y)|$
$= |f(x_n) - f(x) + g(y_n) - g(y)| \leq |f(x_n) - f(x)| + |g(y_n) - g(y)| < \epsilon$ for all $n \geq N$
$\Longrightarrow f(x_n + y_n) \longrightarrow f(x + y)$
$\Longrightarrow x_n + y_n \xrightarrow{w} x + yy - x$
Similarly, we can have $|f(x_n) - f(x)| < \epsilon/|\alpha|$ and $|f(\alpha x_n) - f(\alpha x)|$
$= |\alpha f(x_n) - \alpha f(x)|$
$= |\alpha| |f(x_n) - f(x)| < \epsilon$
$\Longrightarrow \alpha x_n \xrightarrow{w} \alpha x$ ■

Exercise 432 *Show that* $x_n \xrightarrow{w} x_0$ *implies* $\lim_{n\to\infty} \inf \|x_n\| \geq \|x_0\|$

Solution 433 *For any weakly convergent sequence* $x_n \xrightarrow{w} x_0 \neq 0$*, we can choose* n_k *such that the subsequence* $\|x_{n_k}\| \longrightarrow \lim_{n\to\infty} \inf \|x_n\|$*. Note that this does not violate the fact that every subsequence converges weakly to* x_0*. Now, by Hahn-Banach theorem, there exists* $f \in X^{'}$ *such that* $\|f\| = 1$ *and* $f(x_0) = \|x_0\|$*. Then,* $|f(x_{n_k})| \leq \|f\| \|x_{n_k}\| = \|x_{n_k}\|$ *and*

$$\implies \lim_{n_k\to\infty} |f(x_{n_k})| \leq \lim_{n_k\to\infty} \|x_{n_k}\|$$

$$\implies \left| \lim_{n_k\to\infty} f(x_{n_k}) \right| \leq \lim_{n\to\infty} \inf \|x_n\|$$

$$\implies \left| f\left(\lim_{n_k\to\infty} x_{n_k} \right) \right| \leq \lim_{n\to\infty} \inf \|x_n\|$$

$\implies |f(x_0)| \leq \lim_{n\to\infty} \inf \|x_n\|$ *since every subsequence converges weakly to the same limit*

$$\implies \|x_0\| \leq \lim_{n\to\infty} \inf \|x_n\|$$

Exercise 434 ***If*** $x_n \xrightarrow{w} x_0$ *in a normed space* X*, show that* $x_0 \in \bar{Y}$ *where* $Y = span(x_n)$

Solution 435 *Assume that* $x_0 \notin \overline{Y} \implies x_0 \in X - \overline{Y}$*. Then, the conditons satisfy the statement of theorem 4.6-7. Hence there exists* $f \in X^{'}$ *such that* $|f(y)| = 0$ *for all* $y \in \overline{Y}$ *and* $f(x_0) = \delta = \inf_{y\in\overline{Y}} \|y - x_0\|$

Since $Y = span(x_n)$*, then* $x_n \in Y \implies f(x_n) = 0$ *for all* n*. Hence* $f(x_n) \longrightarrow f(x_0)$ *implies* $f(x_0) = 0 = \inf_{y\in\overline{Y}} \|y - x_0\| \implies x_0 \in \overline{Y}$*. Contradiction.*

Exercise 436 ***If*** (x_n) *is a weakly convergent sequence, show that there is a sequence* (y_m) *of linear combinations of elements of* (x_n) *which converges strongly to* x_0

Solution 437 *From the previous exercise, we have that any element* y_m *of* Y *is a linear combination of* (x_n)*. Since* $x_0 \in \overline{Y}$*. therefore either* x_0 *is a limit point or it belongs to* Y*. In the first case* $x_n \longrightarrow x_0$ *strongly. If* x_0 *is not a limit point, then it belongs to* Y *and is, therefore, a linear combination of* (x_n)*, in which case for any linear functional,* $f(x_0) = f(\sum \alpha_{n_k} x_{n_k})$ *implying divergence of the sequence* $f(x_n)$*, which is a contradiction.*

Corollary 438 *Any closed subspace* Y *of a normed space* X *contains the limits of all weakly convergent sequences of elements.*

Definition 439 *A* ***weak Cauchy sequence*** *in a real or complex normed space* X *is a sequence* (x_n) *in* X *such that for every* $f \in X'$*, the sequence* $(f(x_n))$ *is Cauchy in* $\mathbb{R}$ *or* $\mathbb{C}$*.*

ote that $\lim_{n\to\infty} f(x_n)$ exists. A weak Cauchy sequence is bounded

Proof. Let x_n be a weak Cauchy sequence. Then, for any given $\epsilon > 0$, we can find $N \in \mathbb{N}$ such that $|f(x_n) - f(x_m)| < \epsilon$ for $n, m \geq N$. Choose $b = \max\{f(x_1), f(x_2), ..., f(x_{N-1}), \epsilon\}$. Then, $|f(x_n)| \leq b$ ■

