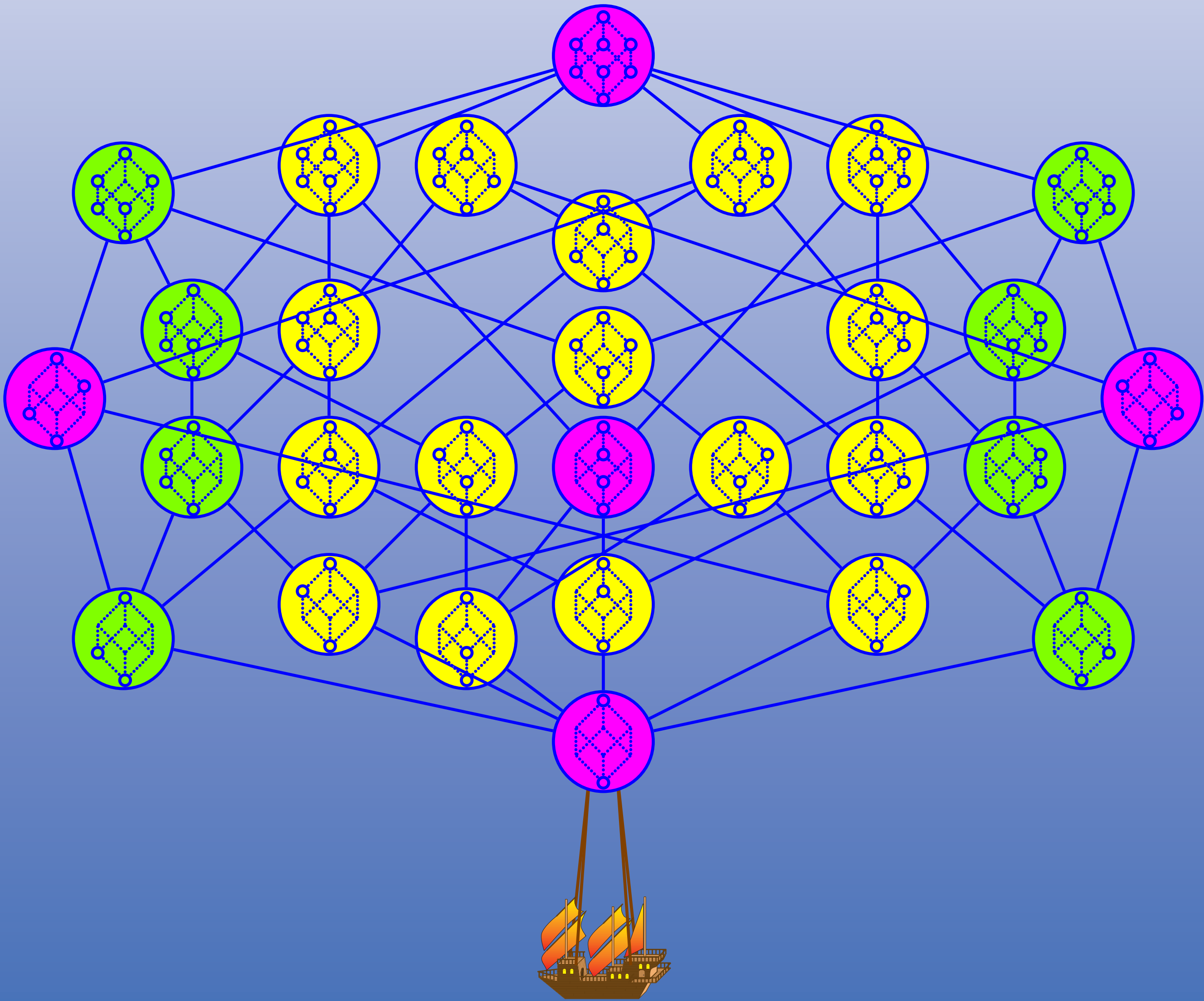
Structure and Analysis of Mathematical Spaces



Daniel J. Greenhoe

Mathematical Structure and Design series

2

title:	*Structure and Analysis of Mathematical Spaces*
document type:	`book`
series:	*Mathematical Structure and Design*
volume:	$\boxed{2}$
author:	Daniel J. Greenhoe
version:	VERSION 0.31X
time stamp:	2019 July 26 (Friday) 07:36pm UTC
copyright:	Copyright © 2019 Daniel J. Greenhoe
license:	Creative Commons license CC BY-NC-ND 4.0
typesetting engine:	XƎLaTeX
document url:	`https://www.researchgate.net/project/Mathematical-Structure-and-Design`

This text was typeset using XƎLaTeX, which is part of the TeXfamily of typesetting engines, which is arguably the greatest development since the Gutenberg Press. Graphics were rendered using the *pstricks* and related packages, and LaTeX graphics support.

The main roman, *italic*, and **bold** font typefaces used are all from the *Heuristica* family of typefaces (based on the *Utopia* typeface, released by *Adobe Systems Incorporated*). The math font is XITS from the XITS font project. The font used in quotation boxes is adapted from *Zapf Chancery Medium Italic*, originally from URW++ Design and Development Incorporated. The font used for the text in the title is Adventor (similar to *Avant-Garde*) from the *TeX-Gyre Project*. The font used for the ISBN in the footer of individual pages is LIQUID CRYSTAL (*Liquid Crystal*) from *FontLab Studio*. The Latin handwriting font is Lavi from the *Free Software Foundation*.

The ship on the cover is loosely based on the *Golden Hind*, a sixteenth century English galleon famous for circumnavigating the globe.[1]

[1] Paine (2000) page 63 ⟨Golden Hind⟩

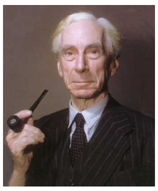
❝Here, on the level sand,
Between the sea and land,
What shall I build or write
Against the fall of night?❞

❝Tell me of runes to grave
That hold the bursting wave,
Or bastions to design
For longer date than mine.❞

Alfred Edward Housman, English poet (1859–1936) [2]

❝The uninitiated imagine that one must await inspiration in order to create. That is a mistake. I am far from saying that there is no such thing as inspiration; quite the opposite. It is found as a driving force in every kind of human activity, and is in no wise peculiar to artists. But that force is brought into action by an effort, and that effort is work. Just as appetite comes by eating so work brings inspiration, if inspiration is not discernible at the beginning.❞

Igor Fyodorovich Stravinsky (1882–1971), Russian-born composer [3]



❝As I think about acts of integrity and grace, I realise that there is nothing in my knowledge to compare with Frege's dedication to truth. His entire life's work was on the verge of completion, much of his work had been ignored to the benefit of men infinitely less capable, his second volume was about to be published, and upon finding that his fundamental assumption was in error, he responded with intellectual pleasure clearly submerging any feelings of personal disappointment. It was almost superhuman and a telling indication of that of which men are capable if their dedication is to creative work and knowledge instead of cruder efforts to dominate and be known.❞

Bertrand Russell (1872–1970), British mathematician, in a 1962 November 23 letter to Dr. van Heijenoort. [4]

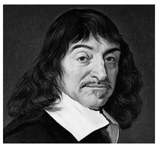
[2] quote: Housman (1936), page 64 ⟨"Smooth Between Sea and Land"⟩, Hardy (1940) ⟨section 7⟩
image: http://en.wikipedia.org/wiki/Image:Housman.jpg

[3] quote: Ewen (1961), page 408, Ewen (1950)
image: http://en.wikipedia.org/wiki/Image:Igor_Stravinsky.jpg

[4] quote: Heijenoort (1967), page 127
image: http://www-history.mcs.st-andrews.ac.uk/PictDisplay/Russell.html

SYMBOLS

❝rugula XVI. Quae vero praesentem mentis attentionem non requirunt, etiamsi ad conclusionem necessaria sint, illa melius est per brevissimas notas designare quam per integras figuras: ita enim memoria non poterit falli, nec tamen interim cogitatio distrahetur ad haec retinenda, dum aliis deducendis incumbit.❞



❝Rule XVI. As for things which do not require the immediate attention of the mind, however necessary they may be for the conclusion, it is better to represent them by very concise symbols rather than by complete figures. It will thus be impossible for our memory to go wrong, and our mind will not be distracted by having to retain these while it is taken up with deducing other matters.❞

René Descartes (1596–1650), French philosopher and mathematician [5]

❝In signs one observes an advantage in discovery which is greatest when they express the exact nature of a thing briefly and, as it were, picture it; then indeed the labor of thought is wonderfully diminished.❞

Gottfried Leibniz (1646–1716), German mathematician, [6]

Symbol list

symbol	description	
numbers:		
$\mathbb{Z}$	integers	$\ldots,-3,-2,-1,0,1,2,3,\ldots$
$\mathbb{W}$	whole numbers	$0,1,2,3,\ldots$
$\mathbb{N}$	natural numbers	$1,2,3,\ldots$
$\mathbb{Z}^{\dashv}$	non-positive integers	$\ldots,-3,-2,-1,0$

...continued on next page...

[5]quote: Descartes (1684a) ⟨rugula XVI⟩, translation: Descartes (1684b) ⟨rule XVI⟩, image: Frans Hals (circa 1650), http://en.wikipedia.org/wiki/Descartes, public domain

[6]quote: Cajori (1993) ⟨paragraph 540⟩, image: http://en.wikipedia.org/wiki/File:Gottfried_Wilhelm_von_Leibniz.jpg, public domain

symbol	description	
$\mathbb{Z}^-$	negative integers	$\ldots, -3, -2, -1$
$\mathbb{Z}_o$	odd integers	$\ldots, -3, -1, 1, 3, \ldots$
$\mathbb{Z}_e$	even integers	$\ldots, -4, -2, 0, 2, 4, \ldots$
$\mathbb{Q}$	rational numbers	$\frac{m}{n}$ with $m \in \mathbb{Z}$ and $n \in \mathbb{Z}\backslash 0$
$\mathbb{R}$	real numbers	completion of $\mathbb{Q}$
$\mathbb{R}^{\vdash}$	non-negative real numbers	$[0, \infty)$
$\mathbb{R}^{\dashv}$	non-positive real numbers	$(-\infty, 0]$
$\mathbb{R}^+$	positive real numbers	$(0, \infty)$
$\mathbb{R}^-$	negative real numbers	$(-\infty, 0)$
$\mathbb{R}^*$	extended real numbers	$\mathbb{R}^* \triangleq \mathbb{R} \cup \{-\infty, \infty\}$
$\mathbb{C}$	complex numbers	
$\mathbb{F}$	arbitrary field	(often either $\mathbb{R}$ or $\mathbb{C}$)
∞	positive infinity	
$-\infty$	negative infinity	
π	pi	$3.14159265\ldots$
relations:		
®	relation	
ⓐ	relational and	
$X \times Y$	Cartesian product of X and Y	
$(\triangle, \nabla)$	ordered pair	
$\lvert z \rvert$	absolute value of a complex number z	
$=$	equality relation	
$\triangleq$	equality by definition	
$\to$	maps to	
$\in$	is an element of	
$\notin$	is not an element of	
$\mathcal{D}(®)$	domain of a relation ®	
$\mathcal{I}(®)$	image of a relation ®	
$\mathcal{R}(®)$	range of a relation ®	
$\mathcal{N}(®)$	null space of a relation ®	
set relations:		
$\subseteq$	subset	
$\subsetneq$	proper subset	
$\supseteq$	super set	
$\supsetneq$	proper superset	
$\nsubseteq$	is not a subset of	
$\not\subset$	is not a proper subset of	
operations on sets:		
$A \cup B$	set union	
$A \cap B$	set intersection	
$A \triangle B$	set symmetric difference	
$A \backslash B$	set difference	
A^c	set complement	
$\lvert \cdot \rvert$	set order	
$\mathbb{1}_A(x)$	set indicator function or characteristic function	
logic:		
1	"true" condition	
0	"false" condition	
$\neg$	logical NOT operation	

...continued on next page...

symbol	description	
$\wedge$	logical AND operation	
$\vee$	logical inclusive OR operation	
$\oplus$	logical exclusive OR operation	
$\implies$	"implies";	"only if"
$\impliedby$	"implied by";	"if"
$\iff$	"if and only if";	"implies and is implied by"
$\forall$	universal quantifier:	"for each"
$\exists$	existential quantifier:	"there exists"
order on sets:		
$\vee$	join or least upper bound	
$\wedge$	meet or greatest lower bound	
$\le$	reflexive ordering relation	"less than or equal to"
$\ge$	reflexive ordering relation	"greater than or equal to"
$<$	irreflexive ordering relation	"less than"
$>$	irreflexive ordering relation	"greater than"
measures on sets:		
$\lvert X \rvert$	order or counting measure of a set X	
distance spaces:		
d	metric or distance function	
linear spaces:		
$\lVert\cdot\rVert$	vector norm	
$\lVert\lvert\cdot\rvert\rVert$	operator norm	
$\langle \triangle \mid \nabla \rangle$	inner-product	
$\operatorname{span}(\boldsymbol{V})$	span of a linear space $\boldsymbol{V}$	
algebras:		
$\Re$	real part of an element in a $*$-algebra	
$\Im$	imaginary part of an element in a $*$-algebra	
set structures:		
$\boldsymbol{T}$	a topology of sets	
$\boldsymbol{R}$	a ring of sets	
$\boldsymbol{A}$	an algebra of sets	
$\varnothing$	empty set	
2^X	power set on a set X	
sets of set structures:		
$\mathcal{T}(X)$	set of topologies on a set X	
$\mathcal{R}(X)$	set of rings of sets on a set X	
$\mathcal{A}(X)$	set of algebras of sets on a set X	
classes of relations/functions/operators:		
2^{XY}	set of *relations* from X to Y	
Y^X	set of *functions* from X to Y	
$\mathcal{S}_{\mathrm{j}}(X,Y)$	set of *surjective* functions from X to Y	
$\mathcal{I}_{\mathrm{j}}(X,Y)$	set of *injective* functions from X to Y	
$\mathcal{B}_{\mathrm{j}}(X,Y)$	set of *bijective* functions from X to Y	
$\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$	set of *bounded* functions/operators from $\boldsymbol{X}$ to $\boldsymbol{Y}$	
$\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	set of *linear bounded* functions/operators from $\boldsymbol{X}$ to $\boldsymbol{Y}$	
$\mathcal{C}(\boldsymbol{X},\boldsymbol{Y})$	set of *continuous* functions/operators from $\boldsymbol{X}$ to $\boldsymbol{Y}$	
specific transforms/operators:		
$\tilde{\mathbf{F}}$	*Fourier Transform* operator	
$\hat{\mathbf{F}}$	*Fourier Series* operator	

...continued on next page...

symbol	description
$\breve{\mathbf{F}}$	*Discrete Time Fourier Series* operator
$\mathbf{Z}$	*Z-Transform* operator
$\tilde{\mathsf{f}}(\omega)$	*Fourier Transform* of a function $\mathsf{f}(x) \in \boldsymbol{L}^2_{\mathbb{R}}$
$\breve{\mathsf{x}}(\omega)$	*Discrete Time Fourier Transform* of a sequence $\left(x_n \in \mathbb{C}\right)_{n\in\mathbb{Z}}$
$\breve{\mathsf{x}}(z)$	*Z-Transform* of a sequence $\left(x_n \in \mathbb{C}\right)_{n\in\mathbb{Z}}$

SYMBOL INDEX

CONTENTS

Part I

Spaces of Analysis

CHAPTER 1

TOPOLOGICAL SPACES

❝*Nevertheless I should not pass over in silence the fact that today the feeling among mathematicians is beginning to spread that the fertility of these abstracting methods is approaching exhaustion. The case is this: that all these nice general concepts do not fall into our laps by themselves. But definite concrete problems were first conquered in their undivided complexity, singlehanded by brute force, so to speak. Only afterwards the axiomaticians came along and stated: Instead of breaking the door with all your might and bruising your hands, you should have constructed such and such a key of skill, and by it you would have been able to open the door quite smoothly. But they can construct the key only because they are able, after the breaking in was successful, to study the lock from within and without. Before you can generalize, formalize, and axiomatize, there must be a mathematical substance.*❞

Hermann Weyl (1885–1955), German mathematician, theoretical physicist, and philosopher [1]

1.1 Set structure

1.1.1 Open sets

Definition 1.1. [2] *Let Γ be a set with an arbitrary (possibly uncountable) number of elements. Let 2^X be the* POWER SET *of a set X.*

[1] quote: Weyl (1935a) ⟨memorial address for Emmy Noether (1882–1935)⟩
Weyl (1935c) ⟨in a book of collected works of Hermann Weyl⟩
Weyl (1935b) pages 140–141 ⟨in a book by Auguste Dick about Emmy Noether⟩
image: http://www.hs.uni-hamburg.de/DE/GNT/hh/biogr/weyl.htm

[2] Munkres (2000) page 76, Riesz (1909), Hausdorff (1914), Tietze (1923), Hausdorff (1937) page 258

DEF

*A family of sets $\mathbf{T} \subseteq 2^X$ is a **topology** on a set X if*

1. $\varnothing \in \mathbf{T}$ ($\varnothing$ is in $\mathbf{T}$) and
2. $X \in \mathbf{T}$ (X is in $\mathbf{T}$) and
3. $U, V \in \mathbf{T} \implies U \cap V \in \mathbf{T}$ (the intersection of a finite number of open sets is open) and
4. $\{U_\gamma | \gamma \in \Gamma\} \subseteq \mathbf{T} \implies \bigcup_{\gamma\in\Gamma} U_\gamma \in \mathbf{T}$ (the union of an arbitrary number of open sets is open).

*A **topological space** is the pair $(X, \mathbf{T})$. An **open set** is any member of $\mathbf{T}$.*
*A **closed set** is any set D such that D^c is open.*
The set of topologies on a set X is denoted $\mathcal{T}(X)$. That is,

$$\mathcal{T}(X) \triangleq \{\mathbf{T} \subseteq 2^X | \mathbf{T} \text{ is a topology}\}$$

Example 1.1. [3] Let $\mathcal{T}(X)$ be the set of topologies on a set X and 2^X the *power set* (Definition A.1 page 265) on X.

EX

$\{\varnothing, X\}$	is a *topology* in	$\mathcal{T}(X)$	(*indiscrete topology* or *trivial topology*)
2^X	is a *topology* in	$\mathcal{T}(X)$	(*discrete topology*)

Example 1.2 (finite complement topology). [4] Let $\mathcal{T}(X)$ be the set of topologies on a set X and 2^X the *power set* (Definition A.1 page 265) on X.

EX

$$\underbrace{\{A \in 2^X | A^c \text{ is finite or } A^c = X\} \in \mathcal{T}(X)}_{\text{is a topology on } X}$$

For examples of topologies on the real line, see the following:
Adams and Franzosa (2008) page 31 ⟨"six topologies on the real line"⟩, Salzmann et al. (2007) pages 64–70 ⟨Weird topologies on the real line⟩, Murdeshwar (1990) page 53 ⟨"often used topologies on the real line"⟩, Joshi (1983) pages 85–91 ⟨§4.2 Examples of Topological Spaces⟩

1.1.2 Order structure of a topology

Theorem 1.1. *Let $\mathbf{L} \triangleq (X, \vee, \wedge; \leq)$ be a* LATTICE.

THM

$$\mathbf{T} \text{ is a TOPOLOGY} \implies (\mathbf{T}, \cup, \cap; \subseteq) \text{ is a DISTRIBUTIVE LATTICE}$$

✎PROOF:

1. By Proposition A.7 (page 280), $(\mathbf{S}, \subseteq)$ is an *ordered set*.
2. By Proposition A.8 (page 281), $\cup$ is *least upper bound* operation on $(\mathbf{S}, \subseteq)$. and $\cap$ is *greatest lower bound* operation on $(\mathbf{S}, \subseteq)$.
3. Therefore, by Definition C.3 (page 305), $(\mathbf{S}, \cup, \cap; \subseteq)$ is a lattice.
4. By Theorem C.3 (page 306), $(\mathbf{S}, \cup, \cap; \subseteq)$ is *idempotent*, *commutative*, *associative*, and *absorptive*.
5. Proof that $(\mathbf{S}, \cup, \cap; \subseteq)$ is *distributive*:

[3] Munkres (2000), page 77, Kubrusly (2011) page 107 ⟨Example 3.J⟩, Steen and Seebach (1978) pages 42–43 ⟨II.4⟩, DiBenedetto (2002) page 18
[4] Munkres (2000), page 77

(a) Proof that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$:

$$
\begin{aligned}
& A \cap (B \cup C) \\
&= \{x \in X | x \in A \wedge x \in (B \cup C)\} && \text{by definition of } \cap \text{ (Definition A.5 page 266)} \\
&= \{x \in X | x \in A \wedge x \in \{x \in X | x \in B \vee x \in C\}\} && \text{by definition of } \cup \text{ (Definition A.5 page 266)} \\
&= \{x \in X | x \in A \wedge (x \in B \vee x \in C)\} \\
&= \{x \in X | (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C)\} \\
&= \{x \in X | x \in A \wedge x \in B\} \cup \{x \in X | x \in A \wedge x \in C\} && \text{by definition of } \cup \text{ (Definition A.5 page 266)} \\
&= (A \cap B) \cup (A \cap C) && \text{by definition of } \cap \text{ (Definition A.5 page 266)}
\end{aligned}
$$

(b) Proof that $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$:
This follows from the fact that $(\boldsymbol{S}, \cup, \cap; \subseteq)$ is a lattice (item (3) page 4), that $\cap$ distributes over $\cup$ (item (5) page 4).

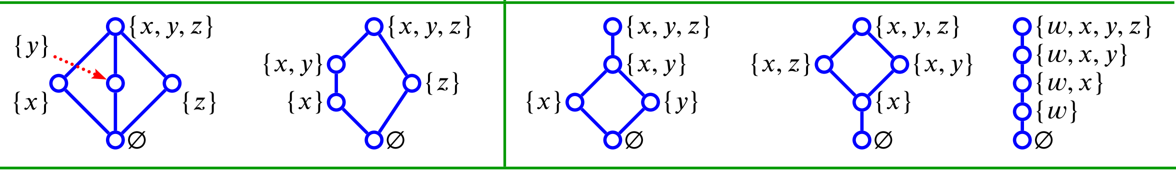
✏

Remark 1.1. Note that in set structures that are *not* closed under the *set union* operation $\cup$ (Definition A.5 page 266), the set union operation $\cup$ is in general *not* equivalent to the *order join* operation $\vee$ with respect to the *set inclusion* relation $\subseteq$ (Definition A.12 page 280). This is illustrated in the next example.

Example 1.3. There are five unlabeled lattices on a five element set (Proposition C.2 page 311). Of these five, three are *distributive*. The following illustrates that the distributive lattices are isomorphic to topologies, while the non-distributive lattices are not.

EX

non-distributive/*not* topologies	*distributive*/*are* topologies



✎PROOF:

1. The first two lattices are non-distributive by *Birkhoff distributivity criterion.*

   (a) This lattice is not a topology because, for example,
   $\{x\} \vee \{y\} = \{x, y, z\} \neq \{x, y\} = \{x\} \cup \{y\}$.
   That is, the set union operation $\cup$ is *not* equivalent to the order join operation $\vee$.

   (b) This lattice is not a topology because, for example,
   $\{x\} \vee \{y\} = \{y\} \neq \{x, y\} = \{x\} \cup \{y\}$

2. The last three lattices are distributive by *Birkhoff distributivity criterion.*

   (a) This lattice is the topology $\boldsymbol{T}_{13}$ of Example 1.6 (page 7). On the set $\{x, y, z\}$, there are a total of three topologies that have this order structure (Example 1.6 page 7):
   $$
   \begin{array}{lcllllll}
   \boldsymbol{T}_{13} &=& \{ & \varnothing, & \{x\}, & \{y\}, & \{x,y\}, & \{x,y,z\} \quad \} \\
   \boldsymbol{T}_{25} &=& \{ & \varnothing, & \{x\}, & \{z\}, & \{x,z\}, & \{x,y,z\} \quad \} \\
   \boldsymbol{T}_{46} &=& \{ & \varnothing, & \{y\}, & \{z\}, & \{y,z\}, & \{x,y,z\} \quad \}
   \end{array}
   $$

   (b) This lattice is the topology $\boldsymbol{T}_{31}$ of Example 1.6 (page 7). On the set $\{x, y, z\}$, there are a total of three topologies that have this order structure (Example 1.6 page 7):
   $$
   \begin{array}{lcllllll}
   \boldsymbol{T}_{31} &=& \{ & \varnothing, & \{x\}, & \{x,y\}, & \{x,z\}, & \{x,y,z\} \quad \} \\
   \boldsymbol{T}_{52} &=& \{ & \varnothing, & \{y\}, & \{x,y\}, & \{y,z\}, & \{x,y,z\} \quad \} \\
   \boldsymbol{T}_{64} &=& \{ & \varnothing, & \{z\}, & \{x,z\}, & \{y,z\}, & \{x,y,z\} \quad \}
   \end{array}
   $$

(c) This lattice is a topology by Definition 1.1 (page 3).

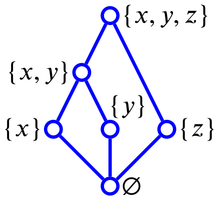
Example 1.4. The set structure

$\boldsymbol{S} \triangleq \{\varnothing, \{x\}, \{y\}, \{z\}, \{x,y\}, \{x,y,z\}\}$

ordered by the set inclusion relation $\subseteq$ is illustrated by the Hasse diagram to the right. Note that

$\{x\} \vee \{z\} = \{x,y,z\} \neq \{x,z\} = \{x\} \cup \{z\}.$

That is, the set union operation $\cup$ is *not* equivalent to the order join operation $\vee$.



1.1.3 Number of topologies

Theorem 1.2. [5]

THM *The **number of topologies** t_n on a finite set X_n with n elements is*

n	0	1	2	3	4	5	6	7	8
t_n	1	1	4	29	355	6942	209,527	9,535,241	642,779,354

n	9	10
t_n	63,260,289,423	8,977,053,873,043

Proposition 1.1. [6] *Let t_n be the number of topologies on a finite set with n elements.*

PRP

$$\lim_{n\to\infty} \frac{t_n}{2^{\frac{n^2}{4}}} = \infty \qquad \text{(lower bound)}$$

$$\lim_{n\to\infty} \frac{t_n}{2^{\left(\frac{1}{2}+\epsilon\right)n^2}} = 0 \quad \forall \epsilon>0 \qquad \text{(upper bound)}$$

$$t_n > n t_{n-1} \qquad \text{(rate of growth)}$$

1.1.4 Closed sets

Theorem 1.3. [7] *Let $(X,\boldsymbol{T})$ be a* TOPOLOGICAL SPACE (Definition 1.1 page 3). *Let $\boldsymbol{T}^*$ be the set of all* CLOSED *sets in $(X,\boldsymbol{T})$. Let Γ be a set with an arbitrary (possibly uncountable) number of elements.*

THM

1. $\varnothing \in \boldsymbol{T}^*$ *($\varnothing$ is* CLOSED *in $(X,\boldsymbol{T})$)*
2. $X \in \boldsymbol{T}^*$ *(X is* CLOSED *in $(X,\boldsymbol{T})$)*
3. $A, B \in \boldsymbol{T}^* \implies A \cup B \in \boldsymbol{T}^*$ *(the union of a finite number of closed sets is* CLOSED*)*
4. $\{D_\gamma | \gamma \in \Gamma\} \subseteq \boldsymbol{T}^* \implies \bigcap_{\gamma\in\Gamma} D_\gamma \in \boldsymbol{T}^*$ *(the intersection of an arbitrary number of closed sets is* CLOSED*)*

[5] Sloane (2014) ⟨http://oeis.org/A000798⟩, Brown and Watson (1996), page 31, Comtet (1974) page 229, Comtet (1966), Chatterji (1967), page 7, Evans et al. (1967), Krishnamurthy (1966), page 157,
[6] Chatterji (1967), pages 6–7, Kleitman and Rothschild (1970)
[7] Aliprantis and Burkinshaw (1998) page 35, Hausdorff (1937) page 258

✎PROOF:

$$
\begin{array}{rcll}
\varnothing \text{ is open} & & & \text{by Definition 1.1 page 3}\\
& \Longrightarrow & \varnothing^{\mathsf{c}} \text{ is closed} & \text{by Definition 1.1 page 3}\\
& \Longrightarrow & X \text{ is closed} & \text{because } \varnothing^{\mathsf{c}} = X\\
X \text{ is open} & & & \text{by Definition 1.1 page 3}\\
& \Longrightarrow & X^{\mathsf{c}} \text{ is closed} & \text{by Definition 1.1 page 3}\\
& \Longleftrightarrow & \varnothing \text{ is closed} & \\
A, B \text{ are closed} & \Longrightarrow & A^{\mathsf{c}}, B^{\mathsf{c}} \text{ are open} & \text{by Definition 1.1 page 3}\\
& \Longrightarrow & A^{\mathsf{c}} \cap B^{\mathsf{c}} \text{ is open} & \text{by Definition 1.1 page 3}\\
& \Longrightarrow & \left(A^{\mathsf{c}} \cap B^{\mathsf{c}}\right)^{\mathsf{c}} \text{ is closed} & \text{by Definition 1.1 page 3}\\
& \Longrightarrow & A \cup B \text{ is closed} & \text{by } \textit{Demorgan's law} \text{ (Theorem A.6 page 278)}\\
\left(\!\left(A_\gamma\right)\!\right)_{\gamma\in\Gamma} \text{ are closed} & \Longrightarrow & \left(\!\left({A^{\mathsf{c}}}_\gamma\right)\!\right)_{\gamma\in\Gamma} \text{ are open} & \text{by Definition 1.1 page 3}\\
& \Longrightarrow & \bigcup_{\gamma\in\Gamma} {A^{\mathsf{c}}}_\gamma \text{ is open} & \text{by Definition 1.1 page 3}\\
& \Longrightarrow & \left(\bigcup_{\gamma\in\Gamma} {A^{\mathsf{c}}}_\gamma\right)^{\mathsf{c}} \text{ is closed} & \text{by Definition 1.1 page 3}\\
& \Longrightarrow & \bigcap_{\gamma\in\Gamma} A_\gamma \text{ is closed} & \text{by } \textit{Demorgan's law} \text{ (Theorem A.6 page 278)}
\end{array}
$$

⇨

Example 1.5. There are four topologies on the set $X \triangleq \{x, y\}$:

topologies on $\{x, y\}$	corresponding closed sets
$T_0 = \{\varnothing, X\}$	$\{\varnothing, X\}$
$T_1 = \{\varnothing, \{x\}, X\}$	$\{\varnothing, \{y\}, X\}$
$T_2 = \{\varnothing, \{y\}, X\}$	$\{\varnothing, \{x\}, X\}$
$T_3 = \{\varnothing, \{x\}, \{y\}, X\}$	$\{\varnothing, \{x\}, \{y\}, X\}$

The topologies (X, T_1) and (X, T_2), as well as their corresponding closed set topological spaces, are all *Serpiński spaces*.

Example 1.6. There are a total of 29 *topologies* (Definition 1.1 page 3) on the set $X \triangleq \{x, y, z\}$:

topologies on $\{x, y, z\}$	corresponding closed sets
$T_{00} = \{\varnothing, X\}$	$\{\varnothing, X\}$
$T_{01} = \{\varnothing, \{x\}, X\}$	$\{\varnothing, \{y, z\}, X\}$
$T_{02} = \{\varnothing, \{y\}, X\}$	$\{\varnothing, \{x, z\}\ X\}$
$T_{04} = \{\varnothing, \{z\}, X\}$	$\{\varnothing, \{x, y\}, X\}$
$T_{10} = \{\varnothing, \{x, y\}, X\}$	$\{\varnothing, \{z\}, X\}$
$T_{20} = \{\varnothing, \{x, z\}, X\}$	$\{\varnothing, \{y\}, X\}$
$T_{40} = \{\varnothing, \{y, z\}, X\}$	$\{\varnothing, \{x\}, X\}$
$T_{11} = \{\varnothing, \{x\}, \{x, y\}, X\}$	$\{\varnothing, \{z\}, \{y, z\}, X\}$
$T_{21} = \{\varnothing, \{x\}, \{x, z\}, X\}$	$\{\varnothing, \{y\}\ \{y, z\}, X\}$
$T_{41} = \{\varnothing, \{x\}, \{y, z\}, X\}$	$\{\varnothing, \{x\}, \{y, z\}, X\}$
$T_{12} = \{\varnothing, \{y\}, \{x, y\}, X\}$	$\{\varnothing, \{z\}, \{x, z\}\ X\}$
$T_{22} = \{\varnothing, \{y\}, \{x, z\}, X\}$	$\{\varnothing, \{y\}, \{x, z\}, X\}$
$T_{42} = \{\varnothing, \{y\}, \{y, z\}, X\}$	$\{\varnothing, \{x\}, \{x, z\}, X\}$
$T_{14} = \{\varnothing, \{z\}, \{x, y\}, X\}$	$\{\varnothing, \{z\}, \{x, y\}, X\}$
$T_{24} = \{\varnothing, \{z\}, \{x, z\}, X\}$	$\{\varnothing, \{y\}, \{x, y\}, X\}$
$T_{44} = \{\varnothing, \{z\}, \{y, z\}, X\}$	$\{\varnothing, \{x\}, \{x, y\}, X\}$
$T_{31} = \{\varnothing, \{x\}, \{x, y\}, \{x, z\}, X\}$	$\{\varnothing, \{y\}, \{z\}, \{y, z\}, X\}$
$T_{52} = \{\varnothing, \{y\}, \{x, y\}, \{y, z\}, X\}$	$\{\varnothing, \{x\}, \{z\}, \{x, z\}, X\}$
$T_{64} = \{\varnothing, \{z\}, \{x, z\}, \{y, z\}, X\}$	$\{\varnothing, \{x\}, \{y\}, \{x, y\}, X\}$

$\boldsymbol{T}_{13} = \{\varnothing, \{x\}, \{y\}, \{x,y\}, X\}$	$\{\varnothing, \{z\}, \{x,z\}, \{y,z\}, X\}$
$\boldsymbol{T}_{25} = \{\varnothing, \{x\}, \{z\}, \{x,z\}, X\}$	$\{\varnothing, \{y\}, \{x,y\}, \{y,z\}, X\}$
$\boldsymbol{T}_{46} = \{\varnothing, \{y\}, \{z\}, \{y,z\}, X\}$	$\{\varnothing, \{x\}, \{x,y\}, \{x,z\}, X\}$
$\boldsymbol{T}_{33} = \{\varnothing, \{x\}, \{y\}, \{x,y\}, \{x,z\}, X\}$	$\{\varnothing, \{y\}, \{z\}, \{x,z\}, \{y,z\}, X\}$
$\boldsymbol{T}_{53} = \{\varnothing, \{x\}, \{y\}, \{x,y\}, \{y,z\}, X\}$	$\{\varnothing, \{x\}, \{z\}, \{x,z\}, \{y,z\}, X\}$
$\boldsymbol{T}_{35} = \{\varnothing, \{x\}, \{z\}, \{x,y\}, \{x,z\}, X\}$	$\{\varnothing, \{y\}, \{z\}, \{x,y\}, \{y,z\}, X\}$
$\boldsymbol{T}_{65} = \{\varnothing, \{x\}, \{z\}, \{x,z\}, \{y,z\}, X\}$	$\{\varnothing, \{x\}, \{y\}, \{x,y\}, \{y,z\}, X\}$
$\boldsymbol{T}_{56} = \{\varnothing, \{y\}, \{z\}, \{x,y\}, \{y,z\}, X\}$	$\{\varnothing, \{x\}, \{z\}, \{x,y\}, \{x,z\}, X\}$
$\boldsymbol{T}_{66} = \{\varnothing, \{y\}, \{z\}, \{x,z\}, \{y,z\}, X\}$	$\{\varnothing, \{x\}, \{y\}, \{x,y\}, \{x,z\}, X\}$
$\boldsymbol{T}_{77} = \{\varnothing, \{x\}, \{y\}, \{z\}, \{x,y\}, \{x,z\}, \{y,z\}, X\}$	$\{\varnothing, \{x\}, \{y\}, \{z\}, \{x,y\}, \{x,z\}, \{y,z\}, X\}$

1.1.5 Bases for topologies

Definition 1.2. [8] *Let* $(X, \boldsymbol{T})$ *be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

DEF

A set $\boldsymbol{B} \subseteq 2^X$ *is a* ***base*** *for a topology* $(X, \boldsymbol{T})$ *if*

1. $\boldsymbol{B} \subseteq \boldsymbol{T}$ *and*
2. $\forall U \in \boldsymbol{T}, \exists \{B_\gamma \in \boldsymbol{B}\}$ such that $U = \bigcup \{B_\gamma \in \boldsymbol{B}\}$

An element $A \in \boldsymbol{B}$ *is called a* ***basic open set***.

Theorem 1.4. [9] *Let* $(X, \boldsymbol{T})$ *be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

THM

$$\{\boldsymbol{B} \text{ is a BASE for } \boldsymbol{T}\} \iff \left\{\begin{array}{l}\text{For every } x \in X \text{ and for every OPEN SET } U \text{ containing } x,\\ \text{there exists } B \in \boldsymbol{B} \text{ such that } x \in B \subseteq U.\end{array}\right\}$$

✎PROOF:

1. Proof for ($\Longrightarrow$) case:

$$\begin{aligned} x \in U \in \boldsymbol{T} &\implies \exists\{B_\gamma \in \boldsymbol{B}\} \text{ such that } U = \bigcup\{B_\gamma \in \boldsymbol{B}\} && \text{by "}\boldsymbol{B}\text{ is a } base\text{" hypothesis}\\ &\implies \exists B_\gamma \text{ such that } x \in B_\gamma \subseteq U && \text{because } B_\gamma \subseteq \bigcup\{B_\gamma \in \boldsymbol{B}\} \end{aligned}$$

2. Proof for ($\Longleftarrow$) case:

$$\begin{aligned} U \in \boldsymbol{T} &\implies \forall x \in U \exists \{B_\gamma \in \boldsymbol{B}\} \text{ such that } U = \bigcup\{B_\gamma \in \boldsymbol{B} \mid x \in B_\gamma\} && \text{by right hypothesis}\\ &\implies \{\boldsymbol{B} \text{ is a } base \text{ for } \boldsymbol{T}\} && \text{by definition of } base\text{: Definition 1.2} \end{aligned}$$

✏

Theorem 1.5. [10] *Let* $(X, \boldsymbol{T})$ *be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3) and* $\boldsymbol{B} \subseteq 2^X$.

THM

$$\boldsymbol{B} \text{ is a base for } (X, \boldsymbol{T}) \iff \left\{\begin{array}{lllll} 1. & x \in X & \implies & \exists B \in \boldsymbol{B} \text{ such that } x \in B & \text{and}\\ 2. & B_1, B_2 \in \boldsymbol{B} & \implies & B_1 \cap B_2 \in \boldsymbol{B} & \end{array}\right.$$

Example 1.7. Consider the 29 topologies on the set $\{x, y, z\}$ (Example 1.6 page 7).

[8] Joshi (1983) page 92 ⟨(3.1) Definition⟩, Davis (2005) page 46 ⟨Definition 4.15⟩
[9] Joshi (1983) pages 92–93 ⟨(3.2) Proposition⟩, Davis (2005) page 46
[10] Bollobás (1999) page 19

EX

This family of sets	is a *base* for these topologies on $\{x, y, z\}$:
$\{\{x\}, \{y, z\}\}$	T_{00}, T_{01}, T_{40}, and T_{41}.
$\{\{y\}, \{x, z\}\}$	T_{00}, T_{02}, T_{20}, and T_{22}.
$\{\{z\}, \{x, y\}\}$	T_{00}, T_{04}, T_{10}, and T_{14}.
$\{\{x\}, \{x, y\}, \{x, z\}\}$	T_{00}, T_{11}, T_{21}, and T_{31}.
$\{\{y\}, \{x, y\}, \{y, z\}\}$	T_{00}, T_{12}, T_{42}, and T_{52}.
$\{\{z\}, \{x, z\}, \{y, z\}\}$	T_{00}, T_{24}, T_{44}, and T_{64}.
$\{\{x\}, \{y\}, \{x, y, z\}\}$	T_{00}, T_{01}, T_{02}, T_{10}, T_{11}, T_{12}, and T_{13}.
$\{\{x\}, \{z\}, \{x, y, z\}\}$	T_{00}, T_{01}, T_{04}, T_{20}, T_{21}, T_{24}, and T_{25}.
$\{\{y\}, \{z\}, \{x, y, z\}\}$	T_{00}, T_{02}, T_{04}, T_{40}, T_{42}, T_{44}, and T_{46}.
$\{\{x\}, \{y\}, \{x, z\}\}$	T_{00}, T_{01}, T_{02}, T_{10}, T_{11}, T_{12}, T_{13}, T_{20}, T_{21}, T_{22}, T_{31}, and T_{33}.
$\{\{x\}, \{y\}, \{y, z\}\}$	T_{00}, T_{01}, T_{02}, T_{10}, T_{11}, T_{12}, T_{13}, T_{40}, T_{41}, T_{42}, T_{52}, and T_{53}.
$\{\{x\}, \{z\}, \{x, y\}\}$	T_{00}, T_{01}, T_{04}, T_{10}, T_{11}, T_{14}, T_{20}, T_{21}, T_{24}, T_{25}, T_{31}, and T_{35}.
$\{\{x\}, \{z\}, \{y, z\}\}$	T_{00}, T_{01}, T_{04}, T_{20}, T_{21}, T_{24}, T_{25}, T_{40}, T_{41}, T_{44}, T_{64}, and T_{65}.
$\{\{y\}, \{z\}, \{x, z\}\}$	T_{00}, T_{02}, T_{04}, T_{20}, T_{22}, T_{24}, T_{40}, T_{42}, T_{44}, T_{46}, T_{64}, and T_{66}.
$\{\{y\}, \{z\}, \{x, z\}\}$	T_{00}, T_{02}, T_{04}, T_{10}, T_{12}, T_{14}, T_{40}, T_{42}, T_{44}, T_{46}, T_{52}, and T_{56}.
$\{\{x\}, \{y\}, \{z\}\}$	all 29 of the topologies.

Example 1.8. [11] Let (X, d) be a *metric space*.

EX The set $\boldsymbol{B} \triangleq \{\mathsf{B}(x, r) \,|\, x \in X,\ r \in \mathbb{N}\}$ (the set of all open balls in (X, d)) is a *base* for a topology on (X, d).

Example 1.9 (the standard topology on the real line). [12]

EX The set $\boldsymbol{B} \triangleq \{(a : b) | a, b \in \mathbb{R},\ a < b\}$ is a *base* for the metric space $(\mathbb{R}, |b - a|)$ (the *usual metric space* on $\mathbb{R}$).

Example 1.10. [13]

EX The set $\boldsymbol{B} \triangleq \{(a : b) | a, b \in \mathbb{Q},\ a < b\}$ is a *base* for the metric space $(\mathbb{R}, |b - a|)$ (the *usual metric space* on $\mathbb{R}$).

The possible advantage of this base over the base of Example 1.9 is that this base is *countable*.

Example 1.11 (lower limit topology/the Sorgenfrey line topology). [14]

EX The set $\boldsymbol{B} \triangleq \{[a : b) | a, b \in \mathbb{R},\ a < b\}$ is a *base* for the metric space $(\mathbb{R}, |b - a|)$ (the *usual metric space* on $\mathbb{R}$).

Under this topology, the *cumulative distribution functions* of probability theory are *continuous*.

Counterexample 1.1. [15] Definition 1.1 (page 3) states that the intersection of a *finite* number of open sets is also open. But under this definition, in general it is *not* true that the intersection of an infinite number of open sets is open. Take for example the *standard topology on the real line* (Example 1.9 page 9):

1. Let $\left(A_n = \left(-\frac{1}{n}, \frac{1}{n}\right)\right)_{n\in\mathbb{N}}$ be a sequence of real intervals. That is

$$\left((-1, 1),\ \left(-\frac{1}{2}, \frac{1}{2}\right),\ \left(-\frac{1}{3}, \frac{1}{3}\right),\ \left(-\frac{1}{4}, \frac{1}{4}\right),\ \left(-\frac{1}{5}, \frac{1}{5}\right),\ \ldots\right)$$

2. Then $\bigcap_{n\in\mathbb{N}} A_n = \{0\}$; that is, $\bigcap_{n\in\mathbb{N}} A_n$ is a set with just one value (0).

[11] Davis (2005) page 46 ⟨Example 4.16⟩
[12] Munkres (2000) page 81, Davis (2005) page 46 ⟨Example 4.16⟩
[13] Davis (2005) page 46 ⟨Example 4.16⟩
[14] Munkres (2000) pages 81–82, Davis (2005) page 48 ⟨Example 4.21⟩
[15] Rosenlicht (1968), page 40

3. A single value is *not* an open set because any ball with radius greater than 0 is not in the set (Lemma 2.3 page 33).

4. Therefore, $\bigcap_{n\in\mathbb{N}} A_n$ is not open.

1.1.6 Order structure of the topologies on a set

In general for a given set X, there is not just one possible topology. Rather, for any sizeable set X, there are myriads of topologies. Some of these topologies are subsets of other topologies; in such a case, Definition 1.3 (page 10) states that we say that the subset topology is *coarser* than the other and that the other superset topology is *finer*. And not only does any individual topology generate a lattice, but as demonstrated by Theorem 1.6 (page 10), all the topologies taken together also form a lattice. Examples of lattices of topologies are provided by the following:

- Example 1.12 (page 10): lattice of the 4 topologies of a 2 element set X.
- Example 1.13 (page 10): lattice of the 29 topologies of a 3 element set X.

Definition 1.3. [16] *Let $(X, \boldsymbol{S})$ and $(X, \boldsymbol{T})$ be two* TOPOLOGICAL SPACES *(Definition 1.1 page 3) on a set X.*

DEF

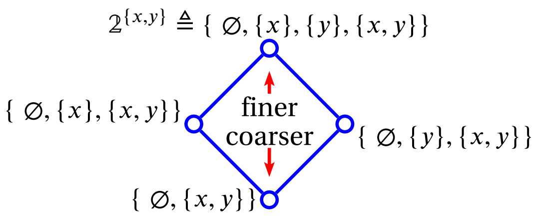
*$\boldsymbol{S}$ is **coarser** than $\boldsymbol{T}$ and $\boldsymbol{T}$ is **finer** than $\boldsymbol{S}$ if*
$\boldsymbol{S} \subseteq \boldsymbol{T}$.
*$\boldsymbol{S}$ is **strictly coarser** than $\boldsymbol{T}$ and $\boldsymbol{T}$ is **strictly finer** than $\boldsymbol{S}$ if*
$\boldsymbol{S} \subsetneq \boldsymbol{T}$.

Theorem 1.6 (Lattice of topologies). [17]

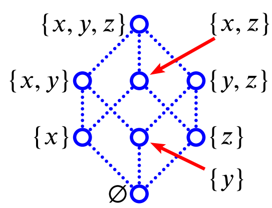
THM

$$\underbrace{\mathcal{T}(X) \triangleq \{T_1, T_2, \ldots | T_n \text{ is a topology on } X\}}_{\text{the set of topologies on } X} \implies (\mathcal{T}(X), \cup, \cap; \subseteq) \text{ is a lattice.}$$

Example 1.12. [18] Example 1.5 (page 7) lists the four topologies on the set $X \triangleq \{x, y\}$. The lattice of these topologies $(\{\boldsymbol{T}_1, \boldsymbol{T}_2, \boldsymbol{T}_3, \boldsymbol{T}_4\}, \cup, \cap; \subseteq)$ is illustrated by the figure below and to the right. Note that there are only four valid topologies out of a total sixteen possible families of sets: $(2^{|2^X|} = 2^{2^{|X|}} = 2^{2^2} = 2^4 = 16)$. Half of the sixteen families are not valid topologies because they do not contain $\varnothing$ and half of the remaining are not valid because they do not contain X. This leaves $16 \times \frac{1}{2} \times \frac{1}{2} = 4$ topologies.



Example 1.13. [19] Let a given topology in $\mathcal{T}(\{x, y, z\})$ be represented by a Hasse diagram as illustrated to the right, where a circle present means the indicated set is in the topology, and a circle absent means the indicated set is not in the topology. Example 1.6 (page 7) lists the 29 topologies $\mathcal{T}(\{x, y, z\})$. The lattice of these 29 topologies $(\mathcal{T}(\{x, y, z\}), \cup, \cap; \subseteq)$ is illustrated in Figure 1.1 (page 11) and Figure 1.2 (page 12). The five topologies $\boldsymbol{T}_1, \boldsymbol{T}_{41}, \boldsymbol{T}_{22}, \boldsymbol{T}_{14}$, and $\boldsymbol{T}_{77}$ are also *algebras of sets* (Definition A.9 page 272); these five sets are shaded in Figure 1.1 and represented as solid dots in Figure 1.2.



[16] Munkres (2000) page 77
[17] Larson and Andima (1975), Vaidyanathaswamy (1960) page 131, Birkhoff (1936a), Stone (1936b), Wallman (1938)
[18] Isham (1999), page 44, Isham (1989), page 1515
[19] Isham (1999), page 44, Isham (1989), page 1516, Steiner (1966), page 386

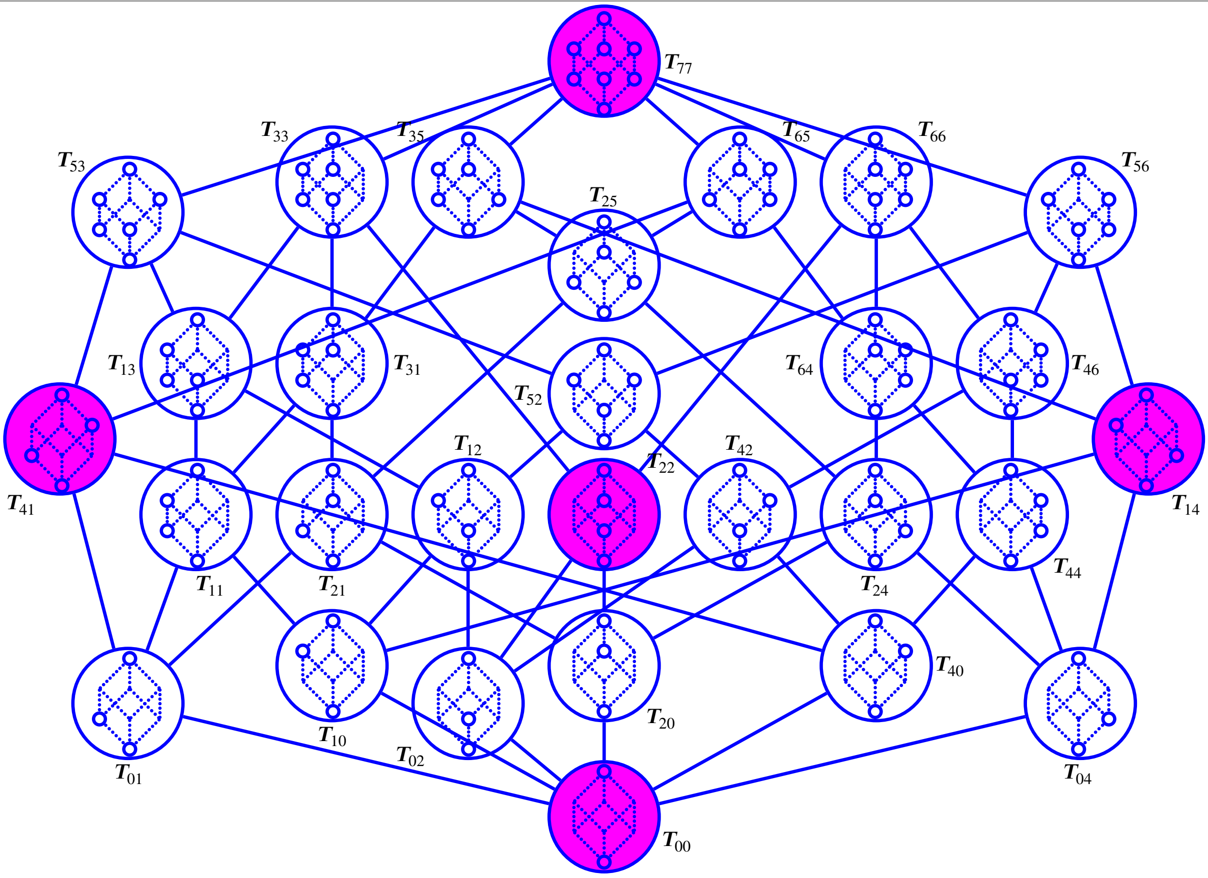


Figure 1.1: Lattice of topologies on $X \triangleq \{x, y, z\}$ (see Example 1.13 page 10)

Theorem 1.7. [20] *Let $\mathcal{T}(X)$ be the lattice of topologies on a set X with $|X|$ elements.*

THM
$$\begin{array}{rcl} |X| \le 2 & \Longrightarrow & \mathcal{T}(X) \text{ is DISTRIBUTIVE} \\ |X| \ge 3 & \Longrightarrow & \mathcal{T}(X) \text{ is NOT MODULAR (and not distributive)} \end{array}$$

Theorem 1.8. [21] *Let $\mathcal{T}(X)$ be the lattice of topologies on a set X.*

THM
$$\mathcal{T}(X) \text{ is SELF-DUAL} \iff |X| \le 3$$

Theorem 1.9. [22]

THM
Every lattice of topologies is complemented.

Theorem 1.10. [23]

THM
Every topology except the discrete and indiscrete topology in the lattice of topologies on a set X has at least $|X| - 1$ complements.

Let $\hat{\Sigma}(X)$ be the set of all topologies on X except for the discrete and indiscrete topologies on X.

Example 1.14. Example 1.6 (page 7) lists the 29 topologies on a set $X \triangleq \{x, y, z\}$. By Theorem 1.10 (page 11), with the exception of T_{00} (the indiscrete topology) and T_{77} (the discrete topology), each

[20] Steiner (1966), page 384
[21] Steiner (1966), page 385
[22] van Rooij (1968), Steiner (1966), page 397, Gaifman (1961), Hartmanis (1958)
[23] Hartmanis (1958), Schnare (1968), page 56, Watson (1994), Brown and Watson (1996), page 32

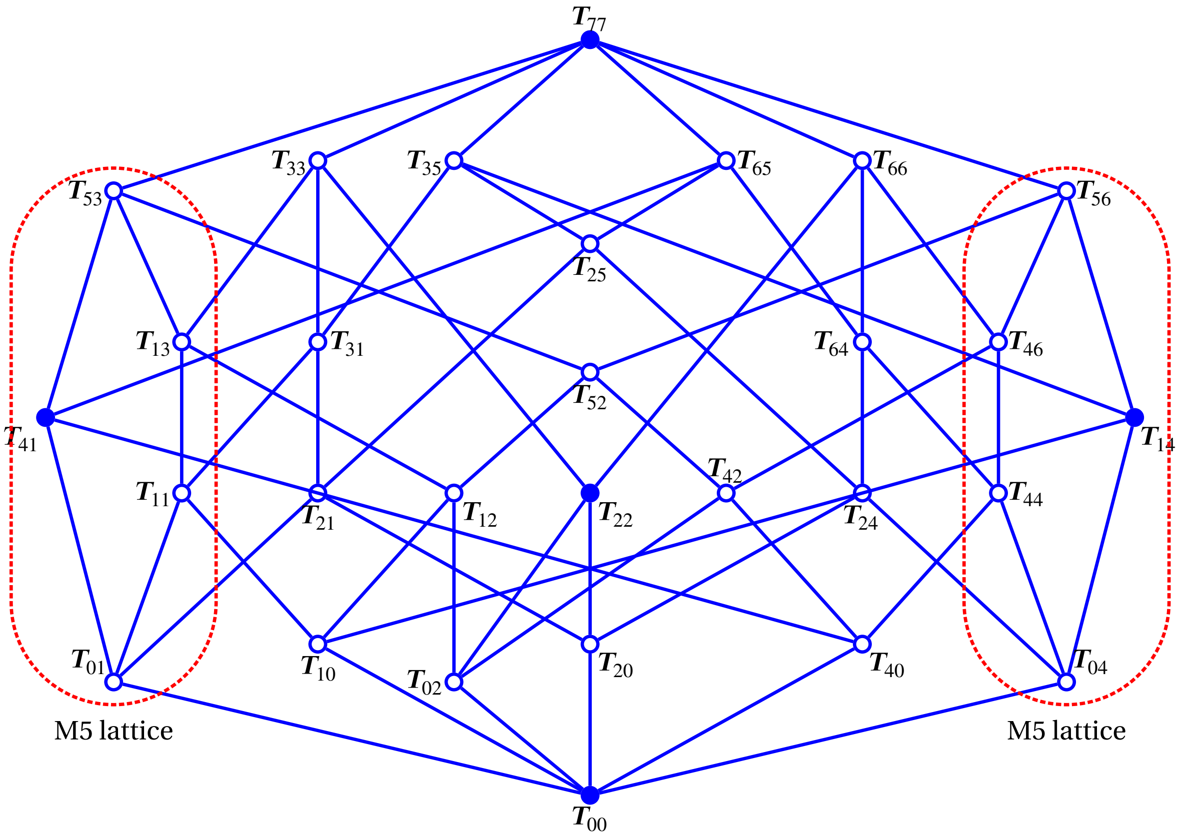


Figure 1.2: Lattice of topologies of $X \triangleq \{x, y, z\}$ (see Example 1.13 page 10)

of those topologies has exactly $|X| - 1 = 3 - 1 = 2$ complements. Listed below are the 29 topologies on $\{x, y, z\}$ along with their respective complements.

topologies on $\{x, y, z\}$	1st complement	2nd compl.
$T_{00} = \{\emptyset, X\}$	T_{77}	
$T_{01} = \{\emptyset, \{x\}, X\}$	T_{56}	T_{66}
$T_{02} = \{\emptyset, \{y\}, X\}$	T_{65}	T_{35}
$T_{04} = \{\emptyset, \{z\}, X\}$	T_{53}	T_{33}
$T_{10} = \{\emptyset, \{x, y\}, X\}$	T_{65}	T_{66}
$T_{20} = \{\emptyset, \{x, z\}, X\}$	T_{53}	T_{56}
$T_{40} = \{\emptyset, \{y, z\}, X\}$	T_{33}	T_{35}
$T_{11} = \{\emptyset, \{x\}, \{x, y\}, X\}$	T_{64}	T_{46}
$T_{21} = \{\emptyset, \{x\}, \{x, z\}, X\}$	T_{52}	T_{46}
$T_{41} = \{\emptyset, \{x\}, \{y, z\}, X\}$	T_{22}	T_{14}
$T_{12} = \{\emptyset, \{y\}, \{x, y\}, X\}$	T_{64}	T_{25}
$T_{22} = \{\emptyset, \{y\}, \{x, z\}, X\}$	T_{41}	T_{14}
$T_{42} = \{\emptyset, \{y\}, \{y, z\}, X\}$	T_{31}	T_{25}
$T_{14} = \{\emptyset, \{z\}, \{x, y\}, X\}$	T_{41}	T_{22}
$T_{24} = \{\emptyset, \{z\}, \{x, z\}, X\}$	T_{52}	T_{13}
$T_{44} = \{\emptyset, \{z\}, \{y, z\}, X\}$	T_{31}	T_{13}
$T_{31} = \{\emptyset, \{x\}, \{x, y\}, \{x, z\}, X\}$	T_{42}	T_{44}
$T_{52} = \{\emptyset, \{y\}, \{x, y\}, \{x, z\}, X\}$	T_{21}	T_{24}
$T_{64} = \{\emptyset, \{z\}, \{x, z\}, \{y, z\}, X\}$	T_{11}	T_{12}
$T_{13} = \{\emptyset, \{x\}, \{y\}, \{x, y\}, X\}$	T_{24}	T_{44}
$T_{25} = \{\emptyset, \{x\}, \{z\}, \{x, z\}, X\}$	T_{12}	T_{42}
$T_{46} = \{\emptyset, \{y\}, \{z\}, \{y, z\}, X\}$	T_{11}	T_{21}
$T_{33} = \{\emptyset, \{x\}, \{y\}, \{x, y\}, \{x, z\}, X\}$	T_{04}	T_{40}
$T_{53} = \{\emptyset, \{x\}, \{y\}, \{x, y\}, \{y, z\}, X\}$	T_{04}	T_{20}
$T_{35} = \{\emptyset, \{x\}, \{z\}, \{x, y\}, \{x, z\}, X\}$	T_{02}	T_{40}
$T_{65} = \{\emptyset, \{x\}, \{z\}, \{x, z\}, \{y, z\}, X\}$	T_{02}	T_{10}
$T_{56} = \{\emptyset, \{y\}, \{z\}, \{x, y\}, \{y, z\}, X\}$	T_{01}	T_{20}
$T_{66} = \{\emptyset, \{y\}, \{z\}, \{x, z\}, \{y, z\}, X\}$	T_{01}	T_{10}
$T_{77} = \{\emptyset, \{x\}, \{y\}, \{z\}, \{x, y\}, \{x, z\}, \{y, z\}, X\}$	T_{00}	

Theorem 1.11. [24]

THM

$$\mathcal{T}(X) \text{ is a topology of sets} \implies \begin{cases} \mathcal{T}(X) \text{ is atomic.} \\ \mathcal{T}(X) \text{ is anti-atomic.} \end{cases}$$

Theorem 1.12. [25] *Let $\mathcal{T}(X)$ be the lattice of topologies on a set X and let $n \triangleq |X|$.*

THM

$\mathcal{T}(X)$ *contains* $2^n - 2$ *atoms*	*for finite* X.
$\mathcal{T}(X)$ *contains* $2^{|X|}$ *atoms*	*for infinite* X.
$\mathcal{T}(X)$ *contains* $n(n-1)$ *anti-atoms*	*for finite* X.
$\mathcal{T}(X)$ *contains* $2^{2^{|X|}}$ *anti-atoms*	*for infinite* X.

[24] Larson and Andima (1975), page 179, Frölich (1964), Vaidyanathaswamy (1960), Vaidyanathaswamy (1947)

[25] Larson and Andima (1975), page 179, Frölich (1964)

1.2 Derived Sets

1.2.1 Definitions

Several useful set structures can be derived from the simple concept of the open set (next definition).

Definition 1.4. [26] *Let (X, T) be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3). Let 2^X be the* POWER SET *of X.*

DEF

*The set A^- is the **closure** of $A \in 2^X$ if* $A^- \triangleq \bigcap \{D \in 2^X | A \subseteq D \text{ and } D \text{ is CLOSED}\}$.

*The set A° is the **interior** of $A \in 2^X$ if* $A^\circ \triangleq \bigcup \{U \in 2^X | U \subseteq A \text{ and } U \text{ is OPEN}\}$.

*A point x is a **closure point** of A if $x \in A^-$.*

*A point x is an **interior point** of A if $x \in A^\circ$.*

*A point x is an **accumulation point** of A if $x \in (A \setminus \{x\})^-$*

*A point x in A^- is a **point of adherence** in A or is **adherent** to A if $x \in A^-$.*

Definition 1.5. [27] *Let (X, T) be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3). Let 2^X be the* POWER SET *of X.*

DEF

*The set ∂A is the **boundary** of $A \in 2^X$ if $\partial A \triangleq A^- \cap \left(A^c\right)^-$.*

*The set A^e is the **exterior** of $A \in 2^X$ if $A^e \triangleq \left(A^c\right)^\circ$.*

*A point x in X is a **boundary point** of A if $x \in \partial A$.*

*A point x in X is an **exterior point** of A if $x \in A^e$.*

*A point x in A^- is a **point of adherence** in A or is **adherent** to A if $x \in A^-$.*

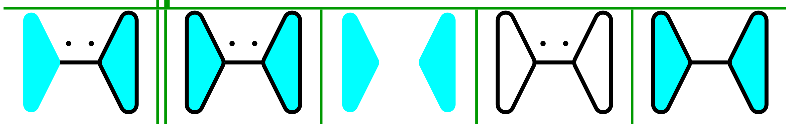
*The set A' is the **derived set** of $A \in 2^X$ if*

$$A' \triangleq \{x \in X | x \text{ is an accumlation point of } A\}.$$

Example 1.15. [28] Let A be the set illustrated as follows in a topogical space (X, T). The sets defined in Definition 1.4 page 14 are illustrated to the right of A.

EX

set A	derived sets			
A	A^-	A°	∂A	A'



1.2.2 Resulting properties

Proposition 1.2. [29] *Let (X, T) be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

PRP $\left\{\begin{array}{l} x \text{ is an ACCUMULATION POINT} \\ \text{of a set } A \text{ in } (X, T) \end{array}\right\} \iff \left\{\begin{array}{l} \text{Every open set containing } x \text{ also contains} \\ \text{another point } y \in A,\ y \neq x). \end{array}\right\}$

Proposition 1.3. [30] *Let (X, T) be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3). Let A^- be the* CLOSURE, *A°*

[26] Gemignani (1972) pages 55–56 ⟨Definition 3.5.7⟩, McCarty (1967) page 90, Munkres (2000) page 95 ⟨§Closure and Interior of a Set⟩, Thron (1966), pages 21–22 ⟨definition 4.8, defintion 4.9⟩, Kelley (1955) page 42, Kubrusly (2001) pages 115–116

[27] Gemignani (1972) pages 55–56 ⟨Definition 3.5.7⟩, McCarty (1967) page 90, Munkres (2000) page 95 ⟨§Closure and Interior of a Set⟩, Thron (1966), pages 21–22 ⟨definition 4.8, defintion 4.9⟩, Kelley (1955) page 42, Kubrusly (2001) pages 115–116, Murdeshwar (1990) page 48 ⟨exterior A^e⟩, Joshi (1983) page 110 ⟨exterior A^e⟩

[28] McCarty (1967) page 90

[29] Kubrusly (2001) pages 115–116 ⟨Proposition 3.26⟩, Murdeshwar (1990) page 48 ⟨1.24 Exercises (19)⟩

[30] McCarty (1967) page 90 ⟨IV.1 THEOREM⟩

the INTERIOR, *and* ∂A *the* BOUNDARY *of a set* A. *Let* 2^X *be the* POWER SET *of* X.

PRP

1. A^- *is* CLOSED $\forall A \in 2^X$.
2. A° *is* OPEN $\forall A \in 2^X$.
3. ∂A *is* CLOSED $\forall A \in 2^X$.

✎PROOF:

$$\begin{aligned}
A^- &\triangleq \bigcap \left\{D \in 2^X | A \subseteq D \text{ and } D \text{ is closed}\right\} && \text{by Definition 1.4 page 14}\\
&\implies A^- \text{ is } \textit{closed} && \text{by Theorem 1.3 page 6}\\
A^\circ &\triangleq \bigcap \left\{U \in 2^X | U \subseteq A \text{ and } U \text{ is open}\right\} && \text{by Definition 1.4 page 14}\\
&\implies A^\circ \text{ is } \textit{open} && \text{by Definition 1.1 page 3}\\
\partial A &\triangleq A^- \cap \left(A^c\right)^- && \text{by Definition 1.4 page 14}\\
&\implies \partial A \text{ is } \textit{closed} && \text{by (1) and Theorem 1.3 page 6}
\end{aligned}$$

✏

Lemma 1.1. [31] *Let* A^- *be the* CLOSURE, A° *the* INTERIOR, *and* ∂A *the* BOUNDARY *of a set* A *in a topological space* (X, T). *Let* 2^X *be the* POWER SET *of* X.

LEM

1. $A^\circ \subseteq A \subseteq A^-$ $\forall A \in 2^X$.
2. $A' \subseteq A^-$ $\forall A \in 2^X$.
3. $A = A^\circ \iff A$ *is* OPEN $\forall A \in 2^X$.
4. $A = A^- \iff A$ *is* CLOSED $\forall A \in 2^X$.
5. $\left\{\begin{array}{ll} D \in 2^X \text{ is CLOSED} & \text{and} \\ A \subseteq D & \end{array}\right\} \implies A^- \subseteq D$ $\left(\begin{array}{l} A^- \text{ is the smallest CLOSED set} \\ \text{containing } A \end{array}\right)$ $\forall A \in 2^X$.
6. $\left\{\begin{array}{ll} U \in 2^X \text{ is OPEN} & \text{and} \\ U \subseteq A & \end{array}\right\} \implies U \subseteq A^\circ$ $\left(\begin{array}{l} A^\circ \text{ is the largest OPEN set} \\ \text{contained in } A \end{array}\right)$ $\forall A \in 2^X$.

✎PROOF:

1. Proof that $A^\circ \subseteq A \subseteq A^-$:

$$\begin{aligned}
A^\circ &\triangleq \bigcup \left\{U \in 2^X | U \subseteq A \text{ and } U \text{ is open}\right\} && \text{by Definition 1.4 page 14}\\
&\subseteq A\\
A &\subseteq \bigcap \left\{D \in 2^X | A \subseteq D \text{ and } D \text{ is closed}\right\}\\
&\triangleq A^- && \text{by Definition 1.4 page 14}
\end{aligned}$$

2. Proof that $A' \subseteq A^-$:

$$\begin{aligned}
A' &\triangleq \{x \in X | x \in (A \setminus \{x\})^-\} && \text{by definition of } A'\text{: Definition 1.5 page 14}\\
&\subseteq \{x \in X | x \in A^-\} && \text{by Theorem 1.16 page 18}\\
&= A^-
\end{aligned}$$

3. Proof that $A = A^\circ \implies A$ is *open*: by Proposition 1.3 page 14.

4. Proof that $A = A^\circ \impliedby A$ is *open*:

$$\begin{aligned}
A^\circ &\triangleq \bigcup \left\{U \in 2^X | U \subseteq A \text{ and } U \text{ is open}\right\} && \text{by Definition 1.4 page 14}\\
&= A && \text{by "} A \text{ is } \textit{open} \text{" hypothesis}
\end{aligned}$$

[31] McCarty (1967) pages 90–91 ⟨IV.1 THEOREM⟩, Aliprantis and Burkinshaw (1998) page 59

5. Proof that $A = A^- \implies A$ is *closed*: by Proposition 1.3 page 14.
6. Proof that $A = A^- \impliedby A$ is *closed*:

$$\begin{aligned} A^- &\triangleq \bigcap \left\{D \in 2^X | A \subseteq D \text{ and } D \text{ is closed}\right\} && \text{by Definition 1.4 page 14} \\ &= A && \text{by "}A\text{ is \textit{closed}" hypothesis} \end{aligned}$$

7. Proof that $\mathit{cls}\, A$ is the smallest *closed* set containing A:

$$\begin{aligned} A^- &\triangleq \bigcap \left\{B \in 2^X | B \text{ is \textit{closed} and } A \subseteq B\right\} && \text{by definition of } A^-\text{: Definition 1.4 page 14} \\ &\subseteq D && \text{because } D \text{ is \textit{closed} by hypothesis} \end{aligned}$$

8. Proof that $\mathit{int}\, A$ is the largest *open* set contained in A:

$$\begin{aligned} U &\subseteq \bigcup \left\{V \in 2^X | V \text{ is \textit{closed} and } V \subseteq A\right\} && \text{by definition of } \bigcap \\ &\triangleq A^\circ && \text{by definition of } A^\circ\text{: Definition 1.4 page 14} \end{aligned}$$

⇨

Theorem 1.13 (Kuratowski closure properties). [32] *Let* (X, T) *be a* TOPOLOGICAL SPACE.

THM

1.	$\varnothing^-$	$=$	$\varnothing$		(NORMALIZED)	*and*
2.	A	$\subseteq$	A^-	$\forall A \in 2^X$	(EXTENSIVE)	*and*
3.	$(A^-)^-$	$=$	A^-	$\forall A \in 2^X$	(IDEMPOTENT)	*and*
4.	$(A \cup B)^-$	$=$	$A^- \cup B^-$	$\forall A, B \in 2^X$	(ADDITIVE).	

✎PROOF:

1. Proof that $\varnothing^- = \varnothing$:

$$\begin{aligned} &\varnothing \text{ is \textit{closed}} && \text{by Theorem 1.3 page 6} \\ &\implies \varnothing^- = \varnothing && \text{by Lemma 1.1 page 15} \end{aligned}$$

2. Proof that $A \subseteq A^-$: by Lemma 1.1 page 15
3. Proof that $(A^-)^- = A^-$:

$$\begin{aligned} (A^-)^- &\triangleq \left(\bigcap \left\{D \in 2^X | (A^-) \subseteq D \text{ and } D \text{ is closed}\right\}\right) && \text{by Definition 1.4} \\ &= A^- && \text{because } (A^-) \text{ is closed by Proposition 1.3} \end{aligned}$$

4. Proof that $A^- \cup B^- = (A \cup B)^-$:

$$\begin{aligned} A^- \cup B^- &= (A^- \cup B^-)^- && \text{by Theorem 1.3 (page 6) } A^- \cup B^- \text{ is \textit{closed}} \\ & && \text{and by Lemma 1.1 (page 15)} \\ &\supseteq (A \cup B)^- && \text{by Lemma 1.1 page 15} \end{aligned}$$

$$\begin{aligned} A^- \cup B^- &\subseteq \Bigg[\underbrace{(A \cup B)^-}_{A \subseteq (A \cup B)^-}\Bigg]^- \cup \Bigg[\underbrace{(A \cup B)^-}_{B \subseteq (A \cup B)^-}\Bigg]^- && \text{because } A \subseteq (A \cup B)^- \text{ and } B \subseteq (A \cup B)^- \\ &= (A \cup B)^- \cup (A \cup B)^- && \text{by item (3)} \\ &\triangleq (A \cup B)^- && \text{by Theorem A.4 page 277} \end{aligned}$$

[32] Kelley (1955) page 43 ⟨1.8 THEOREM⟩, Davis (2005) page 45, Thron (1966), pages 21–22, Hausdorff (1937) page 258, Kuratowski (1922) page 182, Riesz (1906)

Theorem 1.14. *Let* (X, T) *be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

THM						
1.	$\varnothing^\circ$	$=$	$\varnothing$		(NORMALIZED)	and
2.	A°	$\subseteq$	A	$\forall A \in 2^X$	(EXTENSIVE)	and
3.	$(A^\circ)^\circ$	$=$	A°	$\forall A \in 2^X$	(IDEMPOTENT)	and
4.	$(A \cup B)^\circ$	$=$	$A^\circ \cap B^\circ$	$\forall A, B \in 2^X$	(ADDITIVE).	

✎PROOF:

1. Proof that $\varnothing^\circ = \varnothing$:

$$\begin{aligned} &\varnothing \text{ is } open && \text{by Definition 1.1 page 3} \\ &\implies \varnothing^\circ = \varnothing && \text{by Lemma 1.1 page 15} \end{aligned}$$

2. Proof that $A \supseteq A^\circ$: by Lemma 1.1 page 15.

3. Proof that $(A^\circ)^\circ = A^\circ$:

$$\begin{aligned} (A^\circ)^\circ &\triangleq \bigcup \left\{U \in 2^X | U \subseteq A^\circ \text{ and } U \text{ is open}\right\} && \text{by Definition 1.4 page 14} \\ &= A^\circ && \text{by Proposition 1.3 page 14} \end{aligned}$$

4. Proof that $A^\circ \cap B^\circ = (A \cap B)^\circ$:

$$\begin{aligned} A^\circ \cap B^\circ &= (A^\circ \cap B^\circ)^\circ && \text{by Definition 1.1 (page 3) and Lemma 1.1 (page 15)} \\ &\subseteq (A \cap B)^\circ && \text{by Lemma 1.1 page 15} \end{aligned}$$

$$\begin{aligned} A^\circ \cap B^\circ &\supseteq \Bigg[\underbrace{(A \cap B)^\circ}_{A \supseteq (A \cap B)^\circ}\Bigg]^\circ \cap \Bigg[\underbrace{(A \cap B)^\circ}_{B \supseteq (A \cap B)^\circ}\Bigg]^\circ && \text{because } A \supseteq (A \cap B)^\circ \text{ and } B \supseteq (A \cap B)^\circ \\ &= (A \cap B)^\circ \cap (A \cap B)^\circ && \text{by item (3)} \\ &\triangleq (A \cap B)^\circ && \text{by Theorem A.4 page 277} \end{aligned}$$

Theorem 1.15. [33] *Let* (X, T) *be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

THM						
1.	$\varnothing'$	$=$	$\varnothing$		(NORMALIZED)	and
2.	A'	$\subseteq$	A^-	$\forall A \in 2^X$	(EXTENSIVE)	and
3.	$(A \cup B)'$	$=$	$A' \cup B'$	$\forall A, B \in 2^X$	(ADDITIVE).	

✎PROOF:

1. Proof that $\varnothing' = \varnothing$:

$$\begin{aligned} \varnothing' &= \{x \in X | x \in (x \in \varnothing \setminus \{x\})^-\} && \text{by definition of } A'\text{: Definition 1.5 page 14} \\ &= \{x \in X | x \in (x \in \varnothing)^-\} && \text{by definition of } \varnothing \end{aligned}$$

[33] Mukherjee (2005) page 32 ⟨2.2.14 Theorem⟩, Murdeshwar (1990) page 48 ⟨1.24 Exercises (19)⟩

2. Proof that $A' \subseteq A^-$: by Lemma 1.1 page 15.

3. Proof that $A' \cup B' = (A \cup B)'$:

$$\begin{aligned}
A' \cup B' &\triangleq \{x \in X | x \in (x \in A \backslash \{x\})^-\} \cup \{x \in X | x \in (x \in B \backslash \{x\})^-\} && \text{by Definition 1.5 page 14}\\
&= \{x \in X | x \in (x \in A \backslash \{x\})^- \text{ or } x \in (x \in B \backslash \{x\})^-\} && \text{by definition of } \cup\\
&= \{x \in X | x \in (x \in A \backslash \{x\})^- \cup (x \in B \backslash \{x\})^-\} && \text{by definition of } \cup\\
&= \left\{x \in X | x \in \left[x \in (A \cup B) \backslash \{x\}\right]^-\right\} && \text{by Theorem 1.13 page 16}\\
&\triangleq (A \cup B)' && \text{by Definition 1.5 page 14}
\end{aligned}$$

⇨

Theorem 1.16. [34] *Let A^- be the* CLOSURE *and A° the* INTERIOR *(Definition 1.4 page 14) of a set A on the topological space (X, T).*

$$\text{THM}\quad A \subseteq B \implies \left\{\begin{array}{llll} 1. & A^- \subseteq B^- & \text{(isotone)} & \text{and}\\ 2. & A^\circ \subseteq B^\circ & \text{(isotone)} & \text{and}\\ 3. & A' \subseteq B' & \text{(isotone)} & \end{array}\right\} \quad \forall A,B \in 2^X$$

✎PROOF:

$$\begin{aligned}
A^- &\subseteq A^- \cup B^-\\
&= (A \cup B)^- && \text{by Theorem 1.13 page 16 (\textit{additivity})}\\
&= B^- && \text{by } A \subseteq B \text{ hypothesis}
\end{aligned}$$

$$\begin{aligned}
A^\circ &\triangleq \bigcup \left\{U \in 2^X | U \subseteq A \text{ and } U \text{ is open}\right\} && \text{by Definition 1.4 page 14}\\
&\subseteq \bigcup \left\{U \in 2^X | U \subseteq B \text{ and } U \text{ is open}\right\} && \text{by } A \subseteq B \text{ hypothesis}\\
&\triangleq B^\circ && \text{by Definition 1.4 page 14}
\end{aligned}$$

$$\begin{aligned}
A' &\triangleq \{x \in X | x \in (A \backslash \{x\})^-\} && \text{by Definition 1.5 page 14}\\
&\subseteq \{x \in X | x \in (B \backslash \{x\})^-\} && \text{by } A \subseteq B \text{ hypothesis}\\
&\triangleq B' && \text{by Definition 1.5 page 14}
\end{aligned}$$

⇨

Theorem 1.17. [35] *Let (X, T) be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

$$\text{THM}\quad \begin{array}{llllll}
1. & A^{-c} & = & A^{c\circ} & \forall A \in 2^X & \text{(the complement of the closure is the interior of the complement)}\\
2. & A^{\circ c} & = & A^{c-} & \forall A \in 2^X & \text{(the complement of the interior is the closure of the complement)}\\
3. & A^{-} & = & A^{c\circ c} & \forall A \in 2^X & \text{(the complement of the interior of the complement is the closure)}\\
4. & A^{\circ} & = & A^{c-c} & \forall A \in 2^X & \text{(the complement of the closure of the complement is the interior)}\\
5. & \partial A & = & \partial\left(A^c\right) & \forall A \in 2^X & \text{(the boundary of the complement is the boundary)}
\end{array}$$

✎PROOF:

[34] McCarty (1967) page 90 ⟨IV.1 THEOREM⟩, Davis (2005) PAGE 45, Thron (1966), PAGE 42 ⟨THEOREM 8.1⟩, Kubrusly (2001) PAGE 116, Kuratowski (1922) PAGE 183

[35] Murdeshwar (1990) page 43 ⟨Theorem 1.16⟩, McCarty (1967) page 90 ⟨1 THEOREM⟩, Aliprantis and Burkinshaw (1998) PAGES 59–60

1. Proof that $A^\circ = \left(\left(A^c\right)^-\right)^c$:

$$\begin{aligned}
\left(\left(A^c\right)^-\right)^c &= \left(\bigcap\left\{D \in 2^X | A^c \subseteq D \text{ and } D \text{ is } closed\right\}\right)^c && \text{by Definition 1.4} \\
&= \bigcup\left\{D^c \in 2^X | A^c \subseteq D \text{ and } D \text{ is } closed\right\} && \text{by } de\ Morgan's\ law \text{ (Theorem A.5 page 278)} \\
&= \bigcup\left\{D^c \in 2^X | D^c \subseteq A \text{ and } D^c \text{ is } open\right\} && \text{by Definition 1.1 page 3} \\
&\triangleq A^\circ && \text{by Definition 1.4 page 14}
\end{aligned}$$

2. Proof that $A^- = \left(\left(A^c\right)^\circ\right)^c$:

$$\begin{aligned}
\left(\left(A^c\right)^\circ\right)^c &= \left(\bigcup\left\{U \in 2^X | U \subseteq A^c \text{ and } U \text{ is } open\right\}\right)^c && \text{by Definition 1.4} \\
&= \bigcap\left\{U^c \in 2^X | U \subseteq A^c \text{ and } U \text{ is } open\right\} && \text{by } de\ Morgan's\ law \text{ (Theorem A.5 page 278)} \\
&= \bigcup\left\{U^c \in 2^X | A \subseteq U^c \text{ and } U^c \text{ is } closed\right\} && \text{by Definition 1.1 page 3} \\
&\triangleq A^- && \text{by Definition 1.4 page 14}
\end{aligned}$$

3. Proof that $(A^\circ)^c = \left(A^c\right)^-$:

$$\begin{aligned}
(A^\circ)^c &= \left(\left(\left(A^c\right)^-\right)^c\right)^c && \text{by item (1)} \\
&= \left(A^c\right)^- && \text{by Theorem A.5 page 278}
\end{aligned}$$

4. Proof that $(A^-)^c = \left(A^c\right)^\circ$:

$$\begin{aligned}
(A^-)^c &= \left(\left(\left(A^c\right)^\circ\right)^c\right)^c && \text{by item (2)} \\
&= \left(A^c\right)^\circ && \text{by Theorem A.5 page 278}
\end{aligned}$$

5. Proof that $\partial A = \partial\left(A^c\right)$:

$$\begin{aligned}
\partial A &= A^- \cap \left(A^c\right)^- && \text{by Definition 1.5 page 14} \\
&= \left(A^c\right)^- \cap A^- && \text{by Theorem A.4 page 277} \\
&= \left(A^c\right)^- \cap \left(\left(A^c\right)^c\right)^- && \text{by Theorem A.5 page 278} \\
&= \partial\left(A^c\right) && \text{by Definition 1.5 page 14}
\end{aligned}$$

⇨

Theorem 1.18. [36] *Let* (X, T) *be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

THM

1.	A^-	$=$	$A^\circ \cup \partial A$	$=$	$A \cup \partial A$	$=$	$A \cup A'$	$\forall A \in 2^X$
2.	∂A	$=$	$A^- \setminus A^\circ$					$\forall A \in 2^X$
3.	$(A \setminus A^\circ)^\circ$	$=$	$\varnothing$					$\forall A \in 2^X$

✎PROOF:

1. Proof that $A^- = A \cup \partial A$:

(a) lemma: $A \cup \left(A^c\right)^- = X$

$$\begin{aligned}
A \cup \left(A^c\right)^- &\supseteq A \cup A^c && \text{by Lemma 1.1 page 15} \\
&= X && \text{by Theorem A.5 page 278}
\end{aligned}$$

$$A \cup \left(A^c\right)^- \subseteq X$$

[36] Aliprantis and Burkinshaw (1998) pages 59–60, McCarty (1967) page 90 ⟨1 THEOREM⟩, KUBRUSLY (2001) PAGE 116

(b) Proof that $A^- = A \cup \partial A$:

$$
\begin{aligned}
A \cup \partial A &= A \cup \left[A^- \cap \left(A^c\right)^-\right] && \text{by Definition 1.5 page 14}\\
&= [A \cup A^-] \cap \left[A \cup \left(A^c\right)^-\right] && \text{by Theorem A.4 page 277}\\
&= [A \cup A^-] \cap X && \text{by item (1a)}\\
&= A^- \cap X && \text{by Lemma 1.1 page 15}\\
&= A^- && \text{by Theorem A.5 page 278}
\end{aligned}
$$

2. Proof that $A^- = A^\circ \cup \partial A$:

$$
\begin{aligned}
A^\circ \cup \partial A &= A^\circ \cup \left[A^- \cap \left(A^c\right)^-\right] && \text{by definition of } \partial A \text{ (Definition 1.5 page 14)}\\
&= [A^\circ \cup A^-] \cap \left[A^\circ \cup \left(A^c\right)^-\right] && \text{by Theorem A.4 page 277}\\
&= [A^\circ \cup A^-] \cap \left[A^\circ \cup (A^\circ)^c\right] && \text{by Theorem 1.17 page 18}\\
&= [A^\circ \cup A^-] \cap X && \text{by Theorem 1.17 page 18}\\
&= [A^\circ \cup A^-] && \text{by Theorem A.5 page 278}\\
&= A^- && \text{by Lemma 1.1 page 15}
\end{aligned}
$$

3. Proof that $A^- = A \cup A'$:

(a) Proof that $A \cup A' \subseteq A^-$:

$$
\begin{aligned}
A \cup A' &\triangleq A \cup \{x \in X | x \in (A\backslash\{x\})^-\} && \text{by definition of } A'\text{: Definition 1.5 page 14}\\
&\subseteq A \cup \{x \in X | x \in A^-\} && \text{by Theorem 1.16 page 18}\\
&= A \cup A^- \\
&\subseteq A^- \cup A^- && \text{by Lemma 1.1 page 15}\\
&= A^- && \text{by Theorem A.4 page 277}
\end{aligned}
$$

(b) Proof that $A^- \supseteq A \cup A'$:

$$
\begin{aligned}
x \notin A \cup A' &\implies x \notin A \cup \{x \in X | x \in (A\backslash\{x\})^-\} && \text{by definition of } A'\text{: Definition 1.5 page 14}\\
&\implies x \notin \{x \in X | x \in (A\backslash\{x\})^-\}\\
&\implies x \notin \{x \in X | x \in (A)^-\}\\
&\iff x \notin A^-\\
&\implies A^- \subseteq A \cup A'
\end{aligned}
$$

4. Proof that $\partial A = A^-\backslash A^\circ$:

$$
\begin{aligned}
A^-\backslash A^\circ &= A^- \cap (A^\circ)^c && \text{by Theorem A.1 page 267}\\
&= A^- \cap \left[((A^c)^-)^c\right]^c && \text{by Theorem 1.17}\\
&= A^- \cap (A^c)^- && \text{by \textit{idempotent} property (Theorem A.5 page 278)}\\
&= \partial A && \text{by Definition 1.5 page 14}
\end{aligned}
$$

5. Proof that $(A\backslash A^\circ)^\circ = \varnothing$:

$$
\begin{aligned}
(A\backslash A^\circ)^\circ &= \left[A \cap (A^\circ)^c\right]^\circ && \text{by Theorem A.1 page 267}\\
&= \left[\left[\left(A \cap (A^\circ)^c\right)^c\right]^-\right]^c && \text{by Theorem 1.17}\\
&= \left[\left(A^c \cup A^\circ\right)^-\right]^c && \text{by \textit{idempotent} property (Theorem A.5 page 278)}\\
&= \left[\left(A^c\right)^- \cup (A^\circ)^-\right]^c && \text{by Theorem 1.13 page 16}\\
&= \left(\left(A^c\right)^-\right)^c \cap ((A^\circ)^-)^c && \text{by \textit{de Morgan's law} (Theorem A.5 page 278)}\\
&= A^\circ \cap ((A^\circ)^-)^c && \text{by Theorem 1.17}\\
&= \varnothing && \text{because } A^\circ \subseteq (A^\circ)^- \text{ by Lemma 1.1 page 15}
\end{aligned}
$$

Proposition 1.4. [37] *Let* (X, T) *be a* TOPOLOGICAL SPACE (Definition 1.1 page 3). *Let* A° *be the* INTERIOR, ∂A *be the* BOUNDARY, *and* A^e *be the* EXTERIOR *of a set* A.

PRP

$$X = \underbrace{A^\circ \cup \partial A \cup A^e}_{\text{partition of } X} \qquad \forall A \in 2^X$$

✎PROOF:

$$\begin{aligned}
A^\circ \cup \partial A \cup A^e &= A^\circ \cup \left(A^- \cap A^{c-}\right) \cup A^{c\circ} && \text{by Definition 1.5 page 14}\\
&= \left[(A^\circ \cup A^-) \cap \left(A^\circ \cup A^{c-}\right)\right] \cup A^{c\circ} && \text{by Theorem A.4 page 277}\\
&= \left[(A^-) \cap \left(A^\circ \cup A^{c-}\right)\right] \cup A^{c\circ} && \text{because } A^\circ \subseteq A^-\text{: Theorem 1.19 page 21}\\
&= \left[A^- \cap \left(A^\circ \cup A^{\circ c}\right)\right] \cup A^{c\circ} && \text{by Theorem 1.17 page 18}\\
&= [A^- \cap X] \cup A^{c\circ} && \text{by Theorem A.5 page 278}\\
&= A^- \cup A^{c\circ} && \text{by Theorem A.5 page 278}\\
&= A^- \cup A^{-c} && \text{by Theorem 1.17 page 18}\\
&= X && \text{by Theorem A.5 page 278}
\end{aligned}$$

Theorem 1.19. [38] *Let* A^- *be the* CLOSURE, A° *the* INTERIOR, ∂A *the* BOUNDARY, *and* A' *the* DRIVED SET *of a set* A *in a topological space* (X, T). *Let* 2^X *be the* POWER SET *of* X.

THM

1.	$A^\circ \subseteq A \subseteq A^-$									$\forall A \in 2^X$.
2.	$A' \subseteq A^-$									$\forall A \in 2^X$.
3.	A	$=$	A°	$\iff$	A *is* OPEN	$\iff$	$A \cap \partial A$	$=$	$\varnothing$	$\forall A \in 2^X$.
4.	A	$=$	A^-	$\iff$	A *is* CLOSED	$\iff$	$A \cap \partial A$	$=$	$\varnothing$	$\forall A \in 2^X$.
5.	A	$=$	A^-	$\iff$	A *is* CLOSED	$\iff$	A'	$\subseteq$	A	$\forall A \in 2^X$.

✎PROOF:

1. Proof that $A^\circ \subseteq A \subseteq A^-$: by Lemma 1.1 page 15
2. Proof that $A' \subseteq A^-$: by Lemma 1.1 page 15
3. Proof that $A = A^\circ \iff A$ is *open*: by Lemma 1.1 page 15
4. Proof that $A = A^- \iff A$ is *closed*: by Lemma 1.1 page 15
5. Proof that A is *open* $\implies A \cap \partial A = \varnothing$:

$$\begin{aligned}
A \cap \partial A &\triangleq A \cap \left(A^- \cap A^{c-}\right) && \text{by Definition 1.5 page 14}\\
&= A^\circ \cap \left(A^- \cap A^{c-}\right) && \text{by ``}A\text{ is open'' hypothesis}\\
&= (A^\circ \cap A^-) \cap \left(A^\circ \cap A^{c-}\right) && \text{by Theorem A.4 page 277}\\
&= A^\circ \cap \left(A^\circ \cap A^{c-}\right) && \text{by Lemma 1.1}\\
&= A^\circ \cap A^{c-} && \text{by Theorem A.5}\\
&= A^\circ \cap A^{\circ c} && \text{by Theorem 1.17}\\
&= \varnothing && \text{by Theorem A.5}
\end{aligned}$$

[37] Haaser and Sullivan (1991) page 43

[38] Aliprantis and Burkinshaw (1998) page 59, McCarty (1967) page 90 ⟨IV.1 THEOREM⟩, KUBRUSLY (2001) PAGE 116

6. Proof that A is *open* $\Longleftarrow$ $A \cap \partial A = \varnothing$:

$$\begin{aligned}
\varnothing &= A \cap \partial A \\
&\triangleq A \cap \left(A^- \cap A^{c-}\right) && \text{by Definition 1.5 page 14} \\
&= A \cap A^{c-} && \text{by Lemma 1.1 page 15} \\
&= A \cap A^{\circ c} && \text{by Theorem 1.17 page 18} \\
\Longrightarrow \quad A &= A^{\circ} && \text{by Theorem A.5 page 278}
\end{aligned}$$

7. Proof that A is *closed* $\Longrightarrow$ $\partial A \subseteq A$:

$$\begin{aligned}
\partial A &\triangleq A^- \cap A^{c-} && \text{by Definition 1.5 page 14} \\
&= A \cap A^{c-} && \text{by "A is closed" hypothesis and Lemma 1.1 page 15} \\
&\subseteq A
\end{aligned}$$

8. Proof that A is *closed* $\Longleftarrow$ $\partial A \subseteq A$:

$$\begin{aligned}
A^- &= A \cup \partial A && \text{by Theorem 1.18 page 19} \\
&= A && \text{by "$\partial A \subseteq A$ hypothesis} \\
&\Longrightarrow A \text{ is closed} && \text{by Theorem 1.19 page 21}
\end{aligned}$$

9. Proof that $A = A^-$ $\Longrightarrow$ $A' \subseteq A$:

$$\begin{aligned}
A' &\subseteq A^- && \text{by Lemma 1.1 page 15} \\
&= A && \text{by $A = A^-$ hypothesis}
\end{aligned}$$

10. Proof that $A = A^-$ $\Longleftarrow$ $A' \subseteq A$:

$$\begin{aligned}
A^- &= A \cup A' && \text{by Theorem 1.18 page 19} \\
&\subseteq A \cup A && \text{by $A' \subseteq A$ hypothesis} \\
&\subseteq A && \text{by Theorem A.4 page 277}
\end{aligned}$$

⇨

A weakened form of the closure properties of Theorem 1.13 (page 16) can be used to define a topology (next theorem).

Theorem 1.20 (Kuratowski closure axioms). [39] *Let* f *be a set function on* 2^X.

THM

$$\left\{\begin{array}{llllll}
1. & \mathsf{f}(\varnothing) = \varnothing & & & (\text{NORMALIZED}) & \text{and} \\
2. & A \subseteq \mathsf{f}(A) & \forall A \in 2^X & & (\text{EXTENSIVE}) & \text{and} \\
3. & \mathsf{f}(\mathsf{f}(A)) \subseteq \mathsf{f}(A) & \forall A \in 2^X & & & \text{and} \\
4. & \mathsf{f}(A \cup B) = \mathsf{f}(A) \cup \mathsf{f}(B) & \forall A,B \in 2^X & & (\text{ADDITIVE}). &
\end{array}\right\}
\Longrightarrow
\left\{\begin{array}{l}
(X, \boldsymbol{T}(\mathsf{f})) \text{ is a topological space where} \\
\boldsymbol{T}(\mathsf{f}) \triangleq \left\{A \in 2^X | \mathsf{f}(A^c) = A^c\right\}
\end{array}\right\}$$

Lemma 1.2. [40] *Let* $(X, \boldsymbol{T})$ *be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3)*.

LEM

$$\underbrace{\left\{x \in A^-\right\}}_{x \text{ is ADHERENT to } A} \iff \underbrace{\left\{A \cap U \neq \varnothing \quad \forall x \in U \in \boldsymbol{T}\right\}}_{\text{every open set containing } x \text{ meets } A}$$

[39] Thron (1966), pages 42–43, Murdeshwar (1990) pages 45–46

[40] Kubrusly (2001) page 115 ⟨Proposition 3.25⟩

✎PROOF:

1. Proof that $x \in A^- \implies A \cap U \neq \varnothing$:

$$\begin{aligned}
\{x \in U \text{ and } A \cap U = \varnothing\} &\implies \{x \notin U^c \text{ and } A \subseteq U^c\} \\
&\implies A^- \subseteq U^c && \text{by Lemma 1.1 page 15} \\
&\implies x \notin A^- && \text{because } x \in U \iff x \notin U^c \\
&\implies A \cap U \neq \varnothing && \text{because “}x \notin A^-\text{” contradicts “}x \in A^-\text{” hypothesis}
\end{aligned}$$

2. Proof that $x \in A^- \impliedby A \cap U \neq \varnothing$:

$$\begin{aligned}
x \notin A^- &\implies x \in \underbrace{A^{-c}}_{open} && \text{by definition of } A^-\text{: Definition 1.4 page 14} \\
&\implies \varnothing \neq A^- \cap \underbrace{A^{-c}}_{open \text{ set containing } x} && \text{by right hypothesis} \\
&\qquad = \left(A^{-c}\right)^c \cap A^{-c} \\
&\qquad = \varnothing && \text{(contradiction)} \\
&\implies x \in A^-
\end{aligned}$$

⇒

1.3 Supported topological properties

Definition 1.6. [41] *Let (X, T) be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

DEF *A set Y is **dense** in X if $Y^- = X$.*

Definition 1.7. [42] *Let (X, T) be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

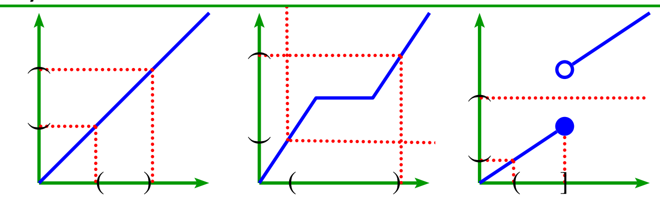
DEF *The set X is **separable** if it contains a* COUNTABLE DENSE *subset.*

Definition 1.8. [43] *Let (X, T_x) and (Y, T_y) be topological spaces. Let f be a function in Y^X.*

DEF
*A function $\mathsf{f} \in Y^X$ is **continuous** if for every* OPEN SET *$U \in T_y$, $\mathsf{f}^{-1}(U)$ is also* OPEN.
*A function is **discontinuous** if it is not* CONTINUOUS.
*The **set of all continuous functions** $C(X, Y)$ in the function space Y^X is*
$$C(X,Y) \triangleq \left\{\mathsf{f} \in Y^X \mid \mathsf{f} \text{ is CONTINUOUS in } X\right\}.$$

Example 1.16.

EX



continuous *continuous* *discontinuous*

---

[41] Murdeshwar (1990) page 248 ⟨2.21 Theorem and Definition⟩, Joshi (1983) page 133 ⟨(5.1.6) Definition⟩
[42] Murdeshwar (1990) page 248 ⟨16.1 Definition⟩, Joshi (1983) page 133 ⟨(6.1.3) Definition⟩
[43] Davis (2005) page 34

Definition 1.8 (previous definition) defines continuity using open sets. Continuity can alternatively be defined using closed sets or closure (next theorem).

Theorem 1.21. [44] *Let $(X, \boldsymbol{T})$ and $(Y, \boldsymbol{S})$ be topological spaces. Let f be a function in Y^X. The following are equivalent:*

THM

1. f *is* CONTINUOUS $\iff$
2. B *is closed in* $(Y, \boldsymbol{S}) \implies \mathsf{f}^{-1}(B)$ *is closed in* $(X, \boldsymbol{T})$ $\quad \forall B \in 2^Y$ $\iff$
3. $\mathsf{f}(A^-) \subseteq \mathsf{f}(A)^-$ $\quad \forall A \in 2^X$ $\iff$
4. $\mathsf{f}^{-1}(B)^- \subseteq \mathsf{f}^{-1}(B^-)$ $\quad \forall B \in 2^Y$

✎PROOF:

1. Proof that (1) $\implies$ (2):

$$\begin{aligned}
B \text{ is } \textit{closed} &\iff B^c \text{ is } \textit{open} && \text{by definition of a } \textit{closed set} \text{ (Definition 1.1 page 3)}\\
&\implies \mathsf{f}^{-1}(B^c) \text{ is } \textit{open} && \text{by (1) and Definition 1.8 page 23}\\
&\iff \left[\mathsf{f}^{-1}(B)\right]^c \text{ is } \textit{open} && \text{because } \mathsf{f}^{-1}(B^c) = \left[\mathsf{f}^{-1}(B)\right]^c\\
&\iff \mathsf{f}^{-1}(B) \text{ is } \textit{closed} && \text{by definition of a } \textit{closed set} \text{ (Definition 1.1 page 3)}
\end{aligned}$$

2. Proof that (2) $\implies$ (3):

   (a) lemma: Proof that $\mathsf{f}^{-1}[\mathsf{f}(A)^-]$ is *closed*:

$$\begin{aligned}
&\mathsf{f}(A)^- \text{ is } \textit{closed} && \text{by Proposition 1.3 page 14}\\
&\implies \mathsf{f}^{-1}[\mathsf{f}(A)^-] \text{ is } \textit{closed} && \text{by (2)}
\end{aligned}$$

   (b) lemma: Proof that $A \subseteq \mathsf{f}^{-1}[\mathsf{f}(A)^-]$:

$$\begin{aligned}
A &\subseteq \mathsf{f}^{-1}[\mathsf{f}(A)] && \text{by result from function theory}\\
&\subseteq \mathsf{f}^{-1}[\mathsf{f}(A)^-] && \text{by Lemma 1.1 page 15}
\end{aligned}$$

   (c) Proof that (2) $\implies$ (3):

$$\begin{aligned}
\mathsf{f}(A^-) &\subseteq \mathsf{f}\left(\mathsf{f}^{-1}[\mathsf{f}(A)^-]^-\right) && \text{by item (2b)}\\
&= \mathsf{f}\left(\mathsf{f}^{-1}[\mathsf{f}(A)^-]\right) && \text{by item (2b)}\\
&\subseteq \mathsf{f}(A)^- && \text{by result from function theory}
\end{aligned}$$

3. Proof that (3) $\implies$ (4):

$$\begin{aligned}
\mathsf{f}^{-1}(B)^- &\subseteq \mathsf{f}^{-1}\mathsf{f}\left[\mathsf{f}^{-1}(B)^-\right] && \text{by result from function theory}\\
&\subseteq \mathsf{f}^{-1}\left(\left[\mathsf{f}\mathsf{f}^{-1}(B)\right]^-\right) && \text{by result from function theory}\\
&\subseteq \mathsf{f}^{-1}(B^-) && \text{by result from function theory}
\end{aligned}$$

[44] McCarty (1967) pages 91–92 ⟨IV.2 THEOREM⟩

4. Proof that (4) $\implies$ (1):

$$\begin{array}{rll}
U \text{ is open} & \implies U^c \text{ is closed} & \text{by Definition 1.1 page 3} \\
& \implies \mathsf{f}^{-1}(U^c) = \mathsf{f}^{-1}(U^{c-}) & \text{by Lemma 1.1 page 15} \\
& \qquad\qquad \supseteq \mathsf{f}^{-1}(U^c)^- & \text{by (4)} \\
& \qquad\qquad \supseteq \mathsf{f}^{-1}(U^c) & \text{by Lemma 1.1 page 15} \\
& \implies \mathsf{f}^{-1}(U^c) = \mathsf{f}^{-1}(U^c)^- & \\
& \iff \mathsf{f}^{-1}(U^c) \text{ is closed} & \text{by Lemma 1.1 page 15} \\
& \iff \left[\mathsf{f}^{-1}(U^c)\right]^c \text{ is open} & \text{by Definition 1.1 page 3} \\
& \iff \mathsf{f}^{-1}(U) \text{ is open} & \text{because } \mathsf{f}^{-1}(U) = \left[\mathsf{f}^{-1}(U^c)\right]^c \\
& \implies \mathsf{f} \text{ is continuous} & \text{by Definition 1.8 page 23}
\end{array}$$

✎

1.4 Neighborhoods

Definition 1.9. [45] *Let (X, T) be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3). Let A° be the* INTERIOR *of a set A (Definition 1.4 page 14).*

DEF

*A set $N_x \in 2^X$ is a **neighborhood** of an element $x \in X$ if*
$x \in N_x{}^\circ$.
*A set N_x is an **open neighborhood** of an element $x \in X$ if*
N_x *is a neighborhood of x and* $N_x \in T$.

Proposition 1.5. [46] *Let (X, T) be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

PRP

$$\left\{ \begin{array}{l}\text{A set } N_x \text{ is a } \textbf{neighborhood} \\ \text{of an element } x \in X\end{array} \right\} \iff \left\{\exists U \in T \quad \text{such that} \quad x \in U \subseteq N_x\right\}$$

Example 1.17. Example 1.6 (page 7) lists the 29 topologies on a set $X \triangleq \{x, y, z\}$. These topologies are listed next along with their open and closed neighborhoods of the element $x \in X$:

topologies on $\{x,y,z\}$	open nbhds. of x	not open nbhds.
$T_{00} = \{\emptyset, X\}$	X	
$T_{01} = \{\emptyset, \{x\}, X\}$	$\{x\}, X$	$\{x,y\}, \{x,z\}$
$T_{02} = \{\emptyset, \{y\}, X\}$	X	$\{x,y\}$
$T_{04} = \{\emptyset, \{z\}, X\}$	X	$\{x,z\}$
$T_{10} = \{\emptyset, \{x,y\}, X\}$	$\{x,y\}, X$	
$T_{20} = \{\emptyset, \{x,z\}, X\}$	$\{x,z\}, X$	
$T_{40} = \{\emptyset, \{y,z\}, X\}$	X	
$T_{11} = \{\emptyset, \{x\}, \{x,y\}, X\}$	$\{x\}, \{x,y\}, X$	$\{x,z\}$
$T_{21} = \{\emptyset, \{x\}, \{x,z\}, X\}$	$\{x\}, \{x,z\}, X$	$\{x,y\}$
$T_{41} = \{\emptyset, \{x\}, \{y,z\}, X\}$	$\{x\}, X$	$\{x,y\}, \{x,z\}$
$T_{12} = \{\emptyset, \{y\}, \{x,y\}, X\}$	$\{x,y\}, X$	
$T_{22} = \{\emptyset, \{y\}, \{x,z\}, X\}$	$\{x,z\}, X$	$\{x,y\}$
$T_{42} = \{\emptyset, \{y\}, \{y,z\}, X\}$	X	$\{x,y\}$
$T_{14} = \{\emptyset, \{z\}, \{x,y\}, X\}$	$\{x,y\}, X$	$\{x,z\}$
$T_{24} = \{\emptyset, \{z\}, \{x,z\}, X\}$	$\{x,z\}, X$	
$T_{44} = \{\emptyset, \{z\}, \{y,z\}, X\}$	X	$\{x,z\}$
$T_{31} = \{\emptyset, \{x\}, \{x,y\}, \{x,z\}, X\}$	$\{x\}, \{x,y\}, \{x,z\}, X$	

[45] Murdeshwar (1990) page 88 ⟨3.1 Definition⟩, Davis (2005) page 43 ⟨Definition 4.7⟩
[46] Murdeshwar (1990) page 88 ⟨3.1 Definition⟩,

$\boldsymbol{T}_{52} = \{\emptyset, \{y\}, \{x,y\}, \{x,z\}, X\}$	$\{x,y\}, \{x,z\}, X$	
$\boldsymbol{T}_{64} = \{\emptyset, \{z\}, \{x,z\}, \{y,z\}, X\}$	$\{x,z\}, X$	
$\boldsymbol{T}_{13} = \{\emptyset, \{x\}, \{y\}, \{x,y\}, X\}$	$\{x\}, \{x,y\}, X$	$\{x,z\}$
$\boldsymbol{T}_{25} = \{\emptyset, \{x\}, \{z\}, \{x,z\}, X\}$	$\{x\}, \{x,z\}, X$	$\{x,y\}$
$\boldsymbol{T}_{46} = \{\emptyset, \{y\}, \{z\}, \{y,z\}, X\}$	X	$\{x,y\}, \{x,z\}$
$\boldsymbol{T}_{33} = \{\emptyset, \{x\}, \{y\}, \{x,y\}, \{x,z\}, X\}$	$\{x\}, \{x,y\}, \{x,z\}, X$	
$\boldsymbol{T}_{53} = \{\emptyset, \{x\}, \{y\}, \{x,y\}, \{y,z\}, X\}$	$\{x\}, \{x,y\}, X$	$\{x,z\}$
$\boldsymbol{T}_{35} = \{\emptyset, \{x\}, \{z\}, \{x,y\}, \{x,z\}, X\}$	$\{x\}, \{x,y\}, X$	$\{x,z\}$
$\boldsymbol{T}_{65} = \{\emptyset, \{x\}, \{z\}, \{x,z\}, \{y,z\}, X\}$	$\{x\}, \{x,z\}, X$	$\{x,y\}$
$\boldsymbol{T}_{56} = \{\emptyset, \{y\}, \{z\}, \{x,y\}, \{y,z\}, X\}$	$\{x,y\}, X$	$\{x,z\}$
$\boldsymbol{T}_{66} = \{\emptyset, \{y\}, \{z\}, \{x,z\}, \{y,z\}, X\}$	$\{x,z\}, X$	$\{x,y\}$
$\boldsymbol{T}_{77} = \{\emptyset, \{x\}, \{y\}, \{z\}, \{x,y\}, \{x,z\}, \{y,z\}, X\}$	$\{x\}, \{x,y\}, \{x,z\}, X$	

Definition 1.10. [47] *Let* $(X, \boldsymbol{T})$ *be a* TOPOLOGICAL SPACE (Definition 1.1 page 3).

DEF

*A set $Y \subseteq X$ is **disconnected** if there exists $A, B \subseteq X$ such that*

1. $A \cup B = Y$ *and*
2. $A \cap B = \emptyset$.

*In this case, Y is said to be **disconnected** by the sets A and B, and the pair A, B is a **separation** of Y.*
*If a set is not disconnected, then it is **connected**.*

Definition 1.11. [48]

DEF

A TOPOLOGICAL SPACE *is a **Hausdorff space** if*

$$\forall x, y \in X, \ \exists N \in \mathbf{N}x \text{ and } M \in \mathbf{N}y \quad \text{such that} \quad N \cap M = \emptyset.$$

Definition 1.12. [49] *Let* $(X, \boldsymbol{T})$ *be a topological space and* $A, B, \{A_i\}, \{B_i\}, \{M_i\} \subseteq X$.

DEF

*A sequence $(A_i)_{i \in I}$ is a **cover** of a set A in the topological space $(X, \boldsymbol{T})$ if*

$$A \subseteq \bigcup_{i \in I} A_i.$$

*A sequence $(B_i)_{i \in J}$ is a **subcover** of set A with respect to a cover $(A_i)_{i \in I}$ if*

$$\{B_i\}_{i \in J} \subsetneq \{A_i\}_{i \in I}.$$

*A sequence $(M_i)_{i \in K}$ is a **minimal cover** of A if $(M_i)_{i \in K}$ is a cover and $(M_i)_{i \in K \setminus \{n\}}$ is not a cover.*
*A cover $(A_i)_{i \in I}$ is a **proper cover** of A if A is not a member.*
*A cover $(A_i)_{i \in I}$ is a **open cover** of A if it consists entirely of open sets.*

Definition 1.13. [50]

DEF

*A set $A \subseteq X$ is **compact** in the topological space $(X, \boldsymbol{T})$ if any open cover of A has a finite subcover.*

[47] Munkres (2000) page 148 ⟨§Connected Spaces⟩, Dieudonné (1969) page 67, Carothers (2000) page 78
[48] Aliprantis and Burkinshaw (1998) page 60, Hausdorff (1914)
[49] Aliprantis and Burkinshaw (1998) page 48
[50] Aliprantis and Burkinshaw (1998) page 62

CHAPTER 2

METRIC SPACES

2.1 Algebraic structure

> ❝ *The Epicureans are wont to ridicule this theorem, saying it is evident even to an ass and needs no proof; it is as much the mark of an ignorant man, they say, to require persuasion of evident truths as to believe what is obscure without question. …That the present theorem is known to an ass they make out from the observation that, if straw is placed at one extremity of the sides, an ass in quest of provender will make his way along the side and not by way of the two others.* ❞
>
> Proclus Lycaeus (412 – 485 AD), Greek philosopher, commenting on the Epicureans opinion regarding the *triangle inequality* property. [1]

A *metric space* is simply a set together with a "*distance*" function, which is called the *metric* of the *metric space* (Definition 2.1 page 27) (next definition). With a metric on a set, we can measure the distance between points in the set.

Definition 2.1. [2] *Let X be a set and $\mathbb{R}^{\vdash}$ the set of non-negative real numbers.*

DEF

A function $\mathsf{d} \in \mathbb{R}^{\vdash X \times X}$ *is a **metric** on X if*

1. $\mathsf{d}(x,y) \geq 0$ $\quad \forall x,y \in X$ (NON-NEGATIVE) *and*
2. $\mathsf{d}(x,y) = 0 \iff x = y$ $\quad \forall x,y \in X$ (NONDEGENERATE) *and*
3. $\mathsf{d}(x,y) = \mathsf{d}(y,x)$ $\quad \forall x,y \in X$ (SYMMETRIC) *and*
4. $\mathsf{d}(x,y) \leq \mathsf{d}(x,z) + \mathsf{d}(z,y)$ $\quad \forall x,y,z \in X$ (SUBADDITIVE/TRIANGLE INEQUALITY).[3]

*A **metric space** is the pair* (X, d).

Actually, it is possible to significantly simplify the definition of a metric to an equivalent statement requiring only half as many conditions. These equivalent conditions (a "*characterization*") are stated in Theorem 2.1 (next).

---

[1] Lycaeus (circa 450), page 251

[2] Dieudonné (1969), page 28, Copson (1968), page 21, Hausdorff (1937) page 109, Fréchet (1928), Fréchet (1906) page 30

[3] Euclid (circa 300BC) ⟨Book I Proposition 20⟩

Theorem 2.1 (metric characterization). [4] *Let* d *be a function in* $(\mathbb{R}^{\vdash})^{X\times X}$.

THM

$\mathsf{d}(x,y)$ *is a metric* $\iff \begin{cases} 1. & \mathsf{d}(x,y) = 0 \iff x=y & \forall x,y\in X \quad \text{and} \\ 2. & \mathsf{d}(x,y) \le \mathsf{d}(z,x)+\mathsf{d}(z,y) & \forall x,y,z\in X \end{cases}$

✎PROOF:

1. Proof that $[\mathsf{d}(x,y)$ is a metric$] \implies [(1)$ and $(2)]$:

   1a. Proof that $\mathsf{d}(x,y)=0 \iff x=y$: by left hypothesis 2 ($\mathsf{d}(x,y)$ is *nondegenerate*)

   1b. Proof that $\mathsf{d}(x,y) \le \mathsf{d}(z,x)+\mathsf{d}(z,y)$:

$$\begin{aligned} \mathsf{d}(x,y) &\le \mathsf{d}(x,z)+\mathsf{d}(z,y) && \text{by right hypothesis 4 (triangle inequality)} \\ &= \mathsf{d}(z,x)+\mathsf{d}(z,y) && \text{by right hypothesis 3 (commutative)} \end{aligned}$$

2. Proof that $[\mathsf{d}(x,y)$ is a metric$] \impliedby [(1)$ and $(2)]$:

   2a. Proof that $\mathsf{d}(x,y) \ge 0$:

$$\begin{aligned} 0 &= \frac{1}{2}\cdot 0 \\ &= \frac{1}{2}\,\mathsf{d}(y,y) && \text{by right hypothesis 1} \\ &= \frac{1}{2}\,\mathsf{d}(y,z)\Big|_{z=y} \\ &\le \frac{1}{2}\,[\mathsf{d}(x,y)+\mathsf{d}(x,z)]_{z=y} && \text{by right hypothesis 2} \\ &= \frac{1}{2}\,[\mathsf{d}(x,y)+\mathsf{d}(x,y)] \\ &= \mathsf{d}(x,y) \end{aligned}$$

   2b. Proof that $\mathsf{d}(x,y)=0 \iff x=y$: by right hypothesis 1

   2c. Proof that $\mathsf{d}(x,y)=\mathsf{d}(y,x)$:

$$\begin{aligned} \mathsf{d}(x,y)|_{z=y} &\le [\mathsf{d}(z,x)+\mathsf{d}(z,y)]_{z=y} && \text{by right hypothesis 2} \\ &= \mathsf{d}(y,x)+\cancelto{0}{\mathsf{d}(y,y)} \\ &= \mathsf{d}(y,x) && \text{by right hypothesis 1} \\ \mathsf{d}(y,x)|_{z=x} &\le [\mathsf{d}(z,y)+\mathsf{d}(z,x)]_{z=x} && \text{by right hypothesis 2} \\ &= \mathsf{d}(x,y)+\cancelto{0}{\mathsf{d}(x,x)} \\ &= \mathsf{d}(x,y) && \text{by right hypothesis 1} \end{aligned}$$

   2d. Proof that $\mathsf{d}(x,y) \le \mathsf{d}(x,z)+\mathsf{d}(z,y)$:

$$\begin{aligned} \mathsf{d}(x,y) &\le \mathsf{d}(z,x)+\mathsf{d}(z,y) && \text{by right hypothesis 2} \\ &= \mathsf{d}(x,z)+\mathsf{d}(z,y) && \text{by result 2c} \end{aligned}$$

⇨

The *triangle inequality* property stated in the definition of metrics (Definition 2.1 page 27) axiomatically endows a metric with an upper bound. Lemma 2.1 (next) demonstrates that there is a complementary lower bound similar in form to the triangle-inequality upper bound.

[4] Busemann (1955a) page 3, Michel and Herget (1993), page 264, Giles (1987) page 18

Lemma 2.1. [5] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27). Let* $|\cdot|$ *be the* ABSOLUTE VALUE *function (Definition F.4 page 346).*

LEM

1. $|\mathsf{d}(x,p) - \mathsf{d}(p,y)| \le \mathsf{d}(x,y) \quad \forall x,y,p \in X$
2. $\mathsf{d}(x,p) - \mathsf{d}(p,y) \le \mathsf{d}(x,y) \quad \forall x,y,p \in X$

✎PROOF:

1. Proof that $|\mathsf{d}(x,p) - \mathsf{d}(p,y)| \le \mathsf{d}(x,y)$:

$$\begin{aligned}
|\mathsf{d}(x,p) - \mathsf{d}(p,y)| &\le |\mathsf{d}(x,y) + \mathsf{d}(y,p) - \mathsf{d}(p,y)| && \text{by } \textit{subadditive} \text{ property (Definition 2.1 page 27)}\\
&= |\mathsf{d}(x,y) + \mathsf{d}(p,y) - \mathsf{d}(p,y)| && \text{by } \textit{symmetry} \text{ property of metrics (Definition 2.1 page 27)}\\
&= |\mathsf{d}(x,y) + 0|\\
&= \mathsf{d}(x,y) && \text{by } \textit{non-negative} \text{ property of metrics (Definition 2.1 page 27)}
\end{aligned}$$

2. Proof that $\mathsf{d}(x,p) \ge \mathsf{d}(p,y) \implies \mathsf{d}(x,p) - \mathsf{d}(p,y) \le \mathsf{d}(x,y)$:

$$\begin{aligned}
\mathsf{d}(x,p) - \mathsf{d}(p,y) &= |\mathsf{d}(x,p) - \mathsf{d}(p,y)| && \text{by left hypothesis and definition of } |\cdot|\\
&\le \mathsf{d}(x,y) && \text{by item (1)}
\end{aligned}$$

3. Proof that $\mathsf{d}(x,p) \le \mathsf{d}(p,y) \implies \mathsf{d}(x,p) - \mathsf{d}(p,y) \le \mathsf{d}(x,y)$:

$$\begin{aligned}
|\mathsf{d}(x,p) - \mathsf{d}(p,y)| &\le 0 && \text{by left hypothesis}\\
&\le \mathsf{d}(x,y) && \text{by } \textit{non-negative} \text{ property of metrics (Definition 2.1 page 27)}
\end{aligned}$$

✏

The *triangle inequality* property stated in the definition of metrics (Definition 2.1 page 27) can be extended from two to any finite number of metrics (next).

Proposition 2.1. [6] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27) and* $(\!(x_n \in X)\!)_1^N$ *an* N-TUPLE *(Definition 8.1 page 127) on* X.

PRP

$$\mathsf{d}(x_1, x_N) \le \sum_{n=1}^{N-1} \mathsf{d}(x_n, x_{n+1}) \qquad \forall N \in \mathbb{N} \setminus 1$$

✎PROOF: Proof by induction.

Proof that the $\{N = 2 \text{ case}\}$ is true:

$$\mathsf{d}(x_1, x_2) \le \sum_{n=1}^{2-1} \mathsf{d}(x_n, x_{n+1})$$

Proof for that the $\{N \text{ case}\} \implies \{N+1 \text{ case}\}$:

$$\begin{aligned}
\mathsf{d}(x_1, x_{N+1}) &\le \mathsf{d}(x_1, x_N) + \mathsf{d}(x_N, x_{N+1}) && \text{by } \textit{subadditive} \text{ property (Definition 2.1 page 27)}\\
&\le \left(\sum_{n=1}^{N-1} \mathsf{d}(x_N, x_{N+1}) \right) + \mathsf{d}(x_N, x_{N+1}) && \text{by } \{N \text{ case}\} \text{ hypothesis}\\
&= \sum_{n=1}^{N} \mathsf{d}(x_n, x_{n+1})
\end{aligned}$$

[5] Dieudonné (1969), page 28, Michel and Herget (1993), page 266
[6] Dieudonné (1969), page 28, Rosenlicht (1968) page 37

Section 2.3.1 (page 32) presents some topological properties of metric spaces. However the property of *boundedness* (next definition) is fundamentally a metric space concept, not a topological one.[7]

Definition 2.2. [8] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27).*

DEF: *A set A is* ***bounded*** *in* (X, d) *if* $A \subseteq X$ *and*
$$\sup\{\mathsf{d}(x, y) \mid x, y \in A\} < \infty$$

In a *metric space* (Definition 2.1 page 27), it is sometimes useful to know the maximum distance between any two points in the set. This maximum distance is called the *diameter* of the set (Definition 2.3, next definition). The *diameter* is an example of a broader class of functions called *set functions*.[9]

Definition 2.3. [10] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27).*

DEF: *The* ***diameter*** *of a set* $A \subseteq X$ *in* (X, d) *is*
$$\operatorname{diam} A \triangleq \begin{cases} 0 & \text{for } A = \varnothing \\ \sup\{\mathsf{d}(x, y) \mid x, y \in A\} & \text{otherwise} \end{cases}$$

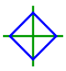
2.2 Open and closed balls

Definition 2.4 (next) defines the *open ball*. In a *metric space* (Definition 2.1 page 27), sets are often specified in terms of an *open ball*; and an open ball is specified in terms of a metric.

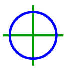
Definition 2.4. [11] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27).*

DEF:
An ***open ball*** *centered at x with radius r is the set* $\mathsf{B}(x, r) \triangleq \{y \in X \mid \mathsf{d}(x, y) < r\}$.
A ***closed ball*** *centered at x with radius r is the set* $\overline{\mathsf{B}}(x, r) \triangleq \{y \in X \mid \mathsf{d}(x, y) \le r\}$.

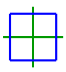
Open balls will often "appear" different in different metric spaces. Some examples include the following (Example 2.1 page 37):



taxi-cab metric



Euclidean metric



sup metric

Lemma 2.2 (next) demonstrates that every point in an open ball is contained in an open ball that is contained in the original open ball (see Figure 2.2 page 31 for an illustration).

Lemma 2.2. [12] *Let* B *be an* OPEN BALL *(Definition 2.4 page 30) in a* METRIC SPACE (X, d).

LEM: $p \in \mathsf{B}(x, r) \iff \exists r_p \text{ such that } \mathsf{B}(p, r_p) \subseteq \mathsf{B}(x, r)$

[7] Munkres (2000), page 121
[8] Thron (1966), page 154 ⟨definition 19.5⟩, Bruckner et al. (1997) page 356
[9] Pap (1995) page 7, Hahn and Rosenthal (1948), Choquet (1954)
[10] Michel and Herget (1993), page 267, Copson (1968), page 23, Molchanov (2005) page 389, Hausdorff (1937), page 166
[11] Aliprantis and Burkinshaw (1998), page 35
[12] Rosenlicht (1968) pages 40–41, Aliprantis and Burkinshaw (1998) page 35

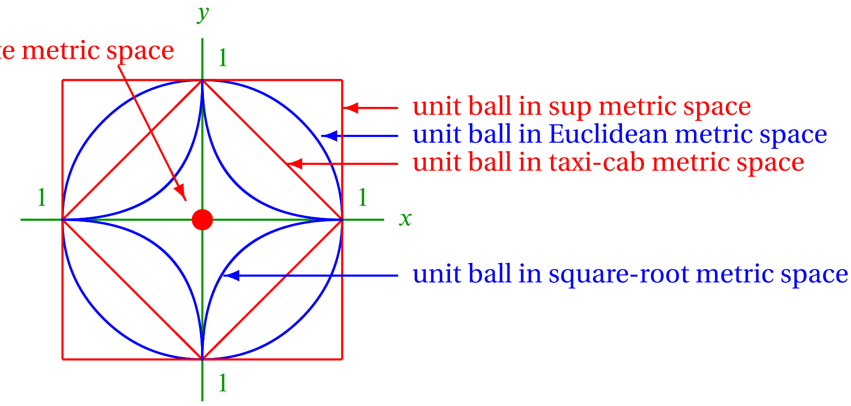


Figure 2.1: Balls on the set $\mathbb{R}^2$ using assorted metrics

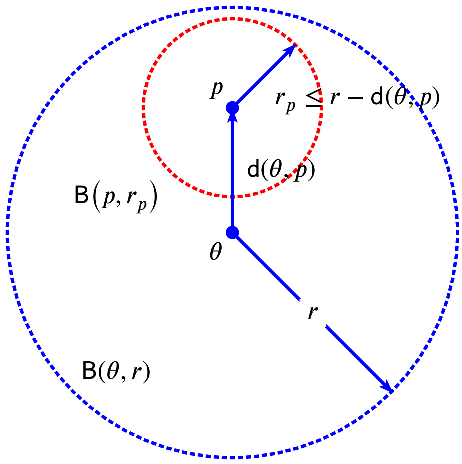


Figure 2.2: Every point in an open ball is contained in an open ball that is contained in the original open ball (Lemma 2.2 page 30)

✎PROOF:

1. lemma: Proof that $p \in \mathsf{B}(x, r) \implies \exists r_p \in \mathbb{R}^+$ such that $r_p < r - \mathsf{d}(\theta, p)$:

$$\begin{aligned}
p \in \mathsf{B}(x,r) &\iff \mathsf{d}(\theta,p) < r && \text{by definition of } \textit{open ball} \text{ (Definition 2.4 page 30)}\\
&\iff 0 < r - \mathsf{d}(\theta,p) && \text{by property of real numbers}\\
&\implies \exists r_p \in \mathbb{R}^+ \text{ such that } 0 < r_p < r - \mathsf{d}(\theta,p) && \text{by property of real numbers}
\end{aligned}$$

2. Proof for ($\implies$) case:

$$\begin{aligned}
\mathsf{B}(p, r_p) &\triangleq \{x \in X | \mathsf{d}(p,x) < r_p \in \mathbb{R}^+\} && \text{by definition of } \textit{open ball} \text{ (Definition 2.4 page 30)}\\
&\subseteq \{x \in X | \mathsf{d}(p,x) < r - \mathsf{d}(\theta,p)\} && \text{by left hypothesis and item (1)}\\
&= \{x \in X | \mathsf{d}(p,x) + \mathsf{d}(\theta,p) < r\} && \text{by property of real numbers}\\
&= \{x \in X | \mathsf{d}(\theta,p) + \mathsf{d}(p,x) < r\} && \text{by } \textit{symmetry} \text{ of metrics (Definition 2.1 page 27)}\\
&\subseteq \{x \in X | \mathsf{d}(\theta,x) < r\} && \text{by } \textit{subadditive} \text{ property (Definition 2.1 page 27),}\\
& && \mathsf{d}(\theta,x) \leq \mathsf{d}(\theta,p) + \mathsf{d}(p,x)
\end{aligned}$$

3. Proof for ($\Longleftarrow$) case:

$$\begin{aligned}
p = \{x \in X | \mathsf{d}(p,x) = 0\} & \quad \text{by } \textit{nondegenerate} \text{ property of metrics: Definition 2.1 page 27} \\
\subseteq \left\{x \in X | \mathsf{d}(p,x) < r_p \in \mathbb{R}^+\right\} & \quad \text{because } 0 < r_p \\
\triangleq \mathsf{B}\left(p, r_p\right) & \quad \text{by definition of } \textit{open ball} \text{ (Definition 2.4 page 30)} \\
\subseteq \mathsf{B}(x,r) & \quad \text{by right hypothesis}
\end{aligned}$$

✏

2.3 Topological structure

2.3.1 Topologies induced by metrics

Theorem 2.2 (page 32) shows that in a *metric space* (Definition 2.1 page 27) (X, d), the metric d always induces a topology $\boldsymbol{T}$ on X. The set X together with topology $\boldsymbol{T}$ is a *topological space*. More specifically, the set of *open balls* in a metric space form a *base* for a *topological space*. Therefore, *every metric space* (Definition 2.1 page 27) *is* a topological space, and everything that is true of a topological space is also true for all *metric space*s.

Theorem 2.2. *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27).*

THM: *The set of all* OPEN BALLS *in* (X, d) *is a* BASE *for the topological space* $(X, \boldsymbol{T})$ *where*
$$\boldsymbol{T} \triangleq \left\{U \in 2^X | U \text{ is the union of balls in } (X, \mathsf{d})\right\}.$$

✎PROOF:

1. The set of all *open balls* in (X, d) is a *base* for $(X, \boldsymbol{T})$ by Lemma 2.2 (page 30) and Theorem 1.4 (page 8).
2. $\boldsymbol{T}$ is a topology on X by Definition 1.2 (page 8).

✏

2.3.2 Open and closed sets

Corollary 2.1 (next) identifies four fundamental properties of open sets in metric spaces. These properties are the same as those defining a topology (Definition 1.1 page 3).

Corollary 2.1. *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27).*

COR:
1. X *is* OPEN.
2. $\varnothing$ *is* OPEN.
3. $\left\{U_n \middle| n=1,2,\ldots,N\right\}$ *are* OPEN $\Longrightarrow \bigcap_{n=1}^{N} U_n$ *is* OPEN.
4. $\left\{U_\gamma \in 2^X | \gamma \in \mathbb{R}\right\}$ *are* OPEN $\Longrightarrow \bigcup_{\gamma \in \Gamma} U_\gamma$ *is* OPEN.

✎PROOF:

1. The *metric space* (X, d) is a *topological space* by Theorem 2.2 (page 32).
2. The four properties are true for any topological space by Theorem 1.14 (page 17).

⇨

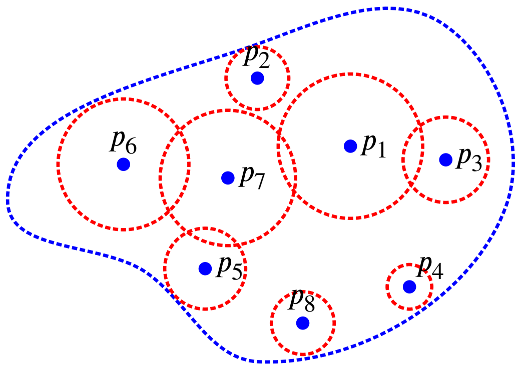


Figure 2.3: Every point in an open set is contained in an open ball that is contained in the original open set (Lemma 2.3 page 33)

Lemma 2.3 (next) demonstrates that every point in an open set is contained in an open ball that is contained in the original open set (see also Figure 2.3 page 33).

Lemma 2.3. *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27).*

LEM

$$\left\{U \in 2^X \text{ is } \textbf{open} \text{ in } (X, \mathsf{d})\right\} \iff \left\{\begin{array}{l} \forall x \in U,\ \exists r \in \mathbb{R}^+ \quad \text{such that} \\ \mathsf{B}(x, r) \subseteq U \end{array}\right\}$$

✎PROOF:

1. Proof for ($\Longrightarrow$) case:

$$\begin{aligned} U &= \bigcup\left\{\mathsf{B}(x_\gamma, r_\gamma) \,|\, \mathsf{B}(x_\gamma, r_\gamma) \subseteq U\right\} && \text{by left hypothesis and Theorem 2.2 page 32} \\ &\supseteq \mathsf{B}(x, r) && \text{because } x \text{ must be in one of those balls in } U \end{aligned}$$

2. Proof for ($\Longleftarrow$) case:

$$\begin{aligned} U &= \bigcup\{x \in X \,|\, x \in U\} && \text{by definition of union operation } \bigcup \\ &= \bigcup\left\{\mathsf{B}(x, r) \,|\, x \in U \text{ and } \mathsf{B}(x, r) \subseteq U\right\} && \text{by right hypothesis} \\ &\Longrightarrow U \text{ is open} && \text{by Theorem 2.2 page 32 and Corollary 2.1 page 32} \end{aligned}$$

⇨

Corollary 2.2. *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27). Let* N *be a finite* WHOLE NUMBER.

COR

1. X *is* CLOSED.
2. $\varnothing$ *is* CLOSED.
3. $\left\{D_\gamma \in 2^X \,|\, \gamma \in \mathbb{R}\right\}$ *are* CLOSED $\Longrightarrow \bigcap_{\gamma \in \mathbb{R}} D_\gamma$ *is* CLOSED.
4. $\left\{D_n \in 2^X \,|\, n = 1, 2, \ldots, N\right\}$ *are* CLOSED $\Longrightarrow \bigcup_{n=1}^{N} D_n$ *is* CLOSED.

✎PROOF:

1. (X, d) is a *topological space* by Theorem 2.2 page 32.
2. The four properties are true of all topological spaces by Theorem 1.3 page 6.

⇨

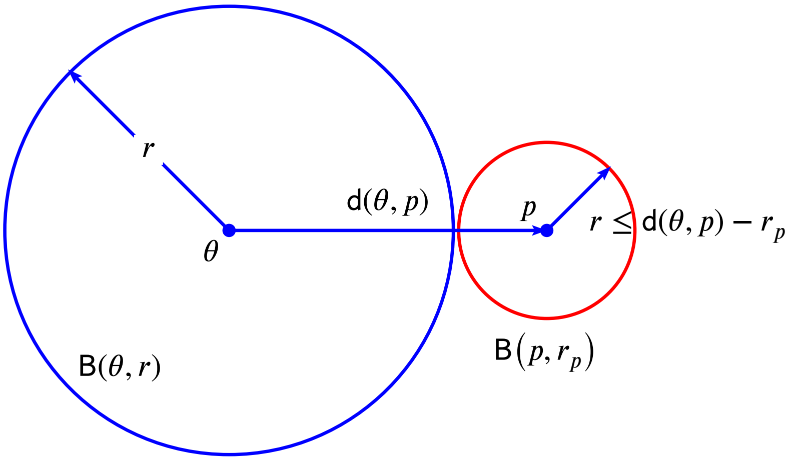


Figure 2.4: Every closed ball is a closed set (Proposition 2.2 page 34)

Proposition 2.2. [13] *Let* B *be an* OPEN BALL *and* $\overline{\mathsf{B}}(\cdot,\cdot)$ *a* CLOSED BALL *(Definition 2.4 page 30) in a metric space* (X, d).

PRP					
Every OPEN BALL	$\mathsf{B}(x, r)$	*in* (X, d)	*is* OPEN	$\forall x \in X$ *and* $\forall r \in \mathbb{R}^+$.
Every CLOSED BALL	$\overline{\mathsf{B}}(x, r)$	*in* (X, d)	*is* CLOSED	$\forall x \in X$ *and* $\forall r \in \mathbb{R}^+$.

✎PROOF:

1. Proof that every open ball is open:

The union of any set of open balls is open — by Theorem 2.2 page 32
$\implies$ the union of a set of just one open ball is open
$\implies$ every open ball is open.

2. lemma: $p \in \left(\overline{\mathsf{B}}(x, r)\right)^{\mathsf{c}} \implies r \leq \mathsf{d}(\theta, p) - r_p$:

[13] Rosenlicht (1968) pages 40–41, Aliprantis and Burkinshaw (1998) page 35

3. Proof that every closed ball is closed (see Figure 2.4 page 34 for illustration):

$$\begin{aligned}
\left(\overline{\mathsf{B}}(x,r)\right)^{\mathsf{c}} &\triangleq \{x \in X | \mathsf{d}(\theta,x) \le r\}^{\mathsf{c}} && \text{by definition of \textit{closed ball} (Definition 2.4 page 30)}\\
&= \{x \in X | \mathsf{d}(\theta,x) > r\} && \text{by definition of set complement}\\
&\supseteq \left\{x \in X | \mathsf{d}(\theta,x) > \mathsf{d}(\theta,p) - r_p\right\} && \text{by item (2)}\\
&= \left\{x \in X | \mathsf{d}(\theta,x) - \mathsf{d}(\theta,p) > -r_p\right\} && \text{by property of real numbers}\\
&= \left\{x \in X | \mathsf{d}(\theta,p) - \mathsf{d}(\theta,x) < r_p\right\} && \text{by property of real numbers}\\
&= \left\{x \in X | \mathsf{d}(p,\theta) - \mathsf{d}(\theta,x) < r_p\right\} && \text{by \textit{symmetric} property of metrics (Definition 2.1 page 27)}\\
&\supseteq \left\{x \in X | \mathsf{d}(p,x) < r_p\right\} && \text{by Lemma 2.1 page 29}\\
&\triangleq \mathsf{B}(p, r_p) && \text{by definition of \textit{open ball} (Definition 2.4 page 30)}\\
&\iff \left(\overline{\mathsf{B}}(\theta,r)\right)^{\mathsf{c}} \text{ is \textit{open}} && \text{by Lemma 2.3 page 33}\\
&\iff \overline{\mathsf{B}}(\theta,r) \text{ is \textit{closed}} && \text{by definition of \textit{closed set} (Definition 1.1 page 3)}
\end{aligned}$$

⇰

In a metric space, all finite sets are *closed* (Proposition 2.3, next). This is *not* in general true for a topological space (Counterexample 2.1 page 36).

Proposition 2.3. [14] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27).*

PRP			
1.	$\{x \in X\}$	*(a single element set)*	*is* CLOSED *in* (X, d).
2.	$\{x_1, x_2, \ldots, x_n \in X\}$	*(set with finite number of elements)*	*is* CLOSED *in* (X, d).
3.	$\{x \in X \vert \mathsf{d}(\theta, x) = r,\ \theta \in X\}$	*(ring centered at θ with radius r)*	*is* CLOSED *in* (X, d).

✎PROOF:

1. Proof that any single element set is closed:
   (a) Let $\left(\overline{\mathsf{B}}(x_n, r_n)\right)_{n\in\mathbb{Z}}$ be a sequence of all the closed balls containing a.
   (b) Then $\{a\} = \bigcap_{n\in\mathbb{Z}} \overline{\mathsf{B}}(x_n, r_n)$.
   (c) By Proposition 2.2 page 34, every closed ball $\overline{\mathsf{B}}(\cdot,\cdot)$ is a closed set.
   (d) By Corollary 2.2 page 33, the infinite intersection of closed sets is also closed. So, $\bigcap_{n\in\mathbb{Z}} \overline{\mathsf{B}}(x_n, r_n)$ is closed.
   (e) Therefore, $\{x\}$ is closed.
2. Proof that any finite element set is closed:
   (a) By the previous result, any single element set $\{x\}$ is closed.
   (b) By Corollary 2.2 (page 33), the finite union of closed sets is also closed.
   (c) Therefore,
   $$\{x_1, x_2, \ldots, x_n\} = \bigcup_{i=1}^{n} \{x_i\}$$
   is closed.
3. Proof that any ring is closed:

[14] Rosenlicht (1968) page 42

(a) By Proposition 2.2 (page 34), the closed ball $\overline{\mathsf{B}}(\theta, r)$ is a closed set.

(b) By Proposition 2.2 (page 34), the open ball $\mathsf{B}(\theta, r)$ is an open set and by Definition 1.1 (page 3), its complement $(\mathsf{B}(\theta, r))^{\mathsf{c}}$ is a closed set.

(c) By Corollary 2.2 (page 33), the intersection of the two closed sets $\overline{\mathsf{B}}(\theta, r) \cap \left((\mathsf{B}(\theta, r))^{\mathsf{c}}\right)$ is also a closed set.

(d) Therefore, the ring is a closed set because

$$\underbrace{\{x \in X | \mathsf{d}(\theta, x) = r\}}_{\text{ring}} = \underbrace{\overline{\mathsf{B}}(\theta, r) \cap \left((\mathsf{B}(\theta, r))^{\mathsf{c}}\right)}_{\text{intersection of two closed sets}} .$$

✏

Counterexample 2.1. Unlike *metric spaces* (Proposition 2.3 page 35), a finite set in a *topological space* (Definition 1.1 page 3) $(X, \boldsymbol{T})$ is *not* in general *closed* (Definition 1.4 page 14).

CNT The finite set $\{x\}$ is *not* closed in the topological space (a *Serpiński space*)

$$\Bigg(\underbrace{\{x, y\}}_{X}, \underbrace{\{\varnothing, \{x\}, \{x, y\}\}}_{\boldsymbol{T}}\Bigg).$$

✎PROOF:

1. A set is *closed* if it is the complement of an open set (Definition 1.1 page 3).
2. The set $\{x\}$ is *not* the complement of any open set in the topology.
3. Therefore, $\{x\}$ is not closed.

✏

2.3.3 Equivalence and Order on metric spaces

Definition 2.5. [15] *Let (X, d_1) be a* METRIC SPACE (Definition 2.1 page 27) *that induces the* TOPOLOGY (Definition 1.1 page 3) $(X, \boldsymbol{T}_1)$ *and* (X, d_2) *be a* METRIC SPACE *that induces the* TOPOLOGY $(X, \boldsymbol{T}_2)$.

DEF d_1 *and* d_2 *are* ***equivalent*** *if*

$$\boldsymbol{T}_1 = \boldsymbol{T}_2.$$

Theorem 2.3. [16] *Let $\{B_1(x, y)\}$ be* OPEN BALLS (Definition 2.4 page 30) *on a* METRIC SPACE (X, d_1) *that induces the* TOPOLOGY $(X, \boldsymbol{T}_1)$ *and* $\{B_2(x, y)\}$ *be* OPEN BALLS *on a* METRIC SPACE (X, d_2) *that induces the* TOPOLOGY $(X, \boldsymbol{T}_2)$.

THM
$$\left.\begin{array}{ll} 1. & \exists \alpha > 0 \quad \text{such that} \quad \mathsf{d}_1(x, y) \le \alpha \mathsf{d}_2(x, y) \\ & \forall x, y \in X \quad \textit{and} \\ 2. & U \textit{ is open in } (X, \mathsf{d}_1) \end{array}\right\} \implies \left\{\begin{array}{ll} 1. & U \textit{ is open in } (X, \mathsf{d}_2) \quad \textit{and} \\ 2. & \boldsymbol{T}_2 \subseteq \boldsymbol{T}_1 \end{array}\right.$$

✎PROOF:

[15] Davis (2005) page 20
[16] Davis (2005) page 20

1. Proof that U is open in (X, d_2):

$$\begin{array}{rll}
& U \text{ is open in } (X, \mathsf{d}_1) & \text{by left hypothesis 2.}\\
\implies & \forall x \in U,\ \exists r > 0 \text{ such that } B_d(x, r) \subseteq X & \text{by Lemma 2.3 page 33}\\
\implies & \forall x \in U,\ \exists r > 0 \text{ such that } \{y \in X | \mathsf{d}_1(x, y) < r\} \subseteq X & \text{by Definition 2.4 page 30}\\
\implies & \forall x \in U,\ \exists r > 0 \text{ such that } \{y \in X | \mathsf{d}_1(x, y) < \alpha r\} \subseteq X & \\
\implies & \forall x \in U,\ \exists r > 0 \text{ such that } \{y \in X | \mathsf{d}_2(x, y) < r\} \subseteq X & \text{by left hypothesis 1.}\\
\implies & \forall x \in U,\ \exists r > 0 \text{ such that } B_2(x, y) \subseteq X & \text{by Definition 2.4 page 30}\\
\implies & U \text{ is open in } (X, \mathsf{d}_2) & \text{by Lemma 2.3 page 33}
\end{array}$$

2. Proof that $\boldsymbol{T}_2 \subseteq \boldsymbol{T}_1$:
Because U is open in $(X, \mathsf{d}_1) \implies U$ is open in (X, d_2), (see above), then $\boldsymbol{T}_2 \subseteq \boldsymbol{T}_1$.

⇨

Example 2.1. [17] Let R be a *commutative ring* and $|\cdot| \in R^R$ be the *absolute value* (Definition F.4 page 346) on R.

EX

The following *metric spaces* are all *equivalent* for any $n \in \mathbb{N}$:

1. $\left(R^n,\ \mathsf{d}_1(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sum_{i=1}^{n} |x_i - y_i|\right)$ (l_1-*metric* or *taxi-cab metric*)
2. $\left(R^n,\ \mathsf{d}_2(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}\right)$ (l_2-*metric* or *Euclidean metric*)
3. $\left(R^n,\ \mathsf{d}_\infty(\boldsymbol{x}, \boldsymbol{y}) \triangleq \max\{|x_i - y_i| | i = 1, 2, \ldots n\}\right)$ (l_∞-*metric* or *sup metric*)

✎PROOF:

1. Proof that (R^n, d_1) and (R^n, d_2) are equivalent:

Let

$$\boldsymbol{z}^{(1)} \triangleq \begin{bmatrix} y_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \qquad \boldsymbol{z}^{(2)} \triangleq \begin{bmatrix} y_1 \\ y_2 \\ x_3 \\ x_4 \\ \vdots \\ x_n \end{bmatrix} \qquad \boldsymbol{z}^{(3)} \triangleq \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ x_4 \\ x_5 \\ \vdots \\ x_n \end{bmatrix} \qquad \boldsymbol{z}^{(k)} \triangleq \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \\ x_{k+1}, x_{k+2}, \vdots \\ x_n \end{bmatrix}$$

$$\begin{aligned}
\mathsf{d}_2(\boldsymbol{x}, \boldsymbol{y}) &\le \mathsf{d}_2(\boldsymbol{x}, \boldsymbol{z}^{(1)}) + \mathsf{d}(\boldsymbol{z}^{(1)}, \boldsymbol{y}) && \text{by definition of } \textit{metric} \text{ (Definition 2.1 page 27)}\\
&\le \mathsf{d}_2(\boldsymbol{x}, \boldsymbol{z}^{(1)}) + \mathsf{d}(\boldsymbol{z}^{(1)}, \boldsymbol{z}^{(2)}) + \mathsf{d}(\boldsymbol{z}^{(2)}, \boldsymbol{y}) && \text{by definition of } \textit{metric} \text{ (Definition 2.1 page 27)}\\
&\le \mathsf{d}_2(\boldsymbol{x}, \boldsymbol{z}^{(1)}) + \mathsf{d}(\boldsymbol{z}^{(1)}, \boldsymbol{z}^{(2)}) + \mathsf{d}(\boldsymbol{z}^{(2)}, \boldsymbol{z}^{(3)}) + \mathsf{d}(\boldsymbol{z}^{(3)}, \boldsymbol{y}) && \text{by definition of } \textit{metric} \text{ (Definition 2.1 page 27)}\\
&\vdots\\
&\le \mathsf{d}_2(\boldsymbol{x}, \boldsymbol{z}^{(1)}) + \mathsf{d}(\boldsymbol{z}^{(1)}, \boldsymbol{z}^{(2)}) + \mathsf{d}(\boldsymbol{z}^{(2)}, \boldsymbol{z}^{(3)}) + \cdots\\
&\qquad + \mathsf{d}(\boldsymbol{z}^{(n-2)}, \boldsymbol{z}^{(n-1)}) + \mathsf{d}(\boldsymbol{z}^{(n-1)}, \boldsymbol{y}) && \text{by definition of } \textit{metric} \text{ (Definition 2.1 page 27)}\\
&= \sqrt{\sum_{i=1}^{n} |x_i - z_i^{(1)}|^2} + \sqrt{\sum_{i=1}^{n} |z_i^{(1)} - z_i^{(2)}|^2} + \cdots
\end{aligned}$$

[17] Davis (2005) pages 20–21

$$
\begin{aligned}
&+ \sqrt{\sum_{i=1}^{n} |z_i^{(n-2)} - z_i^{(n-1)}|^2} + \sqrt{\sum_{i=1}^{n} |z_i^{(n-1)} - y_i|^2} \\
&= \sqrt{|x_1 - y_1|^2} + \sqrt{|x_2 - y_2|^2} + \sqrt{|x_3 - y_3|^2} + \cdots \\
&\quad + \sqrt{|x_{n-1} - y_{n-1}|^2} + \sqrt{|x_n - y_n|^2} \\
&= |x_1 - y_1| + |x_2 - y_2| + |x_3 - y_3| + \cdots + |x_n - y_n| \\
&= \sum_{i=1}^{n} |x_i - y_i| \\
&= \mathsf{d}_1(\boldsymbol{x}, \boldsymbol{y}) && \text{by definition of } \textit{metric} \text{ (Definition 2.1 page 27)}
\end{aligned}
$$

By Theorem 2.3 (page 36), $\mathsf{d}_2(\boldsymbol{x}, \boldsymbol{y}) \le \mathsf{d}_1(\boldsymbol{x}, \boldsymbol{y})$ implies that (R^n, d_1) and (R^n, d_2) are equivalent.