Furthermore, every non-empty subset containing a weak Cauchy sequence is bounded

Proof. Let A be a set in a normed space X such that every nonempty subset

of A contains a weak Cauchy sequence. Assume that A is not bounded. Then, there exists an unbounded sequence in A such that $\|x_n\| \longrightarrow \infty$. Since every subsequence converges to the same limit, we can find a weak Cauchy subsequence which is unbounded, a contradiction. ■

Definition 440 *A normed space X is said to be **weakly complete** if each weak Cauchy sequence in X converges weakly in X.*

Lemma 441 *If X is reflexive, then X is weakly complete.*

Proof. If a normed space is reflexive, then it is complete. It remains to prove that every complete space is weakly complete. This follows from the fact that strong convergence implies weak convergence. ■

1.25 Measure Theory and Hilbert Spaces

One of the most important examples of Hilbert spaces, from the point of view of both theory and applications, is the space of Lebesgue square integrable functions on $\mathbb{R}^n$. Thus, the Lebesgue integral is essential for understanding some of the most important aspects of Hilbert space theory. You're probably familiar with the Riemann integral (ordinary integration) of real-valued functions. However, there are severe limitations on which class of functions can have a Riemann integral. For instance, the function must be smooth. Technically, a smooth function has to do with the existence of derivatives but we will suffice with an intuitive implication of the word. There are functions that are continuous but not smooth. Consider a plotter of white noise. How are we to determine the area under the graph of such a haphazard function? Here, we glimpse the construction of the Lebesgue integral but not be even scratching close to the subject. Interested readers are referred to "Introduction to Lebesgue Integral" by Dr. Abdul Rahim Khan. We will be more focused on a second limitation of the Riemann integral: the domain. As such, the integral is defined for the real numbers. The Lebesgue integral offers a perfect solution to consider more general spaces than the 1D Euclidean space $\mathbb{R}$.

We start with the very basic

Definition 442 *Let $\mathcal{U}$ be a fixed non-empty universal set. The function $f : \mathcal{U} \longrightarrow \{0, 1\}$ is called a **characteristic function** or **indicator function** of $\mathcal{U}$.*

Given any characteristic function f, we can associate a unique subset A of $\mathcal{U}$, namely $A_f = \{x \in \mathcal{U} : f(x) = 1\}$

Conversely, given any subset A of $\mathcal{U}$, we can associate a unique characterisitic function f on $\mathcal{U}$ namely

$$f_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases}$$

This acts as the Boolean/truth-valued operator "belongs to is true" and "belongs to is false". The alternative name for this function, **membership function**, is therefore, justified.

If you've studied logic, then the following theorem will be very englightening. In particular, the relation b is analogous to saying that $p \Longrightarrow q$ is the same as $\sim p \vee q$ (read "not p or q")

Exercise 443 *Let f and g be characteristic functions on $\mathcal{U}$. Define the binary operation $\rightarrow$ by*

$$f \rightarrow g = \begin{cases} 0 \text{ if } f = 1 \text{ and } g = 0 \\ 1 \text{ otherwise} \end{cases}$$

a) Prove that $f \rightarrow g$ is a characteristic function on $\mathcal{U}$

b) Prove that $f \rightarrow g = \max(Nf, g)$ where $Nf = \begin{cases} 0 \text{ if } f = 1 \\ 1 \text{ otherwise} \end{cases}$

c) If $A = \{x \in \mathcal{U} \mid f(x) = 1\}$ and $B = \{x \in \mathcal{U} \mid g(x) = 1\}$, prove that "$f \rightarrow g = 1$" if and only if $A \subseteq B$

Solution 444 *a) We have*

$$(f \rightarrow g)(x) = \begin{cases} 0 \text{ if } f(x) = 1 \text{ and } g(x) = 0 \\ 1 \text{ otherwise} \end{cases}$$

The domain of $(f \rightarrow g)(x)$ relies on the domain of both f and g, which is $\mathcal{U}$. The range is $\{0, 1\}$

b) This can be done by considering every single case for f and g.

c) ($\Longrightarrow$) Let $f \rightarrow g = 1$. We will have three different cases.

Case I

$f(x) = 1$ and $g(x) = 1$

We can rephase this as "if $f(x) = 1$, then $g(x) = 1$" which gives us $A \subseteq B$

Case II

$f(x) = 0$ and $g(x) = 1$

If $f(x) = 0$, then we have the empty set since $f(x) = 0$ for any $x \in \mathcal{U}$. Since the empty set is trivially the subset of everyset, therefore $A \subseteq B$

Case III

$f(x) = 0$ and $g(x) = 0$.