2. Proof that (R^n, d_1) and (R^n, d_∞) are equivalent:

$$
\begin{aligned}
\mathsf{d}_1(\boldsymbol{x}, \boldsymbol{y}) &= \sum_{i=1}^{n} |x_i - y_i| && \text{by definition of } \textit{metric} \text{ (Definition 2.1 page 27)} \\
&\le n \cdot \max\left\{|x_i - y_i| | i = 1, 2, \ldots, N\right\} \\
&= n\, \mathsf{d}_\infty(\boldsymbol{x}, \boldsymbol{y}) && \text{by definition of } \textit{metric} \text{ (Definition 2.1 page 27)}
\end{aligned}
$$

By Theorem 2.3 (page 36), $\mathsf{d}_1(\boldsymbol{x}, \boldsymbol{y}) \le n\mathsf{d}_\infty(\boldsymbol{x}, \boldsymbol{y})$ implies that $(\mathbb{C}^n, \mathsf{d}_1)$ and $(\mathbb{C}^n, \mathsf{d}_\infty)$ are equivalent.

✎

2.3.4 Metrics induced by topologies

There are many topological spaces that are induced by metric spaces, and others that are not. A topology that is induced by a metric is called *metrizable* (next definition).

Definition 2.6. *Let* $(X, \boldsymbol{T})$ *be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3).*

DEF: *A* TOPOLOGY *is* ***metrizable*** *if it is induced by a* METRIC.

Example 2.2. [18] Let $\mathcal{T}(X)$ be the set of *topologies* (Definition 1.1 page 3) on a set X and 2^X the *power set* (Definition A.1 page 265) on X.

EX:

$\{\varnothing, X\}$	is called the *indiscrete topology*.	It is *not* metrizable.
2^X	is called the *discrete topology*.	It *is* metrizable.

Example 2.3. [19]

EX: The **Sierpiński space** $(X, \boldsymbol{T})$ is a topological space with topology $\boldsymbol{T} \triangleq \{\varnothing, \{x\}, \{x, y\}\}$ on the set $X \triangleq \{x, y\}$.
It is *not* metrizable.
The Sierpiński space is also called the **two-point connected space**.

[18] Munkres (2000), page 77, Kubrusly (2011) page 107 ⟨Example 3.J⟩, Steen and Seebach (1978) pages 42–43 ⟨II.4⟩, DiBenedetto (2002) page 18

[19] Joshi (1983) page 90 ⟨example 9⟩, Davis (2005) page 42 ⟨Example 4.4.4⟩

2.4 Additional properties

2.4.1 Separable metric spaces

Definition 1.7 page 23 gives the definition of a separable space.

Theorem 2.4. [20] *Let* (Y, d) *be a subspace of a* METRIC SPACE (X, d).

THM: (X, d) *is* SEPARABLE $\implies$ (Y, d) *is* SEPARABLE *(Definition 1.7 page 23)*

2.4.2 Compact metric spaces

Definition 2.7 (Borel-Lebesgue axiom). [21] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27). Let* I *be an* INFINITE INDEXING SET *and* $J \subsetneq I$ *be a* FINITE INDEXING SET. *Let* $(\!(U_n)\!)_{n\in I}$ *be a* SEQUENCE *(Definition 8.1 page 127) of* OPEN SETS *(Definition 1.1 page 3).*

DEF: *A set* A *is* ***compact*** *if*

$$\underbrace{A \subseteq \bigcup_{n\in I} U_n}_{A \text{ is covered by an infinite union of open sets}} \implies \underbrace{A \subseteq \bigcup_{n\in J} U_n}_{A \text{ is covered by a finite union of open sets}}$$

Proposition 2.4. [22] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27). Let* Y *be a subset of* X.

PRP: $\left\{\begin{array}{lll} 1. & (X, \mathsf{d}) \text{ is COMPACT} & \text{and} \\ 2. & Y \text{ is CLOSED in } (X, \mathsf{d}) & (Y = Y^-) \end{array}\right\} \implies \{(Y, \mathsf{d}) \text{ is COMPACT}\}$

✎PROOF:

1. Because (Y, d) is a metric space, there exists a sequence of open sets $(\!(A_n)\!)_{i\in I}$, where I is an infinite indexing set, such that
$$Y \subseteq \bigcup_{i\in I} A_i.$$
2. By left hypothesis 2, Y is closed, which means its complement Y^c is open.
3. Combining the above two statements, we have
$$X = Y \cup Y^c \subseteq \left(\bigcup_{i\in I} A_i\right) \cup Y^c.$$
4. By left hypothesis 1, X is compact and therefore can be covered by a *finite* number of open sets. Let J be a finite indexing set such that
$$X \subseteq \left(\bigcup_{i\in J} A_i\right) \cup Y^c.$$

[20] Runda (2005) page 32 ⟨Theorem 2.2.17⟩
[21] Dieudonné (1969), pages 57–58, Rosenlicht (1968) page 54
[22] Dieudonné (1969), page 62, Rosenlicht (1968) page 54

5. By left hypothesis 2, $Y \subseteq X$. Therefore

$$Y \subseteq X \subseteq \left(\bigcup_{i\in J} A_i\right) \cup Y^c.$$

6. And so, Y is covered by a finite number of open sets ($(A_i)_{i\in J}$ and Y^c), and Y is therefore *compact*.

⇨

Proposition 2.5 (Nested set property). [23] *Let* (X, d) *be a* METRIC SPACE *and* $(Y_n)_{n\in\mathbb{Z}}$ *a* SEQUENCE *(Definition 8.1 page 127) of sets.*

THM

$$\left.\begin{array}{llll} 1. & (X,\mathsf{d}) \text{ is COMPACT} & & \text{and} \\ 2. & Y_n \subseteq X \quad \forall n\in\mathbb{Z} & (Y_n \text{ are subsets of } X) & \text{and} \\ 3. & Y_n \neq \varnothing \quad \forall n\in\mathbb{Z} & (Y_n \text{ are non-empty}) & \text{and} \\ 4. & Y_n \supseteq Y_{n+1} \quad \forall n\in\mathbb{Z} & (\text{nested subsets}) & \text{and} \\ 5. & Y_n = Y_n^- \quad \forall n\in\mathbb{Z} & (Y \text{ is closed}). & \end{array}\right\} \implies \underbrace{\left|\bigcap_{n\in\mathbb{Z}} Y_n\right| \geq 1}_{\text{all subsets have at least one common element}}$$

✎PROOF: Proof is by contradiction.

1. Note that $Y_n \supseteq Y_{n+1} \iff Y^c{}_n \subseteq Y^c{}_{n+1}$
2. Suppose that the statement is false; that is, $\left|\bigcap_{n\in\mathbb{Z}} Y_n\right| = 0$.

$$\begin{aligned} \left|\bigcap_{n\in\mathbb{Z}} Y_n\right| = 0 &\iff \bigcap_{n\in\mathbb{Z}} Y_n = \varnothing \\ &\iff \bigcup_{n\in\mathbb{Z}} Y^c{}_n = X && \text{by } \textit{de Morgan's law} \text{ (Theorem A.6 page 278)} \\ &\implies \bigcup_{n=1}^{N} Y^c{}_n = X \text{ for some finite } N && \text{by compactness hypothesis} \\ &\implies Y^c{}_N = X && \text{because } Y^c{}_n \subseteq Y^c{}_{n+1} \\ &\iff Y_N = \varnothing \end{aligned}$$

3. But this is a *contradiction*, because by left hypothesis 3, $Y_n \neq \varnothing$.
4. Therefore, $\left|\bigcap_{n\in\mathbb{Z}} Y_n\right| \geq 1$.

⇨

2.4.3 Orthogonality on metric linear spaces

Definition 2.8. [24] *Let* $(\mathbf{V}, \mathsf{d})$ *be a* METRIC LINEAR SPACE. *Let* $[x_1 : x_2]$ *and* $[y_1 : y_2]$ *be* LINE SEGMENTS *in the linear space* $\mathbf{V}$ *that intersect at a point* $p \in [x_1 : x_2]$.

DEF

The line segments $[x_1 : x_2]$ *and* $[y_1 : y_2]$ *are* ***orthogonal*** *in the metric linear space* $(\mathbf{V}, \mathsf{d})$ *if*

$$\underbrace{\mathsf{d}(y_1, p) \leq \mathsf{d}(y_2, q) \qquad \forall q \in [x_1 : x_2]}_{p \text{ is the closest point in } [x_1 : x_2] \text{ to } y_1}$$

[23] Rosenlicht (1968) page 55
[24] Birkhoff (1935), page 169

2.5 Metric transforms

If we know that one or more functions are metrics, then we can use them to generate other metrics. This is demonstrated by the following:

Theorem 2.5 (page 41): generate a metric using an isometry.
Theorem 2.6 (page 41): generate a metric using a monotone function.
Theorem 2.8 (page 43): generate a metric using a *metric preserving function*.
Theorem 2.9 (page 46): generate a metric from a linear combination of metrics.
Theorem 2.10 (page 47): generate an N-dimensional metric from weighted 1-dimensional metrics.

2.5.1 Metric transforms on the domains of metrics

Definition 2.9. [25] *Let* (X, d) *and* (Y, p) *be* METRIC SPACES (Definition 2.1 page 27).

DEF: *The function* $\mathsf{f} \in Y^X$ *is an* ***isometry*** *on* $(Y, \mathsf{p})^{(X,\mathsf{d})}$ *if*

$$\mathsf{d}(x, y) = \mathsf{p}(\mathsf{f}(x), \mathsf{f}(y)) \qquad \forall x,y\in X$$

The spaces (X, d) *and* (Y, p) *are* ***isometric*** *if there exists an isometry on* $(Y, \mathsf{p})^{(X,\mathsf{d})}$.

Theorem 2.5. [26] *Let* (X, d) *and* (Y, p) *be* METRIC SPACES. *Let* f *be a function in* Y^X *and* f^{-1} *its inverse in* X^Y.

THM:
$$\left\{\mathsf{f} \text{ is an } \textbf{isometry} \text{ on } (Y, \mathsf{p})^{(X,\mathsf{d})}\right\} \iff \left\{\mathsf{f}^{-1} \text{ is an } \textbf{isometry} \text{ on } (X, \mathsf{d})^{(Y,\mathsf{p})}\right\}$$

If a function p is a *metric* and a function g is *injective*, then the function $\mathsf{d}(x, y) \triangleq \mathsf{p}(\mathsf{g}(x), \mathsf{g}(y))$ is also a *metric* (next theorem). For an example of this with $\mathsf{p}(x, y) \triangleq |x - y|$ and $\mathsf{g} \triangleq \arctan(x)$, see Example 2.25 (page 60).

Theorem 2.6 (Pullback metric/g-transform metric). [27] *Let* X *and* Y *be sets. Let* g *be a function in* Y^X.

THM:
$$\left\{\begin{array}{lll} 1. & \mathsf{p} \text{ is a metric on } Y & \text{and} \\ 2. & \mathsf{g} \text{ is INJECTIVE} & \end{array}\right\} \implies \left\{\begin{array}{c} \mathsf{d}(x, y) = \mathsf{p}(\mathsf{g}(x), \mathsf{g}(y)) \quad \forall x,y\in X \\ \text{is a metric on } Y \end{array}\right\}$$

✎PROOF:

1. Proof that $x = y \implies \mathsf{d}(x, y) = 0$:

$$\begin{aligned}
\mathsf{d}(x, y) &\triangleq \mathsf{p}(\phi(x), \phi(y)) && \text{by definition of } \mathsf{d} \\
&= \mathsf{p}(\phi(x), \phi(x)) && \text{by } x = y \text{ hypothesis} \\
&= 0 && \text{by } \textit{nondegenerate} \text{ property of metric } \mathsf{p} \text{ (Definition 2.1 page 27)} \\
&= 0
\end{aligned}$$

2. Proof that $x = y \impliedby \mathsf{d}(x, y) = 0$:

$$\begin{aligned}
0 &= \mathsf{d}(x, y) && \text{by right hypothesis} \\
&\triangleq \mathsf{p}(\phi(x), \phi(y)) && \text{by definition of } \mathsf{d} \\
\implies \mathsf{p}(\phi(x), \phi(y)) &= 0 \text{ for } n = 1, 2, \ldots, N && \text{because } \mathsf{p} \text{ is non-negative} \\
\implies x &= y && \text{by left hypothesis 2}
\end{aligned}$$

[25] Thron (1966), page 153 ⟨definition 19.4⟩, Giles (1987) page 124 ⟨Definition 6.22⟩, Khamsi and Kirk (2001) page 15 ⟨Definition 2.4⟩, Kubrusly (2001) page 110
[26] Thron (1966), page 153 ⟨theorem 19.5⟩
[27] Deza and Deza (2009) page 81

3. Proof that $\mathsf{d}(x,y) \le \mathsf{d}(z,x) + \mathsf{d}(z,y)$:

$$\begin{aligned}
\mathsf{d}(x,y) &\triangleq \mathsf{p}(\phi(x),\phi(y)) && \text{by definition of } \mathsf{d}\\
&\le \big(\mathsf{p}(\phi(x),\phi(z)) + \mathsf{d}(\phi(z),\phi(y))\big) && \text{by \textit{subadditive} property of } \mathsf{p} \text{ (Definition 2.1 page 27)}\\
&= \mathsf{p}(\phi(z),\phi(x)) + \mathsf{p}(\phi(z),\phi(y)) && \text{by \textit{symmetry} property of metric } \mathsf{p} \text{ (Definition 2.1 page 27)}\\
&\triangleq \mathsf{d}(z,x) + \mathsf{d}(z,y) && \text{by definition of } \mathsf{d}
\end{aligned}$$

⇒

2.5.2 Metric preserving functions

Definition 2.10. [28] *Let* $\mathbb{M}$ *be the set of all* METRIC SPACES *on a set* X.

DEF *A* FUNCTION $\phi \in \mathbb{R}^{\vdash \mathbb{R}^{\vdash}}$ *is **metric preserving** if*
$\mathsf{d}(x,y) \triangleq \phi \circ \mathsf{p}(x,y)$ *is a metric on* X *for all* $(X,\mathsf{p}) \in \mathbb{M}$

Theorem 2.7 (next theorem) presents some necessary conditions for a function ϕ to be *metric preserving*. Theorem 2.8 (page 43) presents some sufficient conditions. But first some conditions that are *not* necessary:

1. It is *not* necessary for ϕ to be *continuous* (see Example 2.8 page 45).
2. It is *not* necessary for ϕ to be *nondecreasing* (see Example 2.9 page 45).
3. It is *not* necessary for ϕ to be *monotonic* (see Example 2.10 page 46).

Theorem 2.7 (necessary conditions). [29] *Let* $\mathcal{R}\mathsf{f}$ *be the* RANGE *of a function* f.

THM
$$\left\{\begin{array}{l}\phi \text{ is a metric}\\ \text{preserving function}\end{array}\right\} \implies \left\{\begin{array}{lll}1. & \phi^{-1}(0) = \{0\} & \text{and}\\ 2. & \mathcal{R}\phi \subseteq \mathbb{R}^{\vdash} & \text{and}\\ 3. & \phi(x+y) \le \phi(x) + \phi(y) & (\phi \text{ is SUBADDITIVE})\end{array}\right\}$$

✎PROOF:

1. Proof that ϕ is a *metric preserving function* $\implies \phi^{-1}(0) = \{0\}$:

   (a) Suppose that the statement is not true and $\phi^{-1}(0) = \{0, a\}$.

   (b) Then $\phi(a) = 0$ and for some x, y such that $x \ne y$ and $\mathsf{d}(x,y) = a$ we have

   $$\begin{aligned}
   \phi \circ \mathsf{d}(x,y) &= \phi(a)\\
   &= 0\\
   &\implies \phi \circ \mathsf{d} \text{ is \textit{not} a metric}\\
   &\implies \phi \text{ is \textit{not} a metric preserving function}
   \end{aligned}$$

   (c) But this contradicts the original hypothesis, and so it must be that $\phi^{-1}(0) = \{0\}$.

2. Proof that $\mathcal{R}\phi \subseteq \mathbb{R}^{\vdash}$:

$$\begin{aligned}
\mathcal{R}\phi \circ \mathsf{d} &\subseteq \mathcal{R}\mathsf{d}\\
&\subseteq \mathbb{R}^{\vdash}
\end{aligned}$$

[28] Vallin (1999), page 849 ⟨Definition 1.1⟩, Corazza (1999), page 309, Deza and Deza (2009) page 80
[29] Corazza (1999), page 310 ⟨Proposition 2.1⟩, Deza and Deza (2009) page 80

3. Proof that ϕ is a metric preserving function $\implies$ ϕ is *subadditive*:
   (a) For ϕ to be a *metric preserving function*, by definition it must work with *all metric spaces*.
   (b) So to develop necessary conditions, we can pick any metric space we want (because it is necessary that ϕ preserves it as a metric space).
   (c) For this proof we choose the metric space $(\mathbb{R}, \mathsf{d})$ where $\mathsf{d}(x,y) \triangleq |x-y|$ for all $x, y \in \mathbb{R}^{\vdash}$:

$$\begin{aligned}
\phi(x)+\phi(y) &= \phi(|(x+y)-x|)+\phi(|x-0|) && \text{by definition of } |\cdot| \\
&= (\phi\circ\mathsf{d})(x+y,x)+(\phi\circ\mathsf{d})(x,0) && \text{by definition of } \mathsf{d} \\
&\geq (\phi\circ\mathsf{d})(x+y,0) && \text{by left hypothesis and Definition 2.1 page 27} \\
&= \phi(|(x+y)-0|) && \text{by definition of } \mathsf{d} \\
&= \phi(x+y) && \text{because } x, y \in \mathbb{R}^{\vdash}
\end{aligned}$$

✏

Theorem 2.8 (sufficient conditions). [30] *Let ϕ be a function in $\mathbb{R}^{\mathbb{R}}$.*

THM

$$\left.\begin{array}{llll}
1. & x \geq y \implies \phi(x) \geq \phi(y) \quad \forall x,y \in \mathbb{R}^{\vdash} & \text{(NONDECREASING)} & \text{and} \\
2. & \phi(0) = 0 & & \text{and} \\
3. & \phi(x+y) \leq \phi(x)+\phi(y) \quad \forall x,y \in \mathbb{R}^{\vdash} & \text{(SUBADDITIVE).} &
\end{array}\right\} \implies \phi \text{ is a METRIC PRESERVING FUNCTION}$$

✎PROOF:

1. Proof that $\phi\circ\mathsf{d}(x,y) = 0 \implies x = y$:

$$\begin{aligned}
\phi\circ\mathsf{d}(x,y) = 0 &\implies \mathsf{d}(x,y) = 0 && \text{by } \phi \text{ hypothesis 2} \\
&\implies x = y && \text{by } \textit{nondegenerate} \text{ property page 27}
\end{aligned}$$

2. Proof that $\phi\circ\mathsf{d}(x,y) = 0 \impliedby x = y$:

$$\begin{aligned}
\phi\circ\mathsf{d}(x,y) &= \phi\circ\mathsf{d}(x,x) && \text{by } x = y \text{ hypothesis} \\
&= \phi(0) && \text{by } \textit{nondegenerate} \text{ property page 27} \\
&= 0 && \text{by } \phi \text{ hypothesis 2}
\end{aligned}$$

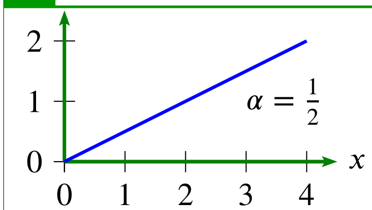
3. Proof that $\phi\circ\mathsf{d}(x,y) \leq \phi\circ\mathsf{d}(z,x)+\phi\circ\mathsf{d}(z,y)$:

$$\begin{aligned}
\phi\circ\mathsf{d}(x,y) &\leq \phi(\mathsf{d}(x,z)+\mathsf{d}(z,y)) && \text{by } \phi \text{ hypothesis 1 and triangle inequality page 27} \\
&\leq \phi(\mathsf{d}(z,x)+\mathsf{d}(z,y)) && \text{by symmetric property of } \mathsf{d} \text{ page 27} \\
&\leq \phi\circ\mathsf{d}(z,x)+\phi\circ\mathsf{d}(z,y) && \text{by } \phi \text{ hypothesis 3}
\end{aligned}$$

✏

Example 2.4 (α-scaled metric/dilated metric). [31] Let (X, d) be a *metric space* (Definition 2.1 page 27).

EX $\phi(x) \triangleq \alpha x,\ \alpha \in \mathbb{R}^+$, is *metric preserving* $\quad \big(\mathsf{p}(x,y) \triangleq \alpha\mathsf{d}(x,y)$ is a metric on $X\big)$



[30] Corazza (1999) ⟨Proposition 2.3⟩, Deza and Deza (2009) page 80, Kelley (1955) page 131 ⟨Problem C⟩
[31] Deza and Deza (2006) page 44

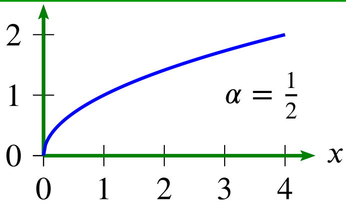
✎PROOF:

1. Note that ϕ satisfies the conditions of Theorem 2.8 (page 43).
2. Therefore, by Theorem 2.8 (page 43), $\mathsf{d}(x, y)$ is a metric on X.

⇒

Example 2.5 (power transform metric/snowflake transform metric). [32] Let (X, d) be a *metric space* (Definition 2.1 page 27).

EX $\phi(x) \triangleq x^\alpha,\ \alpha \in (0:1]$, is *metric preserving* $\left(\mathsf{p}(x, y) \triangleq [\mathsf{d}(x, y)]^\alpha, \text{ is a metric on } X\right)$



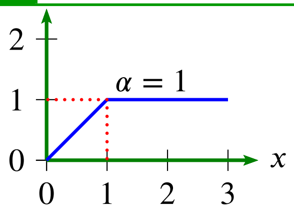
✎PROOF:

1. Note that ϕ satisfies the conditions of Theorem 2.8 (page 43) for $0 < \alpha \leq 1$.
2. Therefore, by Theorem 2.8 (page 43), $\mathsf{d}(x, y)$ is a metric on X.

⇒

Example 2.6 (α-truncated metric/radar screen metric). [33] Let (X, d) be a *metric space* (Definition 2.1 page 27).

EX $\phi(x) \triangleq \min\{\alpha,\ x\},\ \alpha \in \mathbb{R}^+$, is *metric preserving* $\left(\begin{array}{l}\mathsf{p}(x, y) \triangleq \min\{\alpha,\ \mathsf{d}(x, y)\}\\ \text{is a metric on } X\end{array}\right)$



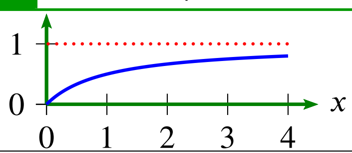
✎PROOF:

1. Note that ϕ satisfies the conditions of Theorem 2.8 (page 43).
2. $\mathsf{d}(x, y) \triangleq \min\{\alpha, \mathsf{p}(x, y)\} = \phi \circ \mathsf{p}(x, y)$
3. Therefore, by Theorem 2.8 (page 43), $\mathsf{d}(x, y)$ is a metric.

⇒

Example 2.7 (bounded metric). [34] Let (X, d) be a *metric space* (Definition 2.1 page 27).

EX $\phi(x) \triangleq \dfrac{x}{1+x}$ is *metric preserving* $\left(\mathsf{p}(x, y) \triangleq \dfrac{\mathsf{d}(x, y)}{1 + \mathsf{d}(x, y)} \text{ is also a metric on } X\right)$



[32] Deza and Deza (2009) page 81, Deza and Deza (2006) page 45
[33] Giles (1987) page 33, Deza and Deza (2006) pages 242–243
[34] Vallin (1999), page 849, Aliprantis and Burkinshaw (1998) page 39

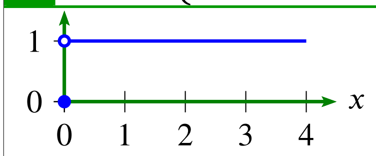
✎PROOF:

1. Note that ϕ satisfies the conditions of Theorem 2.8 (page 43).
2. $\mathsf{d}(x,y) \triangleq \dfrac{\mathsf{p}(x,y)}{1+\mathsf{p}(x,y)} = \phi \circ \mathsf{p}(x,y)$
3. Therefore, by Theorem 2.8 (page 43), $\mathsf{d}(x,y)$ is a metric.

✏

Example 2.8. [35] Let ϕ be a function in $\mathbb{R}^{\mathbb{R}}$.

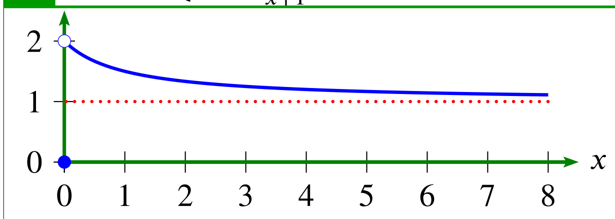
EX $\phi(x) \triangleq \begin{cases} 0 & \text{for } x \le 0 \\ 1 & \text{for } x > 0 \end{cases}$ is a *discontinuous* metric preserving function



✎PROOF: This result follows directly from Theorem 2.8 page 43. ✏

Example 2.9. Let ϕ be a function in $\mathbb{R}^{\mathbb{R}}$.

EX $\phi(x) \triangleq \begin{cases} 0 & \text{for } x = 0 \\ 1 + \frac{1}{x+1} & \text{for } x > 0 \end{cases}$ is a metric preserving function



✎PROOF:

1. Note that $\phi \circ \mathsf{d}(x,x) = 0 \iff x = 0$.
2. Lemma: $\frac{1}{a+b} \le \frac{1}{a} + \frac{1}{b}$ for $a, b \in \mathbb{R}^+$:

$$\begin{aligned} \frac{1}{a+b} &\le \frac{1}{a} \\ &\le \frac{1}{a} + \frac{1}{b} \end{aligned}$$

3. Proof that $\phi \circ \mathsf{d}$ is *subadditive*:

$$\begin{aligned}
\phi \circ \mathsf{d}(x,y) &= 1 + \frac{1}{1+\mathsf{d}(x,y)} && \text{by definition of } \phi \\
&\le 1 + \frac{1}{1+\mathsf{d}(x,z)+\mathsf{d}(z,y)} && \text{by \textit{subadditive} property of metric } \mathsf{d} \\
&\le 1 + \frac{1}{1+\mathsf{d}(x,z)+1+\mathsf{d}(z,y)} \\
&\le 1 + \frac{1}{1+\mathsf{d}(x,z)} + \frac{1}{1+\mathsf{d}(z,y)} && \text{by Example 2} \\
&\le 1 + \frac{1}{1+\mathsf{d}(x,z)} + 1 + \frac{1}{1+\mathsf{d}(z,y)} \\
&= \phi \circ \mathsf{d}(x,z) + \phi \circ \mathsf{d}(z,y) && \text{by definition of } \phi
\end{aligned}$$

[35] Corazza (1999), page 311

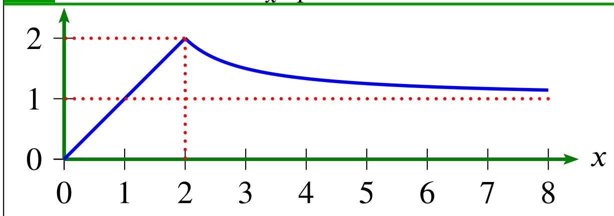
4. Therefore, by Theorem 2.1 (page 27), $\phi \circ \mathsf{d}(x,y)$ is a *metric* and ϕ is a *metric preserving function*.

✏

Example 2.10. [36] Let ϕ be a function in $\mathbb{R}^{\mathbb{R}}$.

EX

$$\phi(x) \triangleq \begin{cases} x & \text{for } x \le 2 \\ 1 + \frac{1}{x-1} & \text{for } x > 2 \end{cases} \quad \text{is a \textit{nonmonotonic} metric preserving function}$$



2.5.3 Product metrics

Theorem 2.9 (Fréchet product metric). [37] *Let X be a set.*

THM

$$\left\{ \begin{array}{lll} 1. & (\mathsf{p}_n) \text{ are METRICS on } X & \text{and} \\ 2. & \alpha_n \ge 0 \quad \forall n = 1,2,\ldots,N & \text{and} \\ 3. & \max\left\{ \alpha_n \middle| n=1,2,\ldots,N \right\} > 0 & \end{array} \right\} \implies \left\{ \begin{array}{c} \mathsf{d}(x,y) = \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(x,y) \\ \text{is a METRIC on } X \end{array} \right\}$$

✎PROOF:

1. Proof that $x = y \implies \mathsf{d}(x,y) = 0$:

$$\begin{aligned} \mathsf{d}(x,y) &= \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(x,y) && \text{by definition of d} \\ &= \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(x,x) && \text{by left hypothesis} \\ &= \sum_{n=1}^{N} 0 && \text{by \textit{nondegenerate} property of metrics (Definition 2.1 page 27)} \\ &= 0 \end{aligned}$$

2. Proof that $x = y \impliedby \mathsf{d}(x,y) = 0$:

$$\begin{aligned} 0 &= \mathsf{d}(x,y) && \text{by right hypothesis} \\ &= \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(x,y) && \text{by definition of d} \\ \implies & \mathsf{p}_n(x,y) = 0 \quad \forall x,y \in \boldsymbol{X} && \text{by metric properties page 27} \\ \implies & x = y \quad \forall x,y \in \boldsymbol{X} && \text{by non-degenerate property of metrics page 27} \end{aligned}$$

[36] Corazza (1999), page 309, Doboš (1998), page 25 ⟨Example 1⟩, Jůza (1956)
[37] Deza and Deza (2006) page 47, Deza and Deza (2009) page 84, Steen and Seebach (1978) pages 64–65 ⟨Example 37.7⟩, Isham (1999) page 10

3. Proof that $\mathsf{d}(x,y) \le \mathsf{d}(z,x) + \mathsf{d}(z,y)$:

$$\begin{aligned}
\mathsf{d}(x,y) &= \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(x,y) && \text{by definition of } \mathsf{d}\\
&\le \sum_{n=1}^{N} \alpha_n \left[\mathsf{p}_n(x,z) + \mathsf{p}_n(z,y)\right] && \text{by } \textit{subadditive} \text{ property (Definition 2.1 page 27)}\\
&= \sum_{n=1}^{N} \alpha_n \left[\mathsf{p}_n(z,x) + \mathsf{p}_n(z,y)\right] && \text{by } \textit{symmetry} \text{ property (Definition 2.1 page 27)}\\
&= \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(z,x) + \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(z,y)\\
&= \mathsf{d}(z,x) + \mathsf{d}(z,y) && \text{by definition of } \mathsf{d}
\end{aligned}$$

✎

Theorem 2.10 (Power mean metrics). *Let X be a set. Let $(\!(x_n \in X)\!)_1^N$ and $(\!(y_n \in X)\!)_1^N$ be N-tuples on X.*

THM

$$\left.\begin{array}{ll} 1. & \mathsf{p} \text{ is a METRIC on } X \quad \textit{and} \\ 2. & \displaystyle\sum_{n=1}^{N} \lambda_n = 1 \end{array}\right\} \implies \left\{ \begin{array}{l} \mathsf{d}\left((\!(x_n)\!), (\!(y_n)\!)\right) \triangleq \left(\displaystyle\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(x_n,y_n)\right)^{\frac{1}{r}}, \\ \qquad r \in [1:\infty], \quad \text{is a METRIC on } X \end{array}\right.$$

Moreover, if $r = \infty$, then $\mathsf{d}\left((\!(x_n)\!), (\!(y_n)\!)\right) = \max\limits_{n=1,\ldots,N} \mathsf{p}\left(x_n, y_n\right)$.

✎PROOF:

1. Proof that $(\!(x_n)\!) = (\!(y_n)\!) \implies \mathsf{d}\left((\!(x_n)\!), (\!(y_n)\!)\right) = 0$ for $r \in [1:\infty)$:

$$\begin{aligned}
\mathsf{d}((\!(x_n)\!), (\!(y_n)\!)) &\triangleq \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(x_n,y_n)\right)^{\frac{1}{r}} && \text{by definition of } \mathsf{d}\\
&= \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(x_n,x_n)\right)^{\frac{1}{r}} && \text{by } (\!(x_n)\!) = (\!(y_n)\!) \text{ hypothesis}\\
&= \left(\sum_{n=1}^{N} 0\right)^{\frac{1}{r}} && \text{because } \mathsf{p} \text{ is } \textit{nondegenerate}\\
&= 0
\end{aligned}$$

2. Proof that $(\!(x_n)\!) = (\!(y_n)\!) \impliedby \mathsf{d}\left((\!(x_n)\!), (\!(y_n)\!)\right) = 0$ for $r \in [1:\infty)$:

$$\begin{aligned}
0 &= \mathsf{d}\left((\!(x_n)\!), (\!(y_n)\!)\right) && \text{by } \mathsf{d}\left((\!(x_n)\!), (\!(y_n)\!)\right) = 0 \text{ hypothesis}\\
&\triangleq \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(x_n,y_n)\right)^{\frac{1}{r}} && \text{by definition of } \mathsf{d}\\
\implies \left(\mathsf{p}\left(x_n,y_n\right)\right)^{\frac{1}{r}} &= 0 \text{ for } n = 1,2,\ldots,N && \text{because } \mathsf{p} \text{ is } \textit{non-negative}\\
\implies (\!(x_n)\!) &= (\!(y_n)\!) && \text{because } \mathsf{p} \text{ is } \textit{nondegenerate}
\end{aligned}$$

3. Proof that d satisfies the triangle inequality property for $r = 1$:

$$\begin{aligned}
\mathsf{d}(\langle x_n \rangle, \langle y_n \rangle) &\triangleq \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(x_n, y_n)\right)^{\frac{1}{r}} && \text{by definition of } \mathsf{d}\\
&= \sum_{n=1}^{N} \lambda_n \mathsf{p}(x_n, y_n) && \text{by } r = 1 \text{ hypothesis}\\
&\le \sum_{n=1}^{N} \lambda_n \left[\mathsf{p}(z_n, x_n) + \mathsf{p}(z_n, y_n)\right] && \text{by \textit{triangle inequality}}\\
&= \sum_{n=1}^{N} \lambda_n \mathsf{p}(z_n, x_n) + \sum_{n=1}^{N} \lambda_n \mathsf{p}(z_n, y_n)\\
&= \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(z_n, x_n)\right)^{\frac{1}{r}} + \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(z_n, y_n)\right)^{\frac{1}{r}} && \text{by } r = 1 \text{ hypothesis}\\
&\triangleq \mathsf{d}(\langle z_n \rangle, \langle x_n \rangle) + \mathsf{d}(\langle z_n \rangle, \langle y_n \rangle) && \text{by definition of } \mathsf{d}
\end{aligned}$$

4. Proof that d satisfies the triangle inequality property for $r \in (1 : \infty)$:

$$\begin{aligned}
&\mathsf{d}(\langle x_n \rangle, \langle y_n \rangle)\\
&\triangleq \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(x_n, y_n)\right)^{\frac{1}{r}} && \text{by definition of } \mathsf{d}\\
&\le \left(\sum_{n=1}^{N} \lambda_n \left[\mathsf{p}(z_n, x_n) + \mathsf{p}(z_n, y_n)\right]^r\right)^{\frac{1}{r}} && \text{by \textit{subadditive} property (Definition 2.1 page 27)}\\
&= \left(\sum_{n=1}^{N} \left[\lambda_n^{\frac{1}{r}} \mathsf{p}(z_n, x_n) + \lambda_n^{\frac{1}{r}} \mathsf{p}(z_n, y_n)\right]^r\right)^{\frac{1}{r}} && \text{by \textit{subadditive} property (Definition 2.1 page 27)}\\
&\le \left(\sum_{n=1}^{N} \left[\lambda_n^{\frac{1}{r}} \mathsf{p}(z_n, x_n)\right]^r\right)^{\frac{1}{r}} + \left(\sum_{n=1}^{N} \left[\lambda_n^{\frac{1}{r}} \mathsf{p}(z_n, y_n)\right]^r\right)^{\frac{1}{r}} && \text{by \textit{Minkowski's inequality} (Theorem 11.5 page 190)}\\
&\le \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(z_n, x_n)\right)^{\frac{1}{r}} + \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(z_n, y_n)\right)^{\frac{1}{r}}\\
&\triangleq \mathsf{d}(\langle z_n \rangle, \langle x_n \rangle) + \mathsf{d}(\langle z_n \rangle, \langle y_n \rangle) && \text{by definition of } \mathsf{d}
\end{aligned}$$

5. Proof for the $r = \infty$ case:

   (a) Proof that $\mathsf{d}(\langle x_n \rangle, \langle y_n \rangle) = \max \langle x_n \rangle$: by Theorem 11.3 page 184

   (b) Proof that $\langle x_n \rangle = \langle y_n \rangle \implies \mathsf{d}(\langle x_n \rangle, \langle y_n \rangle) = 0$:

$$\begin{aligned}
\mathsf{d}(\langle x_n \rangle, \langle y_n \rangle) &\triangleq \max\left\{\mathsf{p}(x_n, y_n) \mid n = 1, 2, \ldots, N\right\} && \text{by definition of } \mathsf{d}\\
&= \max\left\{\mathsf{p}(x_n, x_n) \mid n = 1, 2, \ldots, N\right\} && \text{by } \langle x_n \rangle = \langle y_n \rangle \text{ hypothesis}\\
&= 0 && \text{because } \mathsf{p} \text{ is \textit{nondegenerate}}
\end{aligned}$$

(c) Proof that $\lparen x_n \rparen = \lparen y_n \rparen \impliedby \mathsf{d}(\lparen x_n \rparen, \lparen y_n \rparen) = 0$:

$$\begin{aligned}
0 &= \mathsf{d}(\lparen x_n \rparen, \lparen y_n \rparen) && \text{by } \mathsf{d}(\lparen x_n \rparen, \lparen y_n \rparen) = 0 \text{ hypothesis}\\
&\triangleq \max\{\mathsf{p}(x_n, y_n) \mid n = 1, 2, \ldots, N\} && \text{by definition of } \mathsf{d}\\
\implies \mathsf{p}(x_n, y_n) &= 0 \text{ for } n = 1, 2, \ldots, N\\
\implies \lparen x_n \rparen &= \lparen y_n \rparen && \text{because } \mathsf{p} \text{ is } \textit{nondegenerate}
\end{aligned}$$

(d) Proof that d satisfies the triangle inequality property:

$$\begin{aligned}
&\mathsf{d}(\lparen x_n \rparen, \lparen y_n \rparen)\\
&\triangleq \max\{\mathsf{p}(x_n, y_n) \mid n = 1, 2, \ldots, N\} && \text{by definition of } \mathsf{d}\\
&\le \max\{\mathsf{p}(x_n, z_n) + \mathsf{p}(z_n, y_n) \mid n = 1, 2, \ldots, N\} && \text{by } \textit{subadditive} \text{ property}\\
&\le \max\{\mathsf{p}(x_n, z_n) \mid n = 1, 2, \ldots, N\} + \max\{\mathsf{p}(z_n, y_n) \mid n = 1, 2, \ldots, N\} && \text{by } \textit{non-negative} \text{ property}\\
&= \max\{\mathsf{p}(z_n, x_n) \mid n = 1, 2, \ldots, N\} + \max\{\mathsf{p}(z_n, y_n) \mid n = 1, 2, \ldots, N\} && \text{by } \textit{symmetry} \text{ property}\\
&\triangleq \mathsf{d}(\lparen z_n \rparen, \lparen x_n \rparen) + \mathsf{d}(\lparen z_n \rparen, \lparen y_n \rparen) && \text{by definition of } \mathsf{d}
\end{aligned}$$

✏

Example 2.11 (Generalized Taxi-Cab Metric). Let X be a set. Let $\lparen x_n \in X \rparen_1^N$ and $\lparen y_n \in X \rparen_1^N$ be N-tuples on X.

EX

$$\{\mathsf{p} \text{ is a metric on } X\} \implies \mathsf{d}(\lparen x_n \rparen, \lparen y_n \rparen) \triangleq \sum_{n=1}^{N} \mathsf{p}(x_n, y_n) \quad \forall x_n, y_n \in X \quad \text{is a metric on } X$$

✎PROOF:

$$\begin{aligned}
\mathsf{d}(\lparen x_n \rparen, \lparen y_n \rparen) &= \sum_{n=1}^{N} \mathsf{p}(x_n, y_n)\\
&= \underbrace{(N^r)\underbrace{\left(\sum_{n=1}^{N} \frac{1}{N}\mathsf{p}^r(x_n, y_n)\right)^{\frac{1}{r}}}_{\text{metric by Theorem 2.10 page 47}}}_{\text{metric by Theorem 2.8 page 43 (see also Example 2.4 page 43)}} && \text{where } r \triangleq 1\\
&\implies \mathsf{d}(\lparen x_n \rparen, \lparen y_n \rparen) \text{ is a metric}
\end{aligned}$$

✏

In the French railway system, a large number of railway lines go through Paris. This means that often the distance from city x to city y is $\mathsf{d}(x, p) + \mathsf{d}(p, y)$ where p represents Paris. This situation gives motivation for the *French Railroad Metric* (next).

Proposition 2.6. [38] *Let X be a set and $\mathsf{p} \in \mathbb{R}^{X \times X}$ be a function.*

PRP

If p is a metric, then the following functions are also metrics:

1. $\mathsf{d}_f(x, y; z) = \begin{cases} 0 & \textit{for } x = y \\ \mathsf{p}(x, z) + \mathsf{p}(z, y) & \textit{for } x \neq y \end{cases}$ (FRENCH RAILWAY METRIC)
2. $\mathsf{d}_{\mathsf{p}}(x, y) = \begin{cases} 0 & \textit{for } x = y \\ \mathsf{p}(0, x) + \mathsf{p}(0, y) & \textit{for } x \neq y \end{cases}$ (POST OFFICE METRIC)

[38] Giles (1987), pages 17,34, Runde (2005), page 25

✎PROOF:

1. Proof the $\mathsf{d}_f(x, y; z)$ is a metric:

Proof that $x = y \implies \mathsf{d}_f(x, y; z) = 0$:

$$\begin{aligned}
\mathsf{d}_f(x, y; z) &= \mathsf{d}_f(x, x) && \text{by left hypothesis}\\
&= 0 && \text{by definition of } \mathsf{d}_f
\end{aligned}$$

Proof that $x = y \impliedby \mathsf{d}_f(x, y; z) = 0$:

$$\begin{aligned}
0 &= \mathsf{d}_f(x, y; z) && \text{by right hypothesis}\\
&\triangleq \begin{cases} 0 & \text{for } x = y \\ \mathsf{p}(x, z) + \mathsf{p}(z, y) & \text{for } x \neq y \end{cases} && \text{by definition of } \mathsf{d}_f\\
&\geq \begin{cases} 0 & \text{for } x = y \\ \mathsf{d}_f(x, y; z) & \text{for } x \neq y \end{cases} && \text{by Definition 2.1}\\
&\geq 0 && \text{by Definition 2.1}\\
\implies \mathsf{d}_f(x, y; z) &= 0 \qquad \forall x, y \in \boldsymbol{X}\\
\implies x &= y && \text{by Definition 2.1}
\end{aligned}$$

Proof that $\mathsf{d}_f(x, y; z) \leq \mathsf{d}_f(u, x) + \mathsf{d}_f(u, y)$:

$$\begin{aligned}
\mathsf{d}_f(x, y; z) &\triangleq \begin{cases} 0 & \text{for } x = y \\ \mathsf{p}(x, z) + \mathsf{p}(z, y) & \text{for } x \neq y \end{cases} && \text{by definition of } \mathsf{d}_f\\
&\leq \begin{cases} 0 & \text{for } x = y \\ \mathsf{p}(u, z) + \mathsf{p}(z, x) & \text{for } x \neq y \end{cases}\\
&\quad + \begin{cases} 0 & \text{for } x = y \\ \mathsf{p}(u, z) + \mathsf{p}(z, y) & \text{for } x \neq y \end{cases} && \text{by Definition 2.1}\\
&= \mathsf{d}_f(u, x) + \mathsf{d}_f(u, y) && \text{by definition of } \mathsf{d}_f
\end{aligned}$$

2. Proof for Post Office Metric: this is a special case of the French Railroad metric (with $z = 0$).

⇒

2.6 Examples

Example 2.12. [39] Let $|\cdot| \in \mathbb{R}^{\vdash R}$ be an *absolute value* (Definition F.4 page 346) function on a *ring* (Definition F.2 page 345) R.

EX The function $\mathsf{d}(x, y) \triangleq |x - y|$ is a *metric*. This metric is called the **usual metric**. It is defined on any ring such as the ring of *real numbers*, *rational numbers*, *complex numbers*, etc.

✎PROOF: Proof by use of Theorem 2.1 (page 27) ...

[39] Davis (2005) page 16

Proof that $x = y \implies \mathsf{d}(x,y) = 0$:

$$\begin{aligned}
\mathsf{d}(x,y) &\triangleq |x-y| && \text{by definition of } \mathsf{d}\\
&= |x-x| && \text{by left hypothesis}\\
&= 0
\end{aligned}$$

Proof that $x = y \impliedby \mathsf{d}(x,y) = 0$:

$$\begin{aligned}
0 &= \mathsf{d}(x,y) && \text{by right hypothesis}\\
&= |x-y| && \text{by definition of } \mathsf{d}\\
\implies x &= y && \text{by property of } |\cdot|
\end{aligned}$$

Proof that $\mathsf{d}(x,y) = \mathsf{d}(z,x) + \mathsf{d}(z,y)$:

$$\begin{aligned}
\mathsf{d}(x,y) &= |x-y| && \text{by definition of } \mathsf{d}\\
&= |x-z+z-y|\\
&\le |x-z| + |z-y| && \text{by \textit{subadditive} property of } |\cdot| \text{ (Definition F.4 page 346)}\\
&= |z-x| + |z-y|\\
&= \mathsf{d}(z,x) + \mathsf{d}(z,y) && \text{by definition of } \mathsf{d} \text{ (Definition 2.1 page 27)}
\end{aligned}$$

⇒

Example 2.13 (The discrete metric). [40] Let X be a set and $\mathsf{d} \in \mathbb{R}^{X \times X}$.

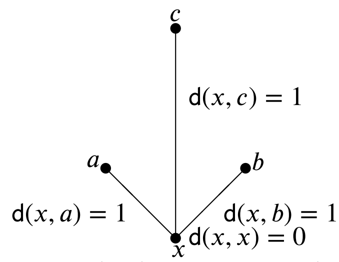
EX

$$\mathsf{d}(x,y) \triangleq \begin{cases} 1 & \text{for } x \ne y \\ 0 & \text{for } x = y \end{cases} \quad \text{is a metric.}$$

d is *not* generated by a norm.

$$\mathsf{B}(0,1) = \{0\}$$

$$\operatorname{diam} \mathsf{B}(0,1) = 0$$



This metric is called the *discrete metric*. It is unusual among metrics because so little is required of the set X. In particular, X does not need to be equipped with any order structure (does not need to be a partially or totally ordered set). The diameter of (X, d) is 1.

PROOF:

1. Proof that $\mathsf{d}(x,y)$ is a metric (using Theorem 2.1 page 27):

Proof that $x = y \implies \mathsf{d}(x,y) = 0$:

$$\begin{aligned}
\mathsf{d}(x,y) &\triangleq \begin{cases} 1 & \text{for } x \ne y \\ 0 & \text{for } x = y \end{cases} && \text{by definition of } \mathsf{d}\\
&= 0 && \text{by } x = y \text{ hypothesis}
\end{aligned}$$

Proof that $x = y \impliedby \mathsf{d}(x,y) = 0$:

$$\begin{aligned}
0 &= \mathsf{d}(x,y) && \text{by } \mathsf{d}(x,y) = 0 \text{ hypothesis}\\
&\triangleq \begin{cases} 1 & \text{for } x \ne y \\ 0 & \text{for } x = y \end{cases} && \text{by definition of } \mathsf{d}\\
\implies x &= y
\end{aligned}$$

[40] Busemann (1955a) page 4 ⟨Comments on the axioms⟩, Giles (1987) page 13, Copson (1968) page 24, Khamsi and Kirk (2001) page 19 ⟨Example 2.1⟩

Proof that $\mathsf{d}(x,y) \le \mathsf{d}(z,x) + \mathsf{d}(z,y)$:

$$\begin{aligned}
\mathsf{d}(x,y) &\triangleq \begin{cases} 1 & \text{for } x \ne y \\ 0 & \text{for } x = y \end{cases} && \text{by definition of } \mathsf{d} \\
&\le \begin{cases} 1 & \text{for } z \ne x \\ 0 & \text{for } z = x \end{cases} + \begin{cases} 1 & \text{for } z \ne y \\ 0 & \text{for } z = y \end{cases} \\
&= \mathsf{d}(z,x) + \mathsf{d}(z,y) && \text{by definition of } \mathsf{d}
\end{aligned}$$

2. Proof that d is not generated by a norm:

$$\begin{aligned}
\|\alpha x\| &= \mathsf{d}(\alpha x, 0) && \text{for some function } \|\cdot\| \\
&= \begin{cases} 1 & \text{for } \alpha x \ne 0 \\ 0 & \text{for } \alpha x = 0 \end{cases} && \text{by definition of } \mathsf{d} \\
&= \begin{cases} 1 & \text{for } x \ne 0 \\ 0 & \text{for } x = 0 \end{cases} \\
&= \mathsf{d}(x,0) && \text{by definition of } \mathsf{d} \\
&\ne |\alpha|\mathsf{d}(x,0) \\
&= |\alpha|\,\|x\|
\end{aligned}$$

3. Proof that $\mathsf{B}(0,1) = \{0\}$:

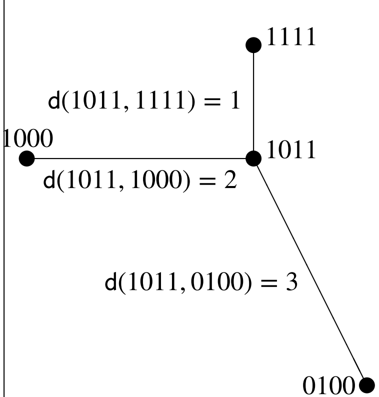
$$\begin{aligned}
\mathsf{B}(0,1) &= \{x \in \boldsymbol{X} | \mathsf{d}(0,x) < 1\} && \text{by definition of open ball } \mathsf{B} \text{ page 30} \\
&= \{0\}
\end{aligned}$$

4. Proof that $\operatorname{diam} \mathsf{B}(0,1) = 0$:

$$\begin{aligned}
\operatorname{diam} \mathsf{B}(0,1) &= \operatorname{diam} \{0\} && \text{by previous result} \\
&= \sup\{\mathsf{d}(x,y) \,|\, x,y \in \{0\}\} && \text{by definition of diam page 30} \\
&= \sup\{\mathsf{d}(0,0)\} \\
&= \sup\{0\} && \text{by } \textit{non-degenerate} \text{ propery of } \mathsf{d} \text{ (Definition 2.1 page 27)} \\
&= 0
\end{aligned}$$

⇨

2.6.1 Metrics on finite sets



Example 2.14 (Hamming distance). Let $\boldsymbol{x} \triangleq (x_1, x_2, \ldots, x_n)$, $\boldsymbol{y} \triangleq (y_1, y_2, \ldots, y_n)$, and $x_i, y_i \in \{0,1\}$.

EX

The *Hamming distance* between $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as

$$\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) \triangleq \sum_{i=1}^{n} \mathsf{p}(x_i, y_i)$$

where

$$\mathsf{p}(x,y) \triangleq \begin{cases} 1 & \text{for } x \ne y \\ 0 & \text{for } x = y \end{cases}.$$

The function d is a metric.

✎PROOF: Example 2.13 (page 51) already showed that p is a metric. And because of this and by Proposition 2.6 (page 49), d is also a metric. ⇨

Example 2.15 (lattice metric). [41] Let $L = (X, \circledcirc, \vee, \wedge)$ be a lattice.

EX

Let $\|x\| : X \to \mathbb{R}$ be a function that satisfies the conditions

1. $x \circledcirc y \implies \|x\| \le \|y\| \qquad \forall x,y \in X$ (*monotonic*)
2. $\|x \vee y\| + \|x \wedge y\| = \|x\| + \|y\| \qquad \forall x,y \in X$

Then $\mathsf{d}(x,y) \triangleq \|x \vee y\| - \|x \wedge y\|$ is a **metric** on L.

✎PROOF:

1. Proof that $\mathsf{d}(x,y) \ge 0$:

$$\begin{aligned}
\mathsf{d}(x,y) &= \|x \vee y\| - \|x \wedge y\| && \text{by definition of } \mathsf{d}(x,y)\\
&\ge 0 && \text{by condition 1 and because } x \vee y \ge x \wedge y
\end{aligned}$$

2. Proof that $\mathsf{d}(x,y) = 0 \implies x = y$:

$$\begin{aligned}
\mathsf{d}(x,y) = 0 &\implies \|x \vee y\| = \|x \wedge y\| && \text{by definition of } \mathsf{d}(x,y)\\
&\implies x \vee y = x \wedge y && \text{by definition of } \|x\| \text{ condition 1}\\
&\implies x = y && \text{by definition of } \vee \text{ and } \wedge
\end{aligned}$$

3. Proof that $\mathsf{d}(x,y) = 0 \impliedby x = y$:

$$\begin{aligned}
\mathsf{d}(x,y) &= \|x \vee y\| - \|x \wedge y\| && \text{by definition of } \mathsf{d}(x,y)\\
&= \|x \vee x\| - \|x \wedge x\| && \text{by right hypothesis}\\
&= \|x\| - \|x\| && \text{by \textit{idempotent} property (Theorem C.3 page 306)}\\
&= 0
\end{aligned}$$

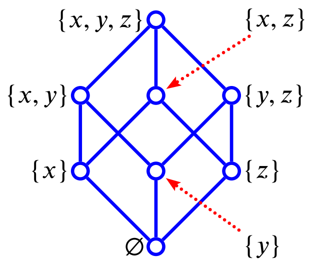
4. Proof that $\mathsf{d}(x,y) = \mathsf{d}(y,x)$:

$$\begin{aligned}
\mathsf{d}(x,y) &= \|x \vee y\| - \|x \wedge y\| && \text{by definition of } \mathsf{d}(x,y)\\
&= \|y \vee x\| - \|y \wedge x\| && \text{by \textit{commutative} property (Theorem C.3 page 306)}\\
&= \mathsf{d}(y,x) && \text{by definition of } \mathsf{d}(x,y)
\end{aligned}$$

5. Proof that $\mathsf{d}(x,y) \le \mathsf{d}(x,z) + \mathsf{d}(z,y)$:

$$\begin{aligned}
&\mathsf{d}(x,z) + \mathsf{d}(z,y)\\
&= \underbrace{(\|x \vee z\| - \|x \wedge z\|)}_{\mathsf{d}(x,y)} + \underbrace{(\|z \vee y\| - \|z \wedge y\|)}_{\mathsf{d}(z,y)} && \text{by definition of } \mathsf{d}(x,y)\\
&= (\|x \vee z\| + \|z \vee y\|) - (\|x \wedge z\| + \|z \wedge y\|)\\
&= (\|(x \vee z) \vee (z \vee y)\| + \|(x \vee z) \wedge (z \vee y)\|)\\
&\quad - (\|(x \wedge z) \vee (z \wedge y)\| + \|(x \wedge z) \wedge (z \wedge y)\|) && \text{by definition of } \|x\|\\
&= (\|(x \vee y) \vee z\| + \|(x \vee z) \wedge (z \vee y)\|)\\
&\quad - (\|(x \wedge z) \vee (z \wedge y)\| + \|(x \wedge y) \wedge z\|) && \text{by Theorem C.3 page 306}\\
&\ge (\|(x \vee y) \vee z\| + \|(x \wedge y) \vee z\|)\\
&\quad - (\|(x \vee y) \wedge z\| + \|(x \wedge y) \wedge z\|) && \text{by \textit{distributive inequality} (Theorem C.6 page 309)}\\
&\ge (\|(x \vee y) \vee z\| + \|(x \vee y) \wedge z\|)\\
&\quad - (\|(x \wedge y) \vee z\| + \|(x \wedge y) \wedge z\|) && \text{by \textit{minimax inequality} (Theorem C.5 page 308)}\\
&= (\|x \vee y\| + \|z\|) - (\|x \wedge y\| + \|z\|) && \text{by definition of } \|\cdot\| \text{ page 83}\\
&= \|x \vee y\| - \|x \wedge y\|\\
&= \mathsf{d}(x,y) && \text{by definition of } \mathsf{d} \text{ page 27}
\end{aligned}$$

[41] Blumenthal (1970) page 25



Example 2.16 (metric on powerset lattice). Let X be a set, 2^X the power set of X and $|A|$ the order of a set A (the number of elements in A). The tupple $(2^X, \cup, \cap; \subseteq)$ is a *lattice*.[42] A metric $\mathsf{d}(A,B) : 2^X \to \mathbb{R}$ can be defined as

$$\mathsf{d}(A,B) \triangleq \|A \cup B\| - \|A \cap B\| \quad \text{where} \quad \|A\| \triangleq |A| \qquad \forall A,B \in 2^X$$

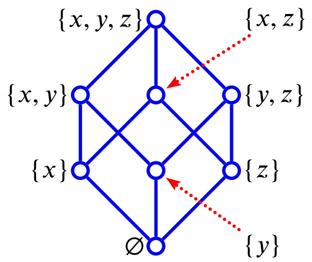
The *Hasse diagram* for $(2^{\{x,y,z\}}, \subseteq)$ is illustrated in the figure to the left.

✎PROOF: Proof that $\|A\|$ satisfies the conditions of a lattice norm:

1. $A \subseteq B \implies |A| \le |B| \implies \|A\| \le \|B\|$
2. $\|A \cup B\| + \|A \cap B\| = |A \cup B| + |A \cap B| = |A| + |B| = \|A\| + \|B\|$

⇰

Example 2.17 (Symmetric difference metric / Fréchet-Nikodym-Aronszayn distance). [43] Let X be a set, 2^X the power set of X, $A \triangle B$ the symmetric difference of sets $A, B \subseteq X$, and $|A|$ the order of a set A (the number of elements in A).[44]



The function $\{\mathsf{d}(A,B) \triangleq |A \triangle B| \qquad \forall A,B \in 2^X\}$ is a metric.

The tupple $(2^X, \subseteq, \cup, \cap)$ is a *lattice*. The *Hasse diagram* for $(2^{\{x,y,z\}}, \subseteq)$ is illustrated in the figure to the left. Notice that the distance (the metric) $\mathsf{d}(A,B)$ between any two sets A and B is just the shortest number of nodes that one must travel to get from A to B.

✎PROOF: The distance between any two sets is simply the number of elements that are different between the two sets. Therefore, this example is essentially the same as Example 2.14 (page 52) (Hamming distance example). ⇰

2.6.2 Metrics on infinite sets

Example 2.18. [45] Let $\mathsf{d} : X \to \mathbb{R}$, $\mathsf{x} : X \to Y$, and $\mathsf{y} : X \to Y$ be functions on a set X. Then (X, d) is a metric space if d is defined as

$$\mathsf{d}(\mathsf{x},\mathsf{y}) = \sup_{t \in A} |\mathsf{x}(t) - \mathsf{y}(t)|$$

✎PROOF:

1. Proof that $\mathsf{d}(\mathsf{x},\mathsf{y}) = 0 \implies \mathsf{x} = \mathsf{y}$:

$$\begin{aligned} 0 &= \mathsf{d}(\mathsf{x},\mathsf{y}) && \text{by left hypothesis} \\ &= \sup_{t \in A} |\mathsf{x}(t) - \mathsf{y}(t)| && \text{by definition of } \mathsf{d} \\ \implies \mathsf{x} &= \mathsf{y} \end{aligned}$$

[42] $2^{\{x,y,z\}}$ *lattice example:* Example C.2 page 310
[44] Deza and Deza (2006) page 25
[45] Dieudonné (1969), page 29

2. Proof that $\mathsf{d}(\mathsf{x},\mathsf{y}) = 0 \impliedby \mathsf{x} = \mathsf{y}$:

$$\begin{aligned}
\mathsf{d}(\mathsf{x},\mathsf{y}) &= \sup_{t\in A} |\mathsf{x}(t) - \mathsf{y}(t)| && \text{by definition of } \mathsf{d}\\
&= \sup_{t\in A} |\mathsf{x}(t) - \mathsf{x}(t)| && \text{by right hypothesis}\\
&= 0
\end{aligned}$$

3. Proof that $\mathsf{d}(\mathsf{x},\mathsf{y}) \le \mathsf{d}(\mathsf{x},\mathsf{z}) + \mathsf{d}(\mathsf{y},\mathsf{z})$:

$$\begin{aligned}
\mathsf{d}(\mathsf{x},\mathsf{y}) &= \sup_{t\in A} |\mathsf{x}(t) - \mathsf{y}(t)| && \text{by definition of } \mathsf{d}\\
&= \sup_{t\in A} |\mathsf{x}(t) - \mathsf{z}(t) + \mathsf{z}(t) - \mathsf{y}(t)|\\
&\le \sup_{t\in A} |\mathsf{x}(t) - \mathsf{z}(t)| + \sup_{t\in A} |\mathsf{z}(t) - \mathsf{y}(t)|\\
&= \sup_{t\in A} |\mathsf{x}(t) - \mathsf{z}(t)| + \sup_{t\in A} |\mathsf{y}(t) - \mathsf{z}(t)|\\
&= \mathsf{d}(\mathsf{x},\mathsf{z}) + \mathsf{d}(\mathsf{y},\mathsf{z}) && \text{by definition of } \mathsf{d}
\end{aligned}$$

✏

Example 2.19 (p-adic metric). [46] For any rational number $x \in \mathbb{Q}$, there exists

1. the sequence of all prime numbers $(p_i)_{i\in\mathbb{N}} = (1, 2, 3, 5, \ldots)$,
2. a sequence of numerator exponents $(n_i \in \mathbb{W})_{i\in\mathbb{N}}$, and ,
3. a sequence of denominator exponents $(m_i \in \mathbb{W})_{i\in\mathbb{N}}$

such that $$x = \frac{\prod_{i\in\mathbb{N}} p_i^{n_i}}{\prod_{i\in\mathbb{N}} p_i^{m_i}}.$$

EX

Then for any prime number p, the pair $(\mathbb{Q}, \mathsf{d}(x,y;p))$ is a metric space where

$$\mathsf{d}(x,y;p) = \begin{cases} 0 & \text{for } x = y \\ \dfrac{1}{p^{\theta(x-y;p)}} & \text{for } x \neq y \end{cases} \qquad \forall x,y\in\mathbb{Q}$$

where the function θ is defined as

$$\theta(x;p) = n_i - m_i \qquad \text{where the value of index } i \text{ is such that } p = p_i.$$

2.6.3 Metrics on n-tuples

Example 2.20. Let $(x_n)_1^N$ and $(y_n)_1^N$ be n-tuples over a set X. Let ϕ be a function in $\mathbb{R}^X$ on X.

EX

$$\left\{\begin{array}{ll}
1. & \phi \text{ is } \textit{convex}\\
2. & \phi \text{ is } \textit{strictly monotonic}\\
3. & \phi(0) = 0\\
4. & \log\circ\phi\circ\exp \text{ is } \textit{convex}\\
5. & \phi(-x) = \phi(x)\ (\textit{even})
\end{array}\right\} \implies \left\{\begin{array}{l}
\mathsf{d}\left((x_n), (y_n)\right) \triangleq \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi(x_n - y_n)\right)\\
\text{is a } \textit{metric}.
\end{array}\right\}$$

In the special case that $\phi(x) \triangleq |x|$ such that

$$\mathsf{d}\left((x_n)_1^N, (y_n)_1^N\right) \triangleq \left(\sum_{n=1}^N \lambda_n |x_n - y_n|^r\right)^{\frac{1}{r}},$$

$\|x\| \triangleq \mathsf{d}(x,x)$ is a *norm*.

[46] Dieudonné (1969), page 30

✎PROOF:

1. Proof that $\left(x_n\right)_1^N = \left(y_n\right)_1^N \implies \mathsf{d}\left(\left(x_n\right)_1^N, \left(y_n\right)_1^N\right) = 0$: by definition of d.
2. Proof that $\left(x_n\right)_1^N = \left(y_n\right)_1^N \impliedby \mathsf{d}\left(\left(x_n\right)_1^N, \left(y_n\right)_1^N\right) = 0$: by *strictly monotonic* property.
3. Proof that $\mathsf{d}\left(\left(x_n\right)_1^N, \left(y_n\right)_1^N\right) \le \mathsf{d}\left(\left(z_n\right)_1^N, \left(x_n\right)_1^N\right) \mathsf{d}\left(\left(z_n\right)_1^N, \left(y_n\right)_1^N\right)$:

$$\begin{aligned}
\mathsf{d}\left(\left(x_n\right), \left(y_n\right)\right) &= \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(x_n - y_n)\right) && \text{by definition of } \mathsf{d} \\
&= \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(x_n - z_n + z_n - y_n)\right) \\
&\le \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(x_n - z_n)\right) + \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(z_n - y_n)\right) && \text{by Theorem 11.2 page 184} \\
&= \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(z_n - x_n)\right) + \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(z_n - y_n)\right) && \text{by } \textit{even} \text{ property} \\
&= \mathsf{d}\left(\left(z_n\right), \left(x_n\right)\right) + \mathsf{d}\left(\left(z_n\right), \left(y_n\right)\right) && \text{by definition of } \mathsf{d}
\end{aligned}$$

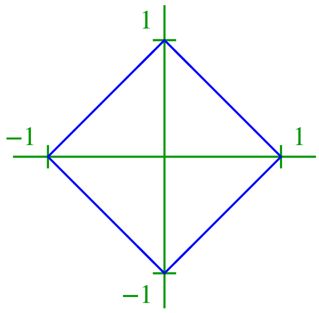
4. Therefore by Theorem 2.1 (page 27), d is a *metric*.
5. $\|\cdot\|$ is a *norm* by Proposition 8.5 (page 137).

✏

Example 2.21 (Taxi-cab metric). [47]

EX

- $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sum_{i=1}^{n} |x_i - y_i|$ *is a metric.*
- d *is generated by a norm.*
- $\mathsf{B}(0,1)$ in $(\mathbb{R}^n, \mathsf{d})$ *is convex.*
- $\operatorname{diam} \mathsf{B}(\boldsymbol{x}, r) = 2r$



✎PROOF:

1. Proof that d is a metric:
   (a) By Example 2.12 (page 50), $\mathsf{p}(x,y) = |x-y|$ is a metric.
   (b) By the definition of d, $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sum_{i=1}^{n} |x_i - y_i|$
   (c) And so d is a *Fréchet product metric* and is a *metric* by Theorem 2.9 (page 46).
2. Proof d is generated by a norm:
   (a) d is generated by a norm if and only if $\|\boldsymbol{x}\| \triangleq \sum_{i=1}^{n} |x_i|$ is a norm.

[47] Deza and Deza (2006), page 240
Dieudonné (1969), page 29

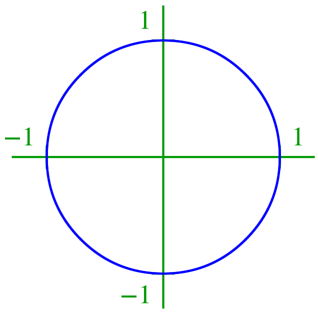
(b) Proof that $\|\boldsymbol{x}\| \triangleq \sum_{i=1}^{n} |x_i|$ is a norm is given by Example 5.2 (page 85).

3. Proof that the ball is convex:
By Theorem 5.4 (page 87), all metrics generated by a norm are convex.

Example 2.22 (Euclidean metric). [48]

EX

- $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}$ *is a metric.*
- d *is generated by a norm.*
- $\mathsf{B}(0, 1)$ in $(\mathbb{R}^2, \mathsf{d})$ *is convex.*
- $\operatorname{diam} \mathsf{B}(\boldsymbol{x}, r) = 2r$



PROOF:

1. Proof that d is a metric:

   (a) By Example 2.12 (page 50), $\mathsf{p}(x, y) = |x - y|$ is a metric.

   (b) By the definition of d, $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}$

   (c) And so d is a *Fréchet product metric* and is a *metric* by Theorem 2.9 (page 46).

2. Proof d is generated by a norm:

   (a) d is generated by a norm if and only if $\|\boldsymbol{x}\| \triangleq \sqrt{\sum_{i=1}^{n} |x_i|^2}$ is a norm.

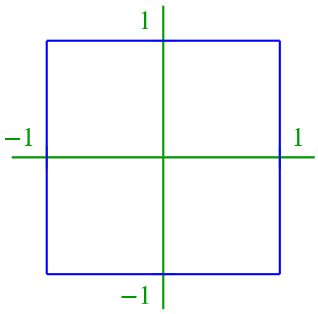
   (b) Proof that $\|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$ is a norm is given by Example 5.2 (page 85).

3. Proof that the ball is convex:
By Theorem 5.4 (page 87), all metrics generated by a norm are convex.

Example 2.23 (Sup metric).

EX

- $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \max\left\{|x_i - y_i| \,|\, i = 1, 2, \ldots, n\right\}$ *is a metric.*
- d *is generated by a norm.*
- $\mathsf{B}(0, 1)$ in $(\mathbb{R}^n, \mathsf{d})$ *is convex.*
- $\operatorname{diam} \mathsf{B}(\boldsymbol{x}, r) = 2\sqrt{2}r$



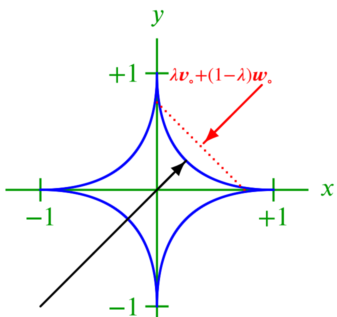
PROOF:

48 Dieudonné (1969), page 29

1. Proof that d is a metric:
   (a) By Example 2.12 (page 50), $\mathsf{p}(x, y) = |x - y|$ is a metric.
   (b) By the definition of d, $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \max\left\{\mathsf{p}(x_i, y_i) | i = 1, 2, \ldots n\right\}$
   (c) And so d is a *Fréchet product metric* and is a *metric* by Theorem 2.9 (page 46).
2. Proof d is generated by a norm:
   (a) d is generated by a norm if and only if $\|\boldsymbol{x}\| \triangleq \max\left\{|x_i| | i = 1, 2, \ldots n\right\}$ is a norm.
   (b) Proof that $\|\boldsymbol{x}\| \triangleq \max\left\{|x_i| | i = 1, 2, \ldots n\right\}$ is a norm is given by Example 5.2 (page 85).
3. Proof that the ball is convex:
   By Theorem 5.4 (page 87), all metrics generated by a norm are convex.

⇨

Example 2.24 (Parabolic metric). [49]



$x^2+y^2-2xy-2x-2y+1=0$
(parabolic equation)

Let X be a set and $\boldsymbol{x} \triangleq (x_k \in X)_1^N$ and $\boldsymbol{y} \triangleq (y_k \in X)_1^N$ be tuples on X.

EX

1. $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sum_{i=1}^{n} \sqrt{|x_i - y_i|}$ *is a metric.*
2. d is *not generated by a norm.*
3. $\mathsf{B}(0, 1)$ in $(\mathbb{R}^n, \mathsf{d})$ is *not convex.*

✎PROOF:

1. Proof that d is a metric:

Proof that $\boldsymbol{x} = \boldsymbol{y} \implies \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) = 0$:

$$\begin{aligned}
\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) &= \mathsf{d}(\boldsymbol{x}, \boldsymbol{x}) && \text{by left hypothesis} \\
&= \sum_{i=1}^{n} \sqrt{|x_i - x_i|} && \text{by definition of d} \\
&= 0
\end{aligned}$$

Proof that $\boldsymbol{x} = \boldsymbol{y} \impliedby \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) = 0$:

$$\begin{aligned}
0 &= \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) && \text{by right hypothesis} \\
&= \sum_{i=1}^{n} \sqrt{|x_i - y_i|} && \text{by definition of d} \\
&\implies x_1 = x_2 \text{ and } y_1 = y_2 && \text{because } |\cdot| \text{ is positive} \\
&\implies \boldsymbol{x} = \boldsymbol{y} && \text{by definitions of } \boldsymbol{v} \text{ and } \boldsymbol{w}
\end{aligned}$$

[49] Norfolk (1991), page 2
http://groups.google.com/group/sci.math/msg/c0eb7e19631c31ea

Proof that $\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) \le \mathsf{d}(\boldsymbol{x},\boldsymbol{z}) + \mathsf{d}(\boldsymbol{z},\boldsymbol{y})$:

$$\begin{aligned}
\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) &= \sum_{i=1}^{n} \sqrt{|x_i - y_i|} && \text{by definition of } \mathsf{d}\\
&\le \sum_{i=1}^{n} \sqrt{|x_i - z_i| + |z_i - y_i|} && \text{by triangle inequality property of usual metric } |\cdot|\\
&= \sum_{i=1}^{n} \sqrt{2}\sqrt{\frac{1}{2}|x_i - z_i| + \frac{1}{2}|z_i - y_i|}\\
&= \sum_{i=1}^{n} \sqrt{2}\left(\frac{1}{2}\sqrt{|x_i - z_i|} + \frac{1}{2}\sqrt{|z_i - y_i|}\right) && \text{by Jensen's inequality page 144}\\
&= \frac{\sqrt{2}}{2}\sum_{i=1}^{n}\left(\sqrt{|x_i - z_i|} + \sqrt{|z_i - y_i|}\right)\\
&\le \sum_{i=1}^{n} \sqrt{|z_i - x_i|} + \sum_{i=1}^{n} \sqrt{|z_i - y_i|}\\
&= \mathsf{d}(\boldsymbol{z},\boldsymbol{x}) + \mathsf{d}(\boldsymbol{z},\boldsymbol{y})
\end{aligned}$$

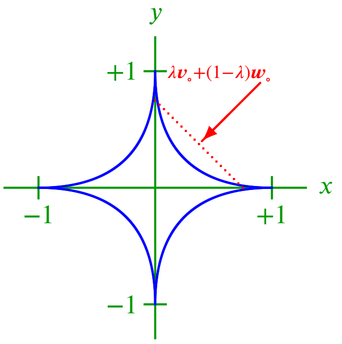
2. Proof d is not generated by a norm:

$$\begin{aligned}
\|\alpha(\boldsymbol{v} - \boldsymbol{w})\| &= \|\alpha\boldsymbol{v} - \alpha\boldsymbol{w}\|\\
&= \mathsf{d}(\alpha\boldsymbol{v}, \alpha\boldsymbol{w}) && \text{for some function } \|\cdot\|\\
&= \sqrt{|\alpha x_1 - \alpha x_2|} + \sqrt{|\alpha y_1 - \alpha y_2|} && \text{by definition of } \mathsf{d}\\
&= \sqrt{|\alpha||x_1 - x_2|} + \sqrt{|\alpha||y_1 - y_2|}\\
&= \sqrt{|\alpha|}\left(\sqrt{|x_1 - x_2|} + \sqrt{|y_1 - y_2|}\right)\\
&= \sqrt{|\alpha|}\mathsf{d}(\boldsymbol{v}, \boldsymbol{w}) && \text{by definition of } \mathsf{d}\\
&= \sqrt{|\alpha|}\;\|\boldsymbol{v} - \boldsymbol{w}\| && \text{by definition of function } \|\cdot\|\\
&\ne |\alpha|\;\|\boldsymbol{v} - \boldsymbol{w}\|\\
\implies & \|\cdot\| \text{ is not a norm.} && \text{by homogeneous property of norms page 83}
\end{aligned}$$

3. Proof that the ball is not convex: Let $\boldsymbol{v} \triangleq \left(\frac{3}{4}, 0\right)$ and $\boldsymbol{w} \triangleq \left(0, \frac{3}{4}\right)$.

$$\begin{aligned}
\mathsf{d}\left(\mathbb{0}, \frac{1}{2}\boldsymbol{v} + (1 - \frac{1}{2})\boldsymbol{w}\right) &= \mathsf{d}\left(\mathbb{0}, \frac{1}{2}\boldsymbol{v} + \frac{1}{2}\boldsymbol{w}\right) && \text{let } \lambda = \frac{1}{2}\\
&= \mathsf{d}\left((0,\,0), \frac{1}{2}\left(\frac{3}{4},\,0\right) + \frac{1}{2}\left(0,\,\frac{3}{4}\right)\right) && \text{by definition of } \boldsymbol{v} \text{ and } \boldsymbol{w}\\
&= \mathsf{d}\left((0,\,0), \left(\frac{3}{8},\,0\right) + \left(0,\,\frac{3}{8}\right)\right)\\
&= \mathsf{d}\left((0,\,0), \left(\frac{3}{8},\,\frac{3}{8}\right)\right)\\
&= \sqrt{\left|0 - \frac{3}{8}\right|} + \sqrt{\left|0 - \frac{3}{8}\right|} && \text{by definition of } \mathsf{d}\\
&= 2\sqrt{\frac{3}{8}}\\
&= \frac{2}{2}\sqrt{\frac{3}{2}}\\
&> 1
\end{aligned}$$

Example 2.25 (Inverse tangent metric). [50]



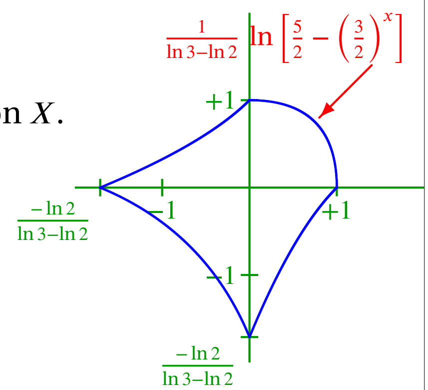
Let X be a set and $\boldsymbol{x} \triangleq (\!(x_k \in X)\!)_1^N$ and $\boldsymbol{y} \triangleq (\!(y_k \in X)\!)_1^N$ be sequences on X.

EX $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sum_{i=1}^{n} \left|\arctan x_i - \arctan y_i\right|$ *is a* METRIC.

✎PROOF:

1. The function $\mathsf{d}(x, y) \triangleq |x - y|$ is a *metric* (the *usual metric*, Example 2.12 page 50).
2. The function $\mathsf{g}(x) \triangleq \arctan(x)$ is *injective* in $\mathbb{R}^{\mathbb{R}}$.
3. Therefore, d is a *Pullback metric* (or g-*transform metric*), and by Theorem 2.6 (page 41), d is a *metric*.

Example 2.26 (Exponential metric).



Let X be a set and $\boldsymbol{x} \triangleq (\!(x_k \in X)\!)_1^N$ and $\boldsymbol{y} \triangleq (\!(y_k \in X)\!)_1^N$ be sequences on X.

EX

1. $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq 2\sum_{i=1}^{n} \left|\left(\frac{3}{2}\right)^{x_i} - \left(\frac{3}{2}\right)^{y_i}\right|$ *is a metric.*
2. d is *not generated by a norm.*
3. $\mathsf{B}(\theta, 1)$ in $(\mathbb{R}^n, \mathsf{d})$ is *not convex.*

✎PROOF:

1. Proof that d is a metric:
   (a) By Example 2.12 (page 50), $\mathsf{p}(x, y) \triangleq |x - y|$ is a metric (the *usual metric*).

[50] Copson (1968), page 25
Khamsi and Kirk (2001) page 14

(b) The function $\mathsf{f}(x) \triangleq 2\left(\left(\frac{3}{2}\right)^x - 1\right)$ is strictly increasing in x. Proof:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\mathsf{f}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}2\left(\left(\frac{3}{2}\right)^x - 1\right) \\
&= 2\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{3}{2}\right)^x \\
&= 2\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\ln\frac{3}{2}}\right)^x \\
&= 2\frac{\mathrm{d}}{\mathrm{d}x}e^{x\ln\frac{3}{2}} \\
&= 2\left(\ln\frac{3}{2}\right)e^{x\ln\frac{3}{2}} \\
&= 2\left(\ln\frac{3}{2}\right)\left(e^{\ln\frac{3}{2}}\right)^x \\
&= 2\left(\ln\frac{3}{2}\right)\left(\frac{3}{2}\right)^x \\
&> 0 \qquad \forall x \in \mathbb{R}
\end{aligned}$$

(c) Therefore, by Theorem 2.6 (page 41), d is a metric.

2. Proof that d is not generated by a norm:

$$\begin{aligned}
\|\alpha(\boldsymbol{x}-\boldsymbol{y})\| &= \|\alpha\boldsymbol{x}-\alpha\boldsymbol{y}\| \\
&= \mathsf{d}(\alpha\boldsymbol{x},\alpha\boldsymbol{y}) && \text{for some function } \|\cdot\| \\
&= 2\sum_{i=1}^{n}\left|\left(\frac{3}{2}\right)^{\alpha x_i} - \left(\frac{3}{2}\right)^{\alpha y_i}\right| && \text{by definition of } \mathsf{d} \\
&\neq 2\sum_{i=1}^{n}\left|\alpha\left(\frac{3}{2}\right)^{x_i} - \alpha\left(\frac{3}{2}\right)^{y_i}\right| \\
&= |\alpha|2\sum_{i=1}^{n}\left|\left(\frac{3}{2}\right)^{x_i} - \left(\frac{3}{2}\right)^{y_i}\right| \\
&= |\alpha|\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) && \text{by definition of } \mathsf{d} \\
&= |\alpha|\ \|\boldsymbol{x}-\boldsymbol{y}\|
\end{aligned}$$

3. Proof that the ball is not convex:

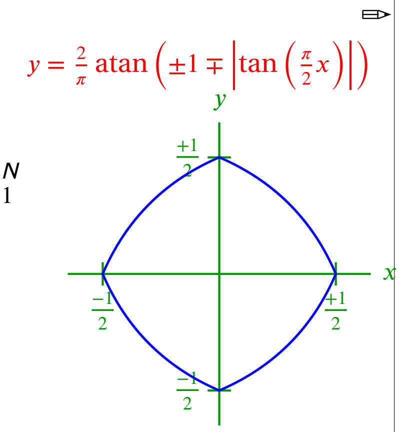
(a) The function $\mathsf{p}(\theta,\boldsymbol{x}) \triangleq 2\sum_{i=1}^{n}\left|\left(\frac{3}{2}\right)^{\theta_i} - \left(\frac{3}{2}\right)^{x_i}\right|$ is not in general convex. Proof:

$$\begin{aligned}
\frac{\partial^2}{\partial x_i^2}\mathsf{p}(\mathbb{0},\boldsymbol{x}) &= \frac{\partial^2}{\partial x_i^2}2\sum_{i=1}^{n}\left|\left(\frac{3}{2}\right)^{0} - \left(\frac{3}{2}\right)^{x_i}\right| \\
&= \frac{\partial^2}{\partial x_i^2}2\left|1 - \left(\frac{3}{2}\right)^{x_i}\right| \\
&= 2\frac{\partial^2}{\partial x_i^2}\left(1 - \left(\frac{3}{2}\right)^{x_i}\right) && \text{for } x_i < 0 \\
&= -2\frac{\partial}{\partial x_i}\left(\ln\frac{3}{2}\right)\left(\frac{3}{2}\right)^{x_i} && \text{for } x_i < 0 \\
&= -2\left(\ln\frac{3}{2}\right)^2\left(\frac{3}{2}\right)^{x_i} && \text{for } x_i < 0 \\
&< 0 && \text{for } x_i < 0
\end{aligned}$$

$\Longrightarrow$ d is not convex

(b) Therefore by Theorem 4.2 (page 81), the ball is not convex.

Example 2.27 (Tangential metric).



Let $X = \{x \in \mathbb{R} | x(-1 : 1)\}$ be a set and $\boldsymbol{x} \triangleq (\!(x_i \in X)\!)_1^N$ and $\boldsymbol{y} \triangleq (\!(y_i \in X)\!)_1^N$ be sequences on X.

EX

1. $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sum_{i=1}^{n} \left| \tan\left(\frac{\pi}{2}x_i\right) - \tan\left(\frac{\pi}{2}y_i\right) \right|$ *is a metric.*
2. d is *not generated by a norm.*
3. $\mathsf{B}(\theta, 1)$ in $(\mathbb{R}^n, \mathsf{d})$ is *convex.*

✎PROOF:

1. Proof that d is a metric:

   (a) By Example 2.12 (page 50), $\mathsf{p}(x, y) \triangleq |x - y|$ is a metric (the *usual metric*).

   (b) The function $\mathsf{f}(x) \triangleq \tan\left(\frac{\pi}{2}x\right)$ is strictly increasing in x. Proof:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\mathsf{f}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\tan\left(\frac{\pi}{2}x\right) \\
&= \frac{\pi}{2}\sec^2\left(\frac{\pi}{2}x\right) \\
&> 0 \qquad \forall x(-1 : 1)
\end{aligned}$$

   (c) Therefore, by Theorem 2.6 (page 41), d is a metric.

2. Proof that d is not generated by a norm:

$$\begin{aligned}
\|\alpha(\boldsymbol{x} - \boldsymbol{y})\| &= \|\alpha\boldsymbol{x} - \alpha\boldsymbol{y}\| \\
&= \mathsf{d}(\alpha\boldsymbol{x}, \alpha\boldsymbol{y}) && \text{for some function } \|\cdot\| \\
&= \sum_{i=1}^{n} \left| \tan\left(\frac{\pi}{2}\alpha x_i\right) - \tan\left(\frac{\pi}{2}\alpha y_i\right) \right| && \text{by definition of } \mathsf{d} \\
&\neq \sum_{i=1}^{n} \left| \alpha\tan\left(\frac{\pi}{2}x_i\right) - \alpha\tan\left(\frac{\pi}{2}y_i\right) \right| \\
&= |\alpha| \sum_{i=1}^{n} \left| \tan\left(\frac{\pi}{2}x_i\right) - \tan\left(\frac{\pi}{2}y_i\right) \right| \\
&= |\alpha|\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) && \text{by definition of } \mathsf{d} \\
&= |\alpha| \;\; \|\boldsymbol{x} - \boldsymbol{y}\|
\end{aligned}$$

3. Proof that the ball is convex:

(a) The function $\mathsf{p}(\theta, \boldsymbol{x}) \triangleq \sum_{i=1}^{n} \left|\tan\left(\frac{\pi}{2}\theta_i\right) - \tan\left(\frac{\pi}{2}x_i\right)\right|$ is convex. Proof:

$$\begin{aligned}
\frac{\partial^2}{\partial x_i^2}\mathsf{d}(\mathbb{0}, \boldsymbol{x}) &= \frac{\partial^2}{\partial x_i^2}\sum_{i=1}^{n}\left|\tan(0) - \tan\left(\frac{\pi}{2}x_i\right)\right| \\
&= \frac{\partial^2}{\partial x_i^2}\left|\tan(0) - \tan\left(\frac{\pi}{2}x_i\right)\right| \\
&= \begin{cases} \frac{\partial^2}{\partial x_i^2}\tan\left(\frac{\pi}{2}x_i\right) & \text{for } x_i \geq 0 \\ \frac{\partial^2}{\partial x_i^2} - \tan\left(\frac{\pi}{2}x_i\right) & \text{for } x_i < 0 \end{cases} \\
&= \begin{cases} \frac{\partial}{\partial x_i}\frac{\pi}{2}\sec^2\left(\frac{\pi}{2}x_i\right) & \text{for } x_i \geq 0 \\ \frac{\partial}{\partial x_i} - \frac{\pi}{2}\sec^2\left(\frac{\pi}{2}x_i\right) & \text{for } x_i < 0 \end{cases} \\
&= \begin{cases} \frac{\pi}{2}\frac{\pi}{2}2\sec^2\left(\frac{\pi}{2}x_i\right)\tan\left(\frac{\pi}{2}x_i\right) & \text{for } x_i \geq 0 \\ -\frac{\pi}{2}\frac{\pi}{2}2\sec^2\left(\frac{\pi}{2}x_i\right)\tan\left(\frac{\pi}{2}x_i\right) & \text{for } x_i < 0 \end{cases} \\
&\geq 0
\end{aligned}$$

(b) Therefore by Theorem 4.2 (page 81), the ball is convex.

⇨

Example 2.28. [51] Let $\mathsf{d}(x, y) = |x - y|^2$ where $|\cdot|$ is the absolute value on $\mathbb{R}$.

- Balls in $(\mathbb{R}, \mathsf{d})$ are *convex* because they are simple intervals.
- But yet d is *not generated by a norm* because

$$\mathsf{d}(ax, ay) = |ax - ay|^2 = |a(x - y)|^2 = |a|^2|x - y|^2 \neq |a||x - y|^2.$$

Example 2.29. [52] Let $\|\cdot\|_2$ be the l_2 norm. Consider the *post office metric*

$$\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \begin{cases} \|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2 & \text{for } \boldsymbol{x} \neq \boldsymbol{y} \\ 0 & \text{for } \boldsymbol{x} = \boldsymbol{y} \end{cases}$$

① The post office metric is *not generated by a norm.*
② The ball generated by the post office metric is in general *not convex.*

✎PROOF:

1. Proof that d is not a norm:

$$\begin{aligned}
\|\mathbb{0}\| &= \|\boldsymbol{x} - \boldsymbol{x}\| \\
&= \mathsf{d}(\boldsymbol{x}, \boldsymbol{x}) && \text{by assumption that } \mathsf{d} \text{ can be generated by a norm } \|\cdot\| \\
&= \|\boldsymbol{x}\|_2 + \|\boldsymbol{x}\|_2 && \text{by definition of the post office metric} \\
&= 2\|\boldsymbol{x}\|_2 \\
&\geq 0 && \text{by positive property of } \|\cdot\| \text{ page 83}
\end{aligned}$$

This implies $\|\cdot\|$ is not a norm because it fails the *non-degenerate* property of norms ($\|\mathbb{0}\| = 0$—see Definition 5.1 page 83) and therefore d is not generated by a norm.

[51] http://groups.google.com/group/sci.math/browse_thread/thread/da44b8a80e97d40f/a977cecea243ad0a
[52] Giles (1987) page 17
http://groups.google.com/group/sci.math/msg/38bb848a9c6d5c29

2. Proof that the ball generated by d is not convex:
   Consider the ball with radius 1 and center $\frac{3}{4}$ generated by the post office metric.

   (a) $\frac{3}{4}$ is in the ball because $\mathsf{d}\left(\frac{3}{4}, \frac{3}{4}\right) = 0 \le 0$

   (b) $\frac{1}{8}$ is in the ball because $\mathsf{d}\left(\frac{3}{4}, \frac{1}{8}\right) = \frac{3}{4} + \frac{1}{8} = \frac{7}{8} \le 1$

   (c) *But* $\frac{1}{2}\cdot\frac{3}{4} + \frac{1}{2}\cdot\frac{1}{8} = \frac{7}{16}$ which is *not* in the ball because $\mathsf{d}\left(\frac{7}{16}, \frac{3}{4}\right) = \frac{7}{16} + \frac{3}{4} = \frac{19}{16} > 1$.

Example 2.30 (The bounded metric). [53] Let X be a set and $\mathsf{d} : X^2 \to \mathbb{R}^{\vdash}$.

EX

$$\mathsf{d}(x,y) \triangleq \frac{\mathsf{p}(x,y)}{1+\mathsf{p}(x,y)} \text{ is a metric.}$$

d is *not* generated by a norm.

$$\mathsf{B}(0,1) = X$$

$$\operatorname{diam}\mathsf{B}(0,1) = \operatorname{diam} X$$

PROOF:

1. Proof that $\mathsf{d}(x,y)$ is a metric (using Theorem 2.1 page 27): Proposition 2.6 (page 49).

2. Proof that d is not generated by a norm:

$$\begin{aligned}
\|\alpha x\| &= \mathsf{d}(\alpha x, 0) && \text{for some function } \|\cdot\| \\
&= \frac{\mathsf{p}(\alpha x,0)}{1+\mathsf{p}(\alpha x,0)} \\
&= \frac{|\alpha|\mathsf{p}(x,0)}{1+|\alpha|\mathsf{p}(x,0)} && \text{assuming } \mathsf{p} \text{ is homogeneous} \\
&= |\alpha|\left[\frac{\mathsf{p}(x,0)}{1+|\alpha|\mathsf{p}(x,0)}\right] \\
&\ne |\alpha|\left[\frac{\mathsf{p}(x,0)}{1+\mathsf{p}(x,0)}\right] \\
&= |\alpha|\mathsf{d}(x,0) \\
&= |\alpha|\,\|x\|
\end{aligned}$$

3. Proof that $\mathsf{B}(0,1) = \{0\}$:

$$\begin{aligned}
\mathsf{B}(0,1) &= \{x \in \boldsymbol{X} | \mathsf{d}(0,x) < 1\} && \text{by definition of open ball } \mathsf{B} \text{ page 30} \\
&= \left\{x \in \boldsymbol{X} | \frac{\mathsf{p}(x,y)}{1+\mathsf{p}(x,y)} < 1\right\} && \text{by definition } \mathsf{d} \\
&= \{x \in \boldsymbol{X}\} \\
&= X
\end{aligned}$$

4. Proof that $\operatorname{diam}\mathsf{B}(0,1) = \operatorname{diam} X$:

$$\operatorname{diam}\mathsf{B}(0,1) = \operatorname{diam} X \qquad \text{by previous result}$$

[53] Copson (1968), page 22

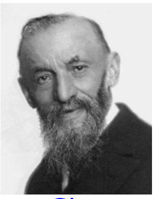
2.7 Literature

Literature survey:

- general reference books about *metric spaces*:
  - Copson (1968)
  - Giles (1987)
- more sophisticated references:
  - Blumenthal (1970)
  - Busemann (1955a)
    - Busemann (1955b)
- *metric spaces* and *convexity*:
  - Khamsi and Kirk (2001)
- "*length spaces*":
  - Burago et al. (2001)
- spaces of *metric spaces*:
  - Burago et al. (2001)
- Very large collections of metric examples:
  - Deza and Deza (2006)
  - Deza and Deza (2009)

CHAPTER 3

LINEAR SPACES



❝*The geometric calculus, in general, consists in a system of operations on geometric entities, and their consequences, analogous to those that algebra has on the numbers. It permits the expression in formulas of the results of geometric constructions, the representation with equations of propositions of geometry, and the substitution of a transformation of equations for a verbal argument.*❞

Giuseppe Peano (1858–1932), Italian mathematician, credited with being one of the first to introduce the concept of the *linear space* (*vector space*).[1]

3.1 Definition and basic results

A *metric space* (Definition 2.1 page 27) is a *set* together with nothing else save a *metric* that gives the space a *topology* (Definition 1.1 page 3). A *linear space* (next definition) in general has no topology but does have some additional *algebraic* structure (APPENDIX F page 345) that is useful in generalizing a number of mathematical concepts. If one wishes to have both algebraic structure and a topology, then this can be accomplished by appending a *topology* to a *linear space* giving a *topological linear space* (Definition 4.1 page 79), a *metric* giving a *metric linear space* (Definition 2.1 page 27), an *inner product* giving an *inner product space* (Definition 6.1 page 95), or a *norm* giving a *normed linear space* (Definition 5.1 page 83).

Definition 3.1. [2] *Let* $(\mathbb{F}, +, \cdot, 0, 1)$ *be a* FIELD *(Definition F.5 page 346). Let* X *be a set, let* $+$ *be an* OPERATOR *(Definition 14.1 page 213) in* X^{X^2}, *and let* $\otimes$ *be an operator in* $X^{\mathbb{F}\times X}$.

---

[1] quote: Peano (1888b), page ix
image http://en.wikipedia.org/wiki/File:Giuseppe_Peano.jpg, public domain

[2] Kubrusly (2001) pages 40–41 ⟨Definition 2.1 and following remarks⟩, Haaser and Sullivan (1991), page 41, Halmos (1948), pages 1–2, Peano (1888a) ⟨Chapter IX⟩, Peano (1888b), pages 119–120, Banach (1922) pages 134–135

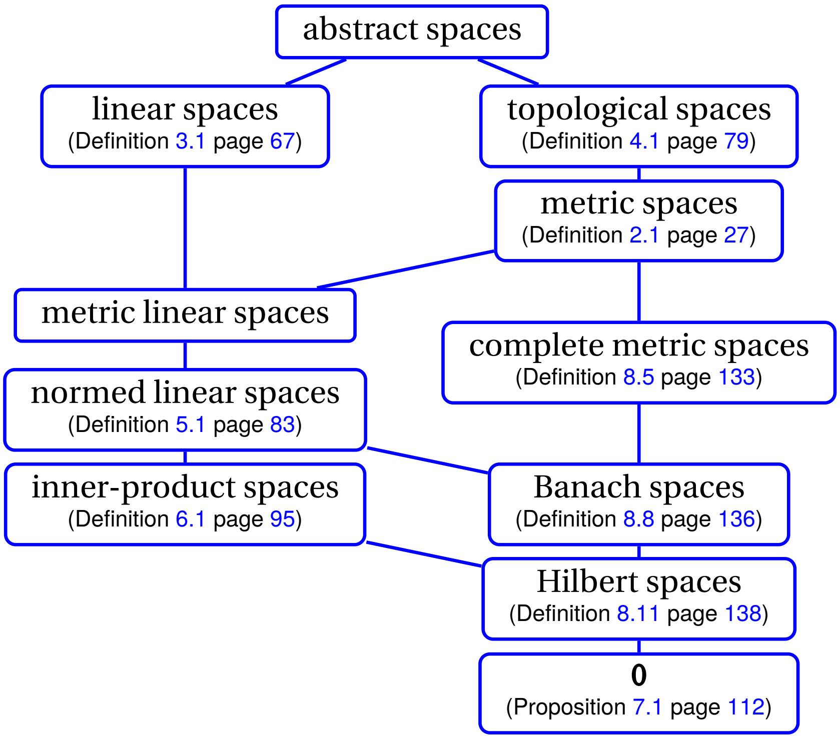


Figure 3.1: Lattice of mathematical spaces

DEF *The structure* $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *is a* ***linear space*** *over* $(\mathbb{F}, +, \cdot, 0, 1)$ *if*

1.	$\exists \mathbb{0} \in X$ such that	$x + \mathbb{0}$	$=$	x	$\forall x \in X$	($+$ IDENTITY) *
2.	$\exists y \in X$ such that	$x + y$	$=$	$\mathbb{0}$	$\forall x \in X$	($+$ INVERSE)
3.		$(x + y) + z$	$=$	$x + (y + z)$	$\forall x,y,z \in X$	($+$ *is* ASSOCIATIVE)
4.		$x + y$	$=$	$y + x$	$\forall x,y \in X$	($+$ *is* COMMUTATIVE)
5.		$1 \cdot x$	$=$	x	$\forall x \in X$	($\cdot$ IDENTITY)
6.		$\alpha \cdot (\beta \cdot x)$	$=$	$(\alpha \cdot \beta) \cdot x$	$\forall \alpha,\beta \in S$ *and* $x \in X$	($\cdot$ ASSOCIATES *with* $\cdot$)
7.		$\alpha \cdot (x + y)$	$=$	$(\alpha \cdot x) + (\alpha \cdot y)$	$\forall \alpha \in S$ *and* $x,y \in X$	($\cdot$ DISTRIBUTES *over* $+$)
8.		$(\alpha + \beta) \cdot x$	$=$	$(\alpha \cdot x) + (\beta \cdot x)$	$\forall \alpha,\beta \in S$ *and* $x \in X$	($\cdot$ PSEUDO-DISTRIBUTES *over* $+$)

The set X *is called the* ***underlying set****. The elements of* X *are called* ***vectors****. The elements of* $\mathbb{F}$ *are called* ***scalars****. A* LINEAR SPACE *is also called a* ***vector space****. If* $\mathbb{F} \triangleq \mathbb{R}$*, then* $\boldsymbol{\Omega}$ *is a* ***real linear space****. If* $\mathbb{F} \triangleq \mathbb{C}$*, then* $\boldsymbol{\Omega}$ *is a* ***complex linear space****.*

Definition 3.2. *Let* $\boldsymbol{L}_1 \triangleq \left(X_1, +, \cdot, (\mathbb{F}_1, \dot{+}, \dot{\times})\right)$ *and* $\boldsymbol{L}_2 \triangleq \left(X_2, +, \cdot, (\mathbb{F}_2, \dot{+}, \dot{\times})\right)$.

DEF $\boldsymbol{\Omega}_2$ *is a* ***linear subspace*** *of* $\boldsymbol{\Omega}_1$ *if*

1. $\boldsymbol{L}_1$ *is a* LINEAR SPACE (*Definition 3.1 page 67*) *and*
2. $\boldsymbol{L}_2$ *is a* LINEAR SPACE (*Definition 3.1 page 67*) *and*
3. $\mathbb{F}_2 \subseteq \mathbb{F}_1$ *and*
4. $X_2 \subseteq X_1$ *and*

Remark 3.1. [3] By the first four conditions ($*$]) listed in Definition 3.1, $(X, +)$ is a **commutative group** (or **abelian group**).

[3] Akhiezer and Glazman (1993), page 1, Haaser and Sullivan (1991), page 41

Often when discussing a linear space, the operator $\cdot$ is simply expressed with juxtaposition (e.g. $\alpha\boldsymbol{x}$ is equivalent to $\alpha\cdot\boldsymbol{x}$). In doing this, there is no risk of ambiguity between scalar-vector multiplication and scalar-scalar multiplication because the operands uniquely identify the precise operator.[4]

Example 3.1 (tuples in $\mathbb{F}^N$). [5] Let $(\!(x_n)\!)_1^N$ be an *N-tuple* (Definition 8.1 page 127) over a *field* (Definition F.5 page 346) $(\mathbb{F},+,\cdot,0,1)$.

EX

Let
$$\begin{array}{rcll} X & \triangleq & \left\{(\!(x_n)\!)_1^N \,|\, x_n\in\mathbb{F}\right\} & \text{and}\\ (\!(x_n)\!)_1^N + (\!(y_n)\!)_1^N & \triangleq & (\!(x_n \dot{+} y_n)\!)_1^N \quad \forall (\!(x_n)\!)\in X & \text{and}\\ \alpha\cdot(\!(x_n)\!)_1^N & \triangleq & (\!(\alpha \dot{\times} x_n)\!)_1^N \quad \forall (\!(x_n)\!)\in X, \alpha\in\mathbb{F}. & \end{array}$$
Then the structure $\left(X,\,+,\,\cdot,\,(\mathbb{F},\,\dot{+},\,\dot{\times})\right)$ is a *linear space*.

Example 3.2 (real numbers). [6] Let $(\mathbb{R},+,\cdot,0,1)$ be the field of real numbers.

EX

The structure $(\mathbb{R},\,+,\,\cdot,\,(\mathbb{R},\,+,\,\cdot))$ is a *linear space*.
That is, the field of real numbers forms a linear space over itself.

Example 3.3 (functions). [7] Let $(\mathbb{F},+,\cdot,0,1)$ be a field. Let Y^X be the set of all functions with domain X and range Y.

EX

Let
$$\begin{array}{rcll} [\mathsf{f}+\mathsf{g}](x) & \triangleq & \mathsf{f}(x)+\mathsf{g}(x) \quad \forall \mathsf{f},\mathsf{g}\in Y^X & (\textit{pointwise addition}) \quad \text{and}\\ [\alpha\cdot\mathsf{f}](x) & \triangleq & \alpha\cdot[\mathsf{f}(x)] \quad \forall \mathsf{f}\in Y^X, \alpha\in\mathbb{F}. & \end{array}$$
Then the structure $\left(Y^X,\,+,\,\cdot,\,(\mathbb{F},\,\dot{+},\,\dot{\times})\right)$ is a *linear space*.

Example 3.4 (functions onto $\mathbb{F}$). [8] Let $(\mathbb{F},+,\cdot,0,1)$ be a field. Let $\mathbb{F}^{\boldsymbol{X}}$ be the set of all functions with domain X and range $\mathbb{F}$.

EX

Let
$$\begin{array}{rcll} [\mathsf{f}+\mathsf{g}](x) & \triangleq & \mathsf{f}(x)+\mathsf{g}(x) \quad \forall \mathsf{f},\mathsf{g}\in \mathbb{F}^{\boldsymbol{X}} & (\textit{pointwise addition}) \quad \text{and}\\ [\alpha\cdot\mathsf{f}](x) & \triangleq & \alpha\cdot[\mathsf{f}(x)] \quad \forall \mathsf{f}\in \mathbb{F}^{\boldsymbol{X}}, \alpha\in\mathbb{F}. & \end{array}$$
Then the structure $\left(\mathbb{F}^{\boldsymbol{X}},\,+,\,\cdot,\,(\mathbb{F},\,+,\,\cdot)\right)$ is a *linear space*.

Theorem 3.1 (Additive identity properties). [9] *Let* $\left(X,\,+,\,\cdot,\,(\mathbb{F},\,\dot{+},\,\dot{\times})\right)$ *be a linear space,* 0 *the* ADDITIVE IDENTITY ELEMENT *(Definition F.1 page 345) with respect to* $\dot{+}$*, and* $\mathbb{0}$ *the* ADDITIVE IDENTITY ELEMENT *with respect to* $+$.

THM

1. $0\boldsymbol{x} = \mathbb{0} \quad \forall \boldsymbol{x}\in X$
2. $\alpha\mathbb{0} = \mathbb{0} \quad \forall \alpha\in\mathbb{F}$
3. $\alpha\boldsymbol{x} = \mathbb{0} \implies \alpha = 0$ *or* $\boldsymbol{x} = \mathbb{0}$
4. $\boldsymbol{x}+\boldsymbol{x} = \boldsymbol{x} \implies \boldsymbol{x} = \mathbb{0}$
5. $\alpha \neq 0$ *and* $\boldsymbol{x}\neq\mathbb{0} \implies \alpha\boldsymbol{x}\neq\mathbb{0}$

✎PROOF:

[4] *Operator overload* is a technique in which two fundamentally different operators or functions share the same symbol or label. It is inherent in the programming language *C++* and is therein called *operator overload*. In C++, you can define two (or more) operators or functions that share the same symbol or name, but yet are completely different. Two such operators (or functions) are distinguished from each other by the type of their operands. So for example, in C++, you can define an $m\times n$ matrix *type* and use operator overload to define a $+$ operator that operates on this new matrix type. So if variables x and y are of floating point type and A and B are of the matrix type, you can then add either type using the same syntax style:

x+y (add two floating point numbers)
A+B (add two matrices)

Even though both of these operations "look" the same, they are of course fundamentally different.

[5] Kubrusly (2001) page 41 ⟨Example 2D⟩

[6] Kubrusly (2001) page 41 ⟨Example 2D⟩, Hamel (1905)

[7] Kubrusly (2001) page 42 ⟨Example 2F⟩

[8] Kubrusly (2001) page 41 ⟨Example 2E⟩

[9] Berberian (1961) page 6 ⟨Theorem 1⟩, Michel and Herget (1993) page 77

1. Proof that $0\boldsymbol{x} = \mathbb{0}$:

$$\begin{aligned}
0\boldsymbol{x} &= 0\boldsymbol{x} + 0\mathbb{0} && \text{by definition of + additive identity element}\\
&= 0\boldsymbol{x} + 0\boldsymbol{x} + (-0\boldsymbol{x}) && \text{by definition of + additive inverse}\\
&= (0+0)\boldsymbol{x} + (-0\cdot\boldsymbol{x}) && \text{by definition of + additive identity element}\\
&= 0\boldsymbol{x} + (-0\boldsymbol{x}) && \text{by Definition 3.1 property 4}\\
&= \mathbb{0} && \text{by definition of + additive identity element}
\end{aligned}$$

2. Proof that $\alpha\mathbb{0} = \mathbb{0}$:

$$\begin{aligned}
\alpha\mathbb{0} &= \alpha(0\boldsymbol{x}) && \text{by item 1}\\
&= (\alpha 0)\boldsymbol{x} && \text{by Definition 3.1 property 6}\\
&= 0\boldsymbol{x}\\
&= \mathbb{0} && \text{by item 1}
\end{aligned}$$

3. Proof that $\alpha \neq 0$ and $\boldsymbol{x} \neq \mathbb{0} \implies \alpha\boldsymbol{x} \neq \mathbb{0}$: Suppose $\alpha\boldsymbol{x} = \mathbb{0}$. Then

$$\begin{aligned}
\boldsymbol{x} &= \left(\frac{1}{\alpha}\alpha\right)\boldsymbol{x}\\
&= \frac{1}{\alpha}(\alpha\boldsymbol{x})\\
&= \frac{1}{\alpha}\mathbb{0}\\
&= \mathbb{0} && \text{by item 2}\\
\implies \boldsymbol{x} &= \mathbb{0}
\end{aligned}$$

This is a *contradiction* and so $\alpha\boldsymbol{x} \neq \mathbb{0}$.

4. Proof that $\alpha\boldsymbol{x} = \mathbb{0} \implies \alpha = 0$ or $\boldsymbol{x} = \mathbb{0}$: contrapositive argument of item 3

5. Proof that $\boldsymbol{x} + \boldsymbol{x} = \boldsymbol{x} \implies \boldsymbol{x} = \mathbb{0}$:

$$\begin{aligned}
\boldsymbol{x} &= \boldsymbol{x} + \mathbb{0} && \text{by \textit{additive identity} property (Definition 3.1 page 67)}\\
&= \boldsymbol{x} + [\boldsymbol{x} + (-\boldsymbol{x})] && \text{by \textit{additive inverse} property (Definition 3.1 page 67)}\\
&= [\boldsymbol{x} + \boldsymbol{x}] + (-\boldsymbol{x}) && \text{by \textit{associative} property (Definition 3.1 page 67)}\\
&= \boldsymbol{x} + (-\boldsymbol{x}) && \text{by left hypothesis}\\
&= \mathbb{0} && \text{by \textit{additive inverse} property (Definition 3.1 page 67)}
\end{aligned}$$

✏

Definition 3.3. [10] *Let $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ be a linear space with vectors $\boldsymbol{x}, \boldsymbol{y} \in X$. Let $-\boldsymbol{y}$ be the additive inverse of $\boldsymbol{y}$ such that $\boldsymbol{y} + (-\boldsymbol{y}) = \mathbb{0}$.*

DEF *The **difference** of $\boldsymbol{x}$ and $\boldsymbol{y}$ is $\boldsymbol{x} + (-\boldsymbol{y})$ and is denoted $\boldsymbol{x} - \boldsymbol{y}$.*

Theorem 3.2 (Additive inverse properties). [11] *Let $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ be a linear space, $\mathbb{0}$ the* ADDITIVE IDENTITY ELEMENT *with respect to +, and $-\boldsymbol{x}$ the* ADDITIVE INVERSE *(Definition F.1 page 345) of $\boldsymbol{x}$ with respect to +.*

THM

1.	$\boldsymbol{x} + \boldsymbol{y}$	$=$	$\mathbb{0} \implies \boldsymbol{x} = -\boldsymbol{y}$	$\forall \boldsymbol{x},\boldsymbol{y} \in X$	*(additive inverse is* UNIQUE*)*
2.	$(-\alpha)\boldsymbol{x}$	$=$	$-(\alpha\boldsymbol{x}) = \alpha(-\boldsymbol{x})$	$\forall \boldsymbol{x} \in X, \alpha \in \mathbb{F}$	
3.	$\alpha(\boldsymbol{x} - \boldsymbol{y})$	$=$	$\alpha\boldsymbol{x} - \alpha\boldsymbol{y}$	$\forall \boldsymbol{x},\boldsymbol{y} \in X, \alpha \in \mathbb{F}$	(DISTRIBUTIVE)
4.	$(\alpha - \beta)\boldsymbol{x}$	$=$	$\alpha\boldsymbol{x} - \beta\boldsymbol{x}$	$\forall \boldsymbol{x} \in X, \alpha,\beta \in \mathbb{F}$	(DISTRIBUTIVE)

[10] Berberian (1961) page 7 ⟨Definition 1⟩

[11] Berberian (1961) page 7 ⟨Corollary 1⟩, Michel and Herget (1993) page 77, Peano (1888a) ⟨Chapter IX⟩, Peano (1888b) pages 119–120, Banach (1922) pages 134–135

✎PROOF:

1. Proof that $x + y = 0 \implies x = -y$:

$$\begin{aligned} x &= x - \mathbb{0} \\ &= x - (x + y) && \text{by left hypothesis} \\ &= (x - x) - y \\ &= \mathbb{0} - y \\ &= -y \end{aligned}$$

2. Proof that $(-\alpha)x = -(\alpha x)$:

$$\begin{aligned} \mathbb{0} &= 0x && \text{by Theorem 3.1 page 69} \\ &= (\alpha - \alpha)x && \text{by field property of } \mathbb{F} \\ &= [\alpha + (-\alpha)]x && \text{by field property of } \mathbb{F} \\ &= \alpha x + (-\alpha)x && \text{by Definition 3.1 page 67} \\ \implies -(\alpha x) &= (-\alpha)x && \text{by item (1) page 71} \end{aligned}$$

3. Proof that $\alpha(-x) = -(\alpha x)$:

$$\begin{aligned} \mathbb{0} &= \alpha\mathbb{0} && \text{by Theorem 3.1 page 69} \\ &= \alpha[x + (-x)] && \text{by definition of additive identity element } -x \\ &= \alpha x + \alpha(-x) && \text{by Definition 3.1 page 67} \\ &= \alpha x + \alpha(-x) \\ \implies -(\alpha x) &= \alpha(-x) && \text{by item (1) page 71} \end{aligned}$$

4. Proof that $\alpha(x - y) = \alpha x - \alpha y$:

$$\begin{aligned} \alpha(x - y) &= \alpha\big[x + (-y)\big] && \text{by Definition 3.3 page 70} \\ &= \alpha x + \alpha(-y) && \text{by Definition 3.1 page 67} \\ &= \alpha x + (-\alpha y) && \text{by item (3) page 71} \\ &= \alpha x - \alpha y && \text{by Definition 3.3 page 70} \end{aligned}$$

5. Proof that $(\alpha - \beta)x = \alpha x - \beta x$:

$$\begin{aligned} (\alpha - \beta)x &= [\alpha + (-\beta)]x && \text{by field properties of } \mathbb{F} \\ &= \alpha x + (-\beta)x && \text{by Definition 3.1} \\ &= \alpha x + [-(\beta x)] && \text{by item (2) page 71} \\ &= \alpha x - (\beta x) && \text{by Definition 3.3 page 70} \end{aligned}$$

⇰

Theorem 3.3. [12] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a linear space,* $\mathbb{0}$ *the additive identity element with respect to* $+$*, and* $-x$ *additive inverse of* x *with respect to* $+$*.*

THM

1.	$\alpha x = \alpha y$ and $\alpha \neq 0$	$\implies$	x	$=$	y	$\forall x,y \in X$
2.	$\alpha x = \beta x$ and $x \neq \mathbb{0}$	$\implies$	α	$=$	β	$\forall x,y \in X,\ \alpha,\beta \in \mathbb{F}$
3.	$z + x = z + y$	$\implies$	x	$=$	y	$\forall x,y,z \in X$

[12] Michel and Herget (1993) page 77, Peano (1888a) ⟨Chapter IX⟩, Peano (1888b) pages 119–120, Banach (1922) pages 134–135

✎PROOF:

1. Proof that $\alpha x = \alpha y$ and $\alpha \neq 0 \implies x = y$:

$$\begin{aligned}
\mathbb{0} &= \frac{1}{\alpha}(\mathbb{0}) && \text{by left hypothesis } (\alpha \neq 0)\\
&= \frac{1}{\alpha}\Big(\alpha x - \alpha y\Big) && \text{by left hypothesis } (\alpha x = \alpha y)\\
&= \frac{1}{\alpha}\alpha\Big(x - y\Big) && \text{by Definition 3.1 page 67}\\
&= x - y
\end{aligned}$$

2. Proof that $\alpha x = \beta x$ and $x \neq \mathbb{0} \implies \alpha = \beta$:

$$\begin{aligned}
\mathbb{0} &= \alpha x + (-\alpha x) && \text{by definition of additive inverse}\\
&= \beta x + (-\alpha x) && \text{by left hypothesis}\\
&= \beta x + (-\alpha)x && \text{by Theorem 3.2 page 70}\\
&= [\beta + (-\alpha)]x && \text{by Definition 3.1 page 67}\\
\implies \beta - \alpha &= 0 && \text{by Theorem 3.1 page 69}\\
\implies \alpha &= \beta && \text{by field properties of } \mathbb{F}
\end{aligned}$$

3. Proof that $z + x = z + y \implies x = y$:

$$\begin{aligned}
\mathbb{0} &= (z + x) - (z + y)\\
&= (x + z) - (z + y) && \text{by Definition 3.1 property 1}\\
&= (x + z) + [(-1)z + (-1)y] && \text{by Definition 3.1 property 3}\\
&= (x + z) + (-z - y) && \text{by previous result 2.}\\
&= x + (z - z) - y\\
&= x - y
\end{aligned}$$

✏

3.2 Order on Linear Spaces

Definition 3.4. [13] *Let $\Omega \triangleq (X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}))$ be a real linear space.*

DEF

*The pair $(\Omega, \leq)$ is an **ordered linear space** if*

1. $x \leq y \implies x + z \leq y + z \quad \forall z \in X$ *and*
2. $x \leq y \implies \alpha x \leq \alpha y \quad \forall \alpha \in \mathbb{F}$

*A vector x is **positive** if $\mathbb{0} \leq x$.*
*The **positive cone** X^+ of $(X, \leq)$ is the set $X^+ \triangleq \{x \in X | \mathbb{0} \leq x\}$.*

Definition 3.5. [14] *Let $(X, \leq)$ be an ordered linear space.*

DEF

*The tuple $L \triangleq (X, \vee, \wedge; \leq)$ is a **Riesz space** if L is a **lattice**.*
*A RIESZ SPACE is also called a **vector lattice**.*

Theorem 3.4. [15] *Let $(X, \vee, \wedge; \leq)$ be a Riesz space* (Definition 3.5 page 72).

THM

$x \vee y = -\big[(-x) \wedge (-y)\big]$	$x \wedge y = -\big[(-x) \vee (-y)\big]$	$\forall x,y \in X$
$x + (y \vee z) = (x + y) \vee (x + z)$	$x + (y \wedge z) = (x + y) \wedge (x + z)$	$\forall x,y,z \in X$
$\alpha(x \vee y) = (\alpha x) \vee (\alpha y)$	$\alpha(x \wedge y) = (\alpha x) \wedge (\alpha y)$	$\forall x,y \in X, \alpha \geq 0$
$x + y = (x \wedge y) + (x \vee y)$		$\forall x,y \in X, \alpha \in \mathbb{F}$

[13] Aliprantis and Burkinshaw (2006) pages 1–2
[14] Aliprantis and Burkinshaw (2006) page 2
[15] Aliprantis and Burkinshaw (2006) page 3 ⟨Theorem 1.2⟩

✎PROOF:

1. Proof that $x \vee y = -\left[(-x) \wedge (-y)\right]$:

$\begin{aligned}(-x)\wedge(-y) &\le -x\\ x &\le -\left[(-x)\wedge(-y)\right]\\ x\vee y &\le -\left[(-x)\wedge(-y)\right]\end{aligned}$	$\begin{aligned}(-x)\wedge(-y) &\le -y\\ y &\le -\left[(-x)\wedge(-y)\right]\end{aligned}$
$\begin{aligned}x &\le x\vee y\\ -(x\vee y) &\le -x\\ -(x\vee y) &\le (-x)\wedge(-y)\\ -\left[(-x)\wedge(-y)\right] &\le x\vee y\end{aligned}$	$\begin{aligned}y &\le x\vee y\\ -(x\vee y) &\le -y\end{aligned}$

2. Proof that $x \wedge y = -\left[(-x) \vee (-y)\right]$:

$x \vee y$	$=$	$-\left[(-x) \wedge (-y)\right]$	by item (1)
$(-x) \vee (-y)$	$=$	$-\left[(-(-x)) \wedge (-(-y))\right]$	replace x with $-x$ and y with y
$(-x) \vee (-y)$	$=$	$-\left[x \wedge y\right]$	$-(-x) = x$
$-\left[x \wedge y\right]$	$=$	$(-x) \vee (-y)$	by symmetry of $=$ relation
$x \wedge y$	$=$	$-\left[(-x) \vee (-y)\right]$	multipy both sides by -1

3. Proof that $x + (y \vee z) = (x + y) \vee (x + z)$:

$\begin{aligned}x+y &\le x+(y\vee z)\\ (x+y)\vee(x+z) &\le x+(y\vee z)\end{aligned}$	$\begin{aligned}x+z &\le x+(y\vee z)\end{aligned}$
$\begin{aligned}y &= -x+(x+y)\\ &\le -x+\left[(x+y)\vee(x+z)\right]\\ y\vee z &\le -x+\left[(x+y)\vee(x+z)\right]\\ x+(y\vee z) &\le (x+y)\vee(x+z)\end{aligned}$	$\begin{aligned}z &= -x+(x+z)\\ &\le -x+\left[(x+y)\vee(x+z)\right]\end{aligned}$

4. Proof that $x + (y \wedge z) = (x + y) \wedge (x + z)$:

$\begin{aligned}x+y &\ge x+(y\wedge z)\\ (x+y)\wedge(x+z) &\ge x+(y\wedge z)\end{aligned}$	$\begin{aligned}x+z &\ge x+(y\wedge z)\end{aligned}$
$\begin{aligned}y &= -x+(x+y)\\ &\ge -x+\left[(x+y)\wedge(x+z)\right]\\ y\wedge z &\ge -x+\left[(x+y)\wedge(x+z)\right]\\ x+(y\wedge z) &\ge (x+y)\wedge(x+z)\end{aligned}$	$\begin{aligned}z &= -x+(x+z)\\ &\ge -x+\left[(x+y)\wedge(x+z)\right]\end{aligned}$

5. Proof that $\alpha(x \vee y) = (\alpha x) \vee (\alpha y)$ for $\alpha \ge 0$:

$\begin{aligned}x &\le x\vee y\\ \alpha x &\le \alpha(x\vee y)\\ (\alpha x)\vee(\alpha y) &\le \alpha(x\vee y)\end{aligned}$	$\begin{aligned}y &\le x\vee y\\ \alpha y &\le \alpha(x\vee y)\end{aligned}$	by Definition 3.4 page 72
$\begin{aligned}\alpha x &\le (\alpha x)\vee(\alpha y)\\ x &\le \alpha^{-1}(\alpha x)\vee(\alpha y)\\ x\vee y &\le \alpha^{-1}(\alpha x)\vee(\alpha y)\\ \alpha(x\vee y) &\le (\alpha x)\vee(\alpha y)\end{aligned}$	$\begin{aligned}\alpha y &\le (\alpha x)\vee(\alpha y)\\ y &\le \alpha^{-1}(\alpha x)\vee(\alpha y)\end{aligned}$	

6. Proof that $\alpha(x \wedge y) = (\alpha x) \wedge (\alpha y)$ for $\alpha \ge 0$:

$\begin{aligned}x &\ge x\wedge y\\ \alpha x &\ge \alpha(x\wedge y)\\ (\alpha x)\wedge(\alpha y) &\ge \alpha(x\wedge y)\end{aligned}$	$\begin{aligned}y &\ge x\wedge y\\ \alpha y &\ge \alpha(x\wedge y)\end{aligned}$	by Definition 3.4 page 72

$\alpha x \geq (\alpha x) \wedge (\alpha y)$ $x \geq \alpha^{-1}(\alpha x) \wedge (\alpha y)$ $x \wedge y \geq \alpha^{-1}(\alpha x) \wedge (\alpha y)$ $\alpha(x \wedge y) \geq (\alpha x) \wedge (\alpha y)$	$\alpha y \geq (\alpha x) \wedge (\alpha y)$ $y \geq \alpha^{-1}(\alpha x) \wedge (\alpha y)$	

7. Proof that $x + y = (x \wedge y) + (x \vee y)$:

$x \leq x \vee y$ $x + y \leq (x \vee y) + y$ $x + y - (x \vee y) \leq y$ $x + y - (x \vee y) \leq x \wedge y$ $x + y \leq (x \vee y) + (x \wedge y)$	$y \leq x \vee y$ $x + vy \leq x + (x \vee y)$ $x + vy - (x \vee y) \leq x$
$x \wedge y \leq x$ $0 \leq x - (x \wedge y)$ $y \leq y + x - (x \wedge y)$ $y \leq x + y - (x \wedge y)$ $x \vee y \leq x + y - (x \wedge y)$ $(x \wedge y) + (x \vee y) \leq x + y$	$x \wedge y \leq y$ $0 \leq y - (x \wedge y)$ $x \leq x + y - (x \wedge y)$ $x \leq x + y - (x \wedge y)$

⇨

Definition 3.6. [16] *Let* $(\boldsymbol{X}, \vee, \wedge; \leq)$ *be a* RIESZ SPACE (Definition 3.5 page 72).

DEF

x^+	*is defined as* $x^+ \triangleq x \vee 0$	*and is called the* ***positive part*** *of* x.
x^-	*is defined as* $x^- \triangleq (-x) \vee 0$	*and is called the* ***negative part*** *of* x.
$\lvert x \rvert$	*is defined as* $\lvert x \rvert \triangleq x \vee (-x)$	*and is called the* ***absolute value*** *of* x.

Theorem 3.5. [17] *Let* $(\boldsymbol{X}, \vee, \wedge; \leq)$ *be a* RIESZ SPACE (Definition 3.5 page 72).

THM

$$\left.\begin{array}{rcl} y - z &=& x \quad \text{and} \\ y \wedge z &=& 0 \end{array}\right\} \iff \left\{\begin{array}{rcl} y &=& x^+ \quad \text{and} \\ z &=& x^- \end{array}\right.$$

✎PROOF:

1. Proof that left hypothesis $\implies$ right hypothesis:

$$\begin{aligned}
x^+ &= x \vee 0 && \text{by definition of } x^+ \text{ Definition 3.6 page 74} \\
&= (y - z) \vee 0 && \text{by left hypothesis} \\
&= (y - z) \vee (z - z) && \\
&= (y \vee z) - z && \text{by Theorem 3.4 page 72} \\
&= \left[y + z - (y \wedge z)\right] - z && \text{by Theorem 3.4 page 72} \\
&= y - (y \wedge z) && \\
&= y - 0 && \text{by left hypothesis} \\
&= y && \\
x^- &= (-x) \vee 0 && \text{by definition of } x^- \text{ Definition 3.6 page 74} \\
&= (z - y) \vee 0 && \text{by left hypothesis} \\
&= (z - y) \vee (y - y) && \\
&= (z \vee y) - y && \text{by Theorem 3.4 page 72}
\end{aligned}$$

[16] Aliprantis and Burkinshaw (2006) page 4, Istrăţescu (1987) page 129
[17] Aliprantis and Burkinshaw (2006) page 4 ⟨Theorem 1.3⟩

$$
\begin{aligned}
&= \left[z + y - (z \wedge y)\right] - z && \text{by Theorem 3.4 page 72}\\
&= z - (z \wedge y)\\
&= z - \mathbb{0} && \text{by left hypothesis}\\
&= z
\end{aligned}
$$

2. Proof that left hypothesis $\Longleftarrow$ right hypothesis:

$$
\begin{aligned}
y - z &= x^{+} - x^{-} && \text{by right hypothesis}\\
&= [x \vee \mathbb{0}] - [(-x) \vee \mathbb{0}] && \text{by Definition 3.6 page 74}\\
&= (x \vee \mathbb{0}) + (x \wedge \mathbb{0}) && \text{by Theorem 3.4 page 72}\\
&= x && \text{by Theorem 3.4 page 72}\\
y \wedge z &= x^{+} \wedge x^{-} && \text{by right hypothesis}\\
&= \left[x^{-} + \left(x^{+} - x^{-}\right)\right] \wedge [x^{-} + \mathbb{0}]\\
& x^{-} + \left[\left(x^{+} - x^{-}\right) \wedge \mathbb{0}\right] && \text{by Theorem 3.4 page 72}\\
& x^{-} + \left[(y - z) \wedge \mathbb{0}\right] && \text{by right hypothesis}\\
& x^{-} + (x \wedge \mathbb{0}) && \text{by previous result}\\
& x^{-} - [-x \vee \mathbb{0}] && \text{by Theorem 3.4 page 72}\\
& x^{-} - x^{-} && \text{by definition of } x^{-} \text{ (Definition 3.6 page 74)}\\
& \mathbb{0}
\end{aligned}
$$

✏

Theorem 3.6. [18] *Let $(\boldsymbol{X}, \vee, \wedge; \leq)$ be a* RIESZ SPACE *(Definition 3.5 page 72). Let x^{+} the* POSITIVE PART *of $x \in \boldsymbol{X}$, x^{-} the* NEGATIVE PART *of $x \in \boldsymbol{X}$, and $|x|$ the* ABSOLUTE VALUE *(Definition 3.6 page 74) of $x \in \boldsymbol{X}$.*

THM

$$
\begin{aligned}
|x| &= x^{+} + x^{-} && \forall x \in \boldsymbol{X}\\
x &= (x - y)^{+} + (x \wedge y) && \forall x \in \boldsymbol{X}
\end{aligned}
$$

✎PROOF:

$$
\begin{aligned}
|x| &= x \vee (-x) && \text{by definition of } |x| \text{ (Definition 3.6 page 74)}\\
&= (2x - x) \vee (\mathbb{0} - x)\\
&= (-x + 2x) \vee (-x + \mathbb{0}) && \text{by \textit{commutative} property (Definition 3.1 page 67)}\\
&= (-x) + (2x \vee \mathbb{0}) && \text{by Theorem 3.4 page 72}\\
&= (2x \vee 2\mathbb{0}) - x && \text{by the \textit{commutative} property (Definition 3.1 page 67)}\\
&= 2(x \vee \mathbb{0}) - x && \text{by Theorem 3.4 page 72}\\
&= 2x^{+} - x && \text{by definition of } x^{+} \text{ (Definition 3.6 page 74)}\\
&= 2x^{+} - \left(x^{+} - x^{-}\right) && \text{by 1}\\
&= x^{+} + x^{-}\\
x &= x + \mathbb{0} && x + y - y\\
&= (x \vee y) + (x \wedge y) - y && \text{by Theorem 3.4 page 72}\\
&= \left[(x - y) \vee (y - y)\right] + (x \wedge y) && \text{by Theorem 3.4 page 72}\\
&= \left[(x - y) \vee \mathbb{0}\right] + (x \wedge y)\\
&= (x - y)^{+} + (x \wedge y) && \text{by definition of } x^{+} \text{ (Definition 3.6 page 74)}
\end{aligned}
$$

✏

[18] Aliprantis and Burkinshaw (2006) page 4

Theorem 3.7. [19] *Let* $(X, \vee, \wedge; \leq)$ *be a* RIESZ SPACE (Definition 3.5 page 72). *Let* x^+ *the* POSITIVE PART *of* $x \in X$, x^- *the* NEGATIVE PART *of* $x \in X$, *and* $|x|$ *the* ABSOLUTE VALUE (Definition 3.6 page 74) *of* $x \in X$.

THM

1. $x \vee y = \frac{1}{2}(x + y + |x - y|) \quad \forall x,y \in X$
2. $x \wedge y = \frac{1}{2}(x + y - |x - y|) \quad \forall x,y \in X$
3. $|x - y| = (x \vee y) - (x \wedge y) \quad \forall x,y \in X$
4. $|x| \vee |y| = \frac{1}{2}(|x + y| + |x - y|) \quad \forall x,y \in X$
5. $|x| \vee |y| = \frac{1}{2}||x + y| - |x - y|| \quad \forall x,y \in X$

✎PROOF:

$$\begin{aligned}
(x + y + |x - y|) &= (x + y) + \big[(x - y) \vee (y - x)\big] && \text{by Definition 3.6 page 74}\\
&= \big[(x + y) + (x - y)\big] \vee \big[(x + y) + (y - x)\big] && \text{by Theorem 3.4 page 72}\\
&= (2x) \vee (2y)\\
&= 2(x \vee y) && \text{by Theorem 3.4 page 72}\\
(x + y - |x - y|) &= (x + y) - \big[(x - y) \vee (y - x)\big] && \text{by Definition 3.6 page 74}\\
&= (x + y) - \big[(-(y - x)) \vee (-(x - y))\big]\\
&= (x + y) + \big[(y - x) \wedge (x - y)\big] && \text{by Theorem 3.4 page 72}\\
&= \big[(x + y) + (y - x)\big] \wedge \big[(x + y) + (x - y)\big] && \text{by Theorem 3.4 page 72}\\
&= (2y) \wedge (2x)\\
&= 2(y \wedge x) && \text{by Theorem 3.4 page 72}\\
&= 2(x \wedge y)\\
|x - y| &= \frac{1}{2}(x + y + |x - y|) - \frac{1}{2}(x + y + |x - y|)\\
&= (x \vee y) - (x \wedge y) && \text{by 1 and 2}\\
|x + y| + |x - y| &= \frac{1}{2}(\mathbb{0} + |2x + 2y|) + |x - y|\\
&= \frac{1}{2}\big[(x + y) + (-x - y) + |(x + y) - (-x - y)|\big] + |x - y|\\
&= \big[(x + y) \vee (-x - y)\big] + |x - y| && \text{by 1}\\
&= \big[(x + y) + |x - y|\big] \vee \big[(-x - y) + |x - y|\big] && \text{by Theorem 3.4 page 72}\\
&= 2(x \vee y) \vee 2\big[(-y) + (-x) + |(-y) - (-x)|\big] && \text{by 1}\\
&= 2(x \vee y) \vee 2\big[(-y) \vee (-x)\big] && \text{by 1}\\
&= 2([x \vee (-x)] \vee (y \vee (-y)))\\
&= 2(|x| \vee |y|) && \text{by Definition 3.6 page 74}\\
||x + y| - |x - y|| &= 2(|x + y| \vee |x - y|) - (|x + y| + |x - y|) && \text{by 1}\\
&= (|x + y + x - y| + |x + y - x + y|) - 2(|x| \vee |y|) && \text{by 3}\\
&= 2(|x| + |y|) - 2(|x| \vee |y|)\\
&= 2(|x| \vee |y|) && \text{by Theorem 3.4 page 72}
\end{aligned}$$

✏

Definition 3.7. [20] *Let* $(X, \vee, \wedge; \leq)$ *be a* RIESZ SPACE (Definition 3.5 page 72). *Let* x^+ *the* POSITIVE PART *of* $x \in X$, x^- *the* NEGATIVE PART *of* $x \in X$, *and* $|x|$ *the* ABSOLUTE VALUE (Definition 3.6 page 74) *of* $x \in X$.

DEF

x *and* y *are* ***disjoint***, *denoted by* $x \perp y$, *if*

$$|x| \wedge |y| = \mathbb{0}.$$

Two subsets U *and* V *of* X *are* ***disjoint***, *denoted by* $U \perp V$ *if*

$$x \perp y \qquad \forall x \in U \text{ and } y \in V$$

[19] Aliprantis and Burkinshaw (2006) page 5 ⟨Theorem 1.4⟩

[20] Aliprantis and Burkinshaw (2006) page 5

Definition 3.8. [21] *Let* $(\boldsymbol{X}, \vee, \wedge; \leq)$ *be a* RIESZ SPACE (Definition 3.5 page 72). *Let* $\boldsymbol{x}^+$ *the* POSITIVE PART *of* $\boldsymbol{x} \in \boldsymbol{X}$, $\boldsymbol{x}^-$ *the* NEGATIVE PART *of* $\boldsymbol{x} \in \boldsymbol{X}$, *and* $|\boldsymbol{x}|$ *the* ABSOLUTE VALUE (Definition 3.6 page 74) *of* $\boldsymbol{x} \in \boldsymbol{X}$. *Let* $\boldsymbol{Y}$ *be a subset of* $\boldsymbol{X}$.

DEF: $\boldsymbol{Y}^d$ *is the* ***disjoint complement*** *of* $\boldsymbol{Y}$ *if* $\quad \boldsymbol{Y}^d \triangleq \{\boldsymbol{x} \in \boldsymbol{X} | \boldsymbol{x} \perp \boldsymbol{y} \quad \forall \boldsymbol{y} \in \boldsymbol{Y}\}$.
The quantity $\boldsymbol{Y}^{dd}$ *is defined as* $\left(\boldsymbol{Y}^d\right)^d$.

Definition 3.9. [22] *Let* $(\boldsymbol{X}, \vee, \wedge; \leq)$ *be a* RIESZ SPACE (Definition 3.5 page 72). *Let* $\boldsymbol{x}^+$ *the* POSITIVE PART *of* $\boldsymbol{x} \in \boldsymbol{X}$, $\boldsymbol{x}^-$ *the* NEGATIVE PART *of* $\boldsymbol{x} \in \boldsymbol{X}$, *and* $|\boldsymbol{x}|$ *the* ABSOLUTE VALUE (Definition 3.6 page 74) *of* $\boldsymbol{x} \in \boldsymbol{X}$.

DEF:
$$\begin{aligned}
|A| &\triangleq \{|\boldsymbol{a}| | \boldsymbol{a} \in A\}\\
A^+ &\triangleq \left\{\boldsymbol{a}^+ | \boldsymbol{a} \in A\right\}\\
A^- &\triangleq \{\boldsymbol{a}^- | \boldsymbol{a} \in A\}\\
A \vee B &\triangleq \left\{\boldsymbol{a} \vee \boldsymbol{b} | \boldsymbol{a} \in A \text{ and } \boldsymbol{b} \in B\right\}\\
A \wedge B &\triangleq \left\{\boldsymbol{a} \wedge \boldsymbol{b} | \boldsymbol{a} \in A \text{ and } \boldsymbol{b} \in B\right\}\\
\boldsymbol{x} \vee A &\triangleq \{\boldsymbol{x} \vee \boldsymbol{a} | \boldsymbol{a} \in A\}\\
\boldsymbol{x} \wedge A &\triangleq \{\boldsymbol{x} \wedge \boldsymbol{a} | \boldsymbol{a} \in A\}
\end{aligned}$$

[21] Aliprantis and Burkinshaw (2006) page 5
[22] Aliprantis and Burkinshaw (2006) page 7

CHAPTER 4

TOPOLOGICAL LINEAR SPACES

4.1 Definitions

A *topologial linear space* (often called a *topological vector space*) is basically a *linear space* (Definition 3.1 page 67) with a *topology* (Definition 1.1 page 3). If the topology is generated by a *metric* (Definition 2.1 page 27), then it is a *metric linear space* (Definition 4.5 page 80). If the topology is generated by a *norm* (Definition 5.1 page 83), then it is a *normed linear space*. If the topology is generated by an *inner product* (Definition 6.1 page 95), then it is an *inner product space*.

Definition 4.1. [1]

DEF *The tuple* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \boldsymbol{T}\right)$ *is a* ***topological linear space*** *if*

1. $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *is a* LINEAR SPACE *and*
2. $\boldsymbol{T}$ *is a* TOPOLOGY *on* 2^X *and*
3. $(\boldsymbol{x}, \boldsymbol{y}) \to \boldsymbol{x} + \boldsymbol{y}$ *is* CONTINUOUS *on* $\boldsymbol{X}^{\boldsymbol{X}\times\boldsymbol{X}}$ *(Definition 1.8 page 23) and*
4. $(\alpha, \boldsymbol{x}) \to \alpha\boldsymbol{x}$ *is* CONTINUOUS *on* $\boldsymbol{X}^{\mathbb{F}\times\boldsymbol{X}}$. *(Definition 1.8 page 23)* .

Definition 4.2. *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \boldsymbol{T}_x\right)$ *be a* TOPOLOGICAL LINEAR SPACE *with topology* $\boldsymbol{T}_x$. *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \boldsymbol{T}_y\right)$ *be a* TOPOLOGICAL LINEAR SPACE *with topology* $\boldsymbol{T}_y$. *Let* Y^X *be the set of all functions (operators) from* X *to* Y.

DEF *The set* $\mathcal{C}(X,Y)$ *is the* ***space of continuous operators*** *from* X *to* Y *and is defined as*
$$\mathcal{C}(X,Y) \triangleq \left\{\mathsf{f} \in Y^X | \mathsf{f} \text{ is continuous with respect to } \left(\boldsymbol{T}_x, \boldsymbol{T}_y\right)\right\}$$

Definition 4.3. *Let* $\boldsymbol{Y} \triangleq \left(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \boldsymbol{T}\right)$ *be a subspace of a* TOPOLOGICAL LINEAR SPACE $\boldsymbol{X} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \boldsymbol{T}\right)$. *Let* Y^- *be the* CLOSURE *of the set* Y *(Definition 1.4 page 14) in the* TOPOLOGICAL SPACE $(X, \boldsymbol{T})$ *(Definition 1.1 page 3)*.

DEF *The subspace* $\boldsymbol{Y}$ *is* ***closed*** *in* $\boldsymbol{X}$ *if*
$$Y = Y^- .$$

Example 4.1. [2] Let A^- be the *closure* (Definition 1.4 page 14) of a set A in a topological space. Let X be the set of all bounded sequences over $\mathbb{R}$. Let Y be the set of all bounded sequences with a finite

[1] Schaefer and Wolff (1999) page 12 ⟨1. Vector Space Topologies⟩, Robertson and Robertson (1980) page 3 ⟨3. Topological Vector Spaces⟩

[2] Kolmogorov and Fomin (1975) page 140 ⟨Example 1⟩

number of zeros. Let $\boldsymbol{T}$ be the standard topology on $\mathbb{R}$ generated by the metric $\mathsf{d}(x,y) \triangleq |x-y|$.

EX

$\boldsymbol{X}$ is a *topological linear space*.
$\boldsymbol{Y}$ is a *topological linear space*.
But Y is *not closed* in $(\boldsymbol{X}, \boldsymbol{T})$ $(Y \neq Y^-)$, because, for example,

$$\left(\left(1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots\right)\right) \text{ is in } Y,$$

but its *closure point*

$$\left(\left(1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots, 0, 0, 0, \cdots\right)\right) \text{ is } \textit{not} \text{ in } Y \text{ (but is in } X).$$

Example 4.2. [3] Let A^- be the *closure* of a set A in a topological space. Let $C_{[a:b]}$ be the set of all *continuous* functions on the real interval $[a:b]$. Let $P_{[a:b]}$ be the set of all *polynomials* on the real interval $[a:b]$. Let $\boldsymbol{T}$ be the standard topology on $\mathbb{R}$ generated by the metric $\mathsf{d}(x,y) \triangleq |x-y|$.

EX

$\left(C_{[a:b]}, \boldsymbol{T}\right)$ is a *topological linear space*.
$\left(P_{[a:b]}, \boldsymbol{T}\right)$ is a *topological linear space*.
But $P_{[a:b]} \neq \left(P_{[a:b]}\right)^- = C_{[a:b]}$, so $P_{[a:b]}$ is *not closed* in $\left(C_{[a:b]}, \boldsymbol{T}\right)$.

✎PROOF: $\left(P_{[a:b]}\right)^- = C_{[a:b]}$ by *Weierstrass' Approximation Theorem*. ✏

4.2 Dual Spaces

Definition 4.4. [4] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \boldsymbol{T}\right)$ *be a topological linear space. Let* $\mathbb{F}^{\boldsymbol{X}}$ *be the set of all functionals from* X *to* $\mathbb{F}$. *Let* $C(\boldsymbol{X}, \mathbb{F})$ *be the space of continuous functionals from* X *to* $\mathbb{F}$.

DEF

The ***algebraic dual space*** $\boldsymbol{X}^\dagger$ *of* $\boldsymbol{X}$ *is* $\quad \boldsymbol{X}^\dagger \triangleq \mathbb{F}^{\boldsymbol{X}}$.
The ***topological dual space*** $\boldsymbol{X}^*$ *of* $\boldsymbol{X}$ *is* $\quad \boldsymbol{X}^* \triangleq C(\boldsymbol{X}, \mathbb{F})$.
The space $\boldsymbol{X}$ *is the* ***predual*** *of* $\boldsymbol{X}^*$. *A topological dual space is also called a* ***dual space, conjugate space*** *or* ***adjoint space***.

Theorem 4.1. *Let* $\boldsymbol{X} = \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *be a linear space with dual space* $\boldsymbol{X}^*$.

THM

$\boldsymbol{X}^*$ *is a linear space.*

4.3 Metric Linear Spaces

Metric space stucture can be added to a linear space resulting in a *metric linear space* (next definition). One key difference between metric linear spaces and normed linear spaces is that the balls in a *normed linear space* (Definition 5.1 page 83) are always *convex* (Definition 9.6 page 142); this is not true for all metric linear spaces.[5]

Definition 4.5. [6] *Let* $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d}\right)$.