If $g(x) = 0$, then $f(x) = 0$. That is, if $x \notin B$, then $x \notin A$. Hence, $B^c \subseteq A^c \iff A \subseteq B$

($\Longleftarrow$) If $A \subseteq B$, then for $x \in A$, we have $x \in B$. Hence, $f(x) = 1$ implies $g(x) = 1$. Thus, $(f \rightarrow g)(x) = 1$. Since this is valid for any x, we have $f \rightarrow g = 1$

Recall that $\vee$ means "or", $\wedge$ means "and".

Proposition 445 *Let $Ch(\mathcal{U})$ be the set of characteristic functions on universal set $\mathcal{U}$. Then, $f, g \in Ch(U) \implies f \vee g,\ f \wedge g, Nf \in Ch(U)$*

Proof. Define $(f \vee g)(x) = f(x) \vee g(x)$, $(f \wedge g)(x) = f(x) \wedge g(x)$ and

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

If $A = \{x \mid f(x) = 1\}$ and $B = \{x \mid g(x) = 1\}$, then $f \vee g$, $f \wedge g$, Nf construct the sets $A \cup B$, $A \cap B$ and A^c, respectively so that they do, indeed, form charactersitic functions. ■

Note that $f(x) \vee g(x)$ can be defined in a multitude of ways. For instance, $\max(f(x), g(x))$, $f(x) + g(x) - f(x)g(x)$. Similarly, $(f \wedge g)(x)$ might correspond to $\min(f(x), g(x))$, $f(x)g(x)$. The characteristic function acts as an intermediary between sets and the Boolean Ring. The following theorem will explain this.

Theorem 446 *If A and B are subsets of $\mathcal{U}$, then*

1. $f_{A \cup B} = \max(f_A, f_B)$
2. $f_{A \cap B} = \min(f_A, f_B)$
3. $f_{A^c} = 1 - f_A$

Proof. $f_{A \cup B}(x) = \begin{cases} 1 \text{ if } x \in A \text{ or } B \\ 0 \text{ if } x \notin \text{either } A \text{ or } B \end{cases}$

Consider the following cases:
1. $f_A(x) = 1$ and $f_B(x) = 0$, then, $f_{A \cup B}(x) = 1$
2. $f_A(x) = 0$ and $f_B(x) = 1$, then, $f_{A \cup B}(x) = 1$
3. $f_A(x) = 0$ and $f_B(x) = 0$, then, $f_{A \cup B}(x) = 0$
4. $f_A(x) = 1$ and $f_B(x) = 1$, then, $f_{A \cup B}(x) = 1$

In all such cases, the definition $\max(f_A, f_B)$ coincides with $f_{A \cup B}$
The proof of part 2 is similar
For part three, consider only the two cases for $f_A(x) = 1$ and 0 ■
All of these ideas can be summed into one theorem:

Theorem 447 *The cardinality of $Ch(\mathcal{U})$ is the same as $\mathcal{P}(U)$*

Therefore, one does get a sense in which one can break down a characteristic function by looking at the underlying subset. What happens if we have additional structure on the set? To use the words of Gerald Folland, the difference between the Riemann and Lebesgue approaches is thus: "to compute the Riemann integral of f, one partitions the domain $[a, b]$ into subintervals", while in the Lebesgue integral, "one is in effect partitioning the range of f ". We hope to give a sense of this approach in the following:

Definition 448 *A **step function** f on the real line $\mathbb{R}$ is a finite linear combination of characteristic functions of semiopen intervals $[a_{k-1}, a_k) \subseteq \mathbb{R}$.*

Thus, for every step function f, there are intervals $[a_0, a_1), ..., [a_{n-1}, a_n)$ and numbers $\lambda_1, ..., \lambda_n \in \mathbb{R}$ such that $f = \sum_{i=1}^{n} \lambda_i f_i$ where f_k is the characteristic function of $[a_{k-1}, a_k)$. The interval on which the step function is defined is partitioned into the given semi-open intervals. We will inherently assumed that the intervals are disjoint and that $a_0 < a_1 < ... < a_n$.

Informally speaking, a step function is a piecewise constant function having only finitely many pieces and these pieces are the intervals $[a_{k-1}, a_k)$. The positioning of the pieces is determined by the values of λ_k.

This definition does not make sense if $f = 0$. Hence from hereon, whenever we speak of step functions, we will mean non-zero step functions.

For all $x \in [a_{k-1}, a_k)$, $f(x) = \lambda_k$ since then the characteristic function for this particular interval is 1 and 0 for the others. In particular, $\lambda_k = f(a_{k-1})$.

You've probably studied Heaviside function in MTH343 Partial Differential Equations. Here are some easy exercises:

Exercise 449 *Show that the Heaviside function is a step function.*

Exercise 450 *The modulus/absolute value of a step function is again the same step function.*

Solution 451 *For all $x \in [a_{k-1}, a_k)$, $|f(x)| = |\lambda_k|$. Thus for each k, $|f(x)| = |\lambda_k| f_k(x)$ since $f_k(x) = 1$ for all $x \in [a_{k-1}, a_k)$ by definition of a characteristic function. Therefore, $|f(x)| = \sum_{i=1}^{n} |\lambda_i| f_i(x)$ or $|f| = \sum_{i=1}^{n} |\lambda_i| f_i$*

Exercise 452 *The sum and product of two step functions is again a step function.*

You are also required to prove this so we can move on without this easy topic taxing us.This exercise is most important. For this, note that for any collection of semi-open intervals, the intersection and union of two semi-open intervals is also a semi-open interval.