[3] Kolmogorov and Fomin (1975) page 140 ⟨Example 2⟩
[4] Hunter and Nachtergaele (2001) page 116 ⟨Definition 5.54⟩, Kurdila and Zabarankin (2005) page 76 ⟨Definitions 2.2.3, 2.2.4⟩, Hewitt and Stomberg (1965) page 211 ⟨Definition 14.6⟩
[5] Bruckner et al. (1997), page 478
[6] Maddox (1989) page 90, Bruckner et al. (1997) page 477 ⟨Definition 12.3⟩, Rolewicz (1985) page 1

DEF *The tuple $\boldsymbol{\Omega}$ is a **metric linear space** if*

1. *if* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *is a* LINEAR SPACE *and*
2. d *is a* METRIC *in* $\mathbb{R}^{\boldsymbol{X}}$ *and*
3. $\mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z}) = \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \quad \forall x,y,z \in X$ (TRANSLATION INVARIANT) *and*
4. $\alpha_n \to \alpha$ *and* $\boldsymbol{x}_n \to \boldsymbol{x} \implies \alpha_n \boldsymbol{x}_n \to \alpha \boldsymbol{x}$

Theorem 4.2. [7] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d})$ *be a metric linear space.*

THM
$$\underbrace{\mathsf{d}(\boldsymbol{\theta}, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda\mathsf{d}(\boldsymbol{\theta}, \boldsymbol{x}) + (1-\lambda)\mathsf{d}(\boldsymbol{\theta}, \boldsymbol{y})}_{\mathsf{d} \text{ is a CONVEX function}} \implies \left\{ \begin{array}{c} \mathsf{B}(\theta, r) \in \boldsymbol{\Omega} \\ \text{is convex} \\ \forall \theta \in X, r \in \mathbb{R}^+ \end{array} \right\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{d}(\boldsymbol{\theta}, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &\leq \lambda\mathsf{d}(\boldsymbol{\theta}, \boldsymbol{x}) + (1-\lambda)\mathsf{d}(\boldsymbol{\theta}, \boldsymbol{y}) && \text{by convexity hypothesis} \\
&\leq \lambda r + (1-\lambda) r && \forall \boldsymbol{x}, \boldsymbol{y} \in \mathsf{B}(\theta, r) \\
&= r \\
&\implies \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in \mathsf{B}(\theta, r) && \forall \boldsymbol{x}, \boldsymbol{y} \in \mathsf{B}(\theta, r) \\
&\implies \mathsf{B}(\theta, r) \in (\boldsymbol{X}, \mathsf{d}) \text{ is convex} && \forall \theta \in \boldsymbol{X}
\end{aligned}$$

✏

Theorem 4.3. [8] *Let* $(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \mathsf{d})$ *be a real metric linear space.*

THM
$$\left\{ \begin{array}{llll} 1. & \mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z}) = \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) & \forall x,y,z \in \boldsymbol{X} & \text{(TRANSLATION INVARIANT) and} \\ 2. & \mathsf{d}(\lambda\boldsymbol{x}, \lambda\boldsymbol{y}) = \lambda\,\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) & \forall x,y \in \boldsymbol{X},\ \lambda \in [0,1] & \text{(HOMOGENEOUS)} \end{array} \right\}$$
$$\implies \{\mathsf{B}(\theta, r) \in (\boldsymbol{X}, \mathsf{d}) \text{ is CONVEX} \quad \forall \theta \in \boldsymbol{X}, r \in \mathbb{R}^+\}$$

✎PROOF:

$$\begin{aligned}
&\mathsf{d}(\boldsymbol{\theta}, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\
&= \mathsf{d}(\mathbb{0}, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} - \boldsymbol{\theta}) && \text{by translation invariance hypothesis} \\
&= \mathsf{d}(\mathbb{0}, \lambda(\boldsymbol{x}-\boldsymbol{\theta}) + (1-\lambda)(\boldsymbol{y}-\boldsymbol{\theta})) \\
&\leq \mathsf{d}(\mathbb{0}, \lambda(\boldsymbol{x}-\boldsymbol{\theta})) + \mathsf{d}(\lambda(\boldsymbol{x}-\boldsymbol{\theta}), \lambda(\boldsymbol{x}-\boldsymbol{\theta}) + (1-\lambda)(\boldsymbol{y}-\boldsymbol{\theta})) && \text{by } \textit{subadditive} \text{ property (Definition 2.1 page 27)} \\
&= \mathsf{d}(\mathbb{0}, \lambda(\boldsymbol{x}-\boldsymbol{\theta})) + \mathsf{d}(\mathbb{0}, \mathbb{0} + (1-\lambda)(\boldsymbol{y}-\boldsymbol{\theta})) && \text{by translation invariance hypothesis} \\
&= \lambda\mathsf{d}(\mathbb{0}, \boldsymbol{x}-\boldsymbol{\theta}) + (1-\lambda)\mathsf{d}(\mathbb{0}, \boldsymbol{y}-\boldsymbol{\theta}) && \text{by homogeneous hypothesis} \\
&= \lambda\mathsf{d}(\boldsymbol{\theta}, \boldsymbol{x}) + (1-\lambda)\mathsf{d}(\boldsymbol{\theta}, \boldsymbol{y}) && \text{by translation invariance hypothesis} \\
&\leq \lambda r + (1-\lambda) r && \forall \boldsymbol{x}, \boldsymbol{y} \in \mathsf{B}(\theta, r) \\
&= r \\
\implies & \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in \mathsf{B}(\theta, r) && \forall \boldsymbol{x}, \boldsymbol{y} \in \mathsf{B}(\theta, r) \\
\implies & \mathsf{B}(\theta, r) \in (\boldsymbol{X}, \mathsf{d}) \text{ is convex} && \forall \theta \in \boldsymbol{X}
\end{aligned}$$

✏

[7] Norfolk (1991), page 5
[8] Norfolk (1991), pages 5–6, http://groups.google.com/group/sci.math/msg/a6f0a7924027957d

CHAPTER 5

NORMED LINEAR SPACES

5.1 Definition and basic results

Definition 5.1. [1] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a* LINEAR SPACE (Definition 3.1 page 67) *and* $|\cdot| \in \mathbb{R}^{\mathbb{F}}$ *the* ABSOLUTE VALUE *function* (Definition F.4 page 346).

D E F

A functional $\|\cdot\|$ *in* $\mathbb{R}^X$ *is a* ***norm*** *if*

1.	$\|x\| \geq 0$	$\forall x \in X$	(STRICTLY POSITIVE)	*and*
2.	$\|x\| = 0 \iff x = \mathbb{0}$	$\forall x \in X$	(NONDEGENERATE)	*and*
3.	$\|\alpha x\| = \|\alpha\|\,\|x\|$	$\forall x \in X,\ \alpha \in \mathbb{C}$	(HOMOGENEOUS)	*and*
4.	$\|x + y\| \leq \|x\| + \|y\|$	$\forall x, y \in X$	(SUBADDITIVE/TRIANGLE INQUALITY).	

A ***normed linear space*** *is the tuple* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$.

The definition of the *norm* (Definition 5.1 page 83) requires that any two vectors in a norm space be *subadditive* (they satisfy the *triangle inequality* property) such that $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$. Acutally, in **any** normed linear space, this property holds true for **any** finite number of vectors—not just two—such that $\|x_1 + x_2 + \cdots + x_N\| \leq \|x_1\| + \|x_2\| + \cdots + \|x_N\|$ (next theorem).

Theorem 5.1 (triangle inequality). [2] *Let* $(\!|x_n \in X|\!)_1^N$ *be an* N-TUPLE (Definition 8.1 page 127) *of vectors in a* NORMED LINEAR SPACE $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$.

T H M

$$\left\|\sum_{n=1}^{N} x_n\right\| \leq \sum_{n=1}^{N} \|x_n\| \qquad \forall N \in \mathbb{N},\ x_n \in V$$

✎PROOF: Proof is by induction:

[1] Aliprantis and Burkinshaw (1998), pages 217–218, Banach (1932a), page 53, Banach (1932b), page 33, Banach (1922) page 135

[2] Michel and Herget (1993), page 344, Euclid (circa 300BC) ⟨Book I Proposition 20⟩

1. Proof for the $N = 1$ case:

$$\begin{aligned}\left\|\sum_{n=1}^{1} \boldsymbol{x}_n\right\| &= \|\boldsymbol{x}_1\| \\ &= \sum_{n=1}^{1} \|\boldsymbol{x}_1\|\end{aligned}$$

2. Proof for the $N = 2$ case:

$$\begin{aligned}\left\|\sum_{n=1}^{2} \boldsymbol{x}_n\right\| &= \left\|\sum_{n=1}^{2} \boldsymbol{x}_n\right\| \\ &= \|\boldsymbol{x}_1 + \boldsymbol{x}_2\| \\ &\le \|\boldsymbol{x}_1\| + \|\boldsymbol{x}_2\| && \text{by Definition 5.1 page 83 (triangle inequality)} \\ &= \sum_{n=1}^{2} \|\boldsymbol{x}_n\|\end{aligned}$$

3. Proof that $[N \text{ case}] \implies [N + 1 \text{ case}]$:

$$\begin{aligned}\left\|\sum_{n=1}^{N+1} \boldsymbol{x}_n\right\| &= \left\|\sum_{n=1}^{N} \boldsymbol{x}_n + \boldsymbol{x}_{n+1}\right\| \\ &\le \left\|\sum_{n=1}^{N} \boldsymbol{x}_n\right\| + \|\boldsymbol{x}_{n+1}\| && \text{by Definition 5.1 page 83 (triangle inequality)} \\ &\le \sum_{n=1}^{N} \|\boldsymbol{x}_n\| + \|\boldsymbol{x}_{n+1}\| && \text{by left hypothesis} \\ &= \sum_{n=1}^{N+1} \|\boldsymbol{x}_n\|\end{aligned}$$

✏

Theorem 5.2 (Reverse Triangle Inequality).[3] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a* NORMED LINEAR SPACE *(Definition 5.1 page 83).*

THM

$$\underbrace{\big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \le \|\boldsymbol{x} - \boldsymbol{y}\|}_{\text{REVERSE TRIANGLE INEQUALITY}} \quad \le \quad \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in X$$

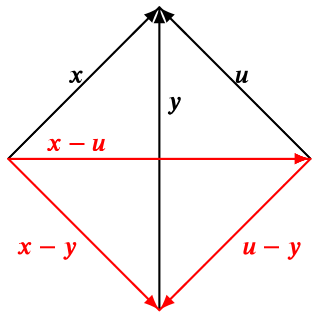
✎PROOF:

$$\begin{aligned}\big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| &= \big|\|(\boldsymbol{x} - \boldsymbol{y}) + \boldsymbol{y}\| - \|\boldsymbol{y}\|\big| \\ &\le \big|\|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y}\| - \|\boldsymbol{y}\|\big| && \text{by Definition 5.1 page 83} \\ &= \big|\|\boldsymbol{x} - \boldsymbol{y}\|\big| \\ &= \|\boldsymbol{x} - \boldsymbol{y}\| && \text{by Definition 5.1 page 83}\end{aligned}$$

$$\begin{aligned}\|\boldsymbol{x} - \boldsymbol{y}\| &\le \|\boldsymbol{x} - \mathbb{0}\| + \|\mathbb{0} - \boldsymbol{y}\| && \text{by previous result with } \boldsymbol{u} = 0 \\ &= \|\boldsymbol{x}\| + |-1|\,\|\boldsymbol{y}\| && \text{by Definition 5.1 page 83} \\ &= \|\boldsymbol{x}\| + \|\boldsymbol{y}\|\end{aligned}$$

✏

[3] Aliprantis and Burkinshaw (1998), page 218, Giles (2000) page 2, Banach (1922) page 136



The shortest distance between two vectors is always the difference of the vectors. This is proven in next and illustrated to the left in the Euclidean space $\mathbb{R}^2$ (the plane)

Proposition 5.1. [4] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a* NORMED LINEAR SPACE *(Definition 5.1 page 83)*.

PRP
$$\|\boldsymbol{x} - \boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{u}\| + \|\boldsymbol{u} - \boldsymbol{y}\| \qquad \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{u} \in X$$

✎PROOF:

$$\begin{aligned}
\|\boldsymbol{x} - \boldsymbol{y}\| &= \|(\boldsymbol{x} - \boldsymbol{u}) + (\boldsymbol{u} - \boldsymbol{y})\| \\
&\leq \|\boldsymbol{x} - \boldsymbol{u}\| + \|\boldsymbol{u} - \boldsymbol{y}\| && \text{by Definition 5.1 page 83}
\end{aligned}$$

⇨

Example 5.1 (The usual norm). [5] Let $\mathbb{R}^{\mathbb{R}}$ be the set of all functions with domain and range the set of *real numbers* $\mathbb{R}$.

EX The **absolute value** (Definition F.4 page 346) $|\cdot| \in \mathbb{R}^{\mathbb{R}}$ is a *norm*.

Example 5.2 (l_p norms). Let $(x_n)_{n\in\mathbb{Z}}$ be a *sequence* (Definition 8.1 page 127) of real numbers. An uncountably infinite number of norms is provided by the $\ell^p_{\mathbb{F}}$ norms $\|(x_n)\|_p$:

EX
$$\|(x_n)\|_p \triangleq \left(\sum_{n\in\mathbb{Z}} |x_n|^p\right)^{\frac{1}{p}} \qquad \text{is a norm for } p \in [1 : \infty]$$

5.2 Relationship between metrics and norms

5.2.1 Metrics generated by norms

The concept of *length* is very closely related to the concept of *distance*. Thus it is not surprising that a *norm* (a “length” function) can be used to define a *metric* (a “distance” function) on any *metric linear space* (Definition 4.5 page 80). Another way to say this is that the norm of a normed linear space *induces* a metric on this space. And so every normed linear space also has a metric. And because every normed linear space has a metric, **every normed linear space is also a metric space**. Actually this can be generalized one step further in that every metric space is also a *topological space*. And so **every normed linear space is also a topological space.** In symbols,

$$\text{normed linear space} \implies \text{metric space} \implies \text{topological space.}$$

[4] Aliprantis and Burkinshaw (1998), page 218
[5] Giles (1987), page 3

Theorem 5.3. [6] *Let* $\mathsf{d} \in \mathbb{R}^{X \times X}$ *be a function on a* REAL *normed linear space* $(X, +, \cdot, (\mathbb{R}, \dotplus, \dot\times), \|\cdot\|)$. *Let* $\mathsf{B}(\boldsymbol{x}, r) \triangleq \{\boldsymbol{y} \in \boldsymbol{X} \mid \|\boldsymbol{y} - \boldsymbol{x}\| < r\}$ *be the* OPEN BALL *(Definition 2.4 page 30) of radius* r *centered at a point* $\boldsymbol{x}$.

THM $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \|\boldsymbol{x} - \boldsymbol{y}\|$ *is a metric on* $\boldsymbol{X}$

✎PROOF: The proof follows directly from the definition of a metric (Definition 2.1 page 27) the definition of *norm* (Definition 5.1 page 83). ✏

The previous theorem defined a metric $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y})$ induced by the norm $\|\boldsymbol{x}\|$. The next definition defines this metric formally.

Definition 5.2. [7] *Let* $(X, +, \cdot, (\mathbb{F}, \dotplus, \dot\times), \|\cdot\|)$ *be a* NORMED LINEAR SPACE *(Definition 5.1 page 83)*.

DEF *The* ***metric induced by the norm*** $\|\cdot\|$ *is the function* $\mathsf{d} \in \mathbb{R}^{X}$ *such that*

$$\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \|\boldsymbol{x} - \boldsymbol{y}\| \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in X$$

Due to its algebraic structure, every norm is *continuous* (Corollary 5.1 page 86). This makes norm spaces very useful in analysis. For a function f be to *continuous*, for every $\epsilon > 0$ there must exist a $\delta > 0$ such that $|\mathsf{f}(x + \delta) - \mathsf{f}(x)| < \epsilon$. The *Reverse Triangle Inequality* (Theorem 5.2 page 84) shows this to be true when $\mathsf{f}(\cdot) \triangleq \|\cdot\|$.

Corollary 5.1. [8] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dotplus, \dot\times), \|\cdot\|)$ *be a* NORMED LINEAR SPACE *(Definition 5.1 page 83)*.

COR *The norm* $\|\cdot\|$ *is* CONTINUOUS *in* $\boldsymbol{\Omega}$.

✎PROOF: This follows from these concepts:

1. The fact that $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \|\boldsymbol{x} - \boldsymbol{y}\|$ is a *metric* (Theorem 5.3 page 86).
2. *Continuity* in a metric space.
3. The *Reverse Triangle Inequality* (Theorem 5.2 page 84).

✏

Theorem 5.4 (next) demonstrates that **all open or closed** balls in **any normed linear space** are *convex*. However, the converse is not true—that is, a metric not generated by a norm may still produce a ball that is convex. Here are some examples:

metric name	example		generated by norm	convex ball
Taxi-cab metric	Example 2.21	page 56	✓	✓
Euclidean metric	Example 2.22	page 57	✓	✓
Sup metric	Example 2.23	page 57	✓	✓
Parabolic metric	Example 2.24	page 58		
exponential metric	Example 2.26	page 60		
Tangential metric	Example 2.27	page 62		✓

[6] Michel and Herget (1993), page 344, Banach (1932a) page 53
[7] Giles (2000) page 1 ⟨1.1 Definition⟩
[8] Giles (2000) page 2

Theorem 5.4. [9] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d})$ *be a* METRIC LINEAR SPACE (Definition 4.5 page 80). *Let* B *be the* OPEN BALL (Definition 2.4 page 30) $\mathsf{B}(\boldsymbol{p}, r) \triangleq \{\boldsymbol{x} \in \boldsymbol{X} | \mathsf{d}(\boldsymbol{p}, \boldsymbol{x}) < r\}$ *(open ball with respect to metric* d *centered at point* $\boldsymbol{p}$ *and with radius* r*).*

THM

$$\left.\begin{array}{l} \underbrace{\exists \|\cdot\| \in \mathbb{R}^{\boldsymbol{X}} \text{ such that } \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{x}\|}_{\mathsf{d} \text{ is generated by a norm}} \end{array}\right\} \implies \left\{\begin{array}{lll} 1. & \mathsf{B}(\boldsymbol{x}, r) = \boldsymbol{x} + B(0, r) & \\ 2. & \mathsf{B}(\mathbb{0}, r) = r\,\mathsf{B}(\mathbb{0}, 1) & \\ 3. & \mathsf{B}(\boldsymbol{x}, r) \text{ is CONVEX} & \\ 4. & \boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) \iff -\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) & \text{(SYMMETRIC)} \end{array}\right.$$

✎PROOF:

1. Proof that $\mathsf{d}(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} + v\boldsymbol{z}) = \mathsf{d}(\boldsymbol{x}, \boldsymbol{y})$ (invariant):

$$\begin{aligned} \mathsf{d}(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} + v\boldsymbol{z}) &= \|(\boldsymbol{y} + v\boldsymbol{z}) - (\boldsymbol{x} + \boldsymbol{z})\| && \text{by left hypothesis} \\ &= \|\boldsymbol{y} - \boldsymbol{x}\| \\ &= \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) && \text{by left hypothesis} \end{aligned}$$

2. Proof that $\mathsf{B}(\boldsymbol{x}, r) = \boldsymbol{x} + B(0, r)$:

$$\begin{aligned} \mathsf{B}(\boldsymbol{x}, r) &= \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) < r\} && \text{by definition of open ball } B \\ &= \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\boldsymbol{y} - \boldsymbol{y}, \boldsymbol{y} - \boldsymbol{x}) < r\} && \text{by right result 1.} \\ &= \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{y} - \boldsymbol{x}) < r\} \\ &= \{\boldsymbol{u} + \boldsymbol{x} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\} && \text{let } \boldsymbol{u} \triangleq \boldsymbol{y} - \boldsymbol{x} \\ &= \boldsymbol{x} + \{\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\} \\ &= \boldsymbol{x} + B(0, r) && \text{by definition of open ball } B \end{aligned}$$

3. Proof that $\mathsf{B}(\mathbb{0}, r) = r\,\mathsf{B}(\mathbb{0}, 1)$:

$$\begin{aligned} \mathsf{B}(\mathbb{0}, r) &= \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{y}) < r\} && \text{by definition of open ball } B \\ &= \left\{\boldsymbol{y} \in \boldsymbol{X} | \frac{1}{r}\mathsf{d}(\mathbb{0}, \boldsymbol{y}) < 1\right\} \\ &= \left\{\boldsymbol{y} \in \boldsymbol{X} | \frac{1}{r}\|\boldsymbol{y} - \mathbb{0}\| < 1\right\} && \text{by left hypothesis} \\ &= \left\{\boldsymbol{y} \in \boldsymbol{X} | \left\|\frac{1}{r}\boldsymbol{y} - \frac{1}{r}\mathbb{0}\right\| < 1\right\} && \text{by homogeneous property of } \|\cdot\| \text{ page 83} \\ &= \left\{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}\left(\frac{1}{r}\mathbb{0}, \frac{1}{r}\boldsymbol{y}\right) < 1\right\} && \text{by left hypothesis} \\ &= \{r\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < 1\} && \text{let } \boldsymbol{u} \triangleq \frac{1}{r}\boldsymbol{y} \\ &= r\,\{\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < 1\} \\ &= r\,B(\mathbb{0}, 1) && \text{by definition of open ball } B \end{aligned}$$

4. Proof that $\mathsf{B}(\boldsymbol{p}, r)$ is convex:
We must prove that for any pair of points $\boldsymbol{x}$ and $\boldsymbol{y}$ in the open ball $\mathsf{B}(\boldsymbol{p}, r)$, any point $\lambda\boldsymbol{x} + (1 - \lambda)\boldsymbol{y}$ is also in the open ball. That is, the distance from any point $\lambda\boldsymbol{x} + (1 - \lambda)\boldsymbol{y}$ to the ball's center $\boldsymbol{p}$ must be less

[9] Giles (2000) page 2 ⟨1.2 Remarks⟩, Giles (1987) pages 22–26 ⟨2.4 Theorem, 2.11 Theorem⟩

than r.

$$\begin{aligned}
\mathsf{d}(\boldsymbol{p}, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \|\boldsymbol{p} - \lambda\boldsymbol{x} - (1-\lambda)\boldsymbol{y}\| && \text{by left hypothesis}\\
&= \Bigg\|\underbrace{\lambda\boldsymbol{p} + (1-\lambda)\boldsymbol{p}}_{\boldsymbol{p}} - \lambda\boldsymbol{x} - (1-\lambda)\boldsymbol{y}\Bigg\|\\
&= \|\lambda\boldsymbol{p} - \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{p} - (1-\lambda)\boldsymbol{y}\|\\
&\le \|\lambda\boldsymbol{p} - \lambda\boldsymbol{x}\| + \|(1-\lambda)\boldsymbol{p} - (1-\lambda)\boldsymbol{y}\| && \text{by subadditivity property of } \|\cdot\| \text{ page 83}\\
&= |\lambda|\,\|\boldsymbol{p} - \boldsymbol{x}\| + |1-\lambda|\,\|\boldsymbol{p} - \boldsymbol{y}\| && \text{by homogeneous property of } \|\cdot\| \text{ page 83}\\
&= \lambda\,\|\boldsymbol{p} - \boldsymbol{x}\| + (1-\lambda)\,\|\boldsymbol{p} - \boldsymbol{y}\| && \text{because } 0 \le \lambda \le 1\\
&\le \lambda r + (1-\lambda) r && \text{because } \boldsymbol{x}, \boldsymbol{y} \text{ are in the ball } \mathsf{B}(\boldsymbol{p}, r)\\
&= r
\end{aligned}$$

5. Proof that $\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) \iff -\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r)$ (symmetric):

$$\begin{aligned}
\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) &\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{y}) < r\} && \text{by definition of open ball } B\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \|\boldsymbol{y} - \mathbb{0}\| < r\} && \text{by left hypothesis}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \|\boldsymbol{y}\| < r\}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \|(-1)(-\boldsymbol{y})\| < r\}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | |-1|\,\|-\boldsymbol{y}\| < r\} && \text{by homogeneous property of } \|\cdot\| \text{ page 83}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \|-\boldsymbol{y} - \mathbb{0}\| < r\}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, -\boldsymbol{y}) < r\} && \text{by left hypothesis}\\
&\iff \boldsymbol{x} \in \{-\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\} && \text{let } \boldsymbol{u} \triangleq -\boldsymbol{y}\\
&\iff \boldsymbol{x} \in (-\{\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\})\\
&\iff \boldsymbol{x} \in (-\mathsf{B}(\mathbb{0}, r))\\
&\iff -\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r)
\end{aligned}$$

✎

Theorem 5.4 (page 87) demonstrates that if a metric d in a metric space $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d})$ is generated by a norm, then the ball $B(x, r)$ in that metric linear space is *convex*. However, the converse is not true. That is, it is possible for the balls in a metric space $(\boldsymbol{Y}, \mathsf{p})$ to be *convex*, but yet the metric p not be generated by a norm. Example 2.28 (page 63) gives one such example.

5.2.2 Norms generated by metrics

Every normed linear space is also a metric linear space (Theorem 5.3 page 86). That is, a metric linear space generates a *normed linear space*. However, the converse is not true—not every metric linear space is a *normed linear space*. A characterization of metric linear spaces that *are* normed linear spaces is given by Theorem 5.5 page 90.

Lemma 5.1. [10] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d})$ *be a* METRIC LINEAR SPACE. *Let* $\|\boldsymbol{x}\| \triangleq \mathsf{d}(\boldsymbol{x}, \mathbb{0})\ \forall \boldsymbol{x} \in X$.

LEM

$$\underbrace{\mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z}) = \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \quad \forall \boldsymbol{x,y,z} \in X}_{\text{TRANSLATION INVARIANT}} \implies \left\{\begin{array}{lrcll}
1. & \|\boldsymbol{x}\| &=& \|-\boldsymbol{x}\| & \forall \boldsymbol{x} \in X \quad \textit{and}\\
2. & \|\boldsymbol{x}\| &=& 0 \iff \boldsymbol{x} = 0 & \forall \boldsymbol{x} \in X \quad \textit{and}\\
3. & \|\boldsymbol{x}+\boldsymbol{y}\| &\le& \|\boldsymbol{x}\| + \|\boldsymbol{y}\| & \forall \boldsymbol{x,y} \in X
\end{array}\right.$$

[10] Oikhberg and Rosenthal (2007), page 599

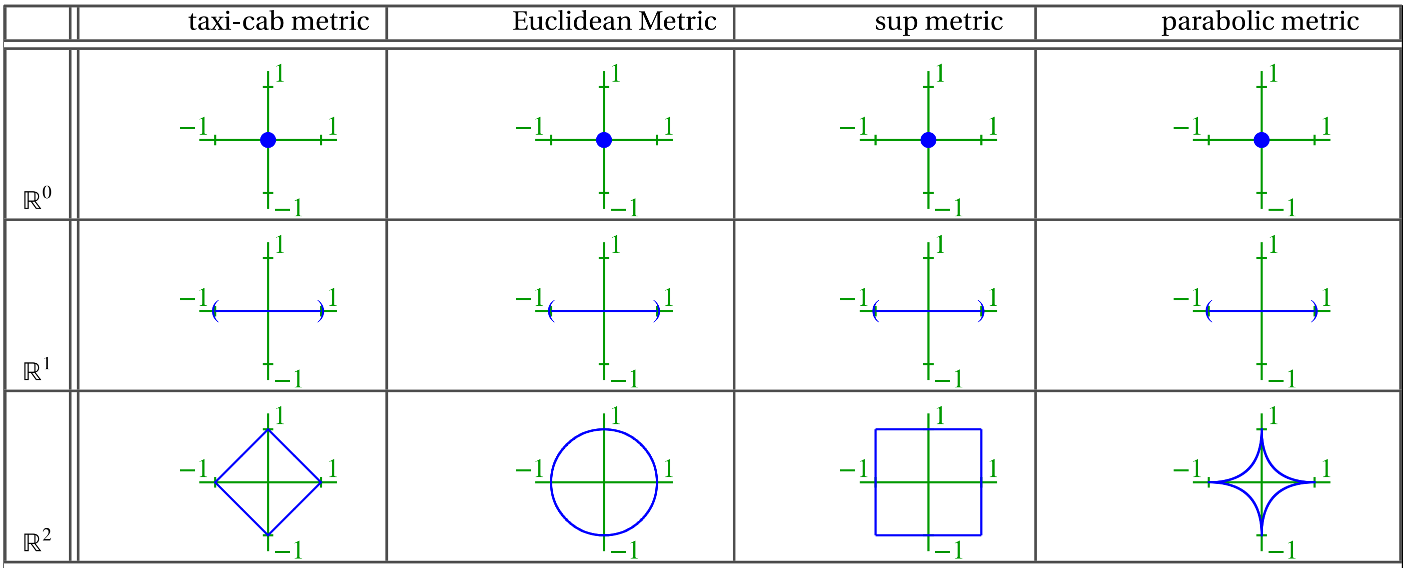


Figure 5.1: Open balls in $(\mathbb{R}^0, \mathsf{d}_n)$, $(\mathbb{R}, \mathsf{d}_n)$, $(\mathbb{R}^2, \mathsf{d}_n)$, and $(\mathbb{R}^3, \mathsf{d}_n)$.

✎PROOF:

1. Proof that $\|\boldsymbol{x}\| = \|-\boldsymbol{x}\|$:

$$\begin{aligned}
\|\boldsymbol{x}\| &= \mathsf{d}(\boldsymbol{x}, \mathbb{0}) && \text{by definition of } \|\cdot\| \\
&= \mathsf{d}(\boldsymbol{x}-\boldsymbol{x}, \mathbb{0}-\boldsymbol{x}) && \text{by translation invariance hypothesis} \\
&= \mathsf{d}(\mathbb{0}, -\boldsymbol{x}) \\
&= \|-\boldsymbol{x}\| && \text{by definition of } \|\cdot\|
\end{aligned}$$

2a. Proof that $\|\boldsymbol{x}\| = 0 \implies \boldsymbol{x} = 0$:

$$\begin{aligned}
0 &= \|\boldsymbol{x}\| && \text{by left hypothesis} \\
&= \mathsf{d}(\boldsymbol{x}, \mathbb{0}) && \text{by definition of } \|\cdot\| \\
&= \mathsf{d}(\boldsymbol{x}, \mathbb{0}) && \text{by definition of } \|\cdot\| \\
&\implies \boldsymbol{x} = \mathbb{0} && \text{by property of metrics page 27}
\end{aligned}$$

2b. Proof that $\|\boldsymbol{x}\| = 0 \impliedby \boldsymbol{x} = 0$:

$$\begin{aligned}
\|\boldsymbol{x}\| &= \mathsf{d}(\boldsymbol{x}, \mathbb{0}) && \text{by definition of } \|\cdot\| \\
&= \mathsf{d}(\mathbb{0}, \mathbb{0}) && \text{by right hypothesis} \\
&= 0 && \text{by property of metrics page 27}
\end{aligned}$$

3. Proof that $\|\boldsymbol{x}+\boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$:

$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\| &= \mathsf{d}(\boldsymbol{x}+\boldsymbol{y}, \mathbb{0}) && \text{by definition of } \|\cdot\| \\
&= \mathsf{d}(\boldsymbol{x}+\boldsymbol{y}-\boldsymbol{y}, \mathbb{0}-\boldsymbol{y}) && \text{by translation invariance hypothesis} \\
&= \mathsf{d}(\boldsymbol{x}, -\boldsymbol{y}) \\
&\le \mathsf{d}(\boldsymbol{x}, \mathbb{0}) + \mathsf{d}(\mathbb{0}, \boldsymbol{y}) && \text{by property of metrics page 27} \\
&= \mathsf{d}(\boldsymbol{x}, \mathbb{0}) + \mathsf{d}(\boldsymbol{y}, \mathbb{0}) && \text{by property of metrics page 27} \\
&= \|\boldsymbol{x}\| + \|\boldsymbol{y}\| && \text{by definition of } \|\cdot\|
\end{aligned}$$

⇨

Theorem 5.5. [11] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *be a* LINEAR SPACE. *Let* $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \|\boldsymbol{x} - \boldsymbol{y}\|\ \forall \boldsymbol{x}, \boldsymbol{y} \in X$.

THM

$$\left.\begin{array}{llll} 1. & \mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z}) = \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) & \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \in X & \text{(TRANSLATION INVARIANT) and} \\ 2. & \mathsf{d}(\alpha\boldsymbol{x}, \alpha\boldsymbol{y}) = |\alpha|\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) & \forall \boldsymbol{x},\boldsymbol{y} \in X, \alpha \in \mathbb{F} & \text{(HOMOGENEOUS)} \end{array}\right\} \iff \|\cdot\| \text{ is a NORM}$$

✎PROOF:

1. Proof of $\Longrightarrow$ assertion:
   (a) Proof that $\|\cdot\|$ is *strictly positive*: This follows directly from the definition of d.
   (b) Proof that $\|\cdot\|$ is *nondegenerate*: This follows directly from Lemma 5.1 (page 88).
   (c) Proof that $\|\cdot\|$ is *homogeneous*: This follows from the second left hypothesis.
   (d) Proof that $\|\cdot\|$ satisfies the *triangle-inequality*: This follows directly from Lemma 5.1 (page 88).
2. Proof of $\Longleftarrow$ assertion:

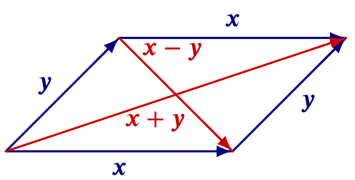
$$\begin{aligned} \mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z}) &= \|(\boldsymbol{x}+\boldsymbol{z}) - (\boldsymbol{y}+\boldsymbol{z})\| && \text{by definition of } \mathsf{d} \\ &= \|\boldsymbol{x} - \boldsymbol{y}\| \\ &= \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) && \text{by definition of } \mathsf{d} \\ \mathsf{d}(\alpha\boldsymbol{x}, \alpha\boldsymbol{y}) &= \|(\alpha\boldsymbol{x}) - (\alpha\boldsymbol{y})\| && \text{by definition of } \mathsf{d} \\ &= \|\alpha(\boldsymbol{x} - \boldsymbol{y})\| \\ &= |\alpha|\ \|\boldsymbol{x} - \boldsymbol{y}\| && \text{by definition of } \|\cdot\| \text{ page 83} \\ &= |\alpha|\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) && \text{by definition of } \mathsf{d} \end{aligned}$$

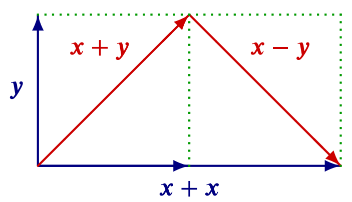
✏

5.3 Orthogonality on normed linear spaces

Traditionally, *orthogonality* (Definition 6.4 page 107) is a property defined in *inner product spaces* (Definition 6.1 page 95) However, the concept of orthogonality can be extended to *normed linear spaces* (Definition 5.1 page 83). Here are some examples:

①	*Isosceles orthogonality*:	Definition 5.3	page 90
②	*Pythagorean orthogonality*:	Definition 5.4	page 92
③	*Birkhoff orthogonality*:	Definition 5.5	page 93





Isosceles orthogonality (Definition 5.3 page 90) can be illustrated using a *parallelogram*, as illustrated in the figure to the upper left. In this case, orthogonality implies that the parallelogram is a rectangle, which in turn implies that the lengths of the two diagonals are equal ($\|\boldsymbol{x}+\boldsymbol{y}\| = \|\boldsymbol{x}-\boldsymbol{y}\|$). Isosceles orthogonality can also be illustrated with a triangle where the sides are of lengths $\|\boldsymbol{x}+\boldsymbol{y}\|$ and $\|\boldsymbol{x}-\boldsymbol{y}\|$ and base of length $\|\boldsymbol{x}+\boldsymbol{x}\|$. In this case if $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal, then the triangle is *isosceles*. This is illustrated in figure to the lower left. Isosceles orthogonality is formally defined next.

Definition 5.3. [12] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *be a* NORMED LINEAR SPACE *(Definition 5.1 page 83)*.

---

[11] Bollobás (1999), page 21

[12] James (1945) page 292 ⟨DEFINITION 2.1⟩, Amir (1986) page 24, Dunford and Schwartz (1957), page 93

DEF *Two vectors x and y are* ***orthogonal in the sense of James*** *if*

$$\|x+y\| = \|x-y\|.$$

This property is also called ***isosceles orthogonality*** *or* ***James orthogonality***.

Theorem 5.6. *Let* $\left(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *be an* INNER-PRODUCT SPACE (Definition 6.1 page 95) *with induced norm* $\|x\| \triangleq \sqrt{\langle x \mid x \rangle}$, ISOSCELES ORTHOGONALITY (Definition 5.3 page 90) *relation* $\obar$, *and inner-product relation* ORTHOGONALITY (Definition 6.4 page 107) *relation* $\perp$.

THM

$$\underbrace{x \obar y}_{\text{orthogonal in the sense of James}} \iff \underbrace{x \perp y}_{\text{orthogonal in the sense of inner-product space}}$$

✎PROOF:

1. Proof that $x \obar y \implies x \perp y$:

$$\begin{aligned}
&4\langle x \mid y \rangle \\
&= \underbrace{\|x+y\|^2 - \|x-y\|^2}_{0 \text{ by } x \obar y \text{ hypothesis}} + i\|x+iy\|^2 - i\|x-iy\|^2 && \text{by \textit{polarization identity} (Theorem 6.6 page 102)} \\
&= 0 + i\|x+iy\|^2 - i\|x-iy\|^2 && \text{by } x \obar y \text{ hypothesis} \\
&= i\left[\|x\|^2 + \|iy\|^2 + 2\Re\langle x \mid iy \rangle\right] \\
&\quad - i\left[\|x\|^2 + \|-iy\|^2 + 2\Re\langle x \mid -iy \rangle\right] && \text{by \textit{Polar Identity} (Lemma 6.1 page 99)} \\
&= i\left[\|x\|^2 + \|y\|^2 + 2\Re\langle x \mid iy \rangle\right] \\
&\quad - i\left[\|x\|^2 + \|y\|^2 + 2\Re\langle x \mid -iy \rangle\right] && \text{by Definition 5.1 page 83 and Definition 6.1 page 95} \\
&= 4i\Re\langle x \mid iy \rangle \\
&= 4i\Re\left[i^*\langle x \mid y \rangle\right] \\
&= 0 && \text{because inner-product space is \textit{real} } (\mathbb{F} = \mathbb{R})
\end{aligned}$$

2. Proof that $x \obar y \impliedby x \perp y$:

$$\begin{aligned}
\|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\Re\langle x \mid y \rangle && \text{by \textit{Polar Identity} (Lemma 6.1 page 99)} \\
&= \|x\|^2 + \|y\|^2 + 0 && \text{by } x \perp y \text{ hypothesis} \\
&= \|x\|^2 + \|y\|^2 - 2\Re\cancel{\langle x \mid y \rangle} \quad 0 \text{ when } x \perp y && \text{by } x \perp y \text{ hypothesis} \\
&= \|x\|^2 + \|-y\|^2 + 2\Re\langle x \mid -y \rangle \\
&= \|x-y\|^2 && \text{by \textit{Polar Identity} (Lemma 6.1 page 99)}
\end{aligned}$$

⇨

Theorem 5.7. [13] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *be a normed linear space and with* ISOSCELES ORTHOGONALITY (Definition 5.3 page 90) *relation* $\obar$.

THM

$$x \obar y \iff y \obar x \iff \alpha x \obar \alpha y \quad \forall \alpha \in \mathbb{F}$$

[13] Amir (1986) page 24

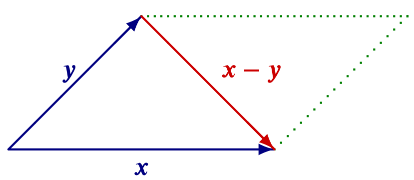
✎PROOF:

$$\begin{aligned}
\boldsymbol{x} ⦹ \boldsymbol{y} &\implies \|\boldsymbol{x}+\boldsymbol{y}\| &&= \|\boldsymbol{x}-\boldsymbol{y}\| && \text{by Definition 5.3 page 90}\\
&\implies \|\boldsymbol{x}+\boldsymbol{y}\| &&= |-1|\ \|\boldsymbol{x}-\boldsymbol{y}\| &&\\
&\implies \|\boldsymbol{x}+\boldsymbol{y}\| &&= \|-(\boldsymbol{x}-\boldsymbol{y})\| && \text{by Definition 5.1 page 83}\\
&\implies \|\boldsymbol{y}+\boldsymbol{x}\| &&= \|\boldsymbol{y}-\boldsymbol{x}\| && \text{by Definition 3.1 page 67}\\
&\implies \boldsymbol{y} ⦹ \boldsymbol{x} && && \text{by Definition 5.3 page 90}
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{y} ⦹ \boldsymbol{x} &\implies \|\boldsymbol{y}+\boldsymbol{x}\| &&= \|\boldsymbol{y}-\boldsymbol{x}\| && \text{by Definition 5.3 page 90}\\
&\implies |\alpha|\ \|\boldsymbol{y}+\boldsymbol{x}\| &&= |\alpha|\ \|\boldsymbol{y}-\boldsymbol{x}\| &&\\
&\implies \|\alpha(\boldsymbol{y}+\boldsymbol{x})\| &&= \|\alpha(\boldsymbol{y}-\boldsymbol{x})\| && \text{by Definition 5.1 page 83}\\
&\implies \|\alpha\boldsymbol{y}+\alpha\boldsymbol{x}\| &&= \|\alpha\boldsymbol{y}-\alpha\boldsymbol{x}\| &&\\
&\implies \|\alpha\boldsymbol{x}+\alpha\boldsymbol{y}\| &&= \|-(\alpha\boldsymbol{x}-\alpha\boldsymbol{y})\| && \text{by Definition 3.1 page 67}\\
&\implies \|\alpha\boldsymbol{x}+\alpha\boldsymbol{y}\| &&= |-1|\ \|\alpha\boldsymbol{x}-\alpha\boldsymbol{y}\| && \text{by Definition 5.1 page 83}\\
&\implies \|\alpha\boldsymbol{x}+\alpha\boldsymbol{y}\| &&= \|\alpha\boldsymbol{x}-\alpha\boldsymbol{y}\| && \text{by Definition F.4 page 346}\\
&\implies \alpha\boldsymbol{x} ⦹ \alpha\boldsymbol{y} && && \text{by Definition 5.3 page 90}
\end{aligned}$$

$$\begin{aligned}
\alpha\boldsymbol{x} ⦹ \alpha\boldsymbol{y} &\implies \|\alpha\boldsymbol{x}+\alpha\boldsymbol{y}\| &&= \|\alpha\boldsymbol{x}-\alpha\boldsymbol{y}\| && \text{by Definition 5.3 page 90}\\
&\implies \|\alpha(\boldsymbol{x}+\boldsymbol{y})\| &&= \|\alpha(\boldsymbol{x}-\boldsymbol{y})\| && \text{by Definition 3.1 page 67}\\
&\implies |\alpha|\ \|\boldsymbol{x}+\boldsymbol{y}\| &&= |\alpha|\ \|\boldsymbol{x}-\boldsymbol{y}\| && \text{by Definition 5.1 page 83}\\
&\implies \|\boldsymbol{x}+\boldsymbol{y}\| &&= \|\boldsymbol{x}-\boldsymbol{y}\| && \text{by Definition 5.1 page 83}\\
&\implies \boldsymbol{x} ⦹ \boldsymbol{y} && && \text{by Definition 5.3 page 90}
\end{aligned}$$

⇰



If a triangle in a plane has two perpendicular sides of lengths a and b and a hypotenuse of length c, then by the *Pythagorean Theorem* (Theorem 6.10 page 108), $a^2 + b^2 = c^2$. This concept of orthogonality can be generalized to normed linear spaces. Two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ (with lengths $\|\boldsymbol{x}\|$ and $\|\boldsymbol{y}\|$) are orthogonal when $\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 = \|\boldsymbol{x}-\boldsymbol{y}\|^2$ ($\boldsymbol{x}-\boldsymbol{y}$ is a kind of "*hypotenuse*"). This kind of orthogonality is defined next and illustrated in the figure to the left.

Definition 5.4. [14] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a* NORMED LINEAR SPACE *(Definition 5.1 page 83).*

DEF

Two vectors $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are* ***orthogonal in the Pythagorean sense*** *if*

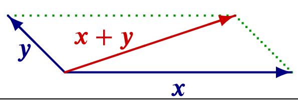
$$\|\boldsymbol{x}-\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.$$

This relationship is also called ***Pythagorean orthogonality****.*

Theorem 5.8. [15] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \nabla \rangle)$ *be an* INNER-PRODUCT SPACE *(Definition 6.1 page 95) with induced norm* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \,|\, \boldsymbol{x}\rangle}$, PYTHAGOREAN ORTHOGONALITY *(Definition 5.4 page 92) relation* ⦹, *and inner-product relation* ORTHOGONALITY *(Definition 6.4 page 107) relation* $\perp$.

THM

$$\underbrace{\boldsymbol{x} ⦹ \boldsymbol{y}}_{\text{orthogonal in the Pythagorean sense}} \iff \underbrace{\boldsymbol{x} \perp \boldsymbol{y}}_{\text{orthogonal in the sense of inner-product space}}$$



Besides *isosceles orthogonality* (Definition 5.3 page 90), orthogonality in normed linear spaces can be defined using *Birkhoff orthogonality*, as defined in Definition 5.5 (next) and illustrated to the left.

[14] James (1945) page 292 ⟨DEFINITION 2.2⟩, Amir (1986) page 57, Drljević (1989) page 232
[15] Amir (1986) page 57

Definition 5.5. [16] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a* NORMED LINEAR SPACE *(Definition 5.1 page 83).*

DEF

Two vectors $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are* ***orthogonal in the sense of Birkhoff*** *if*

$$\|\boldsymbol{x}\| \le \|\boldsymbol{x} + \alpha\boldsymbol{y}\| \qquad \forall \alpha \in \mathbb{F}.$$

This relationship is also called ***Birkhoff orthogonality***.

Theorem 5.9. *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be an* INNER-PRODUCT SPACE *(Definition 6.1 page 95) with induced norm* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$, BIRKHOFF ORTHOGONALITY *relation* ⦶ *(Definition 5.5 page 93), and inner-product relation* ORTHOGONALITY *relation* $\perp$ *(Definition 6.4 page 107).*

THM

$$\underbrace{\boldsymbol{x} \circledcirc \boldsymbol{y}}_{\text{orthogonal in the sense of Birkhoff}} \iff \underbrace{\boldsymbol{x} \perp \boldsymbol{y}}_{\text{orthogonal in the sense of inner-product space}}$$

[16] Amir (1986) page 33, Dunford and Schwartz (1957), page 93, James (1947) page 265

CHAPTER 6

INNER PRODUCT SPACES

6.1 Definition and basic results

Definition 6.1. [1] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE (Definition 3.1 page 67).

DEF

A FUNCTIONAL (Definition 13.1 page 207) $\langle \triangle \mid \triangledown \rangle \in \mathbb{F}^{X \times X}$ *is an* ***inner product*** *on* $\boldsymbol{\Omega}$ *if*

1.	$\langle \alpha \boldsymbol{x} \mid \boldsymbol{y} \rangle$	$=$	$\alpha \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$	$\forall \boldsymbol{x}, \boldsymbol{y} \in X,\ \forall \alpha \in \mathbb{C}$	(HOMOGENEOUS)	*and*
2.	$\langle \boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{u} \rangle$	$=$	$\langle \boldsymbol{x} \mid \boldsymbol{u} \rangle + \langle \boldsymbol{y} \mid \boldsymbol{u} \rangle$	$\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u} \in X$	(ADDITIVE)	*and*
3.	$\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$	$=$	$\langle \boldsymbol{y} \mid \boldsymbol{x} \rangle^*$	$\forall \boldsymbol{x}, \boldsymbol{y} \in X$	(CONJUGATE SYMMETRIC).	*and*
4.	$\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle$	$\geq$	0	$\forall \boldsymbol{x} \in X$	(NON-NEGATIVE)	*and*
5.	$\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle$	$=$	$0 \iff \boldsymbol{x} = \mathbb{0}$	$\forall \boldsymbol{x} \in X$	(NON-ISOTROPIC)	

An inner product is also called a ***scalar product****.*
The tuple $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *is called an* ***inner product space****.*

Theorem 6.1. [2] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be a* LINEAR SPACE (Definition 3.1 page 67).

THM

1.	$\langle \boldsymbol{x} \mid \boldsymbol{y} + \boldsymbol{z} \rangle$	$=$	$\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + \langle \boldsymbol{x} \mid \boldsymbol{z} \rangle$	$\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$		
2.	$\langle \boldsymbol{x} \mid \alpha \boldsymbol{y} \rangle$	$=$	$\alpha^* \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$	$\forall \boldsymbol{x}, \boldsymbol{y} \in X,\ \alpha \in \mathbb{F}$		
3.	$\langle \boldsymbol{x} \mid \mathbb{0} \rangle$	$=$	$\langle \mathbb{0} \mid \boldsymbol{x} \rangle = 0$	$\forall \boldsymbol{x} \in X$		
4.	$\langle \boldsymbol{x} - \boldsymbol{y} \mid \boldsymbol{z} \rangle$	$=$	$\langle \boldsymbol{x} \mid \boldsymbol{z} \rangle - \langle \boldsymbol{y} \mid \boldsymbol{z} \rangle$	$\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$		
5.	$\langle \boldsymbol{x} \mid \boldsymbol{y} - \boldsymbol{z} \rangle$	$=$	$\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle - \langle \boldsymbol{x} \mid \boldsymbol{z} \rangle$	$\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$		
6.	$\langle \boldsymbol{x} \mid \boldsymbol{z} \rangle$	$=$	$\langle \boldsymbol{y} \mid \boldsymbol{z} \rangle$	$\forall \boldsymbol{z} \in X \neq \{\mathbb{0}\}$	$\iff$	$\boldsymbol{x} = \boldsymbol{y}$
7.	$\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$	$=$	0	$\forall \boldsymbol{x} \in X$	$\iff$	$\boldsymbol{y} = \mathbb{0}$

✎PROOF:

$$
\begin{aligned}
\langle \boldsymbol{x} \mid \boldsymbol{y} + \boldsymbol{z} \rangle &= \langle \boldsymbol{y} + \boldsymbol{z} \mid \boldsymbol{x} \rangle^* && \text{by } \textit{conjugate symmetric} \text{ property of } \langle \triangle \mid \triangledown \rangle && \text{(Definition 6.1 page 95)} \\
&= (\langle \boldsymbol{y} \mid \boldsymbol{x} \rangle + \langle \boldsymbol{z} \mid \boldsymbol{x} \rangle)^* && \text{by } \textit{additive} \text{ property of } \langle \triangle \mid \triangledown \rangle && \text{(Definition 6.1 page 95)} \\
&= \langle \boldsymbol{y} \mid \boldsymbol{x} \rangle^* + \langle \boldsymbol{z} \mid \boldsymbol{x} \rangle^* && \text{by } \textit{distributive} \text{ property of } * && \text{(Definition 16.3 page 258)} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + \langle \boldsymbol{x} \mid \boldsymbol{z} \rangle && \text{by } \textit{conjugate symmetric} \text{ property of } \langle \triangle \mid \triangledown \rangle && \text{(Definition 6.1 page 95)} \\
\langle \boldsymbol{x} \mid \alpha \boldsymbol{y} \rangle &= \langle \alpha \boldsymbol{y} \mid \boldsymbol{x} \rangle^* && \text{by } \textit{conjugate symmetric} \text{ property of } \langle \triangle \mid \triangledown \rangle && \text{(Definition 6.1 page 95)}
\end{aligned}
$$

[1] Istrăţescu (1987) page 111 ⟨Definition 4.1.1⟩, Bollobás (1999) pages 130–131, Haaser and Sullivan (1991) page 277, Aliprantis and Burkinshaw (1998), page 276, Peano (1888b) page 72

[2] Berberian (1961) page 27, Haaser and Sullivan (1991) page 277

$$
\begin{aligned}
&= (\alpha \langle \boldsymbol{y} \,|\, \boldsymbol{x} \rangle)^* && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)}\\
&= \alpha^* \langle \boldsymbol{y} \,|\, \boldsymbol{x} \rangle^* && \text{by } \textit{antiautomorphic} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= \alpha^* \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle && \text{by } \textit{conjugate symmetric} \text{ property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)}\\
\langle \boldsymbol{x} \,|\, \mathbb{0} \rangle &= \langle \mathbb{0} \,|\, \boldsymbol{x} \rangle^* && \text{by } \textit{conjugate symmetric} \text{ property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)}\\
&= \langle 0 \cdot \boldsymbol{y} \,|\, \boldsymbol{x} \rangle^* && &&\\
&= (0 \cdot \langle \boldsymbol{y} \,|\, \boldsymbol{x} \rangle)^* && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)}\\
&= 0 && &&\\
\langle \mathbb{0} \,|\, \boldsymbol{x} \rangle &= \langle 0 \cdot \boldsymbol{y} \,|\, \boldsymbol{x} \rangle && &&\\
&= (0 \cdot \langle \boldsymbol{y} \,|\, \boldsymbol{x} \rangle) && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)}\\
&= 0 && &&\\
\langle \boldsymbol{x} - \boldsymbol{y} \,|\, \boldsymbol{z} \rangle &= \langle \boldsymbol{x} + (-\boldsymbol{y}) \,|\, \boldsymbol{z} \rangle && \text{by definition of } + &&\\
&= \langle \boldsymbol{x} \,|\, \boldsymbol{z} \rangle + \langle -\boldsymbol{y} \,|\, \boldsymbol{z} \rangle && \text{by } \textit{additive} \text{ property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)}\\
&= \langle \boldsymbol{x} \,|\, \boldsymbol{z} \rangle - \langle \boldsymbol{y} \,|\, \boldsymbol{z} \rangle && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)}\\
\langle \boldsymbol{x} \,|\, \boldsymbol{y} - \boldsymbol{z} \rangle &= \langle \boldsymbol{y} - \boldsymbol{z} \,|\, \boldsymbol{x} \rangle^* && \text{by } \textit{conjugate symmetric} \text{ property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)}\\
&= (\langle \boldsymbol{y} \,|\, \boldsymbol{x} \rangle - \langle \boldsymbol{z} \,|\, \boldsymbol{x} \rangle)^* && \text{by 4.} &&\\
&= \langle \boldsymbol{y} \,|\, \boldsymbol{x} \rangle^* - \langle \boldsymbol{z} \,|\, \boldsymbol{x} \rangle^* && \text{by } \textit{distributive} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle - \langle \boldsymbol{x} \,|\, \boldsymbol{z} \rangle && \text{by } \textit{conjugate symmetric} \text{ property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)}
\end{aligned}
$$

$$
\begin{aligned}
& & \langle \boldsymbol{x} \,|\, \boldsymbol{z} \rangle = \langle \boldsymbol{y} \,|\, \boldsymbol{z} \rangle & \quad \forall \boldsymbol{z} &&\\
&\iff & \langle \boldsymbol{x} \,|\, \boldsymbol{z} \rangle - \langle \boldsymbol{y} \,|\, \boldsymbol{z} \rangle = 0 & \quad \forall \boldsymbol{z} && \text{by property of complex numbers}\\
&\iff & \langle \boldsymbol{x} - \boldsymbol{y} \,|\, \boldsymbol{z} \rangle = 0 & \quad \forall \boldsymbol{z} && \text{by 4.}\\
&\iff & \boldsymbol{x} - \boldsymbol{y} = \mathbb{0} & \quad \forall \boldsymbol{z} && \text{by } \textit{non-isotropic} \text{ property of } \langle \triangle \,|\, \nabla \rangle \text{ (Definition 6.1 page 95)}
\end{aligned}
$$

Proof that $\langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \implies \boldsymbol{y} = \mathbb{0}$:

1. Suppose $\boldsymbol{y} \neq \mathbb{0}$;
2. Then $\langle \boldsymbol{y} \,|\, \boldsymbol{y} \rangle \neq \mathbb{0}$ by the *non-isotropic* property of $\langle \triangle \,|\, \nabla \rangle$ (Definition 6.1 page 95)
3. But because $\boldsymbol{y} \in X$, the left hypothesis implies that $\langle \boldsymbol{y} \,|\, \boldsymbol{y} \rangle = 0$.
4. This is a *contradiction*.
5. Therefore $\boldsymbol{y} \neq \mathbb{0}$ must be incorrect and $\boldsymbol{y} = \mathbb{0}$ must be correct.

Proof that $\langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \impliedby \boldsymbol{y} = \mathbb{0}$:

$$
\begin{aligned}
\langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle &= \langle \boldsymbol{x} \,|\, \mathbb{0} \rangle && \text{by right hypothesis}\\
&= 0 && \text{by Theorem 6.1 page 95}
\end{aligned}
$$

⇰

One of the most useful and widely used inequalities in analysis is the *Cauchy-Schwarz Inequality* (sometimes also called the *Cauchy-Bunyakovsky-Schwarz Inequality*). In fact, we will use this inequality shortly to prove that every inner product space *has* a norm and therefore every inner product space *is* a normed linear space.

Theorem 6.2 (Cauchy-Schwarz Inequality). [3] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \nabla \rangle)$ *be an* INNER PRODUCT SPACE *and* $|\cdot| \in \mathbb{R}^{\mathbb{C}}$ *an* ABSOLUTE VALUE *function* (Definition F.4 page 346)*. Let* $\|\cdot\|$ *be a function in* $\mathbb{R}^{\mathbb{F}}$ *such*

[3] Haaser and Sullivan (1991) page 278, Aliprantis and Burkinshaw (1998), page 278, Cauchy (1821) page 455, Bunyakovsky (1859) page 6, Schwarz (1885)

that $\|x\| \triangleq \sqrt{\langle x \,|\, x\rangle}$.[4]

THM

$$\begin{array}{rcllll}
|\langle x \,|\, y\rangle|^2 & \le & \langle x \,|\, x\rangle\ \langle y \,|\, y\rangle & & & \forall x,y\in X \\
|\langle x \,|\, y\rangle|^2 & = & \langle x \,|\, x\rangle\ \langle y \,|\, y\rangle & \iff & \exists \alpha \in \mathbb{F} \text{ such that } y = \alpha x & \forall x,y\in X \\
|\langle x \,|\, y\rangle| & \le & \|x\|\ \|y\| & & & \forall x,y\in X \\
|\langle x \,|\, y\rangle| & = & \|x\|\ \|y\| & \iff & \exists \alpha \in \mathbb{F} \text{ such that } y = \alpha x & \forall x,y\in X
\end{array}$$

✎PROOF:

1. Proof that $|\langle x \,|\, y\rangle| \le \|x\|\,\|y\|$: [5]

   (a) $y = \mathbb{0}$ case:

$$\begin{aligned}
|\langle x \,|\, y\rangle|^2 &= |\langle x \,|\, \mathbb{0}\rangle|^2 && \text{by } y = \mathbb{0} \text{ hypothesis} \\
&= |\langle \mathbb{0} \,|\, x\rangle|^2 && \text{by Definition 6.1 page 95} \\
&= |\langle 0\mathbb{0} \,|\, x\rangle|^2 && \text{by Definition 3.1 page 67} \\
&= |0\,\langle \mathbb{0} \,|\, x\rangle|^2 && \text{by Definition 6.1 page 95} \\
&= 0 \\
&= \langle x \,|\, x\rangle\,\langle \mathbb{0} \,|\, \mathbb{0}\rangle \\
&= \langle x \,|\, x\rangle\,\langle y \,|\, y\rangle && \text{by } y = \mathbb{0} \text{ hypothesis}
\end{aligned}$$

   (b) $y \ne \mathbb{0}$ case: Let $\lambda \triangleq \frac{\langle x \,|\, y\rangle}{\langle y \,|\, y\rangle}$.

$$\begin{aligned}
0 &\le \langle x - \lambda y \,|\, x - \lambda y\rangle && \text{by Definition 6.1} \\
&= \langle x \,|\, x - \lambda y\rangle + \langle -\lambda y \,|\, x - \lambda y\rangle && \text{by Definition 6.1} \\
&= \langle x - \lambda y \,|\, x\rangle^* - \lambda\,\langle y \,|\, x - \lambda y\rangle && \text{by Definition 6.1} \\
&= \langle x \,|\, x\rangle^* + \langle -\lambda y \,|\, x\rangle^* - \lambda\,\langle x - \lambda y \,|\, y\rangle^* && \text{by Definition 6.1} \\
&= \langle x \,|\, x\rangle^* - \lambda^*\,\langle y \,|\, x\rangle^* - \lambda\,\langle x \,|\, y\rangle^* - \lambda\,\langle -\lambda y \,|\, y\rangle^* && \text{by Definition 6.1} \\
&= \langle x \,|\, x\rangle - \lambda^*\,\langle x \,|\, y\rangle - \lambda\,\langle x \,|\, y\rangle^* + \lambda\lambda^*\,\langle y \,|\, y\rangle^* && \text{by Definition 6.1} \\
&= \langle x \,|\, x\rangle + \left[\cancelto{0}{\frac{\langle x \,|\, y\rangle}{\langle y \,|\, y\rangle}\lambda^*\,\langle y \,|\, y\rangle - \lambda^*\,\langle x \,|\, y\rangle}\right] - \frac{\langle x \,|\, y\rangle}{\langle y \,|\, y\rangle}\,\langle x \,|\, y\rangle^* && \text{by definition of } \lambda \\
&= \langle x \,|\, x\rangle - \frac{1}{\langle y \,|\, y\rangle}\ |\langle x \,|\, y\rangle|^2 \\
\implies |\langle x \,|\, y\rangle|^2 &\le \langle x \,|\, x\rangle\ \langle y \,|\, y\rangle
\end{aligned}$$

2. Proof that $|\langle x \,|\, y\rangle|^2 = \langle x \,|\, x\rangle\ \langle y \,|\, y\rangle \iff y = ax$:
   Let $\frac{1}{a} \triangleq \lambda \triangleq \frac{\langle x \,|\, y\rangle}{\langle y \,|\, y\rangle}$. Then...

$$\begin{aligned}
& y = ax \\
\iff\ & x = \lambda y \\
\iff\ & x - \lambda y = \mathbb{0} \\
\iff\ & 0 = \langle x - \lambda y \,|\, x - \lambda y\rangle && \text{by Definition 6.1 page 95} \\
& \quad = \langle x \,|\, x - \lambda y\rangle + \langle -\lambda y \,|\, x - \lambda y\rangle && \text{by Definition 6.1 page 95} \\
& \quad = \langle x - \lambda y \,|\, x\rangle^* - \lambda\,\langle y \,|\, x - \lambda y\rangle && \text{by Definition 6.1 page 95} \\
& \quad \vdots && \text{(same steps as in 1(b))}
\end{aligned}$$

[4] The function $\|\cdot\|$ is a *norm* (Theorem 6.4 page 100) and is called the *norm induced by the inner product* $\langle \triangle \,|\, \nabla\rangle$ (Definition 6.2 page 100).

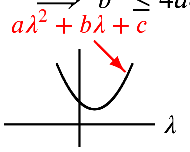
[5] Haaser and Sullivan (1991), page 278

$$= \langle x \mid x\rangle - \frac{1}{\langle y \mid y\rangle} \left|\langle x \mid y\rangle\right|^2$$

$$\iff \qquad \left|\langle x \mid y\rangle\right|^2 = \langle x \mid x\rangle \, \langle y \mid y\rangle$$

3. Alternate proof for $|\langle x \mid y\rangle| \le \|x\| \, \|y\|$: (Note: This is essentially the same proof as used by Schwarz). [6]

   (a) Proof that $\left(a\lambda^2 + b\lambda + c \ge 0 \quad \forall \lambda \in \mathbb{R}\right) \implies b^2 \le 4ac$ (quadratic discriminate inequality):



   Let $k \in (0, \infty)$, and $r_1, r_2 \in \mathbb{C}$ be the roots of $a\lambda^2 + b\lambda + c = 0$. Then

$$\begin{aligned}
0 &\le a\lambda^2 + b\lambda + c && \text{by left hypothesis}\\
&= k(\lambda - r_1)(\lambda - r_2) && \text{by definition of } r_1 \text{ and } r_2\\
&= k(\lambda^2 - r_1\lambda - r_2\lambda + r_1 r_2)
\end{aligned}$$

$$\implies \quad \lambda^2 - r_1\lambda - r_2\lambda + r_1 r_2 \ge 0$$

$$\implies \quad r_1 = r_2^* \qquad \text{because } r_1 r_2 \ge 0 \text{ for } \lambda = 0$$

   The *quadratic equation* places another constraint on r_1 and r_2:

$$\begin{aligned}
\frac{b^2 + \sqrt{b^2 - 4ac}}{2a} &= r_1 && \text{by quadratic equation}\\
&= r_2^* && \text{by previous result}\\
&= \left(\frac{b^2 - \sqrt{b^2 - 4ac}}{2a}\right)^* && \text{by quadratic equation}
\end{aligned}$$

   The only way for this to be true is if $b^2 \le 4ac$ (the **discriminate** is non-positive).

   (b) Proof that $\langle y \mid y\rangle \, \lambda^2 + 2|\langle x \mid y\rangle|\lambda + \langle x \mid x\rangle \ge 0 \quad \forall \lambda \in \mathbb{R}$:

$$\begin{aligned}
0 &\le \langle x + \alpha y \mid x + \alpha y\rangle && \text{by Definition 6.1 page 95}\\
&= \langle x \mid x + \alpha y\rangle + \langle \alpha y \mid x + \alpha y\rangle && \text{by Definition 6.1 page 95}\\
&= \langle x \mid x + \alpha y\rangle + \alpha \, \langle y \mid x + \alpha y\rangle && \text{by Definition 6.1 page 95}\\
&= \langle x + \alpha y \mid x\rangle^* + \alpha \, \langle x + \alpha y \mid y\rangle^* && \text{by Definition 6.1 page 95}\\
&= \langle x \mid x\rangle^* + \langle \alpha y \mid x\rangle^* + \alpha \, \langle x \mid y\rangle^* + \alpha \, \langle \alpha y \mid y\rangle^* && \text{by Definition 6.1 page 95}\\
&= \langle x \mid x\rangle^* + \alpha^* \, \langle y \mid x\rangle^* + \alpha \, \langle x \mid y\rangle^* + \alpha\alpha^* \, \langle y \mid y\rangle^* && \text{by Definition 6.1 page 95}\\
&= \langle x \mid x\rangle + \alpha^* \, \langle x \mid y\rangle + \left(\alpha^* \, \langle x \mid y\rangle\right)^* + |\alpha|^2 \, \langle y \mid y\rangle && \text{by Definition 6.1 page 95}\\
&= \langle x \mid x\rangle + 2\Re\left(\alpha^* \, \langle x \mid y\rangle\right) + |\alpha|^2 \, \langle y \mid y\rangle\\
&\le \langle x \mid x\rangle + 2\left|\alpha^* \, \langle x \mid y\rangle\right| + |\alpha|^2 \, \langle y \mid y\rangle\\
&= \langle x \mid x\rangle + 2|\langle x \mid y\rangle||\alpha| + \langle y \mid y\rangle \, |\alpha|^2\\
&= \langle y \mid y\rangle \, |\alpha|^2 + 2\,|\langle x \mid y\rangle| \, |\alpha| + \langle x \mid x\rangle\\
&= \underbrace{\langle y \mid y\rangle}_{a}\lambda^2 + \underbrace{2|\langle x \mid y\rangle|}_{b}\lambda + \underbrace{\langle x \mid x\rangle}_{c} && \text{because } \lambda \triangleq |\alpha| \in \mathbb{R}
\end{aligned}$$

   (c) The above equation is in the quadratic form used in the lemma of part (a).

$$\left(\underbrace{2|\langle x \mid y\rangle|}_{b}\right)^2 \le 4 \, \underbrace{\langle y \mid y\rangle}_{a} \, \underbrace{\langle x \mid x\rangle}_{c} \qquad \text{by the results of parts (a) and (b)}$$

$$\implies \quad |\langle x \mid y\rangle|^2 \le \langle x \mid x\rangle \, \langle y \mid y\rangle$$

[6] Aliprantis and Burkinshaw (1998), page 278, Steele (2004), page 11

4. Proof that $|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle| \le \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|$:
   This follows directly from the definition $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$.

5. Proof that $|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle| = \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| \iff \exists \alpha \in \mathbb{C}$ such that $\boldsymbol{y} = \alpha \boldsymbol{x}$:
   This follows directly from the definition $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$.

⇨

Corollary 6.1. [7] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *be an* INNER PRODUCT SPACE.

COR: $\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$ *is* CONTINUOUS *(Definition 1.8 page 23) in both* $\boldsymbol{x}$ *and* $\boldsymbol{y}$.

✎PROOF: Let $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$.

$$\begin{aligned}
|\langle \boldsymbol{x} + \boldsymbol{\epsilon} \mid \boldsymbol{y} \rangle - \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 &= |\langle \boldsymbol{x} + \boldsymbol{\epsilon} - \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 && \text{by } \textit{additivity} \text{ of } \langle \triangle \mid \triangledown \rangle && \text{(Definition 6.1 page 95)} \\
&= |\langle \boldsymbol{\epsilon} \mid \boldsymbol{y} \rangle|^2 \\
&\le \|\boldsymbol{\epsilon}\|^2 \, \|\boldsymbol{y}\| && \text{by } \textit{Cauchy-Schwarz Inequality} && \text{(Theorem 6.2 page 96)}
\end{aligned}$$

⇨

6.2 Relationship between norms and inner products

6.2.1 Norms induced by inner products

Lemma 6.1 (Polar Identity). [8] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *be an* INNER PRODUCT SPACE *(Definition 6.1 page 95). Let* $\Re z$ *represent the real part of* $z \in \mathbb{C}$*. Let* $\|\cdot\|$ *be a function in* $\mathbb{R}^{\mathbb{F}}$ *such that* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$.[9]

LEM: $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + 2\mathbf{R}_{\mathrm{e}} \left[\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle\right] + \|\boldsymbol{y}\|^2 \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in X$

✎PROOF:

$$\begin{aligned}
\|\boldsymbol{x} + \boldsymbol{y}\|^2 &= \langle \boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{x} + \boldsymbol{y} \rangle && \text{by definition of } \textit{induced norm} && \text{(Theorem 6.4 page 100)} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} + \boldsymbol{y} \rangle + \langle \boldsymbol{y} \mid \boldsymbol{x} + \boldsymbol{y} \rangle && \text{by Definition 6.1 page 95} \\
&= \langle \boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{x} \rangle^* + \langle \boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{y} \rangle^* && \text{by Definition 6.1 page 95} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle^* + \langle \boldsymbol{y} \mid \boldsymbol{x} \rangle^* + \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle^* + \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle^* && \text{by Definition 6.1 page 95} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle + \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle^* + \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle && \text{by definition of } \textit{inner product} && \text{(Definition 6.1 page 95)} \\
&= \|\boldsymbol{x}\|^2 + 2\Re \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 && \text{by definition of } \textit{induced norm} && \text{(Theorem 6.4 page 100)}
\end{aligned}$$

⇨

[7] Bollobás (1999) page 132, Aliprantis and Burkinshaw (1998) page 279 ⟨Lemma 32.4⟩
[8] Conway (1990) page 4, Heil (2011) page 27 ⟨Lemma 1.36(a)⟩
[9] The function $\|\cdot\|$ is a *norm* (Theorem 6.4 page 100) and is called the *norm induced by the inner product* $\langle \triangle \mid \triangledown \rangle$ (Definition 6.2 page 100).

Theorem 6.3 (Minkowski's inequality). [10] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be an* INNER PRODUCT SPACE. *Let* $\|\cdot\|$ *be a function in* $\mathbb{R}^{\mathbb{F}}$ *such that* $\|x\| \triangleq \sqrt{\langle x \mid x \rangle}$.[11]

THM $$\|x + y\| \le \|x\| + \|y\| \qquad \forall x,y \in X$$

✎PROOF:

$$\begin{aligned}
\|x+y\|^2 &= \|x\|^2 + 2\Re\langle x \mid y\rangle + \|y\|^2 && \text{by } \textit{Polar Identity} \text{ (Lemma 6.1 page 99)}\\
&\le \|x\|^2 + 2|\langle x \mid y\rangle| + \|y\|^2 \\
&\le \|x\|^2 + 2\sqrt{\langle x \mid x\rangle}\sqrt{\langle y \mid y\rangle} + \|y\|^2 && \text{by } \textit{Cauchy-Schwarz Inequality} \text{ (Theorem 6.2 page 96)}\\
&= (\|x\| + \|y\|)^2
\end{aligned}$$

✏

Theorem 6.4 (induced norm). [12] *Let* $\Omega \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be an* INNER PRODUCT SPACE *(Definition 6.1 page 95)*.

THM $$\|x\| \triangleq \sqrt{\langle x \mid x \rangle} \quad \Longrightarrow \quad \|\cdot\| \text{ is a NORM}$$

✎PROOF: For a function to be a norm, it must satisfy the four properties listed in Definition 5.1 (page 83).

1. Proof that $\|\cdot\|$ is a norm:
   (a) Proof that $\|x\| > 0$ for $x \neq 0$ (non-negative):
   By Definition 6.1 page 95, all inner products have this property.
   (b) Proof that $\|x\| = 0 \iff x = 0$ (non-isometric):
   By Definition 6.1, all inner products have this property.
   (c) Prove $\|ax\| = |a|\,\|x\|$ (homogeneous):
   $$\|ax\| \triangleq \sqrt{\langle ax \mid ax\rangle} = \sqrt{aa^* \langle x \mid x\rangle} = \sqrt{|a|^2 \langle x \mid x\rangle} = |a|\,\|x\|$$
   (d) Proof that $\|x + y\| \le \|x\| + \|y\|$ (subadditive): This is true by Minkowski's inequality page 100
2. Proof that every inner product space is a normed linear space:
   Since every inner product induces a norm, so every inner product space has a norm (the norm induced by the inner product) and is therefore a normed linear space.

✏

Theorem 6.4 (previous theorem) demonstrates that in any inner product space $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$, the function $\|x\| \triangleq \sqrt{\langle x \mid x \rangle}$ is a norm. That is, $\|x\|$ is the *norm induced by the inner product.* This norm is formally defined next.

Definition 6.2. *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be an* INNER PRODUCT SPACE *(Definition 6.1 page 95)*.

DEF *The* ***norm induced by the inner product*** $\langle \triangle \mid \triangledown \rangle$ *is defined as*
$$\|x\| \triangleq \sqrt{\langle x \mid x \rangle}$$

[10] Aliprantis and Burkinshaw (1998) pages 278–279 ⟨Theorem 32.3⟩, Maligranda (1995), Minkowski (1910) page 115

[11] The function $\|\cdot\|$ is a *norm* (Theorem 6.4 page 100) and is called the *normed induced by the inner product* $\langle \triangle \mid \triangledown \rangle$ (Definition 6.2 page 100).

[12] Aliprantis and Burkinshaw (1998), pages 278–279, Haaser and Sullivan (1991) page 278

6.2.2 Inner products induced by norms

Theorem 6.4 (page 100) demonstrates that if a *linear space* (Definition 3.1 page 67) has an *inner product* (Definition 6.1 page 95), then that inner product always induces a *norm* (Definition 5.1 page 83), and the relationship between the two is simply $\|x\| = \sqrt{\langle x \mid x\rangle}$ (Definition 6.2 page 100). But what about the converse? What if a linear space has a norm—can that norm also induce an inner product? The answer in general is "no": Not all norms can induce an inner product. But a less harsh answer is "sometimes": Some norms **can** induce inner products. This leads to some important and interesting questions:

1. How many different inner products can be induced from a single norm? The answer turns out to be **at most** one, but maybe none (Theorem 6.5 page 101).
2. When a norm *can* induce an inner product, what is that (unique) inner product? The inner product expressed in terms of the norm is given by the *Polarization Identity* (Theorem 6.6 page 102).
3. Which norms can induce an inner product and which ones cannot? The answer is that norms that satisfy the *parallelogram law* (Theorem 6.7 page 103) **can** induce an inner product; and the ones that don't, cannot (Theorem 6.7 page 103).

Theorem 6.5. [13] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *be a* NORMED LINEAR SPACE *(Definition 5.1 page 83).*

T H M

$$\underbrace{\left.\begin{array}{l} \exists\ \langle \triangle \mid \nabla\rangle \text{ and } (\cdot \mid \cdot) \quad \text{such that} \\ \|x\|^2 = \langle x \mid x\rangle = (x \mid x) \quad \forall x \in X \end{array}\right\}}_{\text{If a norm is induced by two inner products...}} \implies \underbrace{\langle x \mid y\rangle = (x \mid y) \quad \forall x,y \in X}_{\text{...then those two inner products are equivalent.}}$$

✎PROOF:

$$\begin{aligned}
2\langle x \mid y\rangle &= \left[\langle x \mid y\rangle + \langle y \mid x\rangle\right] + \left[\langle x \mid y\rangle - \langle y \mid x\rangle\right] \\
&= \left[\langle x \mid y\rangle + \langle y \mid x\rangle\right] + i\left[-i\langle x \mid y\rangle + i\langle y \mid x\rangle\right] \\
&= \left[\langle x \mid y\rangle + \langle y \mid x\rangle\right] + i\left[-\langle ix \mid y\rangle - \langle y \mid ix\rangle\right] \\
&= \left(\underbrace{\left[\langle x \mid y\rangle + \langle y \mid x\rangle + \langle x \mid x\rangle + \langle y \mid y\rangle\right]}_{\langle x+y \mid x+y\rangle} - \underbrace{\left[\langle x \mid x\rangle + \langle y \mid y\rangle\right]}_{\text{residue}}\right) \\
&\quad - i\left(\underbrace{\left[\langle ix \mid y\rangle + \langle y \mid ix\rangle + \langle ix \mid ix\rangle + \langle y \mid y\rangle\right]}_{\langle ix+y \mid ix+y\rangle} - \underbrace{\left[\langle ix \mid ix\rangle + \langle y \mid y\rangle\right]}_{\text{residue}}\right) \\
&= \left(\left[\langle x+y \mid x+y\rangle\right] - \left[\langle x \mid x\rangle + \langle y \mid y\rangle\right]\right) - i\left(\left[\langle ix+y \mid ix+y\rangle\right] - \left[\langle ix \mid ix\rangle + \langle y \mid y\rangle\right]\right) \\
&= \left(\left[(x+y \mid x+y)\right] - \left[(x \mid x) + (y \mid y)\right]\right) - i\left(\left[(ix+y \mid ix+y)\right] - \left[(ix \mid ix) + (y \mid y)\right]\right) \\
&= \left(\underbrace{\left[(x \mid y) + (y \mid x) + (x \mid x) + (y \mid y)\right]}_{(x+y \mid x+y)} - \underbrace{\left[(x \mid x) + (y \mid y)\right]}_{\text{residue}}\right) \\
&\quad - i\left(\underbrace{\left[(ix \mid y) + (y \mid ix) + (ix \mid ix) + (y \mid y)\right]}_{(ix+y \mid ix+y)} - \underbrace{\left[(ix \mid ix) + (y \mid y)\right]}_{\text{residue}}\right)
\end{aligned}$$

[13] Aliprantis and Burkinshaw (1998), page 280, Bollobás (1999), page 132, Jordan and von Neumann (1935) page 721

$$
\begin{aligned}
&= \big[(x\mid y)+(y\mid x)\big]+i\big[-(ix\mid y)-(y\mid ix)\big]\\
&= \big[(x\mid y)+(y\mid x)\big]+i\big[-i(x\mid y)+i(y\mid x)\big]\\
&= \big[(x\mid y)+(y\mid x)\big]+\big[(x\mid y)-(y\mid x)\big]\\
&= 2(x\mid y)
\end{aligned}
$$

Theorem 6.6 (Polarization Identities). [14] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a linear space,* $\langle \triangle \mid \nabla \rangle \in \mathbb{F}^{X\times X}$ *a function, and* $\|x\| \triangleq \sqrt{\langle x \mid x\rangle}$.

THM

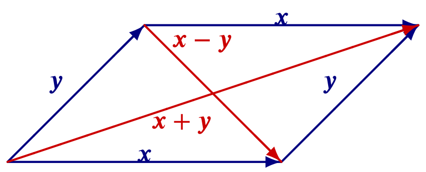
$\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla \rangle\big)$ *is an inner product space* $\implies$

$$
\underbrace{4\langle x\mid y\rangle = \begin{cases} \|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2 & \text{for } \mathbb{F}=\mathbb{C} \quad \forall x,y\in X\\ \|x+y\|^2 - \|x-y\|^2 & \text{for } \mathbb{F}=\mathbb{R} \quad \forall x,y\in X \end{cases}}_{\text{inner product induced by norm}}
$$

✎PROOF:

1. These follow directly from properties of *bilinear functionals* (Theorem 13.2 page 209).
2. Alternative proof for $\mathbb{F}=\mathbb{C}$ case:

$$
\begin{aligned}
&\|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2\\
&= \underbrace{\|x\|^2+\|y\|^2+2\Re\langle x\mid y\rangle}_{\langle x+y\mid x+y\rangle} - \underbrace{\left(\|x\|^2+\|-y\|^2+2\Re\langle x\mid -y\rangle\right)}_{\langle x-y\mid x-y\rangle}\\
&\quad + i\underbrace{\left(\|x\|^2+\|iy\|^2+2\Re\langle x\mid iy\rangle\right)}_{i\langle x+iy\mid x+iy\rangle} - i\underbrace{\left(\|x\|^2+\|-iy\|^2+2\Re\langle x\mid -iy\rangle\right)}_{i\langle x-iy\mid x-iy\rangle} && \text{by Lemma 6.1 page 99}\\
&= \underbrace{\|x\|^2+\|y\|^2+2\Re\langle x\mid y\rangle}_{\langle x+y\mid x+y\rangle} - \underbrace{\left(\|x\|^2+\|y\|^2-2\Re\langle x\mid y\rangle\right)}_{\langle x-y\mid x-y\rangle}\\
&\quad + i\underbrace{\left(\|x\|^2+\|y\|^2+2\Re\langle x\mid iy\rangle\right)}_{i\langle x+iy\mid x+iy\rangle} - i\underbrace{\left(\|x\|^2+\|y\|^2-2\Re\langle x\mid iy\rangle\right)}_{i\langle x-iy\mid x-iy\rangle} && \text{by Definition 6.1 page 95}\\
&= 4\Re\langle x\mid y\rangle + 4i\Re\langle x\mid iy\rangle\\
&= \underbrace{2\left(\langle x\mid y\rangle + \langle x\mid y\rangle^*\right)}_{4\Re\langle x\mid y\rangle} + \underbrace{2i\left(\langle x\mid iy\rangle + \langle x\mid iy\rangle^*\right)}_{4i\Re\langle x\mid iy\rangle}\\
&= 2\left(\langle x\mid y\rangle + \langle x\mid y\rangle^*\right) + 2i\left(i^*\langle x\mid y\rangle + (i^{**})\langle x\mid y\rangle^*\right)\\
&= 2\left(\langle x\mid y\rangle + \langle x\mid y\rangle^*\right) + 2i\left(-i\langle x\mid y\rangle + i\langle x\mid y\rangle^*\right) && \text{by Definition 6.1 page 95}\\
&= 2\langle x\mid y\rangle + 2\langle x\mid y\rangle^* + 2\langle x\mid y\rangle - 2\langle x\mid y\rangle^*\\
&= 4\langle x\mid y\rangle
\end{aligned}
$$



In plane geometry ($\mathbb{R}^2$), the *parallelogram law* states that the sum of the squares of the lengths of the sides of a parallelogram is equal to the sum of the squares of the lengths of its diagonals. This is illustrated in the figure to the left.

[14] Berberian (1961) pages 29–30 ⟨Theorem II.3.3⟩, Istrăţescu (1987) page 110 ⟨Proposition 4.1.5⟩, Bollobás (1999), page 132, Jordan and von Neumann (1935) page 721

Acutally, the parallelogram law can be generalized to *any inner product space* (not just in the plane). And if the parallelogram law happens to hold true in a normed linear space, then that normed linear space is actually an *inner product space*. The parallelogram law and its relation to inner product spaces is stated in the next theorem.

Theorem 6.7 (Parallelogram law). [15] *Let* $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *and* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$.

THM

$\boldsymbol{\Omega}$ *is an inner product space* $\iff$ $\underbrace{2\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{y}\|^2 = \|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2}_{\text{PARALLELOGRAM LAW / VON NEUMANN-JORDAN CONDITION}}$ $\quad \forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{\Omega}$

✎PROOF:

1. Proof that $[\exists \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$ such that $\|\boldsymbol{x}\|^2 = \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle] \implies$ [parallelogram law is true]:

$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 &= \left[\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\mathbf{R_e}\left[2\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle\right]\right] \quad + \left[\|\boldsymbol{x}\|^2 + \|-\boldsymbol{y}\|^2 + 2\mathbf{R_e}\left[2\langle \boldsymbol{x} \mid -\boldsymbol{y} \rangle\right]\right] \\
&\qquad \text{by Lemma 6.1 page 99} \\
&= \left[\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\mathbf{R_e}\left[2\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle\right]\right] \quad + \left[\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\mathbf{R_e}\left[2\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle\right]\right] \\
&= 2\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{y}\|^2
\end{aligned}$$

2. Proof that $[\exists \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$ such that $\|\boldsymbol{x}\|^2 = \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle] \impliedby$ [parallelogram law is true]:
   Note that **if** an inner product exists in the norm linear space $(\boldsymbol{\Omega}, \|\cdot\|)$, **then** that norm linear space is actually an inner product space $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$. And if it is an inner product space, then by Theorem 6.6 page 102 that inner product must be given by the **Polarization Identity**
   $\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = \|a\boldsymbol{x}+\boldsymbol{y}\|^2 - \|a\boldsymbol{x}-\boldsymbol{y}\|^2 + i\|a\boldsymbol{x}+i\boldsymbol{y}\|^2 - i\|a\boldsymbol{x}-i\boldsymbol{y}\|^2$.
   Therefore, here we must use the parallelogram law to show that the bilinear function $\mathsf{f}(\boldsymbol{x},\boldsymbol{y}) \triangleq \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$ given on the left hand side of the "=" relation is indeed an inner product— that is, that it satisfies the requirements of Definition 6.1 page 95.

   (a) Proof that $\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \geq 0$ (non-negative):

$$\begin{aligned}
4\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle &\triangleq \|\boldsymbol{x}+\boldsymbol{x}\|^2 - \cancelto{0}{\|\boldsymbol{x}-\boldsymbol{x}\|^2} + i\|\boldsymbol{x}+i\boldsymbol{x}\|^2 - i\|\boldsymbol{x}-i\boldsymbol{x}\|^2 && \text{by } \textit{Polarization Identity} \\
&= \|2\boldsymbol{x}\|^2 - \mathbb{0} + i\left(\|\boldsymbol{x}+i\boldsymbol{x}\|^2 - \|\boldsymbol{x}-i\boldsymbol{x}\|^2\right) && \text{by Definition 5.1 page 83} \\
&= |2|^2\|\boldsymbol{x}\|^2 + i\left(\|\boldsymbol{x}+i\boldsymbol{x}\|^2 - |i|\,\|\boldsymbol{x}-i\boldsymbol{x}\|^2\right) \\
&= 4\|\boldsymbol{x}\|^2 + i\left(\cancelto{0}{\|\boldsymbol{x}+i\boldsymbol{x}\|^2 - \|i\boldsymbol{x}+\boldsymbol{x}\|^2}\right) && \text{by Definition 5.1 page 83} \\
&= 4\|\boldsymbol{x}\|^2 \\
&\geq 0 && \text{by Definition 5.1 page 83}
\end{aligned}$$

   (b) Proof that $\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle = 0 \iff \boldsymbol{x} = \mathbb{0}$ (non-isotropic):

$$\begin{aligned}
4\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle &= 4\|\boldsymbol{x}\|^2 && \text{by result of part (a)} \\
&= 0 \quad \iff \quad \boldsymbol{x} = \mathbb{0} && \text{by Definition 5.1 page 83}
\end{aligned}$$

[15] Amir (1986), page 8, Istrăţescu (1987) page 110, Day (1973), page 151, Halmos (1998a), page 14, Aliprantis and Burkinshaw (1998) pages 280–281 ⟨Theorem 32.6⟩, Riesz (1934) page 36?, Jordan and von Neumann (1935) pages 721–722

(c) Proof that $\langle x + u \mid y\rangle = \langle x \mid y\rangle + \langle u \mid y\rangle$ (additive):[16]

$$\begin{aligned}
4\langle x+y \mid z\rangle &= 8\left\langle \frac{x+y}{2} \mid z\right\rangle && \text{by Definition 6.1 page 95}\\
&= 2\left\|\frac{x+y}{2}+z\right\|^2 - 2\left\|\frac{x+y}{2}-z\right\|^2\\
&\quad + 2i\left\|\frac{x+y}{2}+z\right\|^2 - 2i\left\|\frac{x+y}{2}-iz\right\|^2 && \text{by \textit{Polarization Identity}}\\
&= \left(2\left\|\frac{x+y}{2}+z\right\|^2 + 2\left\|\frac{x-y}{2}\right\|^2\right)\\
&\quad - \left(2\left\|\frac{x+y}{2}-z\right\|^2 + 2\left\|\frac{x-y}{2}\right\|^2\right)\\
&\quad + \left(2i\left\|\frac{x+y}{2}+z\right\|^2 + 2i\left\|\frac{x-y}{2}\right\|^2\right)\\
&\quad - \left(2i\left\|\frac{x+y}{2}-iz\right\|^2 + 2i\left\|\frac{x-y}{2}\right\|^2\right)\\
&= \left(\|x+z\|^2 + \|y+z\|^2\right) - \left(\|x-z\|^2 + \|y-z\|^2\right)\\
&\quad + \left(i\|x+z\|^2 + i\|y+z\|^2\right) - \left(i\|x-iz\|^2 + i\|y-iz\|^2\right) && \text{by parallelogram law}\\
&= \left(\|x+z\|^2 - \|x-z\|^2 + i\|x+z\|^2 - i\|x-iz\|^2\right)\\
&\quad + \left(\|y+z\|^2 - \|y-z\|^2 + i\|y+z\|^2 - i\|y-iz\|^2\right)\\
&= 4\langle x \mid z\rangle + 4\langle y \mid z\rangle && \text{by Polarization Identity}
\end{aligned}$$

(d) Proof that $\langle x \mid y\rangle = \langle y \mid y\rangle^*$ (*conjugate symmetric*):

$$\begin{aligned}
4\langle x \mid y\rangle &\triangleq \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2 && \text{by Polarization Identity}\\
&= \|y+x\|^2 - \|y-x\|^2 + i\|i(y-ix)\|^2 - i\|-i(y+ix)\|^2 && \text{by Definition 3.1 page 67}\\
&= \|y+x\|^2 - \|y-x\|^2 + i\|y-ix\|^2 - i\|y+ix\|^2 && \text{by Definition 5.1 page 83}\\
&= \left(\|y+x\|^2 - \|y-x\|^2 - i\|y-ix\|^2 + i\|y+ix\|^2\right)^*\\
&= \left(\|y+x\|^2 - \|y-x\|^2 + i\|y+ix\|^2 - i\|y-ix\|^2\right)^*\\
&\triangleq 4\langle y \mid x\rangle^* && \text{by \textit{Polarization Identity}}
\end{aligned}$$

(e) Proof that $\langle \alpha x \mid y\rangle = \alpha\langle x \mid y\rangle$ (*homogeneous*): [17]

i. Proof that $\langle \alpha x \mid y\rangle$ is linear in α:

$$\begin{aligned}
0 &\le \left|\|\alpha x + y\| - \|\beta x + y\|\right| && \text{by Definition F.4 page 346}\\
&\le \left|\|(\alpha x + y) - (\beta x + y)\|\right| && \text{by Theorem 5.2 page 84}\\
&\le \left|\|(\alpha - \beta)x\|\right|
\end{aligned}$$

This implies that as $a \to \beta$, $\|\alpha x + y\| \to \|\beta x + y\|$, which by definition implies that $\|\alpha x + y\|$ linear in α. And by the parallelogram law, $\langle \alpha x \mid y\rangle$ is also linear in α.

ii. Proof that $\langle nx \mid y\rangle = n\langle x \mid y\rangle$ for $n \in \mathbb{Z}$ (integer case):

A. Proof for $n = \pm 1$:

$$\begin{aligned}
\langle nx \mid y\rangle &= \langle \pm 1 x \mid y\rangle && \text{by } n = \pm 1 \text{ hypothesis}\\
&= \pm 1\langle x \mid y\rangle && \text{by definition of \textit{inner product}} && \text{(Definition 6.1 page 95)}\\
&= n\langle x \mid y\rangle && \text{by } n = \pm 1 \text{ hypothesis}
\end{aligned}$$

16 Aliprantis and Burkinshaw (1998), page 281
17 Aliprantis and Burkinshaw (1998), page 138

B. Proof for $n = 0$:

$$\begin{aligned}
\langle n\boldsymbol{x} \mid \boldsymbol{y}\rangle &= \langle 0\boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by } n = 0 \text{ hypothesis}\\
&= \langle \boldsymbol{x} - \boldsymbol{x} \mid \boldsymbol{y}\rangle\\
&= \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle + \langle -1\boldsymbol{x} \mid \boldsymbol{y}\rangle\\
&= \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle - 1\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle\\
&= \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle - \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle\\
&= 0\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle\\
&= n\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by } n = 0 \text{ hypothesis}
\end{aligned}$$

C. Proof for $n = \pm 2$:

$$\begin{aligned}
\langle n\boldsymbol{x} \mid \boldsymbol{y}\rangle &= \langle \pm 2\boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by } n = \pm 1 \text{ hypothesis}\\
&= \langle \pm(\boldsymbol{x} + \boldsymbol{x}) \mid \boldsymbol{y}\rangle\\
&= \pm\,\langle \boldsymbol{x} + \boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by definition of \textit{inner product}} \quad \text{(Definition 6.1 page 95)}\\
&= \pm\,(\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle + \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle) && \text{by additive property}\\
&= \pm 2\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle\\
&= n\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by } n = \pm 2 \text{ hypothesis}
\end{aligned}$$

D. Proof that $[n \text{ case}] \implies [n \pm 1 \text{ case}]$:

$$\begin{aligned}
\langle (n \pm 1)\boldsymbol{x} \mid \boldsymbol{y}\rangle &= \langle n\boldsymbol{x} \pm 1\boldsymbol{x} \mid \boldsymbol{y}\rangle\\
&= \langle n\boldsymbol{x} \mid \boldsymbol{y}\rangle + \langle \pm 1\boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by additive property}\\
&= n\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle \pm 1\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by left hypothesis}\\
&= (n \pm 1)\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle
\end{aligned}$$

iii. Proof that $\langle q\boldsymbol{x} \mid \boldsymbol{y}\rangle = q\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle$ for $q \in \mathbb{Q}$ (rational number case):

$$\begin{aligned}
\frac{n}{m}\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle &= \frac{n}{m}\left\langle \frac{m}{m}\boldsymbol{x} \,\middle|\, \boldsymbol{y}\right\rangle && \text{where } n, m \in \mathbb{Z} \text{ and } m \neq 0\\
&= \frac{nm}{m}\left\langle \frac{1}{m}\boldsymbol{x} \,\middle|\, \boldsymbol{y}\right\rangle && \text{by previous result}\\
&= \frac{m}{m}\left\langle \frac{n}{m}\boldsymbol{x} \,\middle|\, \boldsymbol{y}\right\rangle && \text{by previous result}\\
&= \left\langle \frac{n}{m}\boldsymbol{x} \,\middle|\, \boldsymbol{y}\right\rangle
\end{aligned}$$

iv. Proof that $\langle r\boldsymbol{x} \mid \boldsymbol{y}\rangle = r\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle$ for all $r \in \mathbb{R}$ (real number case):
Because $\mathbb{Q}$ is dense in $\mathbb{R}$ and because $\|\alpha\boldsymbol{x} + \boldsymbol{y}\|$ is continuous in α, so $\langle \alpha\boldsymbol{x} \mid \boldsymbol{y}\rangle = \alpha\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle$ for all $\alpha \in \mathbb{R}$.

v. Proof that $\langle c\boldsymbol{x} \mid \boldsymbol{y}\rangle = c\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle$ for all $c \in \mathbb{C}$ (complex number case):
No proof at this time.

✏

Remark 6.1. [18] The inner product has already been defined in Definition 6.1 (page 95) as a bilinear function that is *non-negative*, *non-isotropic*, *homogeneous*, *additive*, and *conjugate symmetric*. However, given a normed linear space, we could alternatively define the inner product using the *parallelogram law* (Theorem 6.7 page 103) together with the *Polarization Identity* (Theorem 6.6 page 102). Under this new definition, an inner product *exists* if the parallelogram law is satisfied, and is *specified*, in terms of the norm, by the Polarization Identity.

[18] Loomis (1953), pages 23–24, Kubrusly (2001) page 317

Of the uncountably infinite number of $\ell^p_{\mathbb{F}}$ norms, only the norm for $p = 2$ induces an inner product (Proposition 6.1, next).

Proposition 6.1. [19] *Let* $\|(\!(x_n)\!)_{n\in\mathbb{Z}}\|_p$ *be the* $\ell^p_{\mathbb{F}}$ *norm* (Definition 8.10 page 137) *of the sequence* $(\!(x_n)\!)$ *in the space* $\ell^p_{\mathbb{F}}$.

PRP $\|(\!(x_n)\!)\|_p$ *induces an inner product* $\iff$ $p = 2$

✎PROOF:

1. Proof that $\|\cdot\|_p$ induces an inner product $\Longleftarrow$ $p = 2$ (using the *Parallelogram law* page 103):

$$\begin{aligned}
&\|\boldsymbol{x}+\boldsymbol{y}\|_p^2 + \|\boldsymbol{x}-\boldsymbol{y}\|_p^2 \\
&= \|\boldsymbol{x}+\boldsymbol{y}\|_2^2 + \|\boldsymbol{x}-\boldsymbol{y}\|_2^2 && \text{by right hypothesis}\\
&= \left(\sum_{n\in\mathbb{Z}} |x_n+y_n|^2\right)^{\frac{2}{2}} + \left(\sum_{n\in\mathbb{Z}} |x_n-y_n|^2\right)^{\frac{2}{2}} && \text{by definition of } \|\cdot\|_p\\
&= \sum_{n\in\mathbb{Z}} (x_n+y_n)(x_n+y_n)^* + \sum_{n\in\mathbb{Z}} (x_n-y_n)(x_n-y_n)^*\\
&= \sum_{n\in\mathbb{Z}} \left(|x_n|^2 + |y_n|^2 + 2\Re(x_n y_n)\right) + \sum_{n\in\mathbb{Z}} \left(|x_n|^2 + |y_n|^2 - 2\Re(x_n y_n)\right)\\
&= 2\sum_{n\in\mathbb{Z}} |x_n|^2 + 2\sum_{n\in\mathbb{Z}} |y_n|^2\\
&= 2\,\|\boldsymbol{x}\|_2^2 + 2\,\|\boldsymbol{y}\|_2^2 && \text{by definition of } \|\cdot\|_p\\
&= 2\,\|\boldsymbol{x}\|_p^2 + 2\,\|\boldsymbol{y}\|_p^2 && \text{by right hypothesis}\\
&\implies \|\cdot\|_2 \text{ induces an inner product} && \text{by Theorem 6.7 page 103}
\end{aligned}$$

2. Proof that $\|\cdot\|_p$ induces an inner product $\implies$ $p = 2$:

   (a) Let $\boldsymbol{x} \triangleq (1, 0)$ and $\boldsymbol{y} \triangleq (0, 1)$. Then [20]

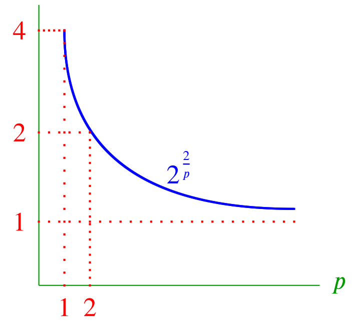
$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|_p^2 + \|\boldsymbol{x}-\boldsymbol{y}\|_p^2 &= \left(\sum_{n\in\mathbb{Z}} |x_n+y_n|^p\right)^{\frac{2}{p}} + \left(\sum_{n\in\mathbb{Z}} |x_n-y_n|^p\right)^{\frac{2}{p}} && \text{by definition of } \|\cdot\|_p\\
&= \left(|1+0|^p + |0+1|^p\right)^{\frac{2}{p}} + \left(|1-0|^p + |0-1|^p\right)^{\frac{2}{p}} && \text{by definitions of } \boldsymbol{x} \text{ and } \boldsymbol{y}\\
&= 2^{\frac{2}{p}} + 2^{\frac{2}{p}}\\
&= 2\cdot 2^{\frac{2}{p}}\\
2\,\|\boldsymbol{x}\|_p^2 + 2\,\|\boldsymbol{y}\|_p^2 &= 2\left(\sum_{n\in\mathbb{Z}} |x_n|^p\right)^{\frac{2}{p}} + 2\left(\sum_{n\in\mathbb{Z}} |y_n|^p\right)^{\frac{2}{p}} && \text{by definition of } \|\cdot\|_p\\
&= 2\left(|1|^p + |0|^p\right)^{\frac{2}{p}} + 2\left(|1|^p + |0|^p\right)^{\frac{2}{p}} && \text{by definitions of } \boldsymbol{x} \text{ and } \boldsymbol{y}\\
&= 2+2\\
&= 4\\
2\cdot 2^{\frac{2}{p}} = 4 = &\iff 2^{\frac{2}{p}} = 2\\
= &\implies p = 2
\end{aligned}$$

[19] Kubrusly (2001) pages 318–319 ⟨Example 5B⟩

[20] http://groups.google.com/group/sci.math/msg/531b1173f08871e9

(b) Proof that $2^{2/p}$ is monotonic decreasing in p (and so $p = 2$ is the only solution):

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}p} 2^{\frac{2}{p}} &= \frac{\mathrm{d}}{\mathrm{d}p} e^{\ln 2^{\frac{2}{p}}} \\
&= \left(e^{\ln 2^{\frac{2}{p}}}\right) \frac{\mathrm{d}}{\mathrm{d}p} \ln 2^{\frac{2}{p}} \\
&= \left(2^{\frac{2}{p}}\right) \frac{\mathrm{d}}{\mathrm{d}p} (2 \ln 2) \frac{1}{p} \\
&= \left(2^{\frac{2}{p}}\right) 2 \ln 2 \left(-\frac{1}{p^2}\right) \\
&< 0 \qquad \forall p \in (0, \infty)
\end{aligned}$$



⇨

6.3 Orthogonality

Definition 6.3.

DEF *The **Kronecker delta function** $\bar{\delta}_n$ is defined as* $\bar{\delta}_n \triangleq \begin{cases} 1 & \text{for } n = 0 \\ 0 & \text{for } n \neq 0. \end{cases}$ *and* $\forall n \in \mathbb{Z}$

Definition 6.4. *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla \rangle\right)$ *be an* INNER PRODUCT SPACE (Definition 6.1 page 95).

DEF
*Two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in X are **orthogonal** if*

$$\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = \begin{cases} 0 & \text{for } \boldsymbol{x} \neq \boldsymbol{y} \\ c \in \mathbb{F} \setminus 0 & \text{for } \boldsymbol{x} = \boldsymbol{y} \end{cases}$$

The notation $\boldsymbol{x} \perp \boldsymbol{y}$ implies $\boldsymbol{x}$ and $\boldsymbol{y}$ are ORTHOGONAL.
*A set $Y \in 2^X$ is **orthogonal** if $\boldsymbol{x} \perp \boldsymbol{y} \quad \forall \boldsymbol{x}, \boldsymbol{y} \in Y$.*
*A set Y is **orthonomal** if it is* ORTHOGONAL *and $\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle = 1 \quad \forall \boldsymbol{y} \in Y$.*
*A sequence $(\boldsymbol{x}_n \in X)_{n \in \mathbb{Z}}$ is **orthogonal** if $\langle \boldsymbol{x}_n \mid \boldsymbol{x}_m \rangle = c\bar{\delta}_{nm}$ for some $c \in \mathbb{R} \setminus 0$.*
*A sequence $(\boldsymbol{x}_n \in X)_{n \in \mathbb{Z}}$ is **orthonormal** if $\langle \boldsymbol{x}_n \mid \boldsymbol{x}_m \rangle = \bar{\delta}_{nm}$.*

The definition of the orthogonality relation $\perp$ has several immediate consequences (next theorem):

Theorem 6.8. [21] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla \rangle\right)$ *be an* INNER PRODUCT SPACE.

THM

1. $\boldsymbol{x} \perp \boldsymbol{x} \iff \boldsymbol{x} = \mathbb{0}$ $\quad \forall \boldsymbol{x} \in X$
2. $\boldsymbol{x} \perp \boldsymbol{y} \implies \alpha \boldsymbol{x} \perp \boldsymbol{y}$ $\quad \forall \alpha \in \mathbb{R}, \boldsymbol{x}, \boldsymbol{y} \in X$ (homogeneous)
3. $\boldsymbol{x} \perp \boldsymbol{y} \iff \boldsymbol{y} \perp \boldsymbol{x}$ $\quad \forall \boldsymbol{x}, \boldsymbol{y} \in X$ (symmetry)
4. $\boldsymbol{x} \perp \boldsymbol{y}$ and $\boldsymbol{y} \perp \boldsymbol{z} \implies \boldsymbol{x} \perp (\boldsymbol{y} + \boldsymbol{z})$ $\quad \forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$ (additive)
5. $\exists \beta \in \mathbb{R}$ such that $\boldsymbol{x} \perp (\beta \boldsymbol{x} + \boldsymbol{y})$ $\quad \forall \boldsymbol{x} \in X \setminus \mathbb{0}, \boldsymbol{y} \in X$

Theorem 6.9. *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla \rangle\right)$ *be an* INNER PRODUCT SPACE.

THM
$$\left.\begin{array}{lrcl} 1. & \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle &=& 0 \quad \text{and} \\ 2. & \boldsymbol{x} + \boldsymbol{y} &=& \mathbb{0} \end{array}\right\} \iff \left\{\begin{array}{lrcl} 1. & \boldsymbol{x} &=& \mathbb{0} \quad \text{and} \\ 2. & \boldsymbol{y} &=& \mathbb{0} \end{array}\right. \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in X$$

✎PROOF:

[21] James (1945), page 292, Drljević (1989) page 232

1. Proof that $x = y = \mathbb{0}$:

$$\begin{aligned}
0 &= \langle \mathbb{0} \,|\, \mathbb{0} \rangle && \text{by *non-isotropic* property of } \langle \triangle \,|\, \nabla \rangle \text{ (Definition 6.1 page 95)} \\
&= \langle x+y \,|\, x+y \rangle && \text{by left hypothesis 2} \\
&= \langle x \,|\, x+y \rangle + \langle y \,|\, x+y \rangle && \text{by *additive* property of } \langle \triangle \,|\, \nabla \rangle \text{ (Definition 6.1 page 95)} \\
&= \langle x \,|\, x \rangle + \langle x \,|\, y \rangle + \langle x \,|\, y \rangle^* + \langle y \,|\, y \rangle && \text{by *conjugate symmetric* and *additive* properties of } \langle \triangle \,|\, \nabla \rangle \\
&= \underbrace{\langle x \,|\, x \rangle}_{\geq 0} + 0 + 0 + \underbrace{\langle y \,|\, y \rangle}_{\geq 0} && \text{by left hypothesis 1} \\
&\implies x = \mathbb{0} \text{ and } y = \mathbb{0} && \text{by *non-negative* and *non-isotropic* properties of } \langle \triangle \,|\, \nabla \rangle
\end{aligned}$$

2. Proof that $\langle x \,|\, y \rangle = 0$:

$$\begin{aligned}
\langle x \,|\, y \rangle &= \langle \mathbb{0} \,|\, \mathbb{0} \rangle && \text{by right hypotheses} \\
&= 0 && \text{by *non-isotropic* property of } \langle \triangle \,|\, \nabla \rangle \text{ (Definition 6.1 page 95)}
\end{aligned}$$

3. Proof that $x + y = \mathbb{0}$:

$$\begin{aligned}
x + y &= \mathbb{0} + \mathbb{0} && \text{by right hypotheses} \\
&= \mathbb{0}
\end{aligned}$$

✏

The *triangle inequality* theorem for vectors in a *normed linear space* (Theorem 5.1 page 83) demonstrates that

$$\left\| \sum_{n=1}^{N} x_n \right\| \leq \sum_{n=1}^{N} \|x_n\|.$$

The *Pythagorean Theorem* (next) demonstrates that this *inequality* becomes *equality* when the set $\{x_n\}$ is orthogonal.

Theorem 6.10 (Pythagorean Theorem). [22] *Let $\{x_n \in X \,|\, n=1,2,\ldots,N\}$ be a set of vectors in an* INNER PRODUCT SPACE *(Definition 6.1 page 95) $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \nabla \rangle)$ and let $\|x\| \triangleq \sqrt{\langle x \,|\, x \rangle}$ (Definition 6.2 page 100).*

THM $\{x_n\}$ *is* ORTHOGONAL $\iff$ $\displaystyle\left\| \sum_{n=1}^{N} x_n \right\|^2 = \sum_{n=1}^{N} \|x_n\|^2 \quad \forall N \in \mathbb{N}$

✎PROOF: 1. Proof for ($\implies$) case:

$$\begin{aligned}
\left\| \sum_{n=1}^{N} x_n \right\|^2 &= \left\langle \sum_{n=1}^{N} x_n \,\middle|\, \sum_{m=1}^{N} x_m \right\rangle && \text{by def. of } \|\cdot\| && \text{(Definition 5.1 page 83)} \\
&= \sum_{n=1}^{N} \sum_{m=1}^{N} \langle x_n \,|\, x_m \rangle && \text{by def. of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)} \\
&= \sum_{n=1}^{N} \sum_{m=1}^{N} \langle x_n \,|\, x_m \rangle \bar{\delta}_{n-m} && \text{by left hypothesis} \\
&= \sum_{n=1}^{N} \langle x_n \,|\, x_n \rangle && \text{by def. of } \bar{\delta} && \text{(Definition 6.3 page 107)} \\
&= \sum_{n=1}^{N} \|x_n\|^2 && \text{by def. of } \|\cdot\| && \text{(Definition 5.1 page 83)}
\end{aligned}$$

[22] Aliprantis and Burkinshaw (1998) pages 282–283 ⟨Theorem 32.7⟩, Kubrusly (2001) page 324 ⟨Proposition 5.8⟩, Bollobás (1999) pages 132–133 ⟨Theorem 3⟩

2. Proof for ($\Longleftarrow$) case:

$$\begin{aligned}
&4\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle \\
&= \|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + i\,\|\boldsymbol{x}+i\boldsymbol{y}\|^2 - i\,\|\boldsymbol{x}-i\boldsymbol{y}\|^2 && \text{by \textit{polarization identity} (Theorem 6.6 page 102)} \\
&= \left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) - \left(\|\boldsymbol{x}\|^2 + \|-\boldsymbol{y}\|^2\right) + i\left(\|\boldsymbol{x}\|^2 + \|i\boldsymbol{y}\|^2\right) - i\left(\|\boldsymbol{x}\|^2 + \|-i\boldsymbol{y}\|^2\right) && \text{by right hypothesis} \\
&= \left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) - \left(\|\boldsymbol{x}\|^2 + |-1|^2\,\|\boldsymbol{y}\|^2\right) + i\left(\|\boldsymbol{x}\|^2 + |i|^2\,\|\boldsymbol{y}\|^2\right) - i\left(\|\boldsymbol{x}\|^2 + |-i|^2\,\|\boldsymbol{y}\|^2\right) && \text{by definition of } \|\cdot\| \\
&= \left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) - \left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) + i\left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) - i\left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) && \text{by def. of } |\cdot| \text{ (Definition F.4 page 346)} \\
&= 0
\end{aligned}$$

⇨

CHAPTER 7

LINEAR SUBSPACES

7.1 Subspaces of a linear space

Linear spaces (Definition 3.1 page 67) can be decomposed into a collection of *linear subspaces* (Definition 7.1 page 112). Often such a collection along with an *order relation* (Definition B.2 page 290) forms a *lattice* (Definition C.3 page 305).

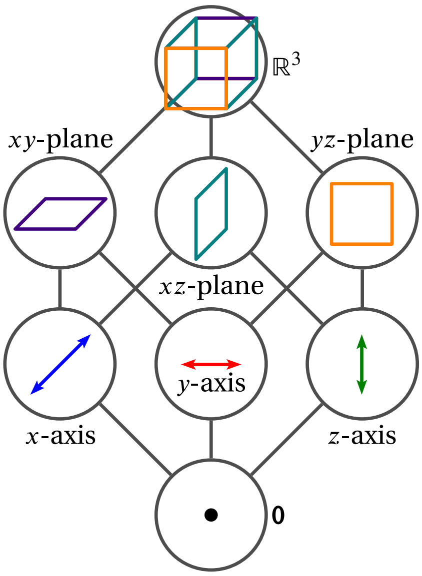
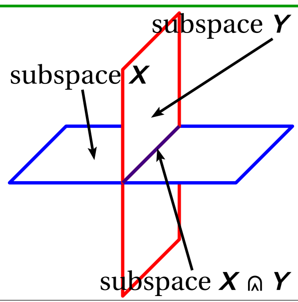


Figure 7.1: lattice of subspaces of $\mathbb{R}^3$ (Example 7.1 page 111)

EX

Example 7.1. The 3-dimensional Euclidean space $\mathbb{R}^3$ contains the 2-dimensional xy-plane and xz-plane subspaces, which in turn both contain the 1-dimensional x-axis subspace. These subspaces are illustrated in the figure to the right and in Figure 7.1 (page 111).



Definition 7.1. [1] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE (Definition 3.1 page 67).

DEF

A tupple $(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *is a* ***linear subspace** of* $\boldsymbol{\Omega}$ *if*

1. $Y \neq \varnothing$ — (*Y must contain at least one element*) *and*
2. $Y \subseteq X$ — (*Y is a subset of X*) *and*
3. $\boldsymbol{x}, \boldsymbol{y} \in Y \implies \boldsymbol{x} + \boldsymbol{y} \in Y$ — (*closed under vector addition*) *and*
4. $\boldsymbol{x} \in Y$ *and* $\alpha \in \mathbb{F} \implies \alpha\boldsymbol{x} \in Y$ — (*closed under scalar-vector multiplication*).

A linear subspace is also called a ***linear manifold***.

Every *linear space* (Definition 3.1 page 67) $\boldsymbol{X}$ has at least two *linear subspaces*—itself and $\mathbf{0}$ (Proposition 7.1 page 112), called the *trivial linear space*. The *linear span* (Definition 10.2 page 151) of every subset of a linear linear space is a subspace (Proposition 7.2 page 113). Every *linear subspace* contains the "zero" vector $\mathbb{0}$, and is *convex* (Definition 9.6 page 142, Proposition 7.3 page 113).

Proposition 7.1. [2] *Let* $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *and* $\mathbf{0} \triangleq (\{\mathbb{0}\}, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$.

PRP

$$\left\{ \boldsymbol{X} \text{ is a LINEAR SPACE (Definition 3.1 page 67)} \right\} \implies \left\{ \begin{array}{lll} 1. & \mathbf{0} \text{ is a LINEAR SUBSPACE of } \boldsymbol{X} & \text{and} \\ 2. & \boldsymbol{X} \text{ is a LINEAR SUBSPACE of } \boldsymbol{X} & \end{array} \right\}$$

✎PROOF: For a structure to be a linear subspace of $\boldsymbol{X}$, it must satisfy the requirements of Definition 7.1 (page 112).

1. Proof that $\{\mathbb{0}\}$ is a linear subspace:

   (a) Note that $\{\mathbb{0}\} \neq \varnothing$.

   (b) Proof that $\boldsymbol{x}, \boldsymbol{y} \in \{\mathbb{0}\} \implies \boldsymbol{x} + \boldsymbol{y} \in \{\mathbb{0}\}$:

   $$\begin{aligned} \boldsymbol{x} + \boldsymbol{y} &= \mathbb{0} + \mathbb{0} && \text{by } \boldsymbol{x}, \boldsymbol{y} \in \{\mathbb{0}\} \text{ hypothesis} \\ &= \mathbb{0} \\ &\in \{\mathbb{0}\} \end{aligned}$$

   (c) Proof that $\boldsymbol{x} \in \{\mathbb{0}\}, \alpha \in \mathbb{F} \implies \alpha\boldsymbol{x} \in \{\mathbb{0}\}$:

   $$\begin{aligned} \alpha\boldsymbol{x} &= \alpha\mathbb{0} && \text{by } \boldsymbol{x} \in \{\mathbb{0}\} \text{ hypothesis} \\ &= \mathbb{0} && \text{by definition of } \mathbb{0} \\ &\in \{\mathbb{0}\} \end{aligned}$$

2. Proof that $\boldsymbol{\Omega}$ is a linear subspace of itself:

   (a) Proof that $X \neq \varnothing$:

   $$X \neq \varnothing$$

   (b) Proof that $\boldsymbol{x}, \boldsymbol{y} \in X \implies \boldsymbol{x} + \boldsymbol{y} \in X$:

   $$\boldsymbol{x} + \boldsymbol{y} \in \{\mathbb{0}\} \qquad \text{because } + : X \times X \to X \text{ (} X \text{ is closed under vector addition)}$$

   (c) Proof that $\boldsymbol{x} \in X, \alpha \in \mathbb{F} \implies \alpha\boldsymbol{x} \in X$:

   $$\alpha\boldsymbol{x} \in X \qquad \text{because } \cdot : \mathbb{F} \times X \to X \text{ (} X \text{ is closed under scalar-vector multiplication)}$$

[1] Michel and Herget (1993) page 81 ⟨Definition 3.2.1⟩, Berberian (1961) page 13 ⟨Definition I.5.1⟩, Halmos (1958), page 16

[2] Michel and Herget (1993) pages 81–83, Haaser and Sullivan (1991) page 43

Proposition 7.2. [3] *Let* $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE (Definition 3.1 page 67). *Let* span *be the* LINEAR SPAN *of a set* Y *in* $\boldsymbol{X}$.

PRP: $\left\{ \begin{array}{l} Y \text{ is a SUBSET of the set } X \\ (Y \subseteq X) \end{array} \right\} \implies \left\{ \operatorname{span} Y \text{ is a LINEAR SUBSPACE of } \boldsymbol{X}. \right\}$

Proposition 7.3. [4] *Let* $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE *and* $\mathbb{0}$ *the zero vector of* $\boldsymbol{X}$.

PRP: $\left\{ \boldsymbol{Y} \text{ is a LINEAR SUBSPACE of } \boldsymbol{X} \right\} \implies \left\{ \begin{array}{lll} 1. & \mathbb{0} \in \boldsymbol{Y} & \text{and} \\ 2. & \boldsymbol{Y} \text{ is CONVEX in } \boldsymbol{X} & \end{array} \right\}$

✎PROOF:

$$\begin{aligned} \boldsymbol{Y} \text{ is a } \textit{subspace} &\implies \exists(\alpha \boldsymbol{y}) \in \boldsymbol{Y} \quad \forall \alpha \in \mathbb{F} && \text{by Definition 7.1 page 112} \\ &\implies \exists \mathbb{0} \in \boldsymbol{Y} && \text{because } \alpha = 0 \in \mathbb{F} \end{aligned}$$

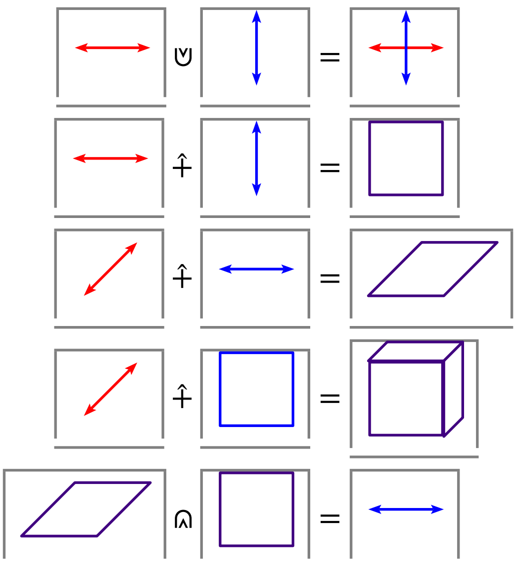
$$\begin{aligned} \boldsymbol{Y} \text{ is a linear subspace} &\implies \boldsymbol{x} + \boldsymbol{y} \in Y \; \forall \boldsymbol{x}, \boldsymbol{y} \in Y \\ &\implies \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in Y \; \forall \boldsymbol{x}, \boldsymbol{y} \in Y \\ &\implies \boldsymbol{Y} \text{ is } \textit{convex} \end{aligned}$$

Definition 7.2. [5] *Let* $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *and* $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be* LINEAR SUBSPACES (Definition 7.1 page 112) *of a* LINEAR SPACE (Definition 3.1 page 67) $\boldsymbol{\Omega} \triangleq (\Omega, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$.

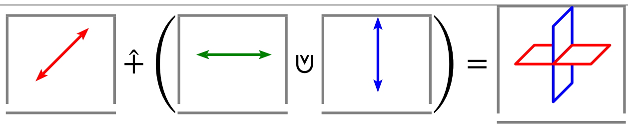
DEF:

$$\begin{array}{rcll} \boldsymbol{X} \hat{+} \boldsymbol{Y} & \triangleq & \left(\left\{\boldsymbol{x}+\boldsymbol{y} | \boldsymbol{x} \in X \text{ and } \boldsymbol{y} \in Y\right\}, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right) & \text{(Minkowski addition)} \\ \boldsymbol{X} \uplus \boldsymbol{Y} & \triangleq & \left(X \cup Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right) & \text{(subspace union)} \\ \boldsymbol{X} \Cap \boldsymbol{Y} & \triangleq & \left(X \cap Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right) & \text{(subspace intersection)} \end{array}$$

Example 7.2. Some examples of operations on subspaces in $\mathbb{R}^3$ are illustrated next:



[3] Michel and Herget (1993) page 86
[4] Michel and Herget (1993) page 81
[5] Wedderburn (1907) page 79

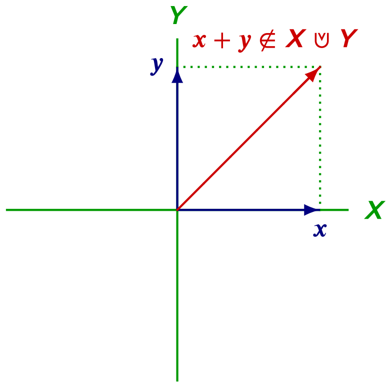


Remark 7.1.

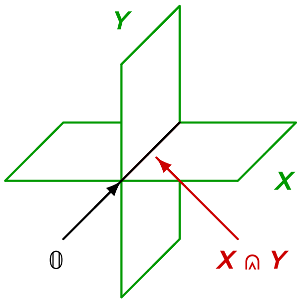
Notice the similarities between the properties of linear subspaces in a linear space (Proposition 7.4 page 114) and the properties of closed sets in a topological space (Theorem 1.3 page 6):

linear subspaces	*closed sets*
$\mathbb{0}$	$\varnothing$
$\boldsymbol{\Omega}$	$\boldsymbol{\Omega}$
$\boldsymbol{X} \hat{+} \boldsymbol{Y}$	$X \cup Y$
$\bigcap_{n=1}^{N} \boldsymbol{X}_n$	$\bigcap_{\gamma\in\Gamma} X_\gamma$

One key difference is that the union of two linear subspaces is not in general a linear subspace. For example, if $\boldsymbol{x}$ is the vector $[1\ 0]$ in the x direction linear subspace of $\mathbb{R}^2$ and $\boldsymbol{y}$ is the vector $[0\ 1]$ in the y direction linear subspace, then $\boldsymbol{x}+\boldsymbol{y}$ is not in the union of the two linear subspaces (it is not on the x axis or y axis but rather at $(1,1)$).[6]



In general, the set of all linear subspaces of a linear space $\boldsymbol{\Omega}$ is *not* closed under the subspace union ($\uplus$) operation; that is, the union of of two linear subspaces is *not* necessarily a linear subspace. However the set *is* closed under Minkowski sum ($\hat{+}$) and subspace intersection ($\Cap$). Proposition 7.4 (next) shows four useful objects are always subspaces. Some of these in Euclidean space $\mathbb{R}^3$ are illustrated to the right.



Proposition 7.4. [7] *Let* $\boldsymbol{X}$ *be a* LINEAR SPACE *(Definition 3.1 page 67).*

PRP
$$\left\{ \begin{array}{l}\{\boldsymbol{X}_n|_{n=1,2,\dots,N}\} \text{ are}\\ \text{LINEAR SUBSPACES of } \boldsymbol{X}\end{array}\right\} \implies \left\{ \begin{array}{lll} 1. & \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \hat{+} \cdots \hat{+} \boldsymbol{X}_N & \text{is a LINEAR SUBSPACE of } \boldsymbol{X}\\ & \text{and} & \\ 2. & \boldsymbol{X}_1 \Cap \boldsymbol{X}_2 \Cap \cdots \Cap \boldsymbol{X}_N & \text{is a LINEAR SUBSPACE of } \boldsymbol{X}\end{array}\right.$$

✎PROOF: For a structure to be a linear subspace of $\boldsymbol{X}$, it must satisfy the requirements of Definition 7.1 (page 112).

1. Proof that $\boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \hat{+} \cdots \hat{+} \boldsymbol{X}_N$ is a *linear subspace* (proof by induction):
   (a) proof for $N=1$ case: by left hypothesis.
   (b) proof for $N=2$ case:

[6] Michel and Herget (1993) page 82
[7] Michel and Herget (1993) pages 81–83

i. proof that $\boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \neq \emptyset$:

$$\begin{aligned}
\boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 &= \left\{\boldsymbol{v}+\boldsymbol{w} | \boldsymbol{v} \in \boldsymbol{X}_1 \text{ and } \boldsymbol{w} \in Y\right\} && \text{by Definition 7.2 page 113}\\
&\supseteq \left\{\boldsymbol{v}+\boldsymbol{w} | \boldsymbol{v} \in \{\mathbb{0}\} \subseteq \boldsymbol{X}_1 \text{ and } \boldsymbol{w} \in \{\mathbb{0}\} \subseteq \boldsymbol{X}_2\right\}\\
&= \{\mathbb{0}+\mathbb{0}\}\\
&= \{\mathbb{0}\}\\
&\neq \emptyset
\end{aligned}$$

ii. proof that $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \implies \boldsymbol{x}+\boldsymbol{y} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2$:

$$\begin{aligned}
\boldsymbol{x}+\boldsymbol{y} &= (\boldsymbol{v}_1+\boldsymbol{w}_1)+(\boldsymbol{v}_2+\boldsymbol{w}_2) && \text{by } \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \text{ hypothesis}\\
&= \underbrace{(\boldsymbol{v}_1+\boldsymbol{v}_2)}_{\text{in } \boldsymbol{X}_1} + \underbrace{(\boldsymbol{w}_1+\boldsymbol{w}_2)}_{\text{in } \boldsymbol{X}_2 \text{ because } \boldsymbol{X}_2 \text{ is a linear subspace}} && \text{by Definition 3.1 page 67}\\
&\in \left\{\boldsymbol{v}+\boldsymbol{w} | \boldsymbol{v} \in \boldsymbol{X}_1 \text{ and } \boldsymbol{w} \in Y\right\}\\
&= \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 && \text{by Definition 7.2 page 113}
\end{aligned}$$

iii. proof that $\boldsymbol{v} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2, \alpha \in F \implies \alpha\boldsymbol{v} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2$:

$$\begin{aligned}
\alpha\boldsymbol{x} &= \alpha(\boldsymbol{v}_1+\boldsymbol{w}_1) && \text{by } \boldsymbol{x} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \text{ hypothesis}\\
&= \underbrace{\alpha\boldsymbol{v}_1}_{\text{in } \boldsymbol{X}_1} + \underbrace{\alpha\boldsymbol{w}_1}_{\text{in } \boldsymbol{X}_2 \text{ because } \boldsymbol{X}_2 \text{ is a linear subspace}} && \text{by Definition 3.1 page 67}\\
&\in \left\{\boldsymbol{v}+\boldsymbol{w} | \boldsymbol{v} \in \boldsymbol{X}_1 \text{ and } \boldsymbol{w} \in Y\right\}\\
&= \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 && \text{by Definition 7.2 page 113}
\end{aligned}$$

(c) Proof that $[N \text{ case}] \implies [N+1 \text{ case}]$:

$$\begin{aligned}
\boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \hat{+} \cdots \hat{+} \boldsymbol{X}_{N+1} &= \underbrace{\left(\boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \hat{+} \cdots \hat{+} \boldsymbol{X}_N\right)}_{\text{linear subspace by } N \text{ case hypothesis}} \hat{+} \boldsymbol{X}_{N+1}\\
&\implies \text{linear subspace by } N = 2 \text{ case (item (1b) page 114)}
\end{aligned}$$

2. Proof that $\boldsymbol{X}_1 \cap \boldsymbol{X}_2 \cap \cdots \cap \boldsymbol{X}_N$ is a *linear subspace* (proof by induction):

(a) proof for $N = 1$ case: $\boldsymbol{X}_1$ is a linear subspace by left hypothesis.

(b) Proof for $N = 2$ case:

i. proof that $\boldsymbol{X} \cap \boldsymbol{Y} \neq \emptyset$:

$$\begin{aligned}
\boldsymbol{X} \cap \boldsymbol{Y} &= \left\{\boldsymbol{x} \in X | \boldsymbol{x} \in X \text{ and } \boldsymbol{w} \in Y\right\}\\
&\supseteq \left\{\boldsymbol{x} \in X | \boldsymbol{x} \in \{\mathbb{0}\} \subseteq \boldsymbol{X} \text{ and } \boldsymbol{x} \in \{\mathbb{0}\} \subseteq \boldsymbol{Y}\right\}\\
&= \{\mathbb{0}+\mathbb{0}\}\\
&= \{\mathbb{0}\}\\
&\neq \emptyset
\end{aligned}$$

ii. proof that $\boldsymbol{x}, \boldsymbol{y} \in X \cap \boldsymbol{Y} \implies \boldsymbol{x}+\boldsymbol{y} \in X \cap \boldsymbol{Y}$:

$$\begin{aligned}
\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X} \cap \boldsymbol{Y} &\implies \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X} \text{ and } \boldsymbol{x}, \boldsymbol{y} \in Y && \text{by Definition A.5 page 266}\\
&\implies \boldsymbol{x}+\boldsymbol{y} \in \boldsymbol{X} \text{ and } \boldsymbol{x}+\boldsymbol{y} \in Y && \text{because } \boldsymbol{X} \text{ and } \boldsymbol{Y} \text{ are linear subspaces}\\
&\implies \boldsymbol{x}+\boldsymbol{y} \in \boldsymbol{X} \cap \boldsymbol{Y} && \text{by Definition A.5 page 266}
\end{aligned}$$

iii. proof that $\boldsymbol{v} \in X \cap \boldsymbol{Y}, \alpha \in F \implies \alpha\boldsymbol{v} \in X \cap \boldsymbol{Y}$:

$$\begin{aligned}
\boldsymbol{x} \in \boldsymbol{X} \cap \boldsymbol{Y} &\implies \boldsymbol{x} \in \boldsymbol{X} \text{ and } \boldsymbol{x} \in Y && \text{by Definition A.5 page 266}\\
&\implies \alpha\boldsymbol{x} \in \boldsymbol{X} \text{ and } \alpha\boldsymbol{x} \in Y && \text{because } \boldsymbol{X} \text{ and } \boldsymbol{Y} \text{ are linear subspaces}\\
&\implies \alpha\boldsymbol{x} \in \boldsymbol{X} \cap \boldsymbol{Y} && \text{by Definition A.5 page 266}
\end{aligned}$$

(c) Proof that [N case] $\implies$ [$N+1$ case]:

$$\begin{aligned}
\boldsymbol{X}_1 \Cap \boldsymbol{X}_2 \Cap \cdots \Cap \boldsymbol{X}_{N+1} &= \underbrace{\left(\boldsymbol{X}_1 \Cap \boldsymbol{X}_2 \Cap \cdots \Cap \boldsymbol{X}_N\right)}_{\text{linear subspace by } N \text{ case hypothsis}} \Cap \boldsymbol{X}_{N+1} \\
&\implies \text{linear subspace by } N = 2 \text{ case (item (2b) page 115)}
\end{aligned}$$

Every linear subspace contains the zero vector $\mathbb{0}$ (Proposition 7.3 page 113). But if a pair of linear subspaces of a linear space $\boldsymbol{X}$ *only* have $\mathbb{0}$ in commmon, then any vector in $\boldsymbol{X}$ can be *uniquely* represented by a single vector from each of the two subspaces (next).

Theorem 7.1. [8] *Let* $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *and* $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be* LINEAR SUBSPACES *(Definition 7.1 page 112) of a* LINEAR SPACE *(Definition 3.1 page 67)* $\boldsymbol{\Omega} \triangleq (\Omega, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$.

THM

$$X \cap Y = \{\mathbb{0}\} \iff \left\{\begin{array}{l} \textit{for every } \boldsymbol{u} \in \boldsymbol{X} \dotplus \boldsymbol{Y} \textit{ there exist } \boldsymbol{x} \in X \textit{ and } \boldsymbol{y} \in Y \textit{ such that} \\ \quad 1.\ \boldsymbol{u} = \boldsymbol{x} + \boldsymbol{y} \qquad \textit{and} \\ \quad 2.\ \boldsymbol{x} \textit{ and } \boldsymbol{y} \textit{ are } \text{UNIQUE}. \end{array}\right\}$$

✎PROOF:

1. Proof that $X \cap Y = \{\mathbb{0}\} \implies$ *unique* $\boldsymbol{x}, \boldsymbol{y}$:
   Suppose that $\boldsymbol{x}$ and $\boldsymbol{y}$ are not unique, but rather $\boldsymbol{u} = \boldsymbol{x}_1 + \boldsymbol{y}_1 = \boldsymbol{x}_2 + \boldsymbol{y}_2$ where $\boldsymbol{x}_1, \boldsymbol{x}_2 \in X$ and $\boldsymbol{y}_1, \boldsymbol{y}_2 \in Y$.

$$\begin{aligned}
\boldsymbol{u} = \boldsymbol{x}_1 + \boldsymbol{y}_1 = \boldsymbol{x}_2 + \boldsymbol{y}_2 &\implies \underbrace{\boldsymbol{x}_1 - \boldsymbol{x}_2}_{\in X} = \underbrace{\boldsymbol{y}_2 - \boldsymbol{y}_1}_{\in Y} \\
&\implies \boldsymbol{x}_1 - \boldsymbol{x}_2,\ \boldsymbol{y}_2 - \boldsymbol{y}_1 \in X \cap Y \\
&\implies \boldsymbol{x}_1 - \boldsymbol{x}_2 = \boldsymbol{y}_2 - \boldsymbol{y}_1 = \mathbb{0} && \text{by left hypothesis} \\
&\implies \boldsymbol{x}_1 = \boldsymbol{x}_2 \quad \text{and} \quad \boldsymbol{y}_2 = \boldsymbol{y}_1 \\
&\implies \boldsymbol{x} \text{ and } \boldsymbol{y} \text{ are } \textit{unique}
\end{aligned}$$

2. Proof that $X \cap Y = \{\mathbb{0}\} \impliedby$ *unique* $\boldsymbol{x}, \boldsymbol{y}$:

$$\begin{aligned}
\boldsymbol{u} &= \boldsymbol{x} + \boldsymbol{y} \\
&= \boldsymbol{x} + \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{y} && \text{for some vector } \boldsymbol{y} \in X \cap Y \\
&= \underbrace{(\boldsymbol{x} + \boldsymbol{y})}_{\in X} + \underbrace{(\boldsymbol{y} - \boldsymbol{y})}_{\in Y} && \text{because } \boldsymbol{x} \in X \text{ and } \boldsymbol{y} \in X \cap Y \ldots \\
&\implies \boldsymbol{x} \text{ and } \boldsymbol{y} \text{ are } \textit{not unique} \text{ if } \boldsymbol{y} \neq \mathbb{0} \\
&\implies \boldsymbol{y} = \mathbb{0} && \text{by right hypothesis} \\
&\implies X \cap Y = \{\mathbb{0}\}
\end{aligned}$$

✏

Theorem 7.2. [9] *Let* $\boldsymbol{\Omega}$ *be a linear subspace and* 2^{Ω} *the set of closed linear subspaces of* $\boldsymbol{\Omega}$.

THM

$(2^{\Omega}, \dotplus, \Cap, \mathbf{0}, \boldsymbol{\Omega}; \subseteq)$ *is a* LATTICE *(Definition C.3 page 305). In particular*

$$\begin{array}{rclrcll}
\boldsymbol{X} \dotplus \boldsymbol{X} &=& \boldsymbol{X} & \boldsymbol{X} \Cap \boldsymbol{X} &=& \boldsymbol{X} & \forall \boldsymbol{X} \in 2^{\Omega} \\
\boldsymbol{X} \dotplus \boldsymbol{Y} &=& \boldsymbol{Y} \dotplus \boldsymbol{X} & \boldsymbol{X} \Cap \boldsymbol{Y} &=& \boldsymbol{Y} \Cap \boldsymbol{X} & \forall \boldsymbol{X}, \boldsymbol{Y} \in 2^{\Omega} \\
(\boldsymbol{X} \dotplus \boldsymbol{Y}) \dotplus \boldsymbol{Z} &=& \boldsymbol{X} \dotplus (\boldsymbol{Y} \dotplus \boldsymbol{Z}) & (\boldsymbol{X} \Cap \boldsymbol{Y}) \Cap \boldsymbol{Z} &=& \boldsymbol{X} \Cap (\boldsymbol{Y} \Cap \boldsymbol{Z}) & \forall \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \in 2^{\Omega} \\
\boldsymbol{X} \dotplus (\boldsymbol{X} \Cap \boldsymbol{Y}) &=& \boldsymbol{X} & \boldsymbol{X} \Cap (\boldsymbol{X} \dotplus \boldsymbol{Y}) &=& \boldsymbol{X} & \forall \boldsymbol{X}, \boldsymbol{Y} \in 2^{\Omega}
\end{array}$$

✎PROOF: These results follow directly from the properties of lattices (Theorem C.3 page 306). ✏

[8] Michel and Herget (1993) page 83 ⟨Theorem 3.2.12⟩, Kubrusly (2001) page 67 ⟨Theorem 2.14⟩
[9] Iturrioz (1985) pages 56–57

7.2 Subspaces of an inner product space

Definition 7.3. [10] *Let* $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \triangledown \rangle\right)$ *be an* INNER PRODUCT SPACE *(Definition 6.1 page 95).*

DEF

*The **orthogonal complement** $A^\perp$ in $\boldsymbol{\Omega}$ of a set $A \subseteq X$ is*

$$A^\perp \triangleq \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in A\}.$$

The expression $A^{\perp\perp}$ is defined as $\left(A^\perp\right)^\perp$.

Proposition 7.5. [11] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \triangledown \rangle\right)$ *be an* INNER PRODUCT SPACE *(Definition 6.1 page 95).*

PRP

$$A \subseteq B \implies B^\perp \subseteq A^\perp \quad \forall A,B \in 2^X \qquad \text{(ANTITONE)}$$

✎PROOF:

$$\begin{aligned}
B^\perp &\triangleq \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in B\} && \text{by definition of } B^\perp \text{ (Definition 7.3 page 117)}\\
&\subseteq \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in A\} && \text{by } A \subseteq B \text{ hypothesis}\\
&= A^\perp && \text{by definition of } A^\perp \text{ (Definition 7.3 page 117)}
\end{aligned}$$

⇰

Every *linear space* $\boldsymbol{X}$ contains $\mathbf{0}$ and $\boldsymbol{X}$ as *linear subspaces* (Proposition 7.1 page 112). If $\boldsymbol{X}$ is also an *inner product space*, then $\mathbf{0}$ and $\boldsymbol{X}$ are *orthogonal complements* of each other (next proposition).

Proposition 7.6. [12] *Let* $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \triangledown \rangle\right)$ *be an* INNER PRODUCT SPACE *(Definition 6.1 page 95) and $\mathbb{0}$ the* VECTOR ADDITIVE IDENTITY ELEMENT *(Definition 3.1 page 67) in $\boldsymbol{\Omega}$.*

PRP

1. $\{\mathbb{0}\}^\perp = X$
2. $X^\perp = \{\mathbb{0}\}$

✎PROOF:

$$\begin{aligned}
\{\mathbb{0}\}^\perp &= \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in \{\mathbb{0}\}\} && \text{by definition of } \perp \text{ (Definition 7.3 page 117)}\\
&= \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \mathbb{0} \rangle = 0\}\\
&= X\\
X^\perp &= \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in X\} && \text{by definition of } \perp \text{ Definition 7.3 page 117}\\
&= \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{x} \rangle = 0\}\\
&= \{\mathbb{0}\}
\end{aligned}$$

⇰

For any set A contained in a linear space $\boldsymbol{X}$, $A^{\perp\perp}$ is a *linear subspace*, and it is the smallest linear subspace containing the set A ($A^{\perp\perp} = \operatorname{span} A$, next theorem). In the case that A is a *linear subspace* rather than just a subset, results simplify significantly (next corollary).

---

[10] Berberian (1961) page 59 ⟨Definition III.2.1⟩, Michel and Herget (1993) page 382, Kubrusly (2001) page 328

[11] Berberian (1961) page 60 ⟨Theorem III.2.2⟩, Kubrusly (2011) page 326

[12] Kubrusly (2011) page 326, Michel and Herget (1993) page 383

Theorem 7.3. [13] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be an* INNER PRODUCT SPACE (Definition 6.1 page 95). *Let* $\operatorname{span} A$ *be the span of a set* A (Definition 10.2 page 151).

THM

$$\left\{\begin{array}{l} A \text{ is a **subset** of } X \\ (A \subseteq X) \end{array}\right\} \implies \left\{\begin{array}{lrcll} 1. & A \cap A^\perp & = & \begin{cases} \{\mathbb{0}\} & \text{if } \mathbb{0} \in A \\ \varnothing & \text{if } \mathbb{0} \notin A \end{cases} & \text{and} \\ 2. & A & \subseteq & A^{\perp\perp} = \operatorname{span} A & \text{and} \\ 3. & A^\perp & = & A^{\perp\perp\perp} = A^{\perp -} = A^{-\perp} = (\operatorname{span} A)^\perp & \text{and} \\ 4. & \multicolumn{4}{l}{A^\perp \text{ is a **subspace** of } \boldsymbol{\Omega}} \end{array}\right\}$$

✎PROOF:

1. Proof that $A \cap A^\perp = \cdots$:

$$\begin{aligned} A \cap A^\perp &= \{\boldsymbol{x} \in X \mid \boldsymbol{x} \in A\} \cap \{\boldsymbol{x} \in X \mid \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle \quad \forall \boldsymbol{y} \in A\} && \text{by definition of } A^\perp \\ &= \{\boldsymbol{x} \in X \mid \boldsymbol{x} \in A \quad \text{and} \quad \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle \quad \forall \boldsymbol{y} \in A\} \\ &= \begin{cases} \{\mathbb{0}\} & \text{if } \mathbb{0} \in A \\ \varnothing & \text{if } \mathbb{0} \notin A \end{cases} \end{aligned}$$

2. Proof that $A \subseteq A^{\perp\perp} = \operatorname{span} A$:

$$\begin{aligned} \boldsymbol{x} \in A &\implies \{\boldsymbol{x}\}^{\perp\perp} \subseteq A^{\perp\perp} \\ &\implies \boldsymbol{x} \in \{\boldsymbol{x}\}^{\perp\perp} \subseteq A^{\perp\perp} \\ &\implies \boldsymbol{x} \in A^{\perp\perp} \end{aligned}$$

but

$$\boldsymbol{x} \in A^{\perp\perp} \not\Longrightarrow \boldsymbol{x} \in A$$

Here is an example for the $\not\Longrightarrow$ part using the linear space $\mathbb{R}^3$:

   (a) Let $A \triangleq \{i\}$, where i is the unit vector on the x-axis.

   (b) Then $A^\perp = \{\boldsymbol{x} \in X \mid \boldsymbol{x} \in \text{yz plane}\}$.

   (c) Then $A^{\perp\perp} = \{\boldsymbol{x} \in X \mid \boldsymbol{x} \in \text{x axis}\}$.

   (d) Therefore, $A \subsetneq A^{\perp\perp}$

3. Proof for $A^\perp$ equivalent expressions:

   (a) Proof that $A^\perp = A^{\perp\perp\perp}$:

$$\begin{aligned} A^\perp &\subseteq (A^\perp)^{\perp\perp} && \text{by item (2)} \\ &= (A^{\perp\perp})^\perp \\ &= A^{\perp\perp\perp} && \text{by Definition 7.3 page 117} \\ A^{\perp\perp\perp} &= (A^{\perp\perp})^\perp && \text{by Definition 7.3 page 117} \\ &\subseteq A^\perp && \text{by item (2) and Proposition 7.5 (page 117)} \end{aligned}$$

   (b) Proof that $A^{\perp\perp\perp} = (\operatorname{span} A)^\perp$: follows directly from item (2) ($A^{\perp\perp} = \operatorname{span} A$).

   (c) Proof that $A^\perp = A^{\perp -}$:

      i. Let $(\!(\boldsymbol{x}_n)\!)$ be an $A^\perp$-valued sequence that converges to the limit $\boldsymbol{x}$ in X.

[13] Michel and Herget (1993) page 383, Kubrusly (2011) page 326

ii. The limit point $\boldsymbol{x}$ must be in $A^\perp$ because for all $\boldsymbol{y} \in A$

$$\begin{aligned}
\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle &= \langle \lim \boldsymbol{x}_n \mid \boldsymbol{y} \rangle && \text{by definition of the sequence } (\!(\boldsymbol{x}_n)\!) \\
&= \lim \langle \boldsymbol{x}_n \mid \boldsymbol{y} \rangle \\
&= 0 && \text{because } (\!(\boldsymbol{x}_n)\!) \text{ is } A^\perp\text{-valued}
\end{aligned}$$

iii. Because $\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in A$, $\boldsymbol{x}$ is in $A^\perp$.

iv. Because $A^\perp$ contains all its limit points, and by the *Closed Set Theorem* (Theorem 8.1 page 128), it must be *closed* ($A^\perp = A^{\perp -}$)

(d) Proof that $A^\perp = A^{-\perp}$:

i. Let $x \in A^\perp$ and $y \in A^-$.

ii. Let $(\!(\boldsymbol{y}_n)\!)$ be an $A^\perp$-valued sequence that converges in X to y.

iii. Thus $A^\perp \perp A^-$ because

$$\begin{aligned}
\langle \boldsymbol{y} \mid \boldsymbol{x} \rangle &= \langle \lim \boldsymbol{y}_n \mid \boldsymbol{x} \rangle && \text{by definition of } (\!(\boldsymbol{y}_n)\!) \\
&= \lim \langle \boldsymbol{y}_n \mid \boldsymbol{x} \rangle \\
&= 0 && \text{because } (\!(\boldsymbol{y}_n)\!) \text{ is } A^\perp\text{-valued}
\end{aligned}$$

iv. Because $A^\perp \perp A^-$, so $A^\perp \subseteq A^{\perp -}$.

v. But $A^{\perp -} \subseteq A^\perp$ because

$$A \subseteq A^- \implies A^{\perp -} \subseteq A^\perp \qquad \text{by } \textit{antitone} \text{ property (Proposition 7.5 page 117)}$$

vi. And so $A^\perp = A^{\perp -}$.

4. Proof that $A^\perp$ is a **subspace** of $\boldsymbol{\Omega}$ (must satisfy the conditions of Definition 7.1 page 112):

(a) Proof that $A^\perp \neq \varnothing$: $A^\perp$ has at least one element, the element $\mathbb{0}$...

$$\begin{aligned}
\langle \mathbb{0} \mid \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in A & && \text{by definition of } \mathbb{0} \\
\implies \mathbb{0} \in A^\perp & && \text{by definition of } A^\perp \text{ (Definition 7.3 page 117)}
\end{aligned}$$

(b) Proof that $A^\perp \subseteq X$:

$$\begin{aligned}
\boldsymbol{u} \in A^\perp &\implies \boldsymbol{u} \in \{\boldsymbol{x} \in X \mid \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in A\} && \text{by definition of } A^\perp \text{ (Definition 7.3 page 117)} \\
&\implies \boldsymbol{u} \in X && \text{by definition of sets}
\end{aligned}$$

(c) Proof that $\boldsymbol{u}, \boldsymbol{v} \in A^\perp \implies (\boldsymbol{u} + \boldsymbol{v}) \in A^\perp$:

$$\begin{aligned}
\boldsymbol{u}, \boldsymbol{v} \in A^\perp &\implies \langle \boldsymbol{u} \mid \boldsymbol{y} \rangle = \langle \boldsymbol{v} \mid \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in A && \text{by definition of } A^\perp \text{ (Definition 7.3 page 117)} \\
&\implies \langle \boldsymbol{u} \mid \boldsymbol{y} \rangle + \langle \boldsymbol{v} \mid \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in A \\
&\implies \langle \boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{y} \rangle = 0 \qquad \forall \boldsymbol{y} \in A && \text{by } \textit{additive} \text{ property of } \langle \triangle \mid \nabla \rangle \text{ (Definition 6.1 page 95)} \\
&\implies \boldsymbol{u} + \boldsymbol{v} \in A^\perp && \text{by definition of } A^\perp \text{ (Definition 7.3 page 117)}
\end{aligned}$$

(d) Proof that $\boldsymbol{v} \in \Omega \implies \alpha\boldsymbol{v} \in A^\perp$:

$$\begin{aligned}
\boldsymbol{v} \in A^\perp &\implies \langle \boldsymbol{v} \mid \boldsymbol{y} \rangle = 0 \qquad \forall \boldsymbol{y} \in A && \text{by definition of } A^\perp \text{ (Definition 7.3 page 117)} \\
&\implies \alpha \langle \boldsymbol{v} \mid \boldsymbol{y} \rangle = \alpha \cdot 0 \quad \forall \boldsymbol{y} \in A \\
&\implies \langle \alpha\boldsymbol{v} \mid \boldsymbol{y} \rangle = 0 \qquad \forall \boldsymbol{y} \in A && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \mid \nabla \rangle \text{ (Definition 6.1 page 95)} \\
&\implies \alpha\boldsymbol{v} \in A^\perp && \text{by definition of } A^\perp \text{ (Definition 7.3 page 117)}
\end{aligned}$$

⇨

Corollary 7.1. *Let* $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *and* $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be* INNER PRODUCT SPACES. *Let* $\operatorname{span} Y$ *be the span of the set* Y (Definition 10.2 page 151).