Exercise 453 *The product of a step function with a number is also a step function.*

In other words, the collection of step functions is closed under scalar multiplication.

Exercise 454 $\min(f, g) = \frac{1}{2}(f + g - |f - g|)$ *and* $\max(f, g) = \frac{1}{2}(f + g + |f - g|)$ *are step functions.*

Since the sum of two step functions is again a step function (exercise), we can have for ourselves a binary operation for addition. The underlying set is the field of real numbers and, therefore, the collection of step functions forms an abelian group under addition. By the last exercise, we can also have a scalar multiplication. Therefore, the collection of step functions forms a vector space.

As you will recall, the Heaviside function can be translated. In general, most of the functions that you're familiar with can be translated to the left or right simply by considering $f(x \pm c)$. For $c > 0$, $x + c$ is a translation of c-units to the left side whereas $x - c$ is a translation to the right. We therefore make no leap when we say that every step function can be translated by a fixed number. Thus, for a step function $f(x) = \sum_{i=1}^{n} \lambda_i f_i(x)$, we can have $f(x \pm c) = \sum_{i=1}^{n} \lambda_i f_i(x \pm c)$. This is still a step function. Let's drop the non-negative restriction on c and have $f(x + c)$. Then, we can define a translation operator τ_c acting on a step function f. Thus, $\tau_c(f(x)) = f(x + c) = \sum_{i=1}^{n} \lambda_i f_i(x + c)$. Effectively, what this is doing in moving each semi-open interval by c units. Thus, $[a_{k-1}, a_k)$ become $[a_{k-1} + c, a_k + c)$. In fact, for each characteristic function $f_i(x)$, this becomes $f_i(x + c)$. That is, $\tau_c(f_i(x))$. Therefore, $\tau_c(f(x)) = \sum_{i=1}^{n} \lambda_i [\tau_c(f_i(x))]$. The translation operator is a linear function!

Let's move to what's our main purpose: integration. Now, each step function can be viewed as a collection of straight-lines, as mentioned. It makes sense, therefore, to call the integral of a such a function as the sum of the rectangles formed this way. We're not considering a continuous function so this integral is not the Riemann integral. It is just the area under the graph of a step function, which we're taking the liberty to call integral to derive an analogy.

Each rectangle has a base width of $a_k - a_{k-1}$. Each rectangle has a height of λ_k. Thus, each rectangle has an area of $\lambda_k(a_k - a_{k-1})$. Thus, we define

$$\int f = \sum_{i=1}^{n} \lambda_i (a_i - a_{i-1})$$

Note the absence of dx. In effect, this would've been $\int f = \int_{-\infty}^{\infty} f(x)\, dx$ had we had for ourselves a continuous function.

It's time for some exercises

Exercise 455 *Let f, g be step functions. Then, $\int \alpha f + \beta g = \alpha \int f + \beta \int g$*

Solution 456 *Let $f = \sum_{i=1}^{n} \mu_i f_i$ and $g = \sum_{i=1}^{m} \upsilon_i g_i$. and assume that the semi-open intervals for f are $[a_{k-1}, a_k)$ and those for g are $[b_{k-1}, b_k)$. In this case, we have $a_0 < a_1 < ... < a_n$ and $b_0 < b_1 < ... < b_m$. Since we have finite numbers, we can*

order both together to get the ordering $c_0 < c_1 < ... < c_{\max\{m,n\}}$. *Technically, this corresponds to taking the union of the partitions. With this, we can form a new set of characteristic functions* h_i. *From a geometric argument, we can have* $\int \alpha f + \beta g = \sum_{i=1}^{\max\{n,m\}} \lambda_i (c_i - c_{i-1})$ *but to make this correspond to our characteristic function* h_i, *we need to take a closer look at* $c_0 < c_1 < ... < c_{\max\{m,n\}}$. *If* $a_j < c_i < b_k$, *then* $h_i = f_j + g_k$

Solving the RHS, we have $\alpha \int f + \beta \int g = \sum_{j=1}^{n} \mu_j (a_j - a_{j-1}) + \sum_{k=1}^{m} \upsilon_k (b_k - b_{k-1})$ *but note that we can also have* $\alpha f + \beta g = \sum_{i=1}^{\max\{m,n\}} \alpha\mu_i f_i + \beta\upsilon_i g_i$ *where the appropriate scalars are assigned a value of zero, if needed. It is then clear that* $\sum_{i=1}^{\max\{n,m\}} \lambda_i (c_i - c_{i-1}) = \sum_{j=1}^{n} \mu_j (a_j - a_{j-1}) + \sum_{k=1}^{m} \upsilon_k (b_k - b_{k-1})$ *where* $\lambda_i = \mu_i f_i (x) + \upsilon_i g_i (x)$ *for* $x \in [c_{k-1}, c_k)$

Exercise 457 $\int f = \int \tau_c f \ \forall c$

The answer lies in the fact that the length of $[a_{k-1}, a_k)$ and $[a_{k-1} + c, a_k + c)$ is the same, that is, $a_k - a_{k-1}$

Exercise 458 $\left|\int f\right| \leq \int |f|$

Solution 459 $\left|\int f\right| = \left|\sum_{i=1}^{n} \lambda_i (a_i - a_{i-1})\right| \leq \sum_{i=1}^{n} |\lambda_i (a_i - a_{i-1})| = \sum_{i=1}^{n} |\lambda_i| (a_i - a_{i-1}) = \int |f|$. *The second last equality holds since* $a_i > a_{i-1}$

For any function f, we say that f is positive or $f \geq 0$ if $f(x) \geq 0$ for all x of the domain. Therefore, we say that $f \leq g \iff 0 \leq g - f$. This is when $0 \leq (g - f)(x) \ \forall x \iff 0 \leq g(x) - f(x) \ \forall x$ by the algebra of functions. Intuitively, a function is said to be less than or equal to another function if the graph drawn of one lies on or above the other.

Exercise 460 $f \leq g \implies \int f \leq \int g$

The support of a non-zero step function is always a finite union of the semi-open intervals. Recall from the introductory chapter that $\text{supp} f = \{x : f(x) \neq 0\}$. For which values is f non-zero? The characteristic function for each respective interval is one. So, each characteristic function's interval has to be included, regardless of the value of λ, unless it is zero.

Let $g = \sum_{i=1}^{m} \mu_i g_i$ be a step function defined on the partition $b_0 < b_1 < ... < b_m$. We can say that $\text{supp} g = [a_0, a_1) \cup [a_1, a_2) \cup ... \cup [a_{n-1}, a_n)$ where $n \leq m$

since we can relable each b_i to a corresponding a_j. We can, therefore, consider a function $f = \sum_{i=1}^{n} \mu_i f_i$ defined only on the support of g. This corresponds to taking away the non-positive scalars. If $|f| < M$, that is, if this function is bounded above by a scalar M, then

Lemma 461 $\left|\int f\right| \leq M \sum_{i=1}^{n} (a_i - a_{i-1})$

Proof. We can look at the constant M as a function $M(x) = M$ for all x. Further, we can also represent M as $M = \sum_{i=1}^{n} \lambda_i f_i$. Then, $|f| < M$ implies $\sum_{i=1}^{n} |\mu_i| f_i < \sum_{i=1}^{n} \lambda_i f_i$ so that $|\mu_i| < \lambda_i \leq M/n$ for all i. Then, $\left|\int f\right| = \sum_{i=1}^{n} |\mu_i| (a_i - a_{i-1}) < \sum_{i=1}^{n} \lambda_i (a_i - a_{i-1}) \leq M \sum_{i=1}^{n} (a_i - a_{i-1})$ ■

Lemma 462 *Let* $[a_1, b_1), [a_2, b_2)$,... *be disjoint subintervals of an interval* $[a, b)$ *such that*

$$\bigcup_{n=1}^{\infty} [a_n, b_n) = [a, b)$$

Then, $\sum_{i=1}^{\infty} [a_n, b_n) = b - a$

Proof. Let $S \subset [a, b)$ consists of all points c such that the lemma holds for the interval $[a, c)$ and the sequence of subintervals $[a_n, b_n) \cap [a, c)$. Therefore, if $c \in S$, then $c - a = \sum_{n=1}^{\infty} (b_{c,n} - a_n)$ where $b_{c,n} = \min\{b_n, c\}$ and the summation is over all those n for which $a_n < b_{c,n}$. It suffices to prove that $b \in S$ so that $S = [a, b)$. To this end we first prove that $\sup S \in S$ and then show that $b = \sup S$. Indeed, if $s = \sup S$ and $\{s_n\}$ is a non-decreasing sequence of elements of S convergent to s, then

$$s_n - a = \sum_{n=1}^{\infty} (b_{s_n,m} - a_m) \leq \sum_{n=1}^{\infty} (b_{s,m} - a_m) \leq s - a$$

Since $s_n - a \longrightarrow s - a$, from the above, we have $\sum_{n=1}^{\infty} (b_{s,m} - a_m) = s - a$ and consequently $s \in S$. Next we show that $s = b$. Suppose $s < b$. Then for some $k \in \mathbb{N}$, $s \in [a_k, b_k)$ and thus $b_k \in S$. Since this contradicts the definition of s, we conclude $s = b$. ■

The above lemma may sound like something obvious, something that does not require a proof. There are cases which the property does not hold, specifically in the rationals. This indicates that some special properties of the reals are essential here.