COR

$$\left\{ \boldsymbol{Y} \text{ is a } \textbf{linear subspace} \text{ of } \boldsymbol{X} \right\} \implies \left\{ \begin{array}{llcll} 1. & Y \cap Y^{\perp} & = & \{\mathbb{0}\} & \text{and} \\ 2. & Y & = & Y^{\perp\perp} = \operatorname{span} Y & \text{and} \\ 3. & Y^{\perp} & = & Y^{\perp\perp\perp} & \text{and} \\ 4. & \multicolumn{4}{l}{Y^{\perp} \text{ is a } \textbf{subspace} \text{ of } \boldsymbol{X}} \end{array} \right\}$$

✎PROOF:

1. Proof that $\boldsymbol{Y} \cap Y^{\perp} = \{\mathbb{0}\}$: This follows from Theorem 7.3 (page 118) and the fact that all subspaces contain the zero vector $\mathbb{0}$ (Proposition 7.3 page 113).
2. Proof that $\boldsymbol{Y} = \boldsymbol{Y}^{\perp\perp} = \operatorname{span} \boldsymbol{Y}$: This follows directly from Theorem 7.3 (page 118).
3. Proof that $\boldsymbol{Y}^{\perp} = \boldsymbol{Y}^{\perp\perp\perp}$: This follows directly from Theorem 7.3 (page 118).
4. Proof that $Y^{\perp}$ is a **subspace** of $\boldsymbol{X}$: This follows directly from Theorem 7.3 (page 118).

⇨

Theorem 7.4. [14] *Let* $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *and* $\boldsymbol{Z} \triangleq (Z, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be* LINEAR SUBSPACES *of an* INNER PRODUCT SPACE $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$.

THM

$$\boldsymbol{Y} \perp \boldsymbol{Z} \quad \implies \quad Y \cap Z = \{\mathbb{0}\}$$

✎PROOF:

$$\begin{aligned} \boldsymbol{x} \in \boldsymbol{Y} \cap \boldsymbol{Z} &\implies \boldsymbol{x} \in \boldsymbol{Y} \text{ and } \boldsymbol{x} \in \boldsymbol{Z} && \text{by definition of } \cap \\ &\implies \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle = 0 && \text{by hypothesis } \boldsymbol{Y} \perp \boldsymbol{Z} \\ &\implies \boldsymbol{x} = \mathbb{0} && \text{by } \textit{non-isotropic} \text{property of } \langle \triangle \mid \triangledown \rangle \text{ (Definition 6.1 page 95)} \end{aligned}$$

⇨

Theorem 7.5. [15] *Let* $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *and* $\boldsymbol{Z} \triangleq (Z, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be linear subspaces of an* INNER PRODUCT SPACE $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$.

THM

$$\left\{ \begin{array}{ll} 1. & \boldsymbol{Y} \perp \boldsymbol{Z} \text{ and} \\ 2. & \boldsymbol{x} \in \boldsymbol{Y} \hat{+} \boldsymbol{Z} \end{array} \right\} \implies \left\{ \begin{array}{ll} 1. & \text{There exists } \boldsymbol{y} \in \boldsymbol{Y} \text{ and } \boldsymbol{z} \in \boldsymbol{Z} \text{ such that } \boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z} \text{ and} \\ 2. & \boldsymbol{y} \text{ and } \boldsymbol{z} \text{ are UNIQUE.} \end{array} \right\}$$

✎PROOF:

1. Proof that $\boldsymbol{y}$ and $\boldsymbol{z}$ exist: by definition of Minkowski addition operator $\hat{+}$ (Definition 7.2 page 113).
2. Proof that $\boldsymbol{y}$ and $\boldsymbol{z}$ are *unique*:
   (a) Suppose $\boldsymbol{x} = \boldsymbol{y}_1 + \boldsymbol{z}_1 = \boldsymbol{y}_1 + \boldsymbol{z}_2$ for $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \boldsymbol{Y}$ and $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \boldsymbol{Z}$.

[14] Kubrusly (2001) page 324
[15] Berberian (1961) page 61 ⟨Theorem III.2.3⟩

(b) This implies

$$\begin{aligned}
\mathbb{0} &= \boldsymbol{x} - \boldsymbol{x} \\
&= (\boldsymbol{y}_1 + \boldsymbol{z}_1) - (\boldsymbol{y}_1 + \boldsymbol{z}_2) \\
&= \underbrace{(\boldsymbol{y}_1 - \boldsymbol{y}_2)}_{\text{in } \boldsymbol{Y}} + \underbrace{(\boldsymbol{z}_1 - \boldsymbol{z}_2)}_{\text{in } \boldsymbol{Z}}
\end{aligned}$$

(c) Because $\boldsymbol{y}_1 - \boldsymbol{y}_2 \in \boldsymbol{Y}$, $\boldsymbol{z}_1 - \boldsymbol{z}_2 \in \boldsymbol{Z}$, $(\boldsymbol{y}_1 - \boldsymbol{y}_2) + (\boldsymbol{z}_1 + \boldsymbol{z}_2) = \mathbb{0}$, and $\langle \boldsymbol{y}_1 - \boldsymbol{y}_2 \,|\, \boldsymbol{z}_1 - \boldsymbol{z}_2 \rangle = 0$, then by Theorem 6.9 (page 107), $\boldsymbol{y}_1 - \boldsymbol{y}_2 = \mathbb{0}$ and $\boldsymbol{z}_1 - \boldsymbol{z}_2 = \mathbb{0}$.

(d) This implies $\boldsymbol{y}_1 = \boldsymbol{y}_2$ and $\boldsymbol{z}_1 = \boldsymbol{z}_2$.

(e) This implies $\boldsymbol{y}$ and $\boldsymbol{z}$ are *unique*.

⇰

7.3 Subspaces of a Hilbert Space

Theorem 7.6. [16] *Let* $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \triangledown \rangle)$ *be a* HILBERT SPACE (Definition 8.11 page 138). *Let* Y *be a* SUBSET *of* X, *and let* $\mathsf{d}(\boldsymbol{x}, Y) \triangleq \inf_{\boldsymbol{y} \in Y} \|\boldsymbol{x} - \boldsymbol{y}\|$.

THM

$$\left\{\begin{array}{lll}
1. & Y \neq \varnothing & \text{and} \\
2. & Y \text{ is CLOSED} \quad \text{(Definition 1.4 page 14)} & \text{and} \\
3. & Y \text{ is CONVEX} \quad \text{(Definition 9.6 page 142)} &
\end{array}\right\} \implies \left\{\begin{array}{l}
\text{There exists } \boldsymbol{p} \in Y \text{ such that} \\
1. \ \mathsf{d}(\boldsymbol{x}, Y) = \|\boldsymbol{x} - \boldsymbol{p}\| \quad \text{and} \\
2. \ \boldsymbol{p} \text{ is UNIQUE.}
\end{array}\right\}$$

✎PROOF:

1. Let $\delta \triangleq \inf\{\boldsymbol{x} - \boldsymbol{y} | \boldsymbol{y} \in Y\}$.

2. Let $(\!(\boldsymbol{y}_n)\!)_{n \in \mathbb{Z}}$ be a sequence such that $\|\boldsymbol{x} - \boldsymbol{y}_n\| \to \delta$.

3. Proof that $(\!(\boldsymbol{y}_n)\!)$ is *Cauchy*:

$$\begin{aligned}
&\lim_{m,n\to\infty} \|\boldsymbol{y}_n - \boldsymbol{y}_m\|^2 \\
&= \lim_{m,n\to\infty} \|(\boldsymbol{y}_n - x) + (x - \boldsymbol{y}_m)\|^2 \\
&= \lim_{m,n\to\infty} \left\{ -\|(\boldsymbol{y}_n - x) - (x - \boldsymbol{y}_m)\|^2 + 2\|\boldsymbol{y}_n - x\|^2 + 2\|x - \boldsymbol{y}_m\|^2 \right\} && \text{by } \textit{parallelogram law} \text{ (page 103)} \\
&= \lim_{m,n\to\infty} \left\{ -4 \left\| \underbrace{\left(\frac{1}{2}\boldsymbol{y}_n + \frac{1}{2}\boldsymbol{y}_m\right)}_{\text{in } \boldsymbol{Y} \text{ by } \textit{convexity}} - x \right\|^2 + 2\|\boldsymbol{y}_n - x\|^2 + 2\|x - \boldsymbol{y}_m\|^2 \right\} \\
&\le \lim_{m,n\to\infty} \left\{ -4\delta^2 + 2\|\boldsymbol{y}_n - x\|^2 + 2\|x - \boldsymbol{y}_m\|^2 \right\} && \text{by definition of } \delta \text{ (item (1))} \\
&= -4\delta^2 + \lim_{m,n\to\infty} \left\{ 2\|\boldsymbol{y}_n - x\|^2 \right\} + \lim_{m,n\to\infty} \left\{ 2\|x - \boldsymbol{y}_m\|^2 \right\} \\
&= -4\delta^2 + 2\delta^2 + 2\delta^2 && \text{by definition of } \delta \text{ (item (1))} \\
&= 0
\end{aligned}$$

[16] Kubrusly (2001) page 330 ⟨Theorem 5.13⟩, Aliprantis and Burkinshaw (1998) page 290 ⟨Theorem 33.6⟩, Berberian (1961) page 68 ⟨Theorem III.5.1⟩

4. Proof that $\mathsf{d}(\boldsymbol{x}, Y) = \|\boldsymbol{x} - \boldsymbol{y}\|$: because $(\!(\boldsymbol{y}_n)\!)$ is *Cauchy* (item (1)) and by the *closed* hypothesis.

5. Proof that $\boldsymbol{y}$ is *unique*: Because in a metric space, the limit of a convergent sequence is *unique* by Theorem 8.4 page 132.

⇨

Theorem 7.7. [17] *Let* $\boldsymbol{H} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *be a* HILBERT SPACE *(Definition 8.11 page 138). Let* $\mathsf{d}(\boldsymbol{x}, Y) \triangleq \inf_{\boldsymbol{y} \in Y} \|\boldsymbol{x} - \boldsymbol{y}\|$*. Let* $\boldsymbol{Y} \triangleq \left(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *and* $Y^{\perp}$ *the* ORTHOGONAL COMPLEMENT *of* Y.

THM
$$\left\{ \boldsymbol{Y} \text{ is a SUBSPACE of } \boldsymbol{H} \right\} \implies \left\{ \begin{array}{lll} \multicolumn{3}{l}{\text{There exists } \boldsymbol{p} \in Y \text{ such that}} \\ 1. & \mathsf{d}(\boldsymbol{x}, Y) = \|\boldsymbol{x} - \boldsymbol{p}\| & \text{and} \\ 2. & \boldsymbol{p} \text{ is UNIQUE} & \text{and} \\ 3. & \boldsymbol{x} - \boldsymbol{p} \in Y^{\perp}. & \end{array} \right\}$$

Theorem 7.8 (Projection Theorem). [18] *Let* $\boldsymbol{H} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *be a Hilbert space.*

THM
$$\left\{ \boldsymbol{Y} \text{ is a SUBSPACE of } \boldsymbol{H} \right\} \implies \left\{ \boldsymbol{Y} \dotplus \boldsymbol{Y}^{\perp} = \boldsymbol{H} \right\}$$

✎PROOF:

$$\begin{aligned}
\boldsymbol{Y} \dotplus \boldsymbol{Y}^{\perp} &= \left[\boldsymbol{Y} \dotplus \boldsymbol{Y}^{\perp}\right]^{\perp\perp} && \text{by Corollary 7.1 page 120} \\
&= \left[\boldsymbol{Y}^{\perp} \cap \boldsymbol{Y}^{\perp\perp}\right]^{\perp} && \text{by Proposition 7.5 (page 117) and Lemma 15.1 (page 247)} \\
&= \{\mathbb{0}\}^{\perp} && \text{by Corollary 7.1 page 120} \\
&= \boldsymbol{H} && \text{by Proposition 7.6 page 117}
\end{aligned}$$

⇨

The inclusion relation $\subseteq$ is an order relation on the set of subspaces of a linear space $\boldsymbol{\Omega}$.

Proposition 7.7. *Let* $\boldsymbol{S}$ *be the set of subspaces of a linear space* $\boldsymbol{\Omega}$*. Let* $\subseteq$ *be the inclusion relation.*

PRP
$(\boldsymbol{S}, \subseteq)$ *is an* ***ordered set***

✎PROOF: $(\boldsymbol{S}, \subseteq)$ is an *ordered set* (Definition B.2 page 290) and because

1.	$\boldsymbol{X} \subseteq \boldsymbol{X}$	$\forall \boldsymbol{X} \in \boldsymbol{S}$	(reflexive)	and	preorder
2.	$\boldsymbol{X} \subseteq \boldsymbol{Y}$ and $\boldsymbol{Y} \subseteq \boldsymbol{Z} \implies \boldsymbol{X} \subseteq \boldsymbol{Z}$	$\forall \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \in \boldsymbol{S}$	(transitive)	and	preorder
3.	$\boldsymbol{X} \subseteq \boldsymbol{Y}$ and $\boldsymbol{Y} \subseteq \boldsymbol{X} \implies \boldsymbol{X} = \boldsymbol{Y}$	$\forall \boldsymbol{X}, \boldsymbol{Y} \in \boldsymbol{S}$	(anti-symmetric)		

⇨

Theorem 7.9. [19] *Let* $\boldsymbol{H}$ *be a Hilbert space and* $2^{\boldsymbol{H}}$ *the set of closed linear subspaces of* $\boldsymbol{H}$.

THM
$\left(2^{\boldsymbol{H}}, \dotplus, \cap, \boldsymbol{0}, \boldsymbol{H}; \subseteq\right)$ *is an* ORTHOMODULAR LATTICE *(Definition 15.3 page 253). In particular*

1.	$\boldsymbol{X} \dotplus \boldsymbol{X}^{\perp} = \boldsymbol{H}$		$\forall \boldsymbol{X} \in \boldsymbol{H}$	(COMPLEMENTED)
2.	$\boldsymbol{X} \cap \boldsymbol{X}^{\perp} = \boldsymbol{0}$		$\forall \boldsymbol{X} \in \boldsymbol{H}$	(COMPLEMENTED)
3.	$\left(\boldsymbol{X}^{\perp}\right)^{\perp} = \boldsymbol{X}$		$\forall \boldsymbol{X} \in \boldsymbol{H}$	(INVOLUTORY)
4.	$\boldsymbol{X} \leq \boldsymbol{Y}$	$\implies \boldsymbol{Y}^{\perp} \leq \boldsymbol{X}^{\perp}$	$\forall \boldsymbol{X}, \boldsymbol{Y} \in \boldsymbol{H}$	(ANTITONE)
5.	$\boldsymbol{X} \leq \boldsymbol{Y}$	$\implies \boldsymbol{X} \dotplus \left(\boldsymbol{X}^{\perp} \cap \boldsymbol{Y}\right) = \boldsymbol{Y}$	$\forall \boldsymbol{X}, \boldsymbol{Y} \in X$	(ORTHOMODULAR IDENTITY)

[17] Kubrusly (2001) page 330 ⟨Theorem 5.13⟩
[18] Bachman and Narici (1966) page 172 ⟨Theorem 10.8⟩, Kubrusly (2001) page 339 ⟨Theorem 5.20⟩
[19] Iturrioz (1985) pages 56–57

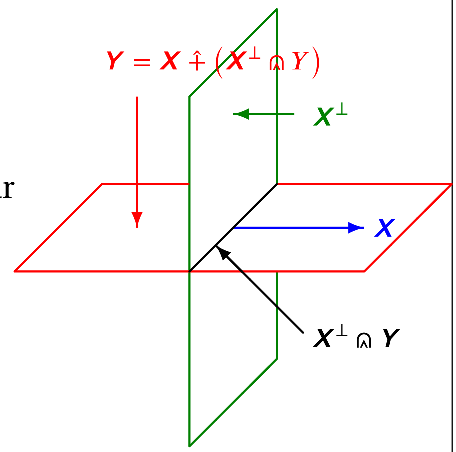
✎PROOF:

1. Proof for *complemented* (1) property: by *Projection Theorem* (Theorem 7.8 page 122).
2. Proof for *complemented* (2) property: by Corollary 7.1 (page 120).
3. Proof for *involutory* property: by Corollary 7.1 (page 120).
4. Proof for *antitone* property: by Proposition 7.5 (page 117).
5. Proof for *orthomodular identity* property:
6. Proof that lattice is *orthomodular*: by 5 properties and definition of *orthomodular lattice* (Definition 15.3 page 253).

✏

This concept is illustrated to the right where $\boldsymbol{X}, \boldsymbol{Y} \in 2^{\boldsymbol{H}}$ are linear subspaces of the linear space $\boldsymbol{H}$ and

$$\boldsymbol{X} \subseteq \boldsymbol{Y} \implies \boldsymbol{Y} = \boldsymbol{X} \hat{+} \left(\boldsymbol{X}^{\perp} \hat{\cap} Y\right).$$



Corollary 7.2. *Let $\boldsymbol{H}$ be a Hilbert space with orthogonality operation $\perp$. Let $\left(2^{\boldsymbol{H}}, \hat{+}, \hat{\cap}, \mathbf{0}, \boldsymbol{H}; \subseteq\right)$ be the lattice of subspaces of $\boldsymbol{H}$.*

COR

$$(\boldsymbol{X} \hat{+} \boldsymbol{Y})^{\perp} = \boldsymbol{X}^{\perp} \hat{\cap} \boldsymbol{Y}^{\perp} \quad \forall \boldsymbol{X}, \boldsymbol{Y} \in 2^{\boldsymbol{H}} \quad \text{(DE MORGAN) and}$$
$$(\boldsymbol{X} \hat{\cap} \boldsymbol{Y})^{\perp} = \boldsymbol{X}^{\perp} \hat{+} \boldsymbol{Y}^{\perp} \quad \forall \boldsymbol{X}, \boldsymbol{Y} \in 2^{\boldsymbol{H}} \quad \text{(DE MORGAN)}$$

✎PROOF: By properties of *orthocomplemented lattices* (Theorem 15.1 page 246). ✏

7.4 Subspace Metrics

Definition 7.4 (Hilbert space gap metric). [20] *Let $\boldsymbol{X}$ be a **Hilbert space** and $\boldsymbol{S}$ the set of subspaces of $\boldsymbol{X}$. Then we define the following metric between subspaces of $\boldsymbol{X}$.*

DEF

$$\mathsf{d}(\boldsymbol{V}, \boldsymbol{W}) \triangleq \|\mathbf{P} - \mathbf{Q}\| \quad \forall \boldsymbol{V}, \boldsymbol{W} \in \boldsymbol{S}$$
(the distance between subspaces $\boldsymbol{V}$ and $\boldsymbol{W}$ is the size of the difference of their projection operators)

$$\text{where } \boldsymbol{V} \triangleq \mathbf{P}\boldsymbol{X}$$
($\mathbf{P}$ is the projection operator that generates the subspace $\boldsymbol{V}$)

$$\text{and } \boldsymbol{W} \triangleq \mathbf{Q}\boldsymbol{X}$$
($\mathbf{Q}$ is the projection operator that generates the subspace $\boldsymbol{W}$).

Definition 7.5 (Banach space gap metric). [21] *Let $\boldsymbol{X}$ be a **Banach space** and $\boldsymbol{S}$ the set of subspaces of $\boldsymbol{X}$. Then we define the following metric between subspaces of $\boldsymbol{X}$.*

DEF

$$\mathsf{d}(\boldsymbol{V}, \boldsymbol{W}) \triangleq \max\left\{ \sup_{\boldsymbol{v} \in \boldsymbol{V}, \|\boldsymbol{v}\|=1} \mathsf{p}(\boldsymbol{v}, \boldsymbol{W}), \sup_{\boldsymbol{w} \in \boldsymbol{W}, \|\boldsymbol{w}\|=1} \mathsf{p}(\boldsymbol{w}, \boldsymbol{V}) \right\} \quad \forall \boldsymbol{V}, \boldsymbol{W} \in \boldsymbol{S}$$

$$\text{where } \mathsf{p}(\boldsymbol{v}, \boldsymbol{W}) \triangleq \inf_{\boldsymbol{w} \in \boldsymbol{W}} \|\boldsymbol{v} - \boldsymbol{w}\|$$
(metric from the point $\boldsymbol{v}$ to the subspace $\boldsymbol{W}$)

[20] Deza and Deza (2006), page 235, Akhiezer and Glazman (1993), page 69, Berkson (1963), page 8, Krein and Krasnoselski (1947)

[21] Akhiezer and Glazman (1993), page 70, Berkson (1963), page 8, Krein et al. (1948)

Definition 7.6 (Schäffer's metric). [22]

DEF

$$\begin{aligned}
\mathsf{d}(V, W) &= \log\left(1 + \max\left\{\mathsf{r}(V, W), \mathsf{r}(W, V)\right\}\right) && \textit{where} \\
\mathsf{r}(V, W) &\triangleq \begin{cases} \inf\left\{\|\mathbf{A} - \mathbf{I}\| \,|\, \mathbf{A}V = W\right\} & \textit{if } \mathbf{A} \textit{ and } \mathbf{A}^{-1} \textit{ both exist} \\ 1 & \textit{otherwise} \end{cases}
\end{aligned}$$

7.5 Literature

Literature survey:

1. Lattice of subspaces
   - Birkhoff and Neumann (1936)
   - Husimi (1937)
   - Sasaki (1954)
   - Loomis (1955)
   - von Neumann (1960)
   - Holland (1970)
   - Halmos (1998b)
   - Amemiya and Araki (1966)
   - Gudder (1979)
   - Gudder (2005)
2. Characterizations of lattice of Hilbert subspaces (cf Iturrioz (1985) page 60):
   - Kakutani and Mackey (1946) ⟨using Banach spaces⟩
   - Piron (1964a) ⟨using pre-Hilbert spaces⟩
     - Piron (1964b) ⟨using pre-Hilbert spaces⟩
   - Amemiya and Araki (1966) ⟨using pre-Hilbert spaces⟩
   - Wilbur (1975) ⟨using locally convex spaces⟩
3. Metrics on subspaces:
   - Burago et al. (2001)

[22] Massera and Schäffer (1958), pages 562–563, Berkson (1963), pages 7–8

Part II

Properties of Spaces

CHAPTER 8

SEQUENCES AND CONVERGENCE

8.1 Definitions

Definition 8.1. [1] *Let X^Y be the set of all functions from a set Y to a set X. Let $\mathbb{Z}$ be the* SET OF INTEGERS.

D E F	*A function* f *in* X^Y *is an* X*-valued* ***sequence*** *if* $Y = \mathbb{Z}$. *A sequence may be denoted in the form* $\left(\!\left(x_n\right)\!\right)_{n\in\mathbb{Z}}$ *or simply as* $\left(\!\left(x_n\right)\!\right)$. *A function* f *in* X^Y *is an* X*-valued* ***n-tuple*** *if* $Y = \{1, 2, \ldots, N\}$. *An n-tuple may be denoted in the form* $\left(\!\left|x_n\right|\!\right)_1^N$ *or simply as* $\left(\!\left|x_n\right|\!\right)$.

Definition 8.2. [2] *Let* $\left(\!\left(x_n\right)\!\right)_{n\in\mathbb{Z}}$ *and* $\left(\!\left(y_n\right)\!\right)_{n\in\mathbb{Z}}$ *be sequences over a field* $\mathbb{F}$.
Let $\left(\!\left|x_n\right|\!\right)_1^N$ *and* $\left(\!\left|y_n\right|\!\right)_1^N$ *be n-tuples over a field* $\mathbb{F}$.

$$\begin{array}{llllll}
\left(\!\left(x_n\right)\!\right) + \left(\!\left(y_n\right)\!\right) & \triangleq & \left(\!\left(x_n + y_n\right)\!\right) & \alpha \left(\!\left(x_n\right)\!\right) & \triangleq & \left(\!\left(\alpha x_n\right)\!\right) \quad \forall\alpha\in\mathbb{F} \\
\left(\!\left|x_n\right|\!\right) + \left(\!\left|y_n\right|\!\right) & \triangleq & \left(\!\left|x_n + y_n\right|\!\right) & \alpha \left(\!\left|x_n\right|\!\right) & \triangleq & \left(\!\left|\alpha x_n\right|\!\right) \quad \forall\alpha\in\mathbb{F}
\end{array}$$

8.2 Sequences in topological spaces

A *topological space* (Definition 1.1 page 3) provides sufficient structure to support the property of *convergence* (next definition) of a sequence. In a *metric space* (Definition 2.1 page 27), a convergent sequence converges to an *unique* limit (Theorem 8.4 page 132). However in a topological space, a convergent sequence may converge to more than one limit (Example 8.1 page 128).

Definition 8.3. [3] *Let* $(X, \boldsymbol{T})$ *be a* TOPOLOGICAL SPACE *(Definition 1.1 page 3)*.

[1] Bromwich (1908), page 1, Thomson et al. (2008) page 23 ⟨Definition 2.1⟩, Joshi (1997) page 31
[2] Haaser and Sullivan (1991) page 42 ⟨2·2 Proposition⟩
[3] Joshi (1983) page 83 ⟨(3.1) Definition⟩, "→" symbol: Leathem (1905) page 13 ⟨section III.11⟩

DEF

A sequence $(x_n)_{n\in\mathbb{Z}}$ ***converges*** *to a point* x *if for each* OPEN SET (Definition 1.1 page 3) U *of* x *there exists* $N \in \mathbb{N}$ *such that*

$x_n \in U$ *for all* $n > N$.

This condition can be expressed in any of the following forms:

1. *The* ***limit*** *of the sequence* (x_n) *is* x.
2. *The sequence* (x_n) *is* ***convergent*** *with limit* x.
3. $\lim_{n\to\infty} (x_n) = x$.
4. $(x_n) \to x$.

A sequence that converges is ***convergent****. A sequence that does not converge is said to* ***diverge****, or is* ***divergent****. An element* $x \in A$ *is a* ***limit point*** *of* A *if it is the limit of some* A*-valued sequence* $(x_n \in A)$.

Example 8.1. [4]

EX

Let $(X, \boldsymbol{T}_{31})$ be a *topological space* where $X \triangleq \{x, y, z\}$ and

$\boldsymbol{T}_{31} \triangleq \{\emptyset, \{x\}, \{x,y\}, \{x,z\}, \{x,y,z\}\}$.

In this space, the sequence $(x, x, x, \ldots)$ converges to x. But this sequence also converges to both y and z because x is in every *neighborhood* (Definition 1.9 page 25) of y and x is in every neighborhood of z. That is, the *limit* (Definition 8.3 page 127) of the sequence is *not* unique.

Example 8.2. In contrast to Example 8.1, note that the limit of the sequence $(x, x, x, \ldots)$ *is* unique in a *topological space* with sufficiently high resolution with respect to y and z such as the following:

EX

Define a *topological space* $(X, \boldsymbol{T}_{56})$ where $X \triangleq \{x, y, z\}$ and

$\boldsymbol{T}_{56} \triangleq \{\emptyset, \{y\}, \{z\}, \{x,y\}, \{y,z\}, \{x,y,z\}\}$.

In this space, the sequence $(x, x, x, \ldots)$ converges to x only. The sequence does *not* converge to y or z because there are *open sets* (Definition 1.1 page 3) containing y or z that do not contain x (the open sets $\{y\}$, $\{z\}$, and $\{y, z\}$).

Theorem 8.1 (The Closed Set Theorem). [5] *Let* $(X, \boldsymbol{T})$ *be a* TOPOLOGICAL SPACE. *Let* A *be a subset of* X $(A \subseteq X)$. *Let* A^- *be the* CLOSURE (Definition 1.4 page 14) *of* A *in* $(X, \boldsymbol{T})$.

THM

$$\underbrace{A \text{ is CLOSED in } (X,\boldsymbol{T})}_{(A = A^-)} \iff \left\{\begin{array}{l}\text{Every } A\text{-valued sequence } (x_n \in A)_{n\in\mathbb{Z}} \\ \text{that CONVERGES in } (X,\boldsymbol{T}) \\ \text{has its LIMIT in } A\end{array}\right\}$$

✎PROOF:

1. Proof that $\lim (x_n) \in A \implies A = A^-$:
   (a) Proof that $A \subseteq A^-$: by Lemma 1.1 page 15.
   (b) Proof that $A^- \subseteq A$:

$$\begin{aligned}
x \in A^- &\implies \text{each open set } U \text{ containing } x \text{ intersects } A && \text{by Lemma 1.2 page 22}\\
&\implies \exists (x_n \in A) \text{ that converges to } x && \text{by Definition 8.3 page 127}\\
&\implies x \in A && \text{by left hypothesis}
\end{aligned}$$

2. Proof that $A = A^- \implies \lim (x_n) \in A$:

$$\begin{aligned}
\lim (x_n) = x &\iff \left\{\begin{array}{l}\text{for each } open\ set\ U \text{ containing } x, \text{ there exists } N\\ \text{such that } x_n \in U, \forall n > N.\end{array}\right\} && \text{by Definition 8.3 page 127}\\
&\implies \text{each open set } U \text{ containing } x \text{ intersects } A && \text{because } x_n \text{ in } A\\
&\iff x \in A^- && \text{by Lemma 1.2 page 22}\\
&\iff x \in A && \text{by } A = A^- \text{ hypothesis}
\end{aligned}$$

[4] Munkres (2000) page 98 ⟨Hausdorff Spaces⟩

[5] Kubrusly (2001) page 118 ⟨Theorem 3.30⟩, Haaser and Sullivan (1991) page 75 ⟨6·9 Proposition⟩, Rosenlicht (1968) pages 47–48

1. Proof that $\lim (\!(x_n)\!) \in A \implies A = A^-$:
2. Proof that $\lim (\!(x_n)\!) \in A \impliedby A = A^-$ (proof by contradiction):

$$
\begin{aligned}
\lim (\!(x_n)\!) \notin A &\implies \lim (\!(x_n)\!) \notin A^- && \text{by } A = A^- \text{ hypothesis}\\
&\implies \lim (\!(x_n)\!) \in (A^-)^c && \text{by definition of set complement: Definition A.5 page 266}\\
&\implies \exists x_n \quad \text{such that} \quad x_n \in (A^-)^c && \text{because } (A^-)^c \text{ is } \textit{open} \text{ and by Definition 8.3 page 127}\\
&\implies \exists x_n \quad \text{such that} \quad x_n \notin A^- && \text{by definition of set complement: Definition A.5 page 266}\\
&\implies \exists x_n \quad \text{such that} \quad x_n \notin A && \text{by } A = A^- \text{ hypothesis}\\
&\implies \text{contradiction} && \text{by definition of } (\!(x_n \in A)\!)\\
&\implies \lim (\!(x_n)\!) \in A
\end{aligned}
$$

1. Proof that $x \in A \implies A$ is closed (proof by contradiction):
   (a) Suppose that A is not closed.
   (b) Then by Definition 1.1 (page 3), A^c is *not* open.
   (c) If A^c is not open, then there is nothing to prevent x to be located in A^c and yet be so close to A that it is *still* a limit point of $(\!(a_n)\!)$. (Note: It is the openness property of A^c that prevents this catastrophe from happening. Without it, as is the case under the "A is not closed supposition", a limit point can get too close to the border and pull in an infinite number of other points from the sequence to the wrong side of the border. That is, the openness property provides a protective buffer on the border that keeps points from being sucked across the border by the limit point.)
   (d) That is, choose an r such that $\mathsf{B}(x,r)\, A^c$.
   (e) Since x is a limit point of the sequence $(\!(a_n)\!)$ and by Theorem 8.2 (page 130),

$$\text{for some } 0 < \varepsilon < r,\ \exists N \quad \text{such that} \quad \mathsf{d}\big(a_n, x\big) < \varepsilon < r.$$

   (f) That is, now all the infinite number of points $(\!(a_n)\!)_{n>N}$ are inside the ball $\mathsf{B}(x,r)$, and thus *inside* A^c.
   (g) Thus, there are points in $(\!(a_n)\!)$ that are in A^c, and not in A where they are by definition supposed to be.
   (h) This is a contradiction, and therefore the original supposition is impossible and A must be closed.

⇨

8.3 Sequences in metric spaces

One of the most important applications of *metric spaces* (Definition 2.1 page 27) in analysis is the concept of *convergence*. Loosely speaking, a sequence that converges somehow implies that its elements are "getting closer and closer" to some value. In a metric space, there are two standard types of sequences often used to describe this:

① *convergent* sequence : The elements of the sequence approach a fixed value x (Theorem 8.2 page 130)

② *Cauchy* sequence : The elements of the sequence approach each other (Definition 8.4 page 130)

The *convergent* condition is "stronger" than the *Cauchy* condition in the sense that all convergent sequences are Cauchy but not all Cauchy sequences are convergent (Theorem 8.3 page 131). If however all the Cauchy sequences in a metric space are also convergent sequences (each sequence converges to a specific point) *and* each of those convergent points is *inside* the metric space, then that metric space is said to be *complete* (Definition 8.5 page 133).

8.3.1 Cauchy sequences

Definition 8.4. [6]

DEF

*A sequence $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ is a **Cauchy sequence** in the metric space (X, d) if*

$$\underbrace{\text{for every } \varepsilon \in \mathbb{R}^+,\ \exists N \in \mathbb{Z} \quad \text{such that} \quad \forall n, m > N,\ \mathsf{d}(x_n, x_m) < \varepsilon.}_{\text{CAUCHY CONDITION}}$$

Lemma 8.1. [7] *Let $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ be a sequence in the metric space (X, d).*

LEM

$$\left\{\begin{array}{l}(\!(x_n)\!) \text{ is CAUCHY}\\ \text{in } (X, \mathsf{d})\end{array}\right\} \implies \left\{\begin{array}{l}(\!(x_n)\!) \text{ is BOUNDED}\\ \text{in } (X, \mathsf{d})\end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
(\!(x_n)\!) \text{ is Cauchy} &\implies \text{for every } \varepsilon \in \mathbb{R}^+,\ \exists N \in \mathbb{N} \quad \text{such that} \quad \forall n, m > N,\ \mathsf{d}(x_n, x_m) < \varepsilon && \text{(Definition 8.4 page 130)}\\
&\implies \exists N \in \mathbb{Z} \quad \text{such that} \quad \forall n, m > N,\ \mathsf{d}(x_n, x_m) < 1 && \text{(arbitrarily choose } \varepsilon \triangleq 1)\\
&\implies \exists N \in \mathbb{Z} \quad \text{such that} \quad \forall n, m \in \mathbb{N},\ \mathsf{d}(x_n, x_{m+1}) < \max\left\{\{1\} \cup \left\{\mathsf{d}(x_p, x_q) \,|\, p, q \ngtr N\right\}\right\}\\
&\implies (\!(x_n)\!) \text{ is bounded} && \text{(by definition of } bounded)
\end{aligned}$$

✏

Proposition 8.1. [8] *Let $(\!(x_n)\!)_{n\in\mathbb{Z}}$ be a sequence in a metric space (X, d). Let $\mathsf{f} : \mathbb{Z} \to \mathbb{Z}$ be a strictly increasing function such that $\mathsf{f}(n) < \mathsf{f}(n+1)$.*

PRP

$$\underbrace{(\!(x_n)\!)_{n\in\mathbb{Z}} \text{ is CAUCHY}}_{\textit{Cauchy sequence}} \implies \underbrace{\left(\!\left(x_{\mathsf{f}(n)}\right)\!\right)_{n\in\mathbb{Z}} \text{ is CAUCHY}}_{\textit{subsequence is also Cauchy}}$$

✎PROOF:

$$\begin{aligned}
&(\!(x_n)\!)_{n\in\mathbb{Z}} \text{ is Cauchy}\\
&\implies \text{for any given } \varepsilon > 0,\ \exists N \quad \text{such that} \quad \forall n, m > N,\ \mathsf{d}(x_n, x_m) < \varepsilon && \text{by Definition 8.4 page 130}\\
&\implies \text{for any given } \varepsilon > 0,\ \exists N' \quad \text{such that} \quad \forall \mathsf{f}(n), \mathsf{f}(m) > N',\ \mathsf{d}(x_{\mathsf{f}(n)}, x_{\mathsf{f}(m)}) < \varepsilon\\
&\implies \left(\!\left(x_{\mathsf{f}(n)}\right)\!\right)_{n\in\mathbb{Z}} \text{ is Cauchy} && \text{by Definition 8.4 page 130}
\end{aligned}$$

✏

8.3.2 Convergence in Metric Space

Theorem 8.2. [9] *Let (X, T) be the TOPOLOGICAL SPACE induced by a metric space (X, d). Let $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ be a sequence in (X, d).*

THM

$$\underbrace{(\!(x_n)\!) \text{ converges to a limit } x}_{\text{(Definition 8.3 page 127)}} \iff \left\{\begin{array}{l}\text{for any } \varepsilon \in \mathbb{R}^+, \text{ there exists } N \in \mathbb{N}\\ \text{such that for all } n > N,\\ \qquad \mathsf{d}(x_n, x) < \varepsilon\end{array}\right\}$$

[6] Apostol (1975) page 73 ⟨4.7⟩, Rosenlicht (1968) page 51
[7] Giles (1987) page 49 ⟨Theorem 3.30⟩
[8] Rosenlicht (1968) page 52
[9] Rosenlicht (1968) page 45, Giles (1987) page 37 ⟨3.2 Definition⟩

✎PROOF:

$$\begin{aligned}
(\!(x_n)\!) \to x &\iff x_n \in U \quad \forall U \in N_x,\, n > N && \text{by Definition 8.3 page 127}\\
&\iff \exists \mathsf{B}(x,\varepsilon) \quad \text{such that} \quad x_n \in \mathsf{B}(x,\varepsilon)\, \forall n > N && \text{by Lemma 2.3 page 33}\\
&\iff \mathsf{d}(x_n, x) < \varepsilon && \text{by Definition 2.4 page 30}
\end{aligned}$$

⇨

A sequence that is *convergent* is always *Cauchy* (next theorem). However, in a metric space, the converse is not true—a sequence that is *convergent* is not in general *Cauchy*. This is in contrast to the special case of a real sequence in the metric space $(\mathbb{R}, |x-y|)$. In this case, all Cauchy sequences are convergent and the Cauchy property is referred to as the *Cauchy condition*.[10]

Theorem 8.3. [11] *Let $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ be a sequence in the metric space (X, d).*

THM: $\left\{ \begin{array}{l} (\!(x_n)\!) \text{ is CONVERGENT} \\ \text{in } (X,\mathsf{d}) \end{array} \right\} \implies \left\{ \begin{array}{l} (\!(x_n)\!) \text{ is CAUCHY} \\ \text{in } (X,\mathsf{d}) \end{array} \right\} \implies \left\{ \begin{array}{l} (\!(x_n)\!) \text{ is BOUNDED} \\ \text{in } (X,\mathsf{d}) \end{array} \right\}$

✎PROOF:

1. Proof that *convergent* $\implies$ *Cauchy*:

$$\begin{aligned}
\mathsf{d}(x_n, x_m) &\le \mathsf{d}(x_n, x) + \mathsf{d}(x, x_m) && \text{by Definition 2.1 page 27 (triangle inequality)}\\
&< \varepsilon + \varepsilon && \text{by left hypothesis}\\
&= 2\varepsilon
\end{aligned}$$

2. Proof that *Cauchy* $\implies$ *bounded*: by Lemma 8.1 (page 130).

⇨

Proposition 8.2. [12] *Let $(\!(x_n)\!)_{n\in\mathbb{Z}}$ be a sequence in a metric space (X, d). Let $\mathsf{f} : \mathbb{Z} \to \mathbb{Z}$ be a strictly increasing function such that $\mathsf{f}(n) < \mathsf{f}(n+1)$.*

PRP: $\left\{ \begin{array}{lll} 1. & (\!(x_n)\!)_{n\in\mathbb{Z}} \text{ is CAUCHY} & \text{and} \\ 2. & \left(\!\left(x_{\mathsf{f}(n)}\right)\!\right)_{n\in\mathbb{Z}} \text{ is CONVERGENT} & \end{array} \right\} \implies (\!(x_n)\!)_{n\in\mathbb{Z}} \text{ is CONVERGENT.}$

✎PROOF:

$$\begin{aligned}
\mathsf{d}(x_n, x) &= \mathsf{d}(x, x_n)\\
&\le \underbrace{\mathsf{d}(x, x_{\mathsf{f}(n)})}_{<\varepsilon \text{ by left hypothesis 2}} + \underbrace{\mathsf{d}(x_{\mathsf{f}(n)}, x_n)}_{<\varepsilon \text{ by left hypothesis 1}}\\
&< \varepsilon + \varepsilon\\
&= 2\varepsilon\\
&\implies (\!(x_n)\!) \text{ is convergent.}
\end{aligned}$$

⇨

[10] Whittaker (1915), pages 13–15 ⟨2.22⟩
[11] Giles (1987) page 49 ⟨Theorem 3.30⟩, Rosenlicht (1968) page 51, Apostol (1975) pages 72–73 ⟨Theorem 4.6⟩
[12] Rosenlicht (1968) page 52

Proposition 8.3. [13] *Let* (X, d) *be a* METRIC SPACE. *Let* $(\mathbb{R}, \mathsf{p})$ *be a metric space of real numbers with the usual metric* $\mathsf{p}(x,y) \triangleq |x-y|$.

PRP
$$\underbrace{(\!(x_n)\!) \to x \text{ and } (\!(y_n)\!) \to y}_{\text{convergence in } (X,\mathsf{d})} \implies \underbrace{(\!(\mathsf{d}(x_n, y_n))\!) \to \mathsf{d}(x,y)}_{\text{convergence in } (\mathbb{R},\mathsf{p})} \qquad \forall x,y,(\!(x_n)\!)_{n\in\mathbb{Z}},(\!(y_n)\!)_{n\in\mathbb{Z}} \in (X,\mathsf{d})$$

✎PROOF:

$$\begin{aligned}
\mathsf{p}\left(\mathsf{d}(x,y), \mathsf{d}(x_n,y_n)\right) &\triangleq \left|\mathsf{d}(x,y) - \mathsf{d}(x_n,y_n)\right| \\
&\le \left[\mathsf{d}(x,x_n) + \mathsf{d}(x_n,y)\right] - \mathsf{d}(x_n,y_n) && \text{by triangle inequality page 27} \\
&\le \mathsf{d}(x,x_n) + \left[\mathsf{d}(x_n,y_n) + \mathsf{d}(y_n,y)\right] - \mathsf{d}(x_n,y_n) && \text{by triangle inequality page 27} \\
&= \mathsf{d}(x,x_n) + \mathsf{d}(y,y_n) && \text{by definition of metric (page 27)} \\
&< \varepsilon + \varepsilon && \text{by left hypothesis} \\
&= 2\varepsilon \\
\implies \mathsf{d}(x_n,y_n) &\to \mathsf{d}(x,y)
\end{aligned}$$

✑

Theorem 8.4 (next) demonstrates that, in a *metric space* (Definition 2.1 page 27), if a sequence *converges* (Definition 8.3 page 127), then the limit it converges to is *unique*—a sequence cannot converge to more than one limit (in a metric space). This is in contrast to the more general topological spaces where a sequence *can* converge to more than one limit (Example 8.1 page 128).

Theorem 8.4 (Uniqueness of limit). [14] *Let* (X, d) *be a* METRIC SPACE. *Let* $x, y, \in X$ *and let* $(\!(x_n)\!)$ *be an* X*-valued sequence.*

THM
$$\underbrace{\left\{(\!(x_n)\!) \to x \text{ and } (\!(x_n)\!) \to y\right\} \implies \{x = y\}}_{\text{the LIMIT of a CONVERGENT sequence is UNIQUE}}$$

✎PROOF:

1. Proof that $\mathsf{d}(x,y) < 2\varepsilon$ for arbitrarily small $\varepsilon > 0$:

$$\begin{aligned}
(\!(x_n)\!) \to x \text{ and } (\!(x_n)\!) \to y &\implies \exists N \text{ such that } \forall n > N,\ \mathsf{d}(x,x_n) < \varepsilon \text{ and } \mathsf{d}(x_n,y) < \varepsilon \\
&\implies \exists N \text{ such that } \forall n > N,\ \mathsf{d}(x,y) = \mathsf{d}(x,x_n) + \mathsf{d}(y_n,y) < 2\varepsilon \\
&\implies \exists N \text{ such that } \forall n > N,\ \mathsf{d}(x,y) < 2\varepsilon \\
&\implies \mathsf{d}(x,y) < 2\varepsilon
\end{aligned}$$

2. Proof that $\mathsf{d}(x,y) = 0$:
   (a) If $\mathsf{d}(x,y) > 0$, then we could choose an arbitrarily small ε such that
   $$\mathsf{d}(x,y) > 2\varepsilon.$$
   (b) But this would contradict the earlier result of $\mathsf{d}(x,y) < 2\varepsilon$.
   (c) Therefore, $\mathsf{d}(x,y) = 0$ (proof by contradiction).

3. Therefore, $x = y$ because by the definition of metrics (Definition 2.1 page 27),
$$\mathsf{d}(x,y) = 0 \iff x = y.$$

[13] Berberian (1961) page 37 ⟨Theorem II.4.1⟩
[14] Rosenlicht (1968) page 46, Thomson et al. (2008) page 32 ⟨Theorem 2.8⟩

Proposition 8.4. [15] *Let* $(\!(x_n)\!)_{n\in\mathbb{Z}}$ *be a sequence in a metric space* (X, d). *Let* $\mathsf{f} : \mathbb{Z} \to \mathbb{Z}$ *be a strictly increasing function such that* $\mathsf{f}(n) < \mathsf{f}(n+1)$.

PRP
$$\underbrace{(\!(x_n)\!)_{n\in\mathbb{Z}} \to x}_{\text{sequence converges to limit } x} \implies \underbrace{\left(\!\left(x_{\mathsf{f}(n)}\right)\!\right)_{n\in\mathbb{Z}} \to x}_{\text{subsequence converges to the same limit } x}$$

✎PROOF:

$$\begin{aligned}
(\!(x_n)\!)_{n\in\mathbb{Z}} \to x &\implies \forall\varepsilon>0,\ \exists N \quad \text{such that} \quad \forall n > N,\ \mathsf{d}(x_n, x) < \varepsilon && \text{by Theorem 8.2 page 130}\\
&\implies \forall\varepsilon>0,\ \exists\mathsf{f}(N) \quad \text{such that} \quad \forall\mathsf{f}(n) > \mathsf{f}(N),\ \mathsf{d}\left(x_{\mathsf{f}(n)}, x\right) < \varepsilon\\
&\implies \left(\!\left(x_{\mathsf{f}(n)}\right)\!\right)_{n\in\mathbb{Z}} \to x && \text{by Theorem 8.2 page 130}
\end{aligned}$$

8.3.3 Complete metric spaces

Definition 8.5. [16] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27)*.

DEF
A sequence $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ *is* ***complete*** *in* (X, d) *if*
$$\underbrace{(\!(x_n)\!) \text{ is CAUCHY in } (X,\mathsf{d}) \implies (\!(x_n)\!) \text{ is convergent in } (X,\mathsf{d}).}_{\text{every CAUCHY SEQUENCE in } (X,\mathsf{d}) \text{ CONVERGES to a limit in } (X,\mathsf{d})}$$

Theorem 8.5. [17] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27). Let* A *be a subset of* X. *Let* A^- *be the* CLOSURE *(Definition 1.4 page 14) of* A *in* (X, d).

THM
$$\{(A, \mathsf{d}) \text{ is COMPLETE}\} \implies \underbrace{A \text{ is CLOSED in } (X, \mathsf{d})}_{A = A^-}$$
$$\left\{\begin{array}{lll} 1. & (X, \mathsf{d}) \text{ is COMPLETE (Definition 8.5 page 133)} & \text{and} \\ 2. & A \text{ is CLOSED in } (X, \mathsf{d}) & (A = A^-) \end{array}\right\} \implies \{(A, \mathsf{d}) \text{ is COMPLETE}\}$$

✎PROOF:

1. Proof that *complete* $\implies$ *closed*:
   (a) Proof that $A \subseteq A^-$: Lemma 1.1 page 15
   (b) Proof that $A^- \subseteq A$ (proof that $x \in A^- \implies x \in A$):
      i. Let x be a point in A^- ($x \in A^-$).
      ii. Define a sequence of open balls $\left(\!\left(\mathsf{B}\left(x, \frac{1}{1}\right), \mathsf{B}\left(x, \frac{1}{2}\right), \mathsf{B}\left(x, \frac{1}{3}\right), \ldots\right)\!\right)$.
      iii. Define a sequence of points $(\!(x_1, x_2, x_3, \ldots)\!)$ such that $x_n \in \mathsf{B}\left(x_n, \frac{1}{n}\right) \cap A$.
      iv. Then $(\!(x_n)\!)$ is *convergent* in X with limit x by Definition 8.3 page 127
      v. and $(\!(x_n)\!)$ is *Cauchy* in A by Definition 8.4 page 130.

[15] Rosenlicht (1968) page 46
[16] Rosenlicht (1968) page 52
[17] Kubrusly (2001) page 128 ⟨Theorem 3.40⟩, Haaser and Sullivan (1991) page 75 ⟨6·10, 6·11 Propositions⟩, Bryant (1985) page 40 ⟨Theorem 3.6, 3.7⟩, Sutherland (1975) pages 123–124

vi. By the left hypothesis ((A, d) is *complete*), $(\!(x_n)\!)$ is therefore also *convergent* in A. Let this limit be y. Note that $y \in A$.
vii. By Theorem 8.4 page 132, limits are *unique*, so $y = x$.
viii. Because $y \in A$ (item (1(b)vi)) and $y = x$ (item (1(b)vii)), so $x \in A$.
ix. Therefore, $x \in A^- \implies x \in A$ and $A^- \subseteq A$.

2. Proof that *complete* and *closed* $\implies$ *complete*:

   (a) By left hypothesis 2, A is closed in (X, d).
   (b) By Theorem 8.1 (page 128) and because A is closed in (X, d), sequences converge in A.
   (c) Therefore by Definition 8.5 (page 133), (A, d) is complete.

⇰

Corollary 8.1. [18] *Let* (X, d) *be a* METRIC SPACE *(Definition 2.1 page 27). Let A be a subset of X. Let A^- be the* CLOSURE *(Definition 1.4 page 14) of A in (X, d).*

COR

$$\{(A, \mathsf{d}) \text{ is COMPLETE}\} \iff \underbrace{A \text{ is CLOSED in } (X, \mathsf{d})}_{A = A^-}$$

✎PROOF: Note that in this corollary, the metric space (X, d) is assumed to be *complete*.

1. Proof that *complete* $\implies$ *closed*: by Theorem 8.5 (1).
2. Proof that *complete* $\impliedby$ *closed*: by *complete* hypothesis and Theorem 8.5 (2).

⇰

Example 8.3. Let $\mathbb{Q}$ be the set of *rational numbers*.

EX

The metric space $(\mathbb{Q}, \mathsf{d}(x, y) = |x - y|)$ is *not complete*.

✎PROOF: Let $(\!(x_n)\!)_{n\in\mathbb{W}}$ be the sequence of values approximating π truncated to n decimal points:

$$(\!(x_n)\!)_{n\in\mathbb{W}} \triangleq (\!(3,\ 3.1,\ 3.14,\ 3.141,\ 3.1415,\ 3.14159,\ 3.141592,\ \ldots)\!)$$

This is a Cauchy sequence. However, this sequence (and all sequences converging to an irrational number) does not converge to a rational number ($\mathbb{Q}$) and thus is not in the metric space $(\mathbb{Q}, \mathsf{d})$ and thus $(\mathbb{Q}, \mathsf{d})$ is *not complete*. ⇰

Example 8.4 (Cauchy's convergence criterion/Cauchy's criterion). [19] Let $(\!(r_n \in \mathbb{R})\!)_{n\in\mathbb{Z}}$ be a **real** sequence.

EX

The metric space $\left((\!(r_n)\!), \left|r_n - r_m\right|\right)$ is *complete*.

Theorem 8.6 (Cantor intersection theorem). [20] *Let (X, d) be a complete* METRIC SPACE, *$(\!(A_n)\!)_{n\in\mathbb{Z}}$ a sequence with each $A_n \in 2^X$, and $|A|$ the number of elements in A.*

THM

$$\left.\begin{array}{llll} 1. & (X, \mathsf{d}) \text{ is COMPLETE} & & \text{and} \\ 2. & A_n \text{ is CLOSED} & \forall n\in\mathbb{N} & \text{and} \\ 3. & \operatorname{diam} A_{n+1} \le \operatorname{diam} A_n & \forall n\in\mathbb{N} & \text{and} \\ 4. & \operatorname{diam} A_n \to 0 & & \end{array}\right\} \implies \left|\bigcap_{n\in\mathbb{N}} A_n\right| = 1$$

[18] Kubrusly (2001) page 128 ⟨Corollary 3.41⟩
[19] Sohrab (2003) page 54 ⟨Theorem 2.2.5⟩
[20] Davis (2005), page 28, Hausdorff (1937), page 150

✎PROOF:

1. Proof that $\left|\bigcap_{n\in\mathbb{Z}} A_n\right| < 2$:
   (a) Let $A \triangleq \cap A_n$.
   (b) $x \neq y \wedge \{x,y\} \in A \implies \mathsf{d}(x,y) > 0 \wedge \{x,y\} \subseteq A_n \forall n$
   (c) $\exists n$ such that diam $A_n < \mathsf{d}(x,y)$ by left hypothesis
   (d) $\implies \exists n$ such that $\sup\left\{\mathsf{d}(x,y) \,|\, x,y \in A_n\right\} < \mathsf{d}(x,y)$
   (e) This is a contradiction, so $\{x,y\} \notin A$ and $\left|\bigcap A_n\right| < 2$.
2. Proof that $|\cap A_n| \geq 1$:
   (a) Let $x_n \in A_n$ and $x_m \in A_m$
   (b) $\forall \varepsilon,\ \exists N \in \mathbb{N}$ such that $A_N < \varepsilon$
   (c) $\forall m,n > N,\ x_n \in A_n \subseteq A_N$ and $x_m \in A_m \subseteq A_N$
   (d) $\mathsf{d}\left(x_n, x_m\right) \leq$ diam $A_N < \varepsilon \implies \{x_n\}$ is a Cauchy sequence
   (e) Because $\{x_n\}$ is complete, $x_n \to x$.
   (f) $\implies x \in cls A_n = A_n$
   (g) $\implies |A_n| \geq 1$

✏

8.4 Sequences on normed linear spaces

8.4.1 Convergence in normed linear spaces

Theorem 8.2 (page 130) defines convergence in a general *metric space* (Definition 2.1 page 27). All *normed linear spaces* (Definition 5.1 page 83) are metric spaces, so they inherit this definition with the *metric induced by the norm* (Definition 5.2 page 86).

Definition 8.6. [21] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *be a normed linear space . Let the metric* d *be defined as* $\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) \triangleq \|\boldsymbol{x}-\boldsymbol{y}\|$.

DEF: *A sequence* $(\boldsymbol{x}_n \in X)_{n\in\mathbb{Z}}$ ***converges in norm*** *or* ***converges strongly*** *to the* ***limit*** $\boldsymbol{x}$ *if* $(\boldsymbol{x}_n)_{n\in\mathbb{Z}}$ *converges to the limit* $\boldsymbol{x}$ *in the metric space* (X,d). *That is, a sequence* $(x_n)_{n\in\mathbb{Z}}$ ***converges strongly*** *in the normed linear space* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *to the* ***limit*** $\boldsymbol{x} \in X$ *if for any* $\varepsilon \in \mathbb{R}^+$ *there exists* $N \in \mathbb{Z}$ *such that*

$$\left\|\boldsymbol{x}_n - \boldsymbol{x}\right\| < \varepsilon \quad \forall n > N.$$

This mode of convergence is called ***strong convergence***.

Definition 8.7. [22] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *be a normed linear space . Let the metric* d *be defined as* $\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) \triangleq \|\boldsymbol{x}-\boldsymbol{y}\|$.

DEF: *A sequence* $(\boldsymbol{x}_n \in X)_{n\in\mathbb{Z}}$ ***converges weakly*** *to the* ***limit*** $\boldsymbol{x}$ *if for every functional* $\mathsf{f} \in \mathbb{F}^X$, $(\mathsf{f}(\boldsymbol{x}_n))_{n\in\mathbb{Z}}$ *converges to the limit* $\mathsf{f}(\boldsymbol{x})$ *in the metric space* (X,d). *That is, a sequence* $(x_n)_{n\in\mathbb{Z}}$ ***converges weakly*** *in the normed linear space* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *to the* ***limit*** $\mathsf{f}(\boldsymbol{x})$ *if for every functional* $\mathsf{f} \in \mathbb{F}^X$ *and for any* $\varepsilon \in \mathbb{R}^+$ *there exists* $N \in \mathbb{Z}$ *such that*

$$\left\|\mathsf{f}(\boldsymbol{x}_n) - \mathsf{f}(\boldsymbol{x})\right\| < \varepsilon \quad \forall n > N.$$

This mode of convergence is called ***weak convergence***.

[21] Bachman and Narici (1966) page 247, Katznelson (2004) page 67 ⟨section 1.1⟩
[22] Bachman and Narici (1966) page 231 ⟨Definition 14.1⟩

8.4.2 Bounded sequences

8.4.3 Complete normed linear spaces

❝*At that time, however, the theory seemed to me to contain for the immediate future nothing but some decades of rather formal and thin work. By this I do not mean to reproach the work of Banach himself but that of the many inferior writers, hungry for easy doctors' theses, who were drawn to it. As I foresaw, it was this class of writers that was first attracted to the theory of Banach spaces.*❞

Norbert Wiener (1894–1964), American mathematician [23]

Definition 8.8. [24] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a normed linear space.*

DEF: *The space* NORMED LINEAR SPACE $\boldsymbol{\Omega}$ *is a* ***Banach space*** *if it is* COMPLETE *with respect to the metric* $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \|\boldsymbol{y} - \boldsymbol{x}\|$.

8.4.4 The l_p spaces

Definition 8.9. [25] *Let* $(\!(x_n \in \mathbb{R})\!)_{n\in\mathbb{Z}}$ *be a real sequence.*

DEF: *The space* $\boldsymbol{\ell}^{\mathsf{p}}_{\mathbb{F}}$ *and space* $\boldsymbol{\ell}^{\infty}_{\mathbb{F}}$ *are defined as*

$$\begin{aligned} \boldsymbol{\ell}^{\mathsf{p}}_{\mathbb{F}} &\triangleq \left\{ (\!(x_n)\!)_{n\in\mathbb{Z}} \,\middle|\, \sum_{n\in\mathbb{Z}} |x_n|^p < \infty \right\} && \forall 1\le p<\infty \\ \boldsymbol{\ell}^{\infty}_{\mathbb{F}} &\triangleq \left\{ (\!(x_n)\!)_{n\in\mathbb{Z}} \,\middle|\, \sup (\!(x_n)\!) < \infty \right\} \end{aligned}$$

Lemma 8.2. *Let* $(\!(x_n \in \mathbb{F})\!)_{n\in\mathbb{Z}}$ *and* $(\!(x_n \in \mathbb{F})\!)_{n\in\mathbb{Z}}$ *be sequences over a field* $\mathbb{F}$.

LEM: $$(\!(x_n)\!), (\!(y_n)\!) \in \boldsymbol{\ell}^{\mathsf{p}}_{\mathbb{F}} \quad\Longrightarrow\quad \left((\!(x_n)\!) + (\!(y_n)\!)\right) \in \boldsymbol{\ell}^{\mathsf{p}}_{\mathbb{F}} \qquad \forall p \in [1 : \infty]$$

✎PROOF:

1. Proof for $p = 1$:

$$\begin{aligned} \sum_{n\in\mathbb{Z}} |x_n + y_n|^1 &= \sum_{n\in\mathbb{Z}} |x_n + y_n| \\ &\le \sum_{n\in\mathbb{Z}} \left(|x_n| + |y_n|\right) && \text{by norm properties of } \|\cdot\| \\ &= \sum_{n\in\mathbb{Z}} |x_n|^1 + \sum_{n\in\mathbb{Z}} |y_n|^1 \\ &< \infty && \text{because } \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\ell}^1_{\mathbb{F}} \end{aligned}$$

[23] quote: Wiener (1956), pages 63–64, Werner (2004), page 41
image: http://www-history.mcs.st-andrews.ac.uk/PictDisplay/Wiener_Norbert.html

[24] Bachman and Narici (1966) page 112 ⟨Definition 8.1⟩, Banach (1932a), page 53 ⟨"espace du type (B)" (space type (B))⟩, Banach (1932b), page 33

[25] Carothers (2000), page 44

2. Proof for $1 < p < \infty$: Let $\|\boldsymbol{x}\|_p \triangleq \left(\sum_{n\in\mathbb{Z}} |x_n|^p\right)^{\frac{1}{p}}$

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} |x_n + y_n|^p &= \Bigg(\underbrace{\left(\sum_{n\in\mathbb{Z}} |x_n + y_n|^p\right)^{\frac{1}{p}}}_{\|\cdot\|_p}\Bigg)^p \\
&= \|\boldsymbol{x}+\boldsymbol{y}\|_p^p && \text{by definition of } \|\cdot\|_p \text{ page 137} \\
&\le \left(\|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p\right)^p && \text{by Minkowski's inequality page 190} \\
&\le \infty && \text{by } \boldsymbol{x},\boldsymbol{y} \in \ell^p_{\mathbb{F}} \text{ hypothesis}
\end{aligned}$$

3. Proof for $p = \infty$:

$$\begin{aligned}
\sup\left\{x_n + y_n | n \in \mathbb{Z}\right\} &\le \sup\left(\!\left(x_n\right)\!\right) + \sup\left(\!\left(y_n\right)\!\right) \\
&\le \infty && \text{by } \boldsymbol{x},\boldsymbol{y} \in \ell^p_{\mathbb{F}} \text{ hypothesis}
\end{aligned}$$

✏

Definition 8.10. *Let $(\!(x_n)\!)_{n\in\mathbb{Z}}$ be a sequence in the space $\ell^p_{\mathbb{F}}$.*

DEF: *The $\ell^p_{\mathbb{F}}$ **norm** $\left\|(\!(x_n)\!)\right\|_p$ of $(\!(x_n)\!)$ is defined as* $\left\|(\!(x_n)\!)\right\|_p \triangleq \left(\sum_{n\in\mathbb{Z}} |x_n|^p\right)^{\frac{1}{p}}$ *for $p \in [1 : \infty]$*

Proposition 8.5. *Let $\left\|(\!(x_n)\!)\right\|_p$ be the $\ell^p_{\mathbb{F}}$ norm of a sequence $(\!(x_n)\!)_{n\in\mathbb{Z}}$ in the space $\ell^p_{\mathbb{F}}$.*

PRP: *$\left\|(\!(x_n)\!)\right\|_p$ is a norm.*

✎PROOF:

Proof that $\|\cdot\|_p \ge 0$:

$$\begin{aligned}
\|\boldsymbol{x}\| &\triangleq \left(\sum_{n\in\mathbb{Z}} |x_n|\right)^{\frac{1}{p}} && \text{by definition of } \|\cdot\| \\
&\ge \sum_{i=1}^{n} 0 \\
&= 0
\end{aligned}$$

Proof that $\|\boldsymbol{x}\| = 0 \implies \boldsymbol{x} = \mathbb{0}$:

$$\begin{aligned}
0 &= \|\boldsymbol{x}\| && \text{by left hypothesis} \\
&= \sum_{i=1}^{n} |x_i| && \text{by definition of } \|\cdot\| \\
&\implies x_i = 0 \quad i = 1, 2, \ldots, n \\
&\implies \boldsymbol{x} = \mathbb{0} && \text{by definition of } \boldsymbol{x}
\end{aligned}$$

Proof that $\|\boldsymbol{x}\| = 0 \impliedby \boldsymbol{x} = \mathbb{0}$:

$$\begin{aligned}
\|\boldsymbol{x}\| &= \sum_{i=1}^{n} |x_i| && \text{by definition of } \|\cdot\| \\
&= \sum_{i=1}^{n} |0| && \text{by right hypothesis} \\
&= 0
\end{aligned}$$

Proof that $\|\alpha x\| = |\alpha|\,\|x\|$:

$$\begin{aligned}
\|\alpha x\| &= \sum_{i=1}^{n} |\alpha x_i| && \text{by definition of } \|\cdot\| \\
&= \sum_{i=1}^{n} |\alpha||x_i| \\
&= |\alpha| \sum_{i=1}^{n} |x_i| \\
&= |\alpha|\,\|x\| && \text{by definition of } x
\end{aligned}$$

Proof that $\|x+y\| \le \|x\| + \|y\|$: by *Minkowski's Inequality* (Theorem 11.5 page 190)

8.5 Complete inner-product spaces

Definition 8.11. [26] *Let* $\Omega \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be an inner-product space.*[27]

DEF: *The inner-product space* Ω *is a **Hilbert space** if it is* COMPLETE *with respect to the metric* $\mathsf{d}(x, y) \triangleq \|x - y\| \triangleq \sqrt{\langle x - y \mid x - y \rangle}$.

Theorem 8.7 (Complemented-subspace theorem). [28] *Let* $B \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a Banach space.*

THM: $$\left\{ \begin{array}{l} \text{Every closed linear subspace } D \text{ in } B \\ \text{has a complement } D^c \text{ in } B \end{array} \right\} \implies \left\{ \begin{array}{l} B \text{ is isomorphic} \\ \text{to a Hilbert space } H \end{array} \right\}$$

8.6 Sequences of functions

For a sequence of real numbers in a metric space, the concept of convergence is well defined and unambiguous (Theorem 8.2 page 130). But for sequences of functions $(\!(\mathsf{f}_n(x))\!)_{n\in\mathbb{Z}}$, on the other hand, there are several different types or "modes" of convergence. Two of the most common modes are *pointwise convergence* (Definition 8.12 page 138) and *uniform convergence* (Definition 8.13 page 139). Both of these are defined in a metric space. In both of these, the value N beyond which the sequence becomes sufficiently "close" to the limit $\mathsf{f}(x)$ depends on a distance parameter ε. The difference between the two modes is that in pointwise convergence, the value N also depends on the value x of the limit $\mathsf{f}(x)$; whereas in uniform convergence, the value N does not depend on x.

Definition 8.12. [29] *Let* $(\!(\mathsf{f}_n(x))\!)_{n\in\mathbb{Z}}$ *be a sequence of functions in a* METRIC SPACE (X, d).

DEF: *The sequence* $(\!(\mathsf{f}_n(x))\!)$ ***converges pointwise** to a **limit*** $\mathsf{f}(x)$ *if for each* $\varepsilon \in \mathbb{R}^+$ *and for each* $x \in X$ *there exists an* $N \in \mathbb{N}$ *(dependent on* x*) such that*
$\mathsf{d}(\mathsf{f}_n(x), \mathsf{f}(x)) < \varepsilon.$

[26] von Neumann (1929) page 55, "Den abstrakten hilbertschen raum nennen wir $\mathcal{H}$" ("we call the abstract Hilbert space $\mathcal{H}$"), Aliprantis and Burkinshaw (1998), page 288

[27] *complete*: Definition 8.5 page 133

[28] Lindenstrauss and Tzafriri (1971), Day (1973), page 157

[29] Tao (2011) pages 94–96 ⟨section 1.5⟩, Thomson et al. (2008) page 368 ⟨Definition 9.3⟩, Tao (2010) page 117 ⟨Example 1.9.3⟩

Definition 8.13. [30] *Let* $(\!(\mathsf{f}_n(x))\!)_{n\in\mathbb{Z}}$ *be a sequence of functions in a* METRIC SPACE (X, d).

DEF

The sequence $(\!(\mathsf{f}_n(x))\!)$ ***converges uniformly*** *to a* ***limit*** $\mathsf{f}(x)$ *if for each* $\varepsilon \in \mathbb{R}^+$ *there exists an* $N \in \mathbb{N}$ *(independent of* x*) such that*

$$\mathsf{d}\big(\mathsf{f}_n(x), \mathsf{f}(x)\big) < \varepsilon \qquad \text{for all } x \in X.$$

Theorem 8.8. *Let* $(\!(\mathsf{f}_n(x))\!)_{n\in\mathbb{Z}}$ *be a sequence of functions in a* METRIC SPACE (X, d).

THM

$(\!(\mathsf{f}_n(x))\!)$ CONVERGES UNIFORMLY $\implies$ $(\!(\mathsf{f}_n(x))\!)$ CONVERGES POINTWISE

✎PROOF: This follows directly from the definition of *uniform convergence* (Definition 8.13 page 139) and the definition of *pointwise convergence* (Definition 8.12 page 138). ✐

[30] Tao (2011) pages 94–96 ⟨section 1.5⟩, Thomson et al. (2008) pages 373–374, Tao (2010) page 117 ⟨Example 1.9.4⟩

CHAPTER 9

INTERVALS AND CONVEXITY

9.1 Intervals

In the real number system, for $a \le b$, the *interval* $[a : b]$ is the set a and b and all the numbers inbetween, as in $[a : b] \triangleq \{x \in \mathbb{R} | a \le x \le b\}$. This concept can be easily generalized:

- In an **ordered set** (Definition B.2 page 290), if two elements x and y are *comparable* and $x \le y$, then we say that x and y and all the elements inbetween, as determined by the ordering relation $\le$, are the interval $[a : b]$ (Definition 9.1 page 141).
- In a **lattice** (Definition C.3 page 305), the concept of the *interval* can be generalized even further. In an arbitrary ordered set, the interval $[x : y]$ of item (9.1) is restricted to the case in which x and y are *comparable* (Definition B.2 page 290). This restriction can be lifted (Definition 9.2 page 141) with the additional structure of upper and lower bounds provided by lattices.
- A **metric space** (Definition 2.1 page 27) in general has no *order relation* $\le$ (Definition B.2 page 290). But intervals can still be defined (Definition 9.4 page 142) in a metric space in terms of the *triangle inequality*.
- A **linear space** (Definition 3.1 page 67) over a real or complex field in general has no *order relation* that compares *vectors* in the space, but the standard order relation $\le$ for real numbers $\mathbb{R}$ can still be used (Definition 9.5 page 142) to define an interval in a linear space.

Definition 9.1 (intervals on ordered sets). [1] *Let* $(X, \le)$ *be an* ORDERED SET *(Definition B.2 page 290).*

DEF					
	The set $[x : y]$	$\triangleq$	$\{z \in X \mid x \le z \le y\}$	*is called a **closed interval***	*and*
	The set $(x : y]$	$\triangleq$	$\{z \in X \mid x < z \le y\}$	*is called a **half-open interval***	*and*
	The set $[x : y)$	$\triangleq$	$\{z \in X \mid x \le z < y\}$	*is called a **half-open interval***	*and*
	The set $(x : y)$	$\triangleq$	$\{z \in X \mid x < z < y\}$	*is called an **open interval**.*	

Definition 9.2 (intervals on lattices). [2] *Let* $(X, \vee, \wedge; \le)$ *be a* LATTICE *(Definition C.3 page 305).*

DEF				
	The set $[x : y]$	$\triangleq$	$\{z \in X \mid x \wedge y \le z \le x \vee y\}$	*is called a **closed interval**.*
	The set $(x : y]$	$\triangleq$	$\{z \in X \mid x \wedge y < z \le x \vee y\}$	*is called a **half-open interval**.*
	The set $[x : y)$	$\triangleq$	$\{z \in X \mid x \wedge y \le z < x \vee y\}$	*is called a **half-open interval**.*
	The set $(x : y)$	$\triangleq$	$\{z \in X \mid x \wedge y < z < x \vee y\}$	*is called an **open interval**.*

[1] Apostol (1975) page 4, Ore (1935) page 409
[2] Duthie (1942) page 2, Ore (1935) page 425 ⟨*quotient structures*⟩

When x and y *are* comparable and $x \le y$, then Definition 9.2 (previous) simplifies to item (9.1) (page 141).

Definition 9.3. [3] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge;\, \le)$ *be a* LATTICE *with dual* $\mathbf{L}^*$. *Let* $[x : y]$ *be a* CLOSED INTERVAL (Definition 9.2 page 141) *on set* X. *The sublattices* $\mathbf{L}[x : y]$ *and* $\mathbf{L}^*[x : y]$ *are defined as follows:*

$$\begin{array}{lcll} \mathbf{L}[x : y] & \triangleq & \{z \in \mathbf{L} \mid z \in [x : y]\} & \forall x,y \in X \\ \mathbf{L}^*[x : y] & \triangleq & \left\{z \in \mathbf{L}^* \mid z \in [x : y]\right\} & \forall x,y \in X \end{array}$$

Definition 9.4. [4]

In a METRIC SPACE (X, d) (Definition 2.1 page 27),
the set $[a : b]$ *is the* ***closed interval*** *from* x *to* y *and is defined as*

$$[x : y] \triangleq \{z \in X \mid \mathrm{d}(x,z) + \mathrm{d}(z,y) = \mathrm{d}(x,y)\}.$$

An element $z \in X$ *is* ***geodesically between*** x *and* y *if* $z \in [x : y]$.

Definition 9.5. [5]

In a LINEAR SPACE $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ (Definition 3.1 page 67),

$$\begin{array}{lcll} [\boldsymbol{x} : \boldsymbol{y}] & \triangleq & \{\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} = \boldsymbol{z} \mid 0 \le \lambda \le 1\} & \text{is called a \textbf{closed interval} and} \\ (\boldsymbol{x} : \boldsymbol{y}] & \triangleq & \{\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} = \boldsymbol{z} \mid 0 < \lambda \le 1\} & \text{is called a \textbf{half-open interval} and} \\ [\boldsymbol{x} : \boldsymbol{y}) & \triangleq & \{\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} = \boldsymbol{z} \mid 0 \le \lambda < 1\} & \text{is called a \textbf{half-open interval} and} \\ (\boldsymbol{x} : \boldsymbol{y}) & \triangleq & \{\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} = \boldsymbol{z} \mid 0 < \lambda < 1\} & \text{is called an \textbf{open interval}.} \end{array}$$

9.2 Convex sets

Using the concept of the *interval* (previous section), we can define the *convex set* (next definition).