Theorem 463 *Let (f_n) be a non-increasing sequence of non-negative step functions such that $\lim_{n\to\infty} f_n(x) = 0$ for every $x \in \mathbb{R}$. Then,*

$$\lim_{n\to\infty} \int f_n = 0$$

Proof. Since the sequence $\left(\int f_n\right)$ is non increasing and bounded from below, it converges (By monotone convergence theorem). Let $\lim_{n\to\infty} \int f_n = \epsilon > 0$ ■

Appendix

1.26 Matrices

Let $\mathbb{F}$ be a field (this requirement can be considerably weakened). Then, an array of elements of the form

$$\begin{bmatrix} a_{11} & a_{12} & ... & a_{1m} \\ a_{21} & a_{22} & ... & a_{2m} \\ . & & & . \\ & . & & . \\ & & . & . \\ a_{n1} & a_{n2} & ... & a_{nm} \end{bmatrix}$$

with $a_{ij} \in \mathbb{F}$ is called a matrix with n rows and m columns. In a short way, we will write such matrices as (a_{ij}). Thus, $A = (a_{ij})_{m\times n}$. Here, $a_{ij} \in \mathbb{F}$. The size of the matrices is sometimes referred to as the **degree** of the matrix.

The addition of two matrices $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$ is defined as $A+B := (a_{ij} + b_{ij})_{m\times n}$. That is, the corresponding entries are added. Similarly, the corresponding entries are subtracted. The addition of two matrices with different degrees is not defined. Let A,B,C be $m \times n$ matrices. Then, we have

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + O = A$ where O is $m \times n$ zero matrix.
4. $A + B = O$ if and only if $B = -A$

The first property holds if addition in the underlying field is commutative. The second is a natural consequence of the associative law for addition in the underlying field. The third can be obtained by letting $O = (0)_{m\times n}$ and the fourth using $-A = (-a_{ij})_{m\times n}$.

Multiplication in matrices is not as simple: if $R = (r_{ij})_{m\times k} = AB$, then

$$r_{ij} := \sum_{t=1}^{n} a_{it} b_{tj}$$

Note that if A is an $m \times n$ matrix, then B has to be an $n \times k$ matrix. So, for instance, if A is 2×3, B is 3×3, and C is 3×1,then $(AB)C = A(BC)$ is not possible. However, this holds for $n \times n$ matrices.

Proof. Let $A = (a_{ij})_{n\times n}$, $B = (b_{ij})_{n\times n}$, $C = (c_{ij})_{n\times n}$ be matrices over a real numbers. That subscript indicates i-th row, j-th column entry. Consider $(AB)C$.

Let $R = (r_{ij})_{n\times n} = AB$, $S = (s_{ij})_{n\times n} = (AB)C$.

Then $s_{ij} = \sum_{k=1}^{n} r_{ik}c_{kj}$ and $r_{ik} = \sum_{k=1}^{n} a_{il}b_{lk}$ by definition of matrix multiplication. From this, we have

$$s_{ij} = \sum_{k=1}^{n} \left(\sum_{k=1}^{n} a_{il}b_{lk} \right) c_{kj} = \sum_{k=1}^{n}\sum_{k=1}^{n} a_{il}b_{lk}c_{kj}$$

This is possible because a_{il}, b_{lk}, c_{kj} are all elements from a field where the distributive law holds. Now consider $A(BC)$.

Let $R = (r_{ij})_{n\times n} = BC, S = (s_{ij})_{n\times n} = A(BC)$. Then, again by definition of matrix multiplication, $s_{ij} = \sum_{k=1}^{n} a_{il}r_{lj}$ and $r_{ik} = \sum_{k=1}^{n} b_{il}c_{lk}$ by definition of matrix multiplication. Therefore,

$$s_{ij} = \sum_{k=1}^{n} a_{il} \left(\sum_{k=1}^{n} b_{il}c_{lk} \right) = s_{ij} = \sum_{k=1}^{n} a_{il} \sum_{k=1}^{n} b_{il}c_{lk}$$

since, again, each enteries are from a field. That is, $(AB)C = A(BC)$. ■

If α is the element of the underlying field $\mathbb{F}$, then $\alpha A = (\alpha a_{ij})_{m\times n}$, which takes care of scalar multiplication. It immediately follows that if α and β are numbers and A is the matrix, then we have $\alpha(\beta A) = (\alpha\beta)A$, $(\alpha+\beta)A = \alpha A+\beta A$ and $\alpha(A+B) = \alpha A+\alpha B$. If α is a number and A and B are the matrices such that product AB is possible then $\alpha(AB) = (\alpha A)B = A(\alpha B)$. If A is a matrix and O is zero matrix then $AO = O$.

Let A,B and C be three matrices then we have $(A + B)C = AC + BC$ and $A(B + C) = AB + AC$.

Proof. $(a(b + c))_{ij} = \sum_{k=1}^{n} (a_{ik}(b + c)_{kj})$

$= \sum_{k=1}^{n} (a_{ik}(b_{kj} + c_{kj}))$

$= \sum_{k=1}^{n} (a_{ik}b_{kj} + a_{ik}c_{kj})$

$= \sum_{k=1}^{n} (ab)_{kj} + \sum_{k=1}^{n} (ac)_{kj}$

$= (ab)_{ij} + (ac)_{ij}$

Similarly we can prove $(A + B)C = AC + BC$. ■

One can note that the distributive properties of the underlying field have been invoked.

The **transpose** A^T of a matrix $A = (a_{ij})_{m\times n}$ is defined as $A^T = (a_{ji})_{n\times m}$

The **trace** of an $n \times n$ square matrix A is defined to be the sum of the elements on the main diagonal (the diagonal from the upper left to the lower right) of A, i.e.

$$tr(A) = a_{11} + a_{22} + a_{33} + ... + a_{nn} = \sum_{i=1}^{n} a_{ii}$$

The trace is only defined for a square matrix (i.e. $n \times n$).

The **identity matrix** I_n is a special matrix $I_n = (i_{ij})_{n\times n}$ such that $i_{ij} = 1$ for $i = j$ and zero otherwise. This is called so because $AI = IA = A$.

Proof. Let $A = (a_{ij})_{n\times n}$, $I = (i_{ij})_{n\times n}$ and $AI = B = (b_{ij})_{n\times n}$. Then, $AI = (b_{ij})_{n\times n}$ where

$$b_{ik} = \sum_{k=1}^{n} a_{il} i_{lk} = \sum_{k=1}^{n} a_{il}$$

Therefore, $(b_{ij})_{n\times n} = (a_{ij})_{n\times n}$. Similarly for IA. ■

The determinant is a single number specific to a matrix which encodes certain properties of matrix that are useful in systems of linear equations, aiding in the provision of the inverse of a matrix. For a square matrix,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The determinant of this matrix is $ad - bc$. The symbol for determinant is two vertical lines either side. $|A|$ means the determinant of the matrix A. This is also written as $\det(A)$. For a 3×3 matrix

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

the determinant is

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= a(ei - fh) - b(di - fg) + c(dh - eg)$$

The determinant of a matrix of arbitrary size can be defined by the Leibniz formula or the Laplace formula.The Leibniz formula for the determinant of an $n \times n$ matrix A is

$$\det(A) = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{i,\sigma_i}$$

Here the sum is computed over all permutations σ of the set $\{1, 2, ..., n\}$. A permutation is a function that reorders this set of integers. The value in the

ith position after the reordering σ is denoted σ_i. For example, for $n = 3$, the original sequence $1, 2, 3$ might be reordered to $\sigma = (2, 3, 1)$, with $\sigma_1 = 2$, $\sigma_2 = 3$, and $\sigma_3 = 1$. For each permutation σ, $sgn(\sigma)$ denotes the signature of σ, a value that is $+1$ whenever the reordering given by σ can be achieved by successively interchanging two entries an even number of times, and -1 whenever it can be achieved by an odd number of such interchanges.

In any of the $n!$ summands, the term

$$\prod_{i=1}^{n} a_{i,\sigma_i}$$

is notation for the product of the entries at positions (i, σ_i), where i ranges from 1 to n.

$a_{1,\sigma_1} .a_{2,\sigma_2} ...a_{n,\sigma_n}$.

For example, the determinant of a 3×3 matrix A ($n = 3$) is

$$\begin{aligned} & \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{i,\sigma i} \\ = \quad & sgn\,(1,2,3) \prod_{i=1}^{n} a_{i,(1,2,3)} + sgn\,(1,3,2) \prod_{i=1}^{n} a_{i,(1,3,2)} + sgn\,(2,1,3) \prod_{i=1}^{n} a_{i,(2,1,3)} \\ & + sgn\,(2,3,1) \prod_{i=1}^{n} a_{i,(2,3,1)} + sgn\,(3,1,2) \prod_{i=1}^{n} a_{i,(3,1,2)} + sgn\,(3,2,1) \prod_{i=1}^{n} a_{i,(3,2,1)} \end{aligned}$$

$$= \prod_{i=1}^{n} a_{i,(1,2,3)} - \prod_{i=1}^{n} a_{i,(1,3,2)} - \prod_{i=1}^{n} a_{i,(2,1,3)} - \prod_{i=1}^{n} a_{i,(2,3,1)} - \prod_{i=1}^{n} a_{i,(3,1,2)} - \prod_{i=1}^{n} a_{i,(3,2,1)}$$

$$= a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{1,2}a_{2,3}a_{3,1} - a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{2,2}a_{3,1}$$

The following properties of the determinant of a matrix can be proved:

- $\det(A) \in \mathbb{F}$
- $\det(I_n) = 1$ where I_n is the $n \times n$ identity matrix.
- $\det(A^T) = \det(A)$
- $\det(A^{-1}) = \frac{1}{\det(A)} = \det(A)^{-1}$
- For square matrices A and B of equal size, $\det(AB) = \det(A)\det(B)$
- $\det(cA) = c^n \det(A)$ for an $n \times n$ matrix.
- If A is a triangular matrix i.e. $a_{ij} = 0$ whenever $i > j$ or $i < j$ then its determinant is equal to the product of its diagonal entries. That is, $\det(A) = a_{1,1,}a_{1,2}, a_{1,3}, ...a_{n,n} = \prod_{i=1}^{n} a_{i,i}$

- If A is a unitriangular matrix i.e. $a_{ij} = 0$ whenever $i > j$ or $i < j$ and $a_{ij} = 1$ for $i = j$, then its determinant is equal to 1.
- $\det(A^*) = \det(A)^*$

A **cofactor** of an $n \times n$ matrix A is a matrix $\left(\det(A_{ij})_{ji}\right)_{n\times n}$. Here, A_{ij} are smaller matrices formed by deleting the i-th row and the j-th column, called the (i,j)-**minor** of A. The matrix formed by taking the transpose of the cofactor matrix of a given original matrix is called the **adjoint** of matrix A and is often written adj(A).

Example 464 *Consider a matrix* $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 0 & 6 \end{bmatrix}$

first find the cofactors of each element

$$A_{11} = \begin{vmatrix} 4 & 5 \\ 0 & 6 \end{vmatrix} = 24$$
$$A_{12} = -\begin{vmatrix} 0 & 5 \\ 1 & 6 \end{vmatrix} = 5$$
$$A_{13} = \begin{vmatrix} 0 & 4 \\ 1 & 0 \end{vmatrix} = -4$$
$$A_{21} = -\begin{vmatrix} 2 & 3 \\ 0 & 6 \end{vmatrix} = -12$$
$$A_{22} = \begin{vmatrix} 1 & 3 \\ 1 & 6 \end{vmatrix} = 3$$
$$A_{23} = -\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} = 2$$
$$A_{31} = \begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = -2$$
$$A_{32} = -\begin{vmatrix} 1 & 3 \\ 0 & 5 \end{vmatrix} = -5$$
$$A_{33} = \begin{vmatrix} 1 & 2 \\ 0 & 4 \end{vmatrix} = 4$$

so the cofactor matrix of A *is*

$$\begin{bmatrix} 24 & 5 & -4 \\ -12 & 3 & 2 \\ -2 & -5 & 4 \end{bmatrix}$$

Finally the adjoint of A *is the tanspose of the cofactor matrix.*

$$adjA = \begin{bmatrix} 24 & -12 & -2 \\ 5 & 3 & -5 \\ -4 & 2 & 4 \end{bmatrix}$$

The multiplicative inverse of a matrix A *is the matrix is defined as* $A^{-1} :=$

$[\det(A)]^{-1} adjA$. *This holds provided that* $\det(A)^{-1} = \det\left(A^{-1}\right) \neq 0$. *In fact, the converse also holds i.e.* A^{-1} *exists if and only if* $\det(A) \neq 0$.

It can be shown that $AA^{-1} = A^{-1}A = I$

Proof. Let $A = (a_{ij})_{n\times n}$ be a square matrix and $\det(A) \neq 0$. We note that

$$\text{adj}(A) = (A_{ij})^T$$

where A_{ij} is the matrix of cofactors of A and is as

$$A_{ij} = (-1)^{i+j} M_{ij}$$

where M_{ij} is the (i, j) minor of A.i.e

$$\text{adj}(A)_{ij} = A_{ji}$$

An easy calculation shows that

$$A\text{adj}(A) = \det(A)I$$

A is invertible if and only if $\det(A)$ is an invertible element of $\mathbb{F}$, and in that case the equation yeilds

$$\text{adj}(A) = \det(A)A^{-1}$$

$\Longrightarrow$

$$A^{-1} = \frac{\text{adj}(A)}{\det(A)}$$

To show

$$AA^{-1} = A^{-1}A = I$$

we have

$$A\text{adj}(A) = \det(A)I$$

so

$$\begin{aligned} AA^{-1} &= A\frac{\text{adj}(A)}{\det(A)} \\ &= [\det(A)]^{-1}[A\text{adj}(A)] \\ &= [\det(A)]^{-1}[\det(A)I] \\ &= I \end{aligned}$$

similarly

$$A^{-1}A = I$$

i.e. $A^{-1}A = AA^{-1} = I$ ■

The operation of a matrix on a vector can be treated as matrix multiplication if a vector is considered to be a $n \times 1$ matrix.