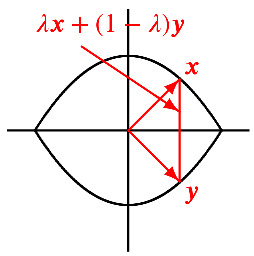
Definition 9.6. [6] *Let* X *be a* SET *in an* ORDERED SET $(X, \le)$, *a* LATTICE $(X, \vee, \wedge;\, \le)$, *a* METRIC SPACE (X, d), *or a* LINEAR SPACE $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$.

A subset $D \subseteq X$ *is a* ***convex set*** *in* X *if*

$$x, y \in D \implies [x : y] \subseteq D.$$

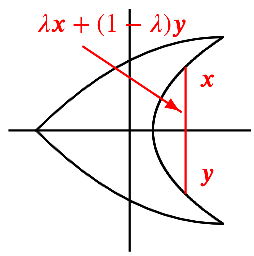
A set that is ***not*** *convex is* ***concave***.

Example 9.1. Consider the *Euclidean space* $\mathbb{R}^2$ (a special case of a *linear space*).



$\Leftarrow$ {The figure to the left is a *convex* set in $\mathbb{R}^2$.

The figure to the right is a *concave* set in $\mathbb{R}^2$.} $\Rightarrow$



Example 9.2. In a *metric space* (Definition 2.1 page 27), examples of *convex sets* are convex *balls*. Examples include those balls generated by the following metrics:

Taxi-cab metric	Example 2.21	page 56
Euclidean metric	Example 2.22	page 57
Sup metric	Example 2.23	page 57
Tangential metric	Example 2.27	page 62

[3] Maeda and Maeda (1970), page 1
[4] van de Vel (1993) page 8
[5] Barvinok (2002) page 2
[6] Barvinok (2002) page 5

Examples of metrics generating balls which are *not* convex include the following:

- Parabolic metric Example 2.24 page 58
- Exponential metric Example 2.26 page 60

9.3 Convex functions

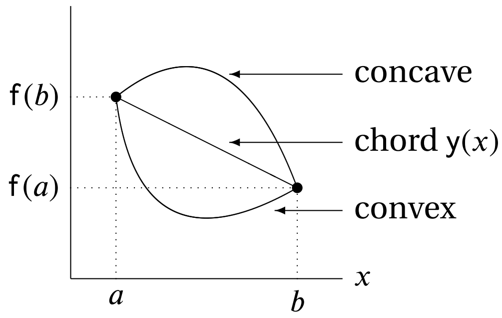


Figure 9.1: Convex and concave functions

Definition 9.7. [7] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE (Definition 3.1 page 67) *and* D *a* CONVEX SET (Definition 9.6 page 142) *in* X.

DEF

A function $\mathsf{f} \in F^D$ *is **convex** if*

$$\mathsf{f}\big(\lambda\boldsymbol{x} + [1-\lambda]\boldsymbol{y}\big) \;\le\; \lambda\mathsf{f}(\boldsymbol{x}) + (1-\lambda)\,\mathsf{f}(\boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in D \text{ and } \forall \lambda \in (0,1)$$

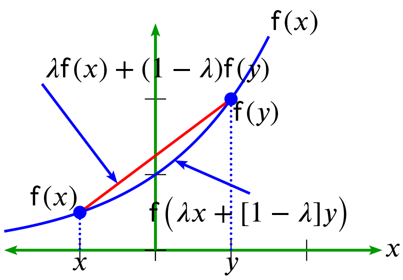
A function $\mathsf{g} \in F^D$ *is **strictly convex** if*

$$\mathsf{g}\big(\lambda\boldsymbol{x} + [1-\lambda]\boldsymbol{y}\big) \;=\; \lambda\mathsf{g}(\boldsymbol{x}) + (1-\lambda)\,\mathsf{g}(\boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in D,\; \boldsymbol{x} \neq \boldsymbol{y},\; \text{and } \forall \lambda \in (0,1)$$

A function $\mathsf{f} \in F^D$ *is **concave** if* $-\mathsf{f}$ *is* CONVEX.

A function $\mathsf{f} \in F^D$ *is **affine** if* f *is* CONVEX *and* CONCAVE.

Example 9.3. The function $\mathsf{f}(x) = 2^x$ is a *convex function* (Definition 9.7 page 143), as illustrated to the right.



Definition 9.8. [8] *Let* $(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE (Definition 3.1 page 67).

DEF

*The **epigraph*** $\operatorname{epi}(\mathsf{f})$ *and **hypograph*** $\operatorname{hyp}(\mathsf{f})$ *of a functional* $\mathsf{f} \in \mathbb{R}^X$ *are defined as*

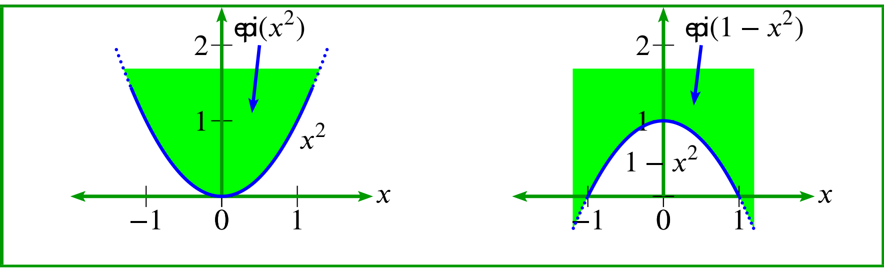
$$\begin{aligned} \operatorname{epi}(\mathsf{f}) &\triangleq \{(x,y) \in X \times \mathbb{R} \,|\, y \ge \mathsf{f}(x)\} \\ \operatorname{hyp}(\mathsf{f}) &\triangleq \{(x,y) \in X \times \mathbb{R} \,|\, y \le \mathsf{f}(x)\} \end{aligned}$$

Example 9.4.

[7] Simon (2011) page 2, Barvinok (2002) page 2, Bollobás (1999), page 3, Jensen (1906), page 176, Clarkson (1936) ⟨strictly convex⟩

[8] Beer (1993) page 13 ⟨§1.3⟩, Aubin and Frankowska (2009) page 222, Aubin (2011) page 223

EX



Proposition 9.1. [9] *Let* $(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE *(Definition 3.1 page 67). Let* f *be a* FUNCTIONAL *in* $\mathbb{R}^X$.

PRP

$$\left\{\begin{array}{l}\mathsf{f} \textit{ is a}\\ \text{CONVEX FUNCTION}\end{array}\right\} \iff \left\{\begin{array}{l}\mathrm{epi}(\mathsf{f}) \textit{ is a}\\ \text{CONVEX SET}\end{array}\right\}$$

Often a function can be proven to be *convex* or *concave*. *Convex* and *concave* functions are defined in Definition 9.9 (page 144) (next) and illustrated in Figure 9.1 (page 143).

Definition 9.9. *Let*

$$\mathsf{y}(x) = \frac{\mathsf{f}(b) - \mathsf{f}(a)}{b-a}(x-a) + \mathsf{f}(a).$$

A function $\mathsf{f} : \mathbb{R} \to \mathbb{R}$ *is*

DEF

(1).	***convex***	*in* $(a:b)$ *if*	$\mathsf{f}(x)$	$\le$	$\mathsf{y}(x)$	*for* $x \in (a:b)$
(2).	***concave***	*in* $(a:b)$ *if*	$\mathsf{f}(x)$	$\ge$	$\mathsf{y}(x)$	*for* $x \in (a:b)$
(3).	***strictly convex***	*in* $(a:b)$ *if*	$\mathsf{f}(x)$	$<$	$\mathsf{y}(x)$	*for* $x \in (a:b)$
(4).	***strictly concave***	*in* $(a:b)$ *if*	$\mathsf{f}(x)$	$>$	$\mathsf{y}(x)$	*for* $x \in (a:b)$

Theorem 9.1 (Jensen's Inequality). [10] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE *(Definition 3.1 page 67), D a subset of X, and* f *a functional in* $\mathbb{F}^D$. *Let* $\sum$ *be the* SUMMATION OPERATOR *(Definition 11.1 page 181).*

THM

$$\left\{\begin{array}{lll}1. & D \text{ is CONVEX} & \textit{and}\\ 2. & \mathsf{f} \text{ is CONVEX} & \textit{and}\\ 3. & \displaystyle\sum_{n=1}^{N}\lambda_n = 1 & \text{(WEIGHTS)}\end{array}\right\} \implies \mathsf{f}\left(\sum_{n=1}^{N}\lambda_n \boldsymbol{x}_n\right) \le \sum_{n=1}^{N}\lambda_n \mathsf{f}(\boldsymbol{x}_n) \qquad \forall \boldsymbol{x}_n \in D,\ N \in \mathbb{N}$$

✎PROOF: Proof is by induction:

1. Proof that statement is true for $N = 1$:

$$\begin{aligned}\mathsf{f}\left(\sum_{n=1}^{N=1}\lambda_n \boldsymbol{x}_n\right) &= \mathsf{f}(\lambda_1 \boldsymbol{x}_1)\\ &\le \mathsf{f}(\lambda_1 \boldsymbol{x}_1)\\ &= \sum_{n=1}^{N=1}\lambda_n \mathsf{f}(\boldsymbol{x}_n)\end{aligned}$$

[9] Udriste (1994) page 63, Kurdila and Zabarankin (2005) page 178 ⟨Proposition 6.1.1⟩, Rockafellar (1970) page 23 ⟨Section 4 Convex Functions⟩, Çinlar and Vanderbei (2013) page 86 ⟨5.4 Theorem⟩

[10] Mitrinović et al. (2010) page 6, Bollobás (1999) page 3, Lay (1982) page 7, Jensen (1906), pages 179–180

2. Proof that statement is true for $N = 2$:

$$\begin{aligned}
\mathsf{f}\left(\sum_{n=1}^{N=2} \lambda_n \boldsymbol{x}_n\right) &= \mathsf{f}(\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2) \\
&\le \lambda_1 \mathsf{f}(\boldsymbol{x}_1) + \lambda_2 \mathsf{f}(\boldsymbol{x}_2) && \text{by convexity hypothesis} \\
&= \sum_{n=1}^{N=2} \lambda_n \mathsf{f}(\boldsymbol{x}_n)
\end{aligned}$$

3. Proof that if the statement is true for N, then it is also true for $N + 1$:

$$\begin{aligned}
\mathsf{f}\left(\sum_{n=1}^{N+1} \lambda_n \boldsymbol{x}_n\right) &= \mathsf{f}\left(\sum_{n=1}^{N} \lambda_n \boldsymbol{x}_n + \lambda_{N+1} \boldsymbol{x}_{N+1}\right) \\
&= \mathsf{f}\left([1-\lambda_{N+1}] \sum_{n=1}^{N} \frac{\lambda_n}{1-\lambda_{N+1}} \boldsymbol{x}_n + \lambda_{N+1} \boldsymbol{x}_{N+1}\right) \\
&\le [1-\lambda_{N+1}]\mathsf{f}\left(\sum_{n=1}^{N} \frac{\lambda_n}{1-\lambda_{N+1}} \boldsymbol{x}_n\right) + \lambda_{N+1}\mathsf{f}(\boldsymbol{x}_{N+1}) && \text{by convexity hypothesis} \\
&\le [1-\lambda_{N+1}] \sum_{n=1}^{N} \frac{\lambda_n}{1-\lambda_{N+1}} \mathsf{f}(\boldsymbol{x}_n) + \lambda_{N+1}\mathsf{f}(\boldsymbol{x}_{N+1}) && \text{by “true for } N\text{” hypothesis} \\
&= \sum_{n=1}^{N} \lambda_n \mathsf{f}(\boldsymbol{x}_n) + \lambda_{N+1}\mathsf{f}(\boldsymbol{x}_{N+1}) \\
&= \sum_{n=1}^{N+1} \lambda_n \mathsf{f}(\boldsymbol{x}_n)
\end{aligned}$$

4. Since the statement is true for $N = 1$, $N = 2$, and true for $N \implies$ true for $N + 1$, then it is true for $N = 1, 2, 3, 4, \ldots$

✏

The next theorem gives another form of convex functions that is a little less intuitive but provides powerful analytic results.

Theorem 9.2. *Let* $\mathsf{f} : \mathbb{R} \to \mathbb{R}$. *For every* $x_1, x_2 \in (a, b)$ *and* $\lambda \in [0, 1]$

THM: f *is convex in* (a, b) $\iff$ $\mathsf{f}(\lambda x_1 + (1-\lambda)x_2) \le \lambda\mathsf{f}(x_1) + (1-\lambda)\mathsf{f}(x_2)$

✎PROOF:

1. prove f is convex $\implies \mathsf{f}(\lambda x_1 + (1-\lambda)x_2) \le \lambda\mathsf{f}(x_1) + (1-\lambda)\mathsf{f}(x_2)$:

$$\begin{aligned}
\mathsf{f}(\lambda x_1 + (1-\lambda)x_2) &\le \frac{\mathsf{f}(b)-\mathsf{f}(a)}{b-a}[\lambda x_1 + (1-\lambda)x_2 - a] + \mathsf{f}(a) \\
&= \frac{\mathsf{f}(x_2)-\mathsf{f}(x_1)}{x_2-x_1}[\lambda x_1 + (1-\lambda)x_2 - x_1] + \mathsf{f}(x_1) \\
&= \frac{\mathsf{f}(x_2)-\mathsf{f}(x_1)}{x_2-x_1}[(x_2-x_1)(1-\lambda)] + \mathsf{f}(x_1) \\
&= (1-\lambda)\mathsf{f}(x_2) - (1-\lambda)\mathsf{f}(x_1) + \mathsf{f}(x_1) \\
&= \lambda\mathsf{f}(x_1) + (1-\lambda)\mathsf{f}(x_2)
\end{aligned}$$

2. prove f is convex $\Longleftarrow$ $f(\lambda x_1 + (1-\lambda)x_2) \le \lambda \mathsf{f}(x_1) + (1-\lambda)\mathsf{f}(x_2)$:

   Let $x = \lambda(b-a)+a$ Notice that as λ varies from 0 to 1, x varies from b to a. So free variable λ works as a change of variable for free variable x.

$$\lambda = \frac{x-a}{b-a}$$

$$\begin{aligned}
\mathsf{f}(x) &= \mathsf{f}(\lambda(b-a)+a) \\
&\le \lambda\mathsf{f}(b) + (1-\lambda)\mathsf{f}(a) \\
&= \lambda[\mathsf{f}(b)-\mathsf{f}(a)] + \mathsf{f}(a) \\
&= \frac{\mathsf{f}(b)-\mathsf{f}(a)}{b-a}(x-a) + \mathsf{f}(a)
\end{aligned}$$

✏

Taking the second derivative of a function provides a convenient test for whether that function is convex.

Theorem 9.3. [11]

THM $\mathsf{f}''(x) > 0 \implies \mathsf{f}$ *is convex*

✎PROOF:

$$\begin{aligned}
\mathsf{f}(x) &= \mathsf{f}(x_0) + \mathsf{f}'(x_0)(x-x_0) + \mathsf{f}''(c)(x-x_0)^2 \\
&\ge \mathsf{f}(x_0) + \mathsf{f}'(x_0)(x-x_0) \\
&= \mathsf{f}(x_0) + \mathsf{f}'(x_0)(x-\lambda x_1 - (1-\lambda)x_2)
\end{aligned}$$

$$\begin{aligned}
\mathsf{f}(x_1) &\ge \mathsf{f}(x_0) + \mathsf{f}'(x_0)(x_1-\lambda x_1 - (1-\lambda)x_2) \\
&= \mathsf{f}(x_0) + \mathsf{f}'(x_0)(1-\lambda)(x_1-x_2) \\
&= \mathsf{f}(x_0) - \mathsf{f}'(x_0)(1-\lambda)(x_2-x_1)
\end{aligned}$$

$$\begin{aligned}
\mathsf{f}(x_2) &\ge \mathsf{f}(x_0) + \mathsf{f}'(x_0)(x_2-\lambda x_1 - (1-\lambda)x_2) \\
&= \mathsf{f}(x_0) + \mathsf{f}'(x_0)\lambda(x_2-x_1)
\end{aligned}$$

$$\begin{aligned}
\lambda\mathsf{f}(x_1) + (1-\lambda)\mathsf{f}(x_2) &\ge \lambda\left[\mathsf{f}(x_0) - \mathsf{f}'(x_0)(1-\lambda)(x_2-x_1)\right] + (1-\lambda)\left[\mathsf{f}(x_0) + \mathsf{f}'(x_0)\lambda(x_2-x_1)\right] \\
&= \lambda\left[\mathsf{f}(x_0) - \mathsf{f}'(x_0)(1-\lambda)(x_2-x_1)\right] + \left[\mathsf{f}(x_0) + \mathsf{f}'(x_0)\lambda(x_2-x_1)\right] - \lambda\left[\mathsf{f}(x_0) + \mathsf{f}'(x_0)\lambda(x_2-x_1)\right] \\
&= \mathsf{f}(x_0) \\
&= \mathsf{f}(\lambda x_1 + (1-\lambda)x_2)
\end{aligned}$$

By Theorem 9.2 (page 145), $\mathsf{f}(x)$ is convex. ✏

9.4 Literature

Literature survey:

[11] Cover and Thomas (1991), pages 24–25

1. Abstract convexity:
   Edelman and Jamison (1985)
   van de Vel (1993)
   Hörmander (1994)
2. Order convexity (lattice theory):
   Edelman (1986)
3. Metric convexity:
   Menger (1928)
   Blumenthal (1970) page 41 ⟨?⟩
   Khamsi and Kirk (2001) pages 35–38

Part III

Structures on Spaces

CHAPTER 10

LINEAR COMBINATIONS

10.1 Linear combinations in linear spaces

A *linear space* (Definition 3.1 page 67) in general is not equipped with a *topology* (Definition 1.1 page 3). Without a topology, it is not possible to determine whether an *infinite sum* (Definition 12.1 page 196) of vectors converges (Definition 8.3 page 127). Therefore in this section (dealing with linear spaces), all definitions related to sums of vectors will be valid for *finite* sums(Definition 11.1 page 181) only (finite "N").

Definition 10.1. [1] *Let* $\{x_n \in X | _{n=1,2,\ldots,N}\}$ *be a set of vectors in a* LINEAR SPACE $(X, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}))$.

DEF

A vector $x \in X$ *is a* ***linear combination*** *of the vectors in* $\{x_n\}$ *if there exists* $\{\alpha_n \in \mathbb{F} | _{n=1,2,\ldots,N}\}$ *such that*

$$x = \sum_{n=1}^{N} \alpha_n x_n .$$

Definition 10.2. [2] *Let* $(X, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}))$ *be a linear space and* Y *be a subset of* X.

DEF

The ***linear span*** *of* Y *is defined as* $\operatorname{span} Y \triangleq \left\{ \sum_{\gamma \in \Gamma} \alpha_\gamma y_\gamma \,\middle|\, \alpha_\gamma \in \mathbb{F},\ y_\gamma \in Y \right\}$.

The set Y ***spans*** *a set* A *if* $A \subseteq \operatorname{span} Y$.

Proposition 10.1. [3] *Let* $\{x_n \in X | _{n=1,2,\ldots,N}\}$ *be a set of vectors in a* LINEAR SPACE $\mathbf{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}))$.

PRP

1. $\operatorname{span}\{x_n\}$ *is a* LINEAR SPACE *(Definition 3.1 page 67)* *and*
2. $\operatorname{span}\{x_n\}$ *is a* LINEAR SUBSPACE *of* $\mathbf{L}$ *(Definition 3.2 page 68).*

Definition 10.3. [4] *Let* $\mathbf{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}))$ *be a* LINEAR SPACE.

DEF

The set $Y \triangleq \{x_n \in X | _{n=1,2,\ldots,N}\}$ *is* ***linearly independent*** *in* $\mathbf{L}$ *if*

$$\left\{ \sum_{n=1}^{N} \alpha_n x_n = 0 \right\} \implies \{\alpha_1 = \alpha_2 = \cdots = \alpha_N = 0\}.$$

The set Y *is* ***linearly dependent*** *in* $\mathbf{L}$ *if* Y *is not linearly independent in* $\mathbf{L}$.

[1] Berberian (1961) page 11 ⟨Definition I.4.1⟩, Kubrusly (2001) page 46

[2] Michel and Herget (1993) page 86 ⟨3.3.7 Definition⟩, Kurdila and Zabarankin (2005) page 44, Searcóid (2002) page 71 ⟨Definition 3.2.5—more general definition⟩

[3] Kubrusly (2001) page 46

[4] Bachman and Narici (1966) pages 3–4, Christensen (2003) page 2, Heil (2011) page 156 ⟨Definition 5.7⟩

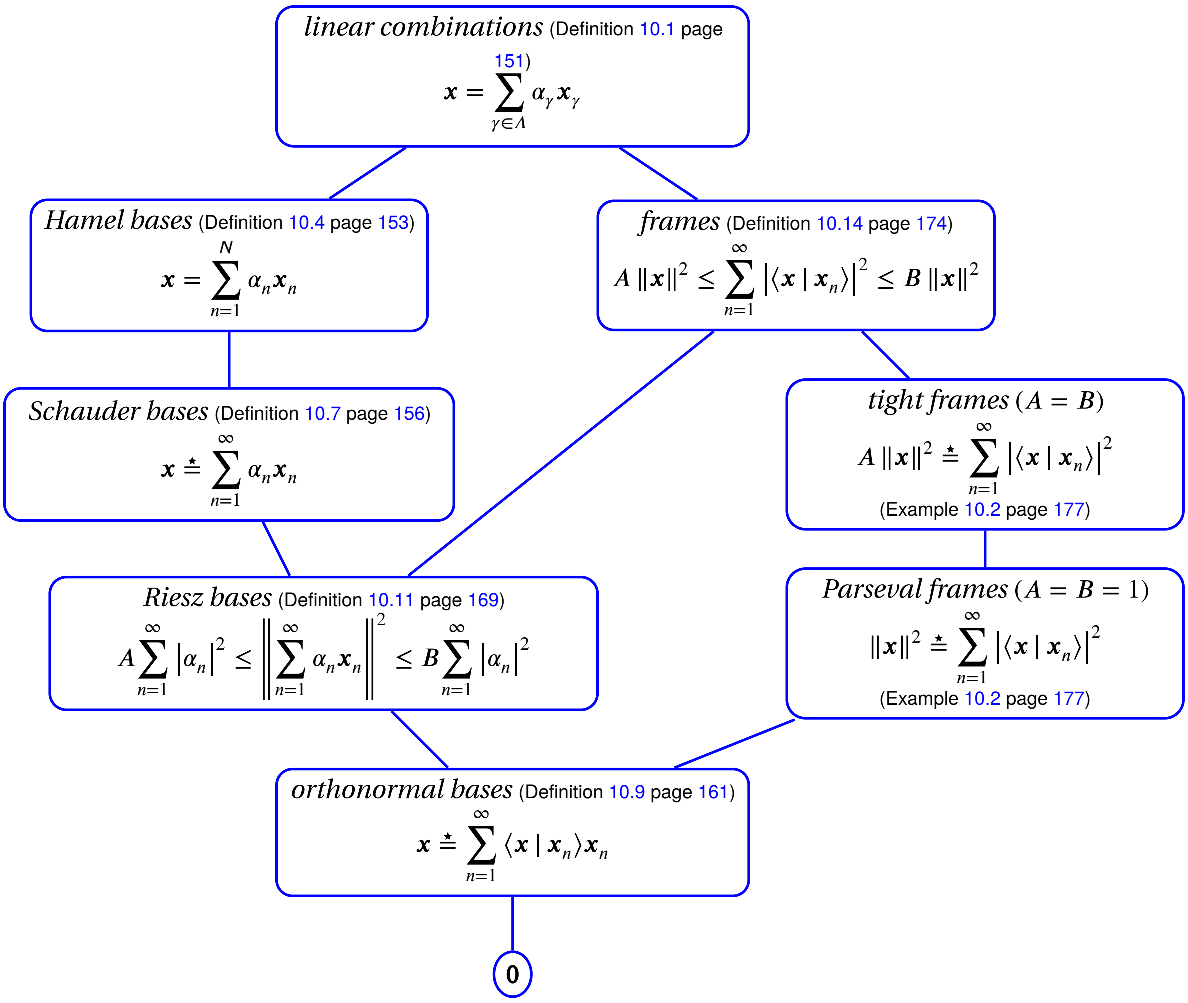


Figure 10.1: Lattice of *linear combinations*

Definition 10.4. [5] *Let* $\{x_n \in X | n=1,2,\ldots,N\}$ *be a set of vectors in a* LINEAR SPACE $\mathbf{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}))$.

DEF

The set $\{x_n\}$ *is a* ***Hamel basis*** *for* $\mathbf{L}$ *if*

1. $\{x_n\}$ SPANS $\mathbf{L}$ (Definition 10.2 page 151) *and*
2. $\{x_n\}$ *is* LINEARLY INDEPENDENT *in* $\mathbf{L}$ (Definition 10.1 page 151) .

A HAMEL BASIS *is also called a* ***linear basis***.

Definition 10.5. [6] *Let* $\mathbf{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}))$ *be a* LINEAR SPACE. *Let* x *be a* VECTOR *in* $\mathbf{L}$ *and* $Y \triangleq \{x_n \in X | n=1,2,\ldots,N\}$ *be a set of vectors in* $\mathbf{L}$.

DEF

The expression $\sum_{n=1}^{N} \alpha_n x_n$ *is the* ***expansion*** *of* x *on* Y *in* $\mathbf{L}$ *if* $x = \sum_{n=1}^{N} \alpha_n x_n$.
In this case, the sequence $(\alpha_n)_{n=1}^{N}$ *is the* ***coordinates*** *of* x *with respect to* Y *in* $\mathbf{L}$.
If $\alpha_N \neq 0$, *then* N *is the* ***dimension*** $\dim \mathbf{L}$ *of* $\mathbf{L}$.

Theorem 10.1. [7] *Let* $\{x_n | n=1,2,\ldots,N\}$ *be a* HAMEL BASIS (Definition 10.4 page 153) *for a* LINEAR SPACE $(X, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}))$.

THM

$$\left\{ x = \sum_{n=1}^{N} \alpha_n x_n = \sum_{n=1}^{N} \beta_n x_n \right\} \implies \underbrace{\alpha_n = \beta_n \quad \forall n=1,2,\ldots,N}_{\text{coordinates of } x \text{ are UNIQUE}} \qquad \forall x \in X$$

✎PROOF:

$$\begin{aligned}
\mathbb{0} &= x - x \\
&= \sum_{n=1}^{N} \alpha_n x_n - \sum_{n=1}^{N} \beta_n x_n \\
&= \sum_{n=1}^{N} (\alpha_n - \beta_n) x_n \\
\implies & \{x_n\} \text{ is } \textit{linearly dependent} \text{ if } (\alpha_n - \beta_n) \neq 0 \qquad \forall n = 1,2,\ldots,N \\
\implies & (\alpha_n - \beta_n) = 0 \qquad \forall n = 1,2,\ldots,N \qquad \text{(because } \{x_n\} \text{ is a } \textit{basis} \text{ and therefore must be } \textit{linearly independent}\text{)} \\
\implies & \alpha_n = \beta_n \text{ for } n = 1,2,\ldots,N
\end{aligned}$$

⇨

Theorem 10.2. [8] *Let* $\mathbf{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}))$ *be a* LINEAR SPACE.

THM

$$\left\{ \begin{array}{ll} 1. & \{x_n \in X | n=1,2,\ldots,N\} \text{ is a HAMEL BASIS for } \mathbf{L} \quad \text{and} \\ 2. & \{y_n \in X | n=1,2,\ldots,M\} \text{ is a set of LINEARLY INDEPENDENT vectors in } \mathbf{L} \end{array} \right\}$$
$$\implies \left\{ \begin{array}{lll} 1. & M \leq N & \text{and} \\ 2. & M = N \implies \{y_n | n=1,2,\ldots,M\} \text{ is a BASIS for } \mathbf{L} & \text{and} \\ 3. & M \neq N \implies \{y_n | n=1,2,\ldots,M\} \text{ is NOT a basis for } \mathbf{L} & \end{array} \right\}$$

✎PROOF:

[5] Hamel (1905), Bachman and Narici (1966) page 4, Kubrusly (2001) pages 48–49 ⟨Section 2.4⟩, Young (2001) page 1, Carothers (2005) page 25, Heil (2011) page 125 ⟨Definition 4.1⟩

[6] Hamel (1905), Bachman and Narici (1966) page 4, Kubrusly (2001) pages 48–49 ⟨Section 2.4⟩, Young (2001) page 1, Carothers (2005) page 25, Heil (2011) page 125 ⟨Definition 4.1⟩

[7] Michel and Herget (1993) pages 89–90 ⟨Theorem 3.3.25⟩

[8] Michel and Herget (1993) pages 90–91 ⟨Theorem 3.3.26⟩

1. Proof that $\{\boldsymbol{y}_1, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}\}$ is a *basis* for $\boldsymbol{L}$:
   (a) Proof that $\{\boldsymbol{y}_1, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}\}$ *spans* $\boldsymbol{L}$:
      i. Because $\{\boldsymbol{x}_n | n=1,2,\ldots,N\}$ is a *basis* for $\boldsymbol{L}$, there exists $\beta \in \mathbb{F}$ and $\{\alpha_n \in \mathbb{F} | n=1,2,\ldots,N\}$ such that $\beta \boldsymbol{y}_1 + \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n = 0.$
      ii. Select an n such that $\alpha_n \neq 0$ and renumber (if necessary) the above indices such that $\boldsymbol{x}_n = -\frac{\beta}{\alpha_n}\boldsymbol{y}_1 + \sum_{n=1}^{N-1} \frac{\alpha_n}{\alpha_n}\boldsymbol{x}_n.$
      iii. Then, for any $\boldsymbol{y} \in X$, we can write

$$\begin{aligned}
\boldsymbol{y} &= \sum_{n=1}^{N} \gamma_{n\in\mathbb{Z}} \boldsymbol{x}_n \\
&= \left(\sum_{n=1}^{N-1} \gamma_{n\in\mathbb{Z}} \boldsymbol{x}_n\right) + \gamma_{n\in\mathbb{Z}}\left(-\frac{\beta}{\alpha_n}\boldsymbol{y}_1 - \sum_{n=1}^{N-1} \frac{\alpha_n}{\alpha_n}\boldsymbol{x}_n\right) \\
&= -\frac{\beta\gamma_n}{\alpha_n}\boldsymbol{y}_1 + \sum_{n=1}^{N-1}\left(\gamma_n - \frac{\alpha_n\gamma_n}{\alpha_n}\right)\boldsymbol{x}_n \\
&= \delta \boldsymbol{y}_1 + \sum_{n=1}^{N-1} \delta_{n\in\mathbb{Z}} \boldsymbol{x}_n
\end{aligned}$$

      iv. This implies that $\{\boldsymbol{y}_1, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}\}$ *spans* $\boldsymbol{L}$:
   (b) Proof that $\{\boldsymbol{y}_1, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}\}$ is *linearly independent*:
      i. If $\{\boldsymbol{y}_1, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}\}$ is *linearly dependent*, then there exists $\{\epsilon, \epsilon_1, \ldots, \epsilon_{N-1}\}$ such that $\epsilon \boldsymbol{y}_1 + \left(\sum_{n=1}^{N-1} \epsilon_{n\in\mathbb{Z}} \boldsymbol{x}_n\right) + 0\boldsymbol{x}_n = 0.$
      ii. item (1(b)i) implies that the coordinate of $\boldsymbol{y}_1$ associated with $\boldsymbol{x}_n$ *is* 0. $\boldsymbol{y}_1 = -\left(\sum_{n=1}^{N-1} \frac{\epsilon_n}{\epsilon}\boldsymbol{x}_n\right) + 0\boldsymbol{x}_n = 0.$
      iii. item (1(a)i) implies that the coordinate of $\boldsymbol{y}_1$ associated with $\boldsymbol{x}_n$ is *not* 0. $\boldsymbol{y}_1 = -\sum_{n=1}^{N} \frac{\alpha_n}{\beta}\boldsymbol{x}_n.$
      iv. This implies that item (1(b)i) (that the set is linearly dependent) is *false* because item (1(b)ii) and item (1(b)iii) *contradict* each other.
      v. This implies $\{\boldsymbol{y}_1, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}\}$ is *linearly independent*.
2. Proof that $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-2}\}$ is a *basis*: Repeat item (1).
3. Suppose $m = n$. Proof that $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_M\}$ is a *basis*: Repeat item (1) $M - 1$ times.
4. Proof that $M \not> N$:
   (a) Suppose that $M = N + 1$.
   (b) Then because $\{\boldsymbol{y}_n | n=1,2,\ldots,N\}$ is a *basis*, there exists $\{\zeta_n | n=1,2,\ldots,N+1\}$ such that $\sum_{n=1}^{N+1} \zeta_{n\in\mathbb{Z}} \boldsymbol{y}_{n\in\mathbb{Z}} = 0.$
   (c) This implies that $\{\boldsymbol{y}_n | n=1,2,\ldots,N+1\}$ is *linearly dependent*.

(d) This implies that $\{y_n|_{n=1,2,\ldots,N+1}\}$ is *not* a basis.

(e) This implies that $M \not> N$.

5. Proof that $M \neq N \implies \{y_n|_{n=1,2,\ldots,M}\}$ is *not* a basis for $\boldsymbol{L}$:

(a) Proof that $M > N \implies \{y_n|_{n=1,2,\ldots,M}\}$ is *not* a basis for $\boldsymbol{L}$: same as in item (4).

(b) Proof that $M < N \implies \{y_n|_{n=1,2,\ldots,M}\}$ is *not* a basis for $\boldsymbol{L}$:

i. Suppose $m = N - 1$.

ii. Then $\{y_n|_{n=1,2,\ldots,N-1}\}$ is a *basis* and there exists λ such that
$$\sum_{n=1}^{N} \lambda_{n\in\mathbb{Z}} y_{n\in\mathbb{Z}} = 0.$$

iii. This implies that $\{y_n|_{n=1,2,\ldots,N}\}$ is *linearly dependent* and is *not* a basis.

iv. But this contradicts item (3), therefore $M \neq N - 1$.

v. Because $M = N$ yields a basis but $M = N - 1$ does not, $M < N - 1$ also does not yield a basis.

✏

Corollary 10.1. [9] *Let* $\boldsymbol{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a linear space.*

COR
$$\underbrace{\left\{\begin{array}{ll} 1. & \{x_n \in X|_{n=1,2,\ldots,N}\} \text{ is a HAMEL BASIS for } \boldsymbol{L} \quad \text{and} \\ 2. & \{y_n \in X|_{n=1,2,\ldots,M}\} \text{ is a HAMEL BASIS for } \boldsymbol{L} \end{array}\right\} \implies \{N = M\}}_{\text{(all Hamel bases for } \boldsymbol{L} \text{ have the same number of vectors)}}$$

✎PROOF: This follows from Theorem 10.2 (page 153). ✏

10.2 Bases in topological linear spaces

A linear space supports the concept of the *span* of a set of vectors (Definition 10.2 page 151). In a topological linear space $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), T)$, a set A is said to be *total* in $\boldsymbol{\Omega}$ if the span of A is *dense* in $\boldsymbol{\Omega}$. In this case, A is said to be a *total set* or a *complete set*. However, this use of "complete" in a "*complete set*" is not equivalent to the use of "complete" in a "*complete metric space*" (Definition 8.5 page 133).[10] In this text, except for these comments and Definition 10.6, "complete" refers to the metric space definition (Definition 8.5 page 133) only.

If a set is both *total* and *linearly independent* (Definition 10.3 page 151) in $\boldsymbol{\Omega}$, then that set is a *Hamel basis* (Definition 10.4 page 153) for $\boldsymbol{\Omega}$.

Definition 10.6. [11] *Let A^- be the* CLOSURE *(Definition 1.4 page 14) of a A in a* TOPOLOGICAL LINEAR SPACE *(Definition 4.1 page 79)* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), T)$. *Let* $\operatorname{span} A$ *be the* SPAN *(Definition 10.2 page 151) of a set A.*

DEF
A set of vectors A is **total** *(or* **complete** *or* **fundamental***) in $\boldsymbol{\Omega}$ if*
$$(\operatorname{span} A)^- = \boldsymbol{\Omega} \qquad (\text{SPAN of } A \text{ is DENSE (Definition 1.6 page 23) in } \boldsymbol{\Omega}).$$

[9] Kubrusly (2001) page 52 ⟨Theorem 2.7⟩, Michel and Herget (1993) page 91 ⟨Theorem 3.3.31⟩
[10] Haaser and Sullivan (1991) pages 296–297 ⟨6·Orthogonal Bases⟩, Rynne and Youngson (2008) page 78 ⟨Remark 3.50⟩, Heil (2011) page 21 ⟨Remark 1.26⟩
[11] Young (2001) page 19 ⟨Definition 1.5.1⟩, Sohrab (2003) page 362 ⟨Definition 9.2.3⟩, Gupta (1998) page 134 ⟨Definition 2.4⟩, Bachman and Narici (1966) pages 149–153 ⟨Definition 9.3, Theorems 9.9 and 9.10⟩

10.3 Schauder bases in Banach spaces

Definition 10.7. [12] *Let* $\boldsymbol{B} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a* BANACH SPACE (Definition 8.8 page 136). *Let* $\stackrel{\star}{=}$ *represent* STRONG CONVERGENCE (Definition 12.5 page 203) *in* $\boldsymbol{B}$.

DEF

The countable set $\{\boldsymbol{x}_n \in X | n\in\mathbb{N}\}$ *is a* ***Schauder basis*** *for* $\boldsymbol{B}$ *if for each* $\boldsymbol{x} \in X$

1. $\exists (\!(\alpha_n \in \mathbb{F})\!)_{n\in\mathbb{N}}$ such that $\boldsymbol{x} \stackrel{\star}{=} \sum_{n=1}^{\infty} \alpha_n \boldsymbol{x}_n$ (STRONG CONVERGENCE *in* $\boldsymbol{B}$) *and*
2. $\left\{ \sum_{n=1}^{\infty} \alpha_n \boldsymbol{x}_n \stackrel{\star}{=} \sum_{n=1}^{\infty} \beta_n \boldsymbol{x}_n \right\} \implies \{(\!(\alpha_n)\!) = (\!(\beta_n)\!)\}$ (COEFFICIENT FUNCTIONALS *are* UNIQUE)

In this case, $\sum_{n=1}^{\infty} \alpha_n \boldsymbol{x}_n$ *is the* ***expansion*** *of* $\boldsymbol{x}$ *on* $\{\boldsymbol{x}_n | n\in\mathbb{N}\}$ *and the elements of* $(\!(\alpha_n)\!)$ *are the* ***coefficient functionals*** *associated with the basis* $\{\boldsymbol{x}_n\}$. *Coefficient functionals are also called* ***coordinate functionals***.

In a Banach space, the existence of a Schauder basis implies that the space is *separable* (Theorem 10.3 page 156). The question of whether the converse is also true was posed by Banach himself in 1932,[13] and became know as "*The basis problem*". This remained an open question for many years. The question was finally answered some 41 years later in 1973 by Per Enflo (University of California at Berkley), with the answer being "no". Enflo constructed a counterexample in which a separable Banach space does *not* have a Schauder basis.[14] Life is simpler in Hilbert spaces where the converse *is* true: a Hilbert space has a Schauder basis *if and only if* it is separable (Theorem 10.11 page 168).

Theorem 10.3. [15] *Let* $\boldsymbol{B} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a* BANACH SPACE. *Let* $\mathbb{Q}$ *be the field of rational numbers.*

THM

$$\left\{ \begin{array}{ll} 1. & \boldsymbol{B} \text{ has a SCHAUDER BASIS} \quad \text{and} \\ 2. & \mathbb{Q} \text{ is DENSE in } \mathbb{F}. \end{array} \right\} \implies \{ \boldsymbol{B} \text{ is SEPARABLE} \}$$

✎PROOF:

1. lemma:

$$\begin{aligned} \left| \left\{ \boldsymbol{x} | \exists (\!(\alpha_n \mathbb{Q})\!)_{n\in\mathbb{N}} \text{ such that } \lim_{N\to\infty} \left\| \boldsymbol{x} - \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n \right\| = 0 \right\} \right| &= |\mathbb{Q} \times \mathbb{N}| \\ &= |\mathbb{Z} \times \mathbb{Z}| \\ &= |\mathbb{Z}| \\ &= \textit{countably infinite} \end{aligned}$$

[12] Carothers (2005) pages 24–25, Christensen (2003) pages 46–49 ⟨Definition 3.1.1 and page 49⟩, Young (2001) page 19 ⟨Section 6⟩, Singer (1970), page 17, Schauder (1927), Schauder (1928)

[13] Banach (1932a), page 111

[14] Enflo (1973), Lindenstrauss and Tzafriri (1977) pages 84–95 ⟨Section 2.d⟩

[15] Bachman et al. (2000) page 112 ⟨3.4.8⟩, Giles (2000) page 17, Heil (2011) page 21 ⟨Theorem 1.27⟩

2. remainder of proof:

$\boldsymbol{B}$ has a *Schauder basis* $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$

$$\begin{aligned}
&\implies \text{for every } \boldsymbol{x}\in\boldsymbol{B}, \text{ there exists } (\alpha_n\in\mathbb{F})_{n\in\mathbb{N}} \text{ such that } \boldsymbol{x} \doteq \sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n && \text{by Definition 10.7 page 156}\\
&\implies \text{for every } \boldsymbol{x}\in\boldsymbol{B}, \text{ there exists } (\alpha_n\in\mathbb{F})_{n\in\mathbb{N}} \text{ such that } \lim_{N\to\infty}\left\|\boldsymbol{x}-\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right\| = 0 && \text{by Definition 12.5}\\
&\implies \text{for every } \boldsymbol{x}\in\boldsymbol{B}, \text{ there exists } (\alpha_n\in\mathbb{Q})_{n\in\mathbb{N}} \text{ such that } \lim_{N\to\infty}\left\|\boldsymbol{x}-\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right\| = 0 && \text{because } \mathbb{Q}^- = \mathbb{F}\\
&\implies \boldsymbol{B} = \left\{\boldsymbol{x}|\exists(\alpha_n\mathbb{Q})_{n\in\mathbb{N}} \text{ such that } \lim_{N\to\infty}\left\|\boldsymbol{x}-\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right\| = 0\right\}\\
&\implies \boldsymbol{B} = \left\{\boldsymbol{x}|\exists(\alpha_n\mathbb{Q})_{n\in\mathbb{N}} \text{ such that } \boldsymbol{x} = \lim_{N\to\infty}\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right\}^-\\
&\implies \boldsymbol{B} \text{ is } \textit{separable} && \text{by (1) lemma page 156}
\end{aligned}$$

✏

Definition 10.8. [16] *Let* $\{\boldsymbol{x}_n|\, n\in\mathbb{N}\}$ *and* $\{\boldsymbol{y}_n|\, n\in\mathbb{N}\}$ *be* SCHAUDER BASES *of a* BANACH SPACE $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$.

DEF $\{\boldsymbol{x}_n\}$ ***is equivalent*** *to* $\{\boldsymbol{y}_n\}$ *if there exists a* BOUNDED INVERTIBLE *operator* $\mathbf{R}$ *in* $\boldsymbol{X}^{\boldsymbol{X}}$ *such that* $\mathbf{R}\boldsymbol{x}_n = \boldsymbol{y}_n \quad \forall n\in\mathbb{Z}$

Theorem 10.4. [17] *Let* $\{\boldsymbol{x}_n|\, n\in\mathbb{N}\}$ *and* $\{\boldsymbol{y}_n|\, n\in\mathbb{N}\}$ *be* SCHAUDER BASES *of a* BANACH SPACE $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$.

THM
$$\left\{\{\boldsymbol{x}_n\} \text{ is EQUIVALENT to } \{\boldsymbol{y}_n\}\right\} \iff \left\{\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n \text{ is CONVERGENT} \iff \sum_{n=1}^{\infty}\alpha_n\boldsymbol{y}_n \text{ is CONVERGENT}\right\}$$

Lemma 10.1. [18] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \boldsymbol{T})$ *be a topological linear space. Let* span$A$ *be the* SPAN *of a set* A (Definition 10.2 page 151). *Let* $\tilde{\mathsf{f}}(\omega)$ *and* $\tilde{\mathsf{g}}(\omega)$ *be the* FOURIER TRANSFORMS (Definition ?? page ??) *of the functions* $\mathsf{f}(x)$ *and* $\mathsf{g}(x)$*, respectively, in* $\boldsymbol{L}^2_{\mathbb{R}}$ (Definition ?? page ??). *Let* $\breve{\mathsf{a}}(\omega)$ *be the* DTFT (Definition ?? page ??) *of a sequence* $(a_n)_{n\in\mathbb{Z}}$ *in* $\boldsymbol{\ell}^2_{\mathbb{R}}$ (Definition ?? page ??).

LEM
$$\left\{\begin{array}{ll}(1). & \{\mathbf{T}^n\mathsf{f}|\, n\in\mathbb{Z}\} \text{ is a SCHAUDER BASIS for } \boldsymbol{\Omega} \quad \text{and}\\ (2). & \{\mathbf{T}^n\mathsf{g}|\, n\in\mathbb{Z}\} \text{ is a SCHAUDER BASIS for } \boldsymbol{\Omega}\end{array}\right\} \implies \left\{\begin{array}{l}\exists\,(a_n)_{n\in\mathbb{Z}} \text{ such that}\\ \tilde{\mathsf{f}}(\omega) = \breve{\mathsf{a}}(\omega)\tilde{\mathsf{g}}(\omega)\end{array}\right\}$$

✎PROOF: Let V_0' be the space spanned by $\{\mathbf{T}^n\phi|\, n\in\mathbb{Z}\}$.

$$\begin{aligned}
\tilde{\mathsf{f}}(\omega) &\triangleq \tilde{\mathbf{F}}\mathsf{f} && \text{by definition of } \tilde{\mathbf{F}} \quad \text{(Definition ?? page ??)}\\
&= \tilde{\mathbf{F}}\sum_{n\in\mathbb{Z}} a_n\mathbf{T}\mathsf{g} && \text{by (2)}\\
&= \sum_{n\in\mathbb{Z}} a_n\tilde{\mathbf{F}}\mathbf{T}\mathsf{g}
\end{aligned}$$

[16] Young (2001) page 25 ⟨Definition 1.8.1, Theorem 1.8.7⟩
[17] Young (2001) page 25 ⟨Definition 1.8.1, Theorem 1.8.7⟩
[18] Daubechies (1992), page 140

$$= \underbrace{\sum_{n\in\mathbb{Z}} a_n e^{-i\omega n}}_{\breve{a}(\omega)} \underbrace{\tilde{\mathbf{F}}\mathsf{g}}_{\tilde{\mathsf{g}}(\omega)} \qquad \text{by Corollary ?? page ??}$$

$$= \breve{a}(\omega)\tilde{\mathsf{g}}(\omega) \qquad \text{by definition of } \check{\mathbf{F}} \text{ and } \tilde{\mathbf{F}} \quad \text{by (Definition ?? page ??, Definition ?? page ??)}$$

$$\begin{aligned}
V_0 &\triangleq \left\{ \mathsf{f}(x) | \mathsf{f}(x) = \sum_{n\in\mathbb{Z}} b_n \mathbf{T}^n \mathsf{g}(x) \right\} \\
&= \left\{ \mathsf{f}(x) | \tilde{\mathbf{F}}\mathsf{f}(x) = \tilde{\mathbf{F}} \sum_{n\in\mathbb{Z}} b_n \mathbf{T}^n \mathsf{g}(x) \right\} \\
&= \left\{ \mathsf{f}(x) | \tilde{\mathsf{f}}(\omega) = \tilde{\mathsf{b}}(\omega)\tilde{\mathsf{g}}(\omega) \right\} \\
&= \left\{ \mathsf{f}(x) | \tilde{\mathsf{f}}(\omega) = \tilde{\mathsf{b}}(\omega)\breve{a}(\omega)\tilde{\mathsf{f}}(\omega) \right\} \\
&= \left\{ \mathsf{f}(x) | \tilde{\mathsf{f}}(\omega) = \tilde{\mathsf{c}}(\omega)\tilde{\mathsf{f}}(\omega) \right\} && \text{where } \tilde{\mathsf{c}}(\omega) \triangleq \tilde{\mathsf{b}}(\omega)\breve{a}(\omega) \\
&= \left\{ \mathsf{f}(x) | \mathsf{f}(x) = \sum_{n\in\mathbb{Z}} \mathsf{c}_n \mathsf{f}(x-n) \right\} \\
&\triangleq V_0'
\end{aligned}$$

⇒

10.4 Linear combinations in inner product spaces

In an *inner product space*, *orthogonality* is a special case of *linear independence*; or alternatively, linear independence is a generalization of orthogonality (next theorem).

Theorem 10.5. [19] *Let* $\{ \mathbf{x}_n \in X | n=1,2,\ldots,N \}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition 6.1 page 95)* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle | \nabla \rangle)$.

THM
$$\left\{ \{\mathbf{x}_n\} \text{ is ORTHOGONAL} \atop \text{(Definition 6.4 page 107)} \right\} \implies \left\{ \{\mathbf{x}_n\} \text{ is LINEARLY INDEPENDENT} \atop \text{(Definition 10.1 page 151)} \right\}$$

✎PROOF:

1. Proof using *Pythagorean theorem* (Theorem 6.10 page 108):
   Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence with at least one nonzero element.

$$\begin{aligned}
\left\| \sum_{n=1}^N \alpha_n \mathbf{x}_n \right\|^2 &= \sum_{n=1}^N \|\alpha_n \mathbf{x}_n\|^2 && \text{by left hypoth. and } \textit{Pythagorean Theorem} \text{ (Theorem 6.10 page 108)} \\
&= \sum_{n=1}^N |\alpha_n|^2 \|\mathbf{x}_n\|^2 && \text{by definition of } \|\cdot\| \text{ (Definition 5.1 page 83)} \\
&> 0 \\
\implies \sum_{n=1}^N \alpha_n \mathbf{x}_n &\neq 0 \\
\implies (\mathbf{x}_n)_{n\in\mathbb{N}} &\text{ is } \textit{linearly independent} && \text{by definition of } \textit{linear independence} \text{ (Definition 10.3 page 151)}
\end{aligned}$$

[19] Aliprantis and Burkinshaw (1998) page 283 ⟨Corollary 32.8⟩, Kubrusly (2001) page 352 ⟨Proposition 5.34⟩

2. Alternative proof:

$$\begin{aligned}
\sum_{n=1}^{N} \alpha_n x_n = \mathbb{0} &\implies \left\langle \sum_{n=1}^{N} \alpha_n x_n \,\middle|\, x_m \right\rangle = \langle \mathbb{0} \,|\, x_m \rangle \\
&\implies \sum_{n=1}^{N} \alpha_n \langle x_n \,|\, x_m \rangle = 0 \\
&\implies \sum_{n=1}^{N} \alpha_n \bar{\delta}(k-m) = 0 \\
&\implies \alpha_m = 0 \qquad \text{for } m = 1, 2, \ldots, N
\end{aligned}$$

✏

Theorem 10.6 (Bessel's Equality).[20] *Let* $\left\{ x_n \in X \,\middle|\, {}_{n=1,2,\ldots,N} \right\}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition 6.1 page 95)* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \nabla \rangle)$ *and with* $\|x\| \triangleq \sqrt{\langle x \,|\, x \rangle}$ *(Definition 6.2 page 100).*

THM

$$\left\{ \begin{array}{c} \{x_n\} \text{ is ORTHONORMAL} \\ \text{(Definition 6.4 page 107)} \end{array} \right\} \implies \left\{ \underbrace{\left\| x - \sum_{n=1}^{N} \langle x \,|\, x_n \rangle x_n \right\|^2}_{\text{approximation error}} = \|x\|^2 - \sum_{n=1}^{N} |\langle x \,|\, x_n \rangle|^2 \quad \forall x \in X \right\}$$

✎PROOF:

$$\begin{aligned}
&\left\| x - \sum_{n=1}^{N} \langle x \,|\, x_n \rangle x_n \right\|^2 \\
&= \|x\|^2 + \left\| \sum_{n=1}^{N} \langle x \,|\, x_n \rangle x_n \right\|^2 - 2\Re \left\langle x \,\middle|\, \sum_{n=1}^{N} \langle x \,|\, x_n \rangle x_n \right\rangle && \text{by \textit{polar identity}} && \text{(Lemma 6.1 page 99)} \\
&= \|x\|^2 + \left\| \sum_{n=1}^{N} \langle x \,|\, x_n \rangle x_n \right\|^2 - 2\Re \left[\left(\sum_{n=1}^{N} \langle x \,|\, x_n \rangle \right)^* \langle x \,|\, x_n \rangle \right] && \text{by property of } \langle \triangle \,|\, \nabla \rangle && \text{(Definition 6.1 page 95)} \\
&= \|x\|^2 + \sum_{n=1}^{N} \| \langle x \,|\, x_n \rangle x_n \|^2 - 2\Re \left[\left(\sum_{n=1}^{N} \langle x \,|\, x_n \rangle \right)^* \langle x \,|\, x_n \rangle \right] && \text{by \textit{Pythagorean Theorem}} && \text{(Theorem 6.10 page 108)} \\
&= \|x\|^2 + \sum_{n=1}^{N} \| \langle x \,|\, x_n \rangle x_n \|^2 - 2\Re \left(\sum_{n=1}^{N} \langle x \,|\, x_n \rangle^* \langle x \,|\, x_n \rangle \right) \\
&= \|x\|^2 + \sum_{n=1}^{N} |\langle x \,|\, x_n \rangle|^2 \underbrace{\|x_n\|^2}_{1} - 2\Re \left(\sum_{n=1}^{N} \langle x \,|\, x_n \rangle^* \langle x \,|\, x_n \rangle \right) && \text{by property of } \|\cdot\| && \text{(Definition 5.1 page 83)} \\
&= \|x\|^2 + \sum_{n=1}^{N} |\langle x \,|\, x_n \rangle|^2 \cdot 1 - 2\Re \left(\sum_{n=1}^{N} \langle x \,|\, x_n \rangle^* \langle x \,|\, x_n \rangle \right) && \text{by def. of \textit{orthonormality}} && \text{(Definition 6.4 page 107)} \\
&= \|x\|^2 + \sum_{n=1}^{N} |\langle x \,|\, x_n \rangle|^2 - 2\Re \sum_{n=1}^{N} |\langle x \,|\, x_n \rangle|^2 \\
&= \|x\|^2 + \sum_{n=1}^{N} |\langle x \,|\, x_n \rangle|^2 - 2 \sum_{n=1}^{N} |\langle x \,|\, x_n \rangle|^2 && \text{because } |\cdot| \text{ is real}
\end{aligned}$$

[20] Bachman et al. (2000) page 103, Pedersen (2000) pages 38–39

$$= \|\boldsymbol{x}\|^2 - \sum_{n=1}^{N} |\langle \boldsymbol{x} \,|\, \boldsymbol{x}_n \rangle|^2$$

⇨

Theorem 10.7 (Bessel's inequality). [21] *Let* $\{ \boldsymbol{x}_n \in X | _{n=1,2,\ldots,N} \}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition 6.1 page 95)* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \nabla \rangle)$ *and with* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \,|\, \boldsymbol{x} \rangle}$ *(Definition 6.2 page 100).*

THM
$$\left\{ \begin{array}{c} \{\boldsymbol{x}_n\} \text{ is ORTHONORMAL} \\ \text{(Definition 6.4 page 107)} \end{array} \right\} \implies \left\{ \sum_{n=1}^{N} |\langle \boldsymbol{x} \,|\, \boldsymbol{x}_n \rangle|^2 \le \|\boldsymbol{x}\|^2 \quad \forall x \in X \right\}$$

✎PROOF:

$$\begin{aligned} 0 &\le \left\| \boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \,|\, \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 && \text{by definition of } \|\cdot\| && \text{(Definition 5.1 page 83)} \\ &= \|\boldsymbol{x}\|^2 - \sum_{n=1}^{N} |\langle \boldsymbol{x} \,|\, \boldsymbol{x}_n \rangle|^2 && \text{by } \textit{Bessel's Equality} && \text{(Theorem 10.6 page 159)} \end{aligned}$$

⇨

The *Best Approximation Theorem* (next) shows that

- the best sequence for representing a vector is the sequence of projections of the vector onto the sequence of basis functions
- the error of the projection is orthogonal to the projection.

Theorem 10.8 (Best Approximation Theorem). [22] *Let* $\{ \boldsymbol{x}_n \in X | _{n=1,2,\ldots,N} \}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition 6.1 page 95)* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \nabla \rangle)$ *and with* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \,|\, \boldsymbol{x} \rangle}$ *(Definition 6.2 page 100).*

THM
$$\left\{ \begin{array}{c} \{\boldsymbol{x}_n\} \text{ is} \\ \text{ORTHONORMAL} \\ \text{(Definition 6.4 page 107)} \end{array} \right\} \implies \left\{ \begin{array}{lll} 1. & \arg\min\limits_{(\alpha_n)_{n=1}^{N}} \left\| \boldsymbol{x} - \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n \right\| = \underbrace{(\langle \boldsymbol{x} \,|\, \boldsymbol{x}_n \rangle)_{n=1}^{N}}_{\text{best } \alpha_n = \langle \boldsymbol{x} | \boldsymbol{x}_n \rangle} & \forall x \in X \quad \text{and} \\ 2. & \underbrace{\left(\sum_{n=1}^{N} \langle \boldsymbol{x} \,|\, \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right)}_{\text{approximation}} \perp \underbrace{\left(\boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \,|\, \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right)}_{\text{approximation error}} & \forall x \in X \end{array} \right\}$$

✎PROOF:

[21] Giles (2000) pages 54–55 ⟨3.13 Bessel's inequality⟩, Bollobás (1999) page 147, Aliprantis and Burkinshaw (1998) page 284

[22] Walter and Shen (2001), pages 3–4, Pedersen (2000), page 39, Edwards (1995), pages 94–100, Weyl (1940)

1. Proof that $(\langle x \mid x_n \rangle)$ is the best sequence:

$$
\begin{aligned}
&\left\| x - \sum_{n=1}^{N} \alpha_n x_n \right\|^2 \\
&= \|x\|^2 - 2\Re\left\langle x \mid \sum_{n=1}^{N} \alpha_n x_n \right\rangle + \left\| \sum_{n=1}^{N} \alpha_n x_n \right\|^2 \\
&= \|x\|^2 - 2\Re\left(\sum_{n=1}^{N} \alpha_n^* \langle x \mid x_n \rangle \right) + \sum_{n=1}^{N} \|\alpha_n x_n\|^2 && \text{by Pythagorean Theorem} && \text{(Theorem 6.10 page 108)} \\
&= \|x\|^2 - 2\Re\left(\sum_{n=1}^{N} \alpha_n^* \langle x \mid x_n \rangle \right) + \sum_{n=1}^{N} |\alpha_n|^2 + \underbrace{\left[\sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 - \sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 \right]}_{0} \\
&= \left[\|x\|^2 - \sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 \right] + \sum_{n=1}^{N} \left[|\langle x \mid x_n \rangle|^2 - 2\mathbf{R_e}\left[\alpha_n^* \langle x \mid x_n \rangle\right] + |\alpha_n|^2 \right] \\
&= \left[\|x\|^2 - \sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 \right] + \sum_{n=1}^{N} \left[|\langle x \mid x_n \rangle|^2 - \alpha_n^* \langle x \mid x_n \rangle - \alpha_n \langle x \mid x_n \rangle^* + |\alpha_n|^2 \right] \\
&= \left\| x - \sum_{n=1}^{N} \langle x \mid x_n \rangle x_n \right\|^2 + \sum_{n=1}^{N} |\langle x \mid x_n \rangle - \alpha_n|^2 && \text{by Bessel's Equality} && \text{(Theorem 10.6 page 159)} \\
&\geq \left\| x - \sum_{n=1}^{N} \langle x \mid x_n \rangle x_n \right\|^2
\end{aligned}
$$

2. Proof that the approximation and approximation error are orthogonal:

$$
\begin{aligned}
\left\langle \sum_{n=1}^{N} \langle x \mid x_n \rangle x_n \mid x - \sum_{n=1}^{N} \langle x \mid x_n \rangle x_n \right\rangle
&= \left\langle \sum_{n=1}^{N} \langle x \mid x_n \rangle x_n \mid x \right\rangle - \left\langle \sum_{n=1}^{N} \langle x \mid x_n \rangle x_n \mid \sum_{n=1}^{N} \langle x \mid x_n \rangle x_n \right\rangle \\
&= \sum_{n=1}^{N} \langle x \mid x_n \rangle^* \langle x \mid x_n \rangle - \sum_{n=1}^{N} \sum_{m=1}^{N} \langle x \mid x_n \rangle \langle x \mid x_m \rangle^* \langle x_n \mid x_m \rangle \\
&= \sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 - \sum_{n=1}^{N} \sum_{m=1}^{N} \langle x \mid x_n \rangle \langle x \mid x_m \rangle^* \bar{\delta}_{nm} \\
&= \sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 - \sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 \\
&= 0
\end{aligned}
$$

✏

10.5 Orthonormal bases in Hilbert spaces

Definition 10.9. *Let* $\left\{ x_n \in X \mid n=1,2,\ldots,N \right\}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition 6.1 page 95)* $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla \rangle\right)$.

DEF

The set $\{x_n\}$ *is an* ***orthogonal basis*** *for* $\boldsymbol{\Omega}$ *if* $\{x_n\}$ *is* ORTHOGONAL *and is a* SCHAUDER BASIS *for* $\boldsymbol{\Omega}$.

The set $\{x_n\}$ *is an* ***orthonormal basis*** *for* $\boldsymbol{\Omega}$ *if* $\{x_n\}$ *is* ORTHONORMAL *and is a* SCHAUDER BASIS *for* $\boldsymbol{\Omega}$.

Definition 10.10. [23] *Let* $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be a Hilbert space.*

DEF

Suppose there exists a set $\left\{ x_n \in X \,\middle|\, n\in\mathbb{N} \right\}$ *such that* $x \doteq \sum_{n=1}^{\infty} \langle x \mid x_n \rangle x_n$.

Then the quantities $\langle x \mid x_n \rangle$ *are called the* ***Fourier coefficients*** *of* x *and the sum* $\sum_{n=1}^{\infty} \langle x \mid x_n \rangle x_n$ *is called the* ***Fourier expansion*** *of* x *or the* ***Fourier series*** *for* x.

Lemma 10.2 (Perfect reconstruction). *Let* $\left\{ x_n \in X \,\middle|\, n\in\mathbb{N} \right\}$ *be a set of vectors in a* HILBERT SPACE *(Definition 8.11 page 138)* $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$.

LEM

$$\left\{ \begin{array}{lll} (1). & (\!(x_n)\!) \text{ is a BASIS for } \boldsymbol{H} & \text{and} \\ (2). & (\!(x_n)\!) \text{ is ORTHONORMAL} & \end{array} \right\} \implies x \doteq \overbrace{\sum_{n=1}^{\infty} \underbrace{\langle x \mid x_n \rangle}_{\text{Fourier coefficient}} x_n}^{\text{Fourier expansion}} \quad \forall x\in X$$

✎PROOF:

$$\begin{aligned}
\langle x \mid x_n \rangle &= \left\langle \sum_{m\in\mathbb{Z}} \alpha_m x_m \,\middle|\, x_n \right\rangle && \text{by left hypothesis (1)} \\
&= \sum_{m\in\mathbb{Z}} \alpha_m \langle x_m \mid x_n \rangle && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \mid \triangledown \rangle && \text{(Definition 6.1 page 95)} \\
&= \sum_{m\in\mathbb{Z}} \alpha_m \bar{\delta}_{n-m} && \text{by left hypothesis (2)} && \text{(Definition 6.4 page 107)} \\
&= \alpha_n
\end{aligned}$$

✏

Proposition 10.2. [24] *Let* $\left\{ x_n \in X \,\middle|\, n\in\mathbb{N} \right\}$ *be a set of vectors in a* HILBERT SPACE *(Definition 8.11 page 138)* $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$.

PRP

$$\underbrace{\|x\|^2 \doteq \sum_{n=1}^{\infty} \left| \langle x \mid x_n \rangle \right|^2}_{\text{PARSEVAL FRAME}} \iff \underbrace{x \doteq \sum_{n=1}^{\infty} \langle x \mid x_n \rangle x_n}_{\text{FOURIER EXPANSION (Definition 10.10 page 162)}} \quad \forall x\in X$$

✎PROOF:

[23] Fabian et al. (2010) page 27 ⟨Theorem 1.55⟩, Young (2001) page 6, Young (1980) page 6

[24] Han et al. (2007) pages 93–94 ⟨Proposition 3.11⟩

1. Proof that *Parseval frame* $\Longleftarrow$ *Fourier expansion*

$$
\begin{aligned}
\|x\|^2 &\triangleq \langle x \mid x \rangle && \text{by definition of } \|\cdot\| && \text{(Definition 5.1 page 83)}\\
&= \left\langle \sum_{n=1}^{\infty} \langle x \mid x_n \rangle x \mid x_n \right\rangle && \text{by right hypothesis}\\
&\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle x \mid x_n \rangle \langle x \mid x_n \rangle && \text{by property of } \langle \triangle \mid \nabla \rangle && \text{(Definition 6.1 page 95)}\\
&\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle x \mid x_n \rangle \langle x \mid x_n \rangle^* && \text{by property of } \langle \triangle \mid \nabla \rangle && \text{(Definition 6.1 page 95)}\\
&\stackrel{\star}{=} \sum_{n=1}^{\infty} |\langle x \mid x_n \rangle|^2 && \text{by property of } \mathbb{C} && \text{(Definition 16.7 page 261)}
\end{aligned}
$$

2. Proof that *Parseval frame* $\Longrightarrow$ *Fourier expansion*

   (a) Let $(\!(e_n)\!)_{n\in\mathbb{N}}$ be the *standard othornormal basis* such that the nth element of e_n is 1 and all other elements are 0.

   (b) Let $\mathbf{M}$ be an operator in $\boldsymbol{H}$ such that $\mathbf{M}x \triangleq \sum_{n=1}^{\infty} \langle x \mid x_n \rangle e_n$.

   (c) lemma: $\mathbf{M}$ is *isometric*. Proof:

$$
\begin{aligned}
\|\mathbf{M}x\|^2 &= \left\| \sum_{n=1}^{\infty} \langle x \mid x_n \rangle e_n \right\|^2 && \text{by definition of } \mathbf{M} && \text{(item (2b) page 163)}\\
&= \sum_{n=1}^{\infty} \|\langle x \mid x_n \rangle e_n\|^2 && \text{by \textit{Pythagorean Theorem}} && \text{(Theorem 6.10 page 108)}\\
&= \sum_{n=1}^{\infty} |\langle x \mid x_n \rangle|^2 \, \|e_n\|^2 && \text{by \textit{homogeneous} property of } \|\cdot\| && \text{(Definition 5.1 page 83)}\\
&= \sum_{n=1}^{\infty} |\langle x \mid x_n \rangle|^2 && \text{by definition of \textit{orthonormal}} && \text{(Definition 6.4 page 107)}\\
&= \|x\|^2 && \text{by Parseval frame hypothesis}\\
\Longrightarrow &\quad \mathbf{M} \text{ is \textit{isometric}} && \text{by definition of \textit{isometric}} && \text{(Definition 14.10 page 232)}
\end{aligned}
$$

   (d) Let $(\!(u_n)\!)_{n\in\mathbb{N}}$ be an *orthornormal basis* for $\boldsymbol{H}$.

   (e) Proof for *Fourier expansion*:

$$
\begin{aligned}
x &= \sum_{n=1}^{\infty} \langle x \mid u_n \rangle u_n && \text{by \textit{Fourier expansion} (Proposition 10.3 page 166)}\\
&= \sum_{n=1}^{\infty} \langle \mathbf{M}x \mid \mathbf{M}u_n \rangle u_n && \text{by (2c) lemma page 163 and Theorem 14.21 page 233}\\
&= \sum_{n=1}^{\infty} \left\langle \sum_{m=1}^{\infty} \langle x \mid x_m \rangle e_m \mid \sum_{k=1}^{\infty} \langle u_n \mid x_k \rangle e_k \right\rangle u_n && \text{by item (2b) page 163}\\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \langle x \mid x_m \rangle \sum_{k=1}^{\infty} \langle u_n \mid x_k \rangle^* \langle e_m \mid e_k \rangle u_n && \text{by Definition 6.1 page 95}\\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \langle x \mid x_m \rangle \langle u_n \mid x_m \rangle^* u_n && \text{by item (2a) page 163 and Definition 6.4 page 107}
\end{aligned}
$$

$$= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \langle x \mid x_m \rangle \langle x_m \mid u_n \rangle u_n \qquad \text{by Definition 6.1 page 95}$$

$$= \sum_{m=1}^{\infty} \langle x \mid x_m \rangle \sum_{n=1}^{\infty} \langle x_m \mid u_n \rangle u_n$$

$$= \sum_{m=1}^{\infty} \langle x \mid x_m \rangle x_m \qquad \text{by item (2d) page 163}$$

✏

When is a set of orthonormal vectors in a Hilbert space $\boldsymbol{H}$ *total*? Theorem 10.9 (next) offers some help.

Theorem 10.9 (The Fourier Series Theorem). [25] *Let* $\{x_n \in X \mid n \in \mathbb{N}\}$ *be a set of vectors in a* HILBERT SPACE *(Definition 8.11 page 138)* $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *and let* $\|x\| \triangleq \sqrt{\langle x \mid x \rangle}$ *(Definition 6.2 page 100).*

THM

(A) $\{x_n\}$ *is* ORTHONORMAL *in* $\boldsymbol{H}$ $\implies$

$$\left\{\begin{array}{llrcll}
 & (1). & (\operatorname{span}\{x_n\})^- & = & \boldsymbol{H} & & (\{x_n\} \text{ is TOTAL in } \boldsymbol{H}) \\
\iff & (2). & \langle x \mid y \rangle & \stackrel{\star}{=} & \displaystyle\sum_{n=1}^{\infty} \langle x \mid x_n \rangle \langle y \mid x_n \rangle^* & \forall x,y \in X & (\text{GENERALIZED PARSEVAL'S IDENTITY}) \\
\iff & (3). & \|x\|^2 & \stackrel{\star}{=} & \displaystyle\sum_{n=1}^{\infty} |\langle x \mid x_n \rangle|^2 & \forall x \in X & (\text{PARSEVAL'S IDENTITY}) \\
\iff & (4). & x & \stackrel{\star}{=} & \displaystyle\sum_{n=1}^{\infty} \langle x \mid x_n \rangle x_n & \forall x \in X & (\text{FOURIER SERIES EXPANSION})
\end{array}\right\}$$

✎PROOF:

1. Proof that (1) $\implies$ (2):

$$\begin{aligned}
\langle x \mid y \rangle &\stackrel{\star}{=} \left\langle \sum_{n=1}^{\infty} \langle x \mid x_n \rangle x_n \;\middle|\; \sum_{m=1}^{\infty} \langle y \mid x_m \rangle x_m \right\rangle && \text{by (A) and (1)} \\
&\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle x \mid x_n \rangle \left\langle x_n \;\middle|\; \sum_{m=1}^{\infty} \langle y \mid x_m \rangle x_m \right\rangle && \text{by property of } \langle \cdot \mid \cdot \rangle \quad \text{(Definition 6.1 page 95)} \\
&\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle x \mid x_n \rangle \sum_{m=1}^{\infty} \langle y \mid x_m \rangle^* \langle x_n \mid x_m \rangle && \text{by property of } \langle \cdot \mid \cdot \rangle \quad \text{(Definition 6.1 page 95)} \\
&\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle x \mid x_n \rangle \sum_{m=1}^{\infty} \langle y \mid x_m \rangle^* \bar{\delta}_{mn} && \text{by (A)} \\
&\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle x \mid x_n \rangle \langle y \mid x_n \rangle^* && \text{by definition of } \bar{\delta}_n \quad \text{(Definition 6.3 page 107)}
\end{aligned}$$

[25] Bachman and Narici (1966) pages 149–155 ⟨Theorem 9.12⟩, Kubrusly (2001) pages 360–363 ⟨Theorem 5.48⟩, Aliprantis and Burkinshaw (1998) pages 298–299 ⟨Theorem 34.2⟩, Christensen (2003) page 57 ⟨Theorem 3.4.2⟩, Berberian (1961) pages 52–53 ⟨Theorem II§8.3⟩, Heil (2011) pages 34–35 ⟨Theorem 1.50⟩, Bracewell (1978) page 112 ⟨Rayleigh's theorem⟩

2. Proof that (2) $\Longrightarrow$ (3):

$$\begin{aligned}
\|x\|^2 &\triangleq \langle x \mid x \rangle && \text{by definition of } \textit{induced norm} && \text{(Theorem 6.4 page 100)}\\
&= \sum_{n=1}^{\infty} \langle x \mid x_n \rangle \langle x \mid x_n \rangle^* && \text{by (2)}\\
&= \sum_{n=1}^{\infty} |\langle x \mid x_n \rangle|^2
\end{aligned}$$

3. Proof that (3) $\Longleftrightarrow$ (4) *not* using (A): by Proposition 10.2 page 162

4. Proof that (3) $\Longrightarrow$ (1) (proof by contradiction):

   (a) Suppose $\{x_n\}$ is *not total*.

   (b) Then there must exist a vector y in H such that the set $B \triangleq \{x_n\} \cup y$ is *orthonormal*.

   (c) Then $1 = \|y\|^2 \neq \sum_{n=1}^{\infty} |\langle y \mid x_n \rangle|^2 = 0$.

   (d) But this contradicts (3), and so $\{x_n\}$ must be *total* and (3) $\Longrightarrow$ (1).

5. Extraneous proof that (3) $\Longrightarrow$ (4) (this proof is not really necessary here):

$$\begin{aligned}
\left\| x - \sum_{n=1}^{\infty} \langle x \mid x_n \rangle x_n \right\|^2 &= \|x\|^2 - \sum_{n=1}^{\infty} |\langle x \mid x_n \rangle|^2 && \text{by } \textit{Bessel's Equality} && \text{(Theorem 10.6 page 159)}\\
&= 0 && \text{by (3)}\\
\Longrightarrow \quad & x \doteq \sum_{n=1}^{\infty} \langle x \mid x_n \rangle x_n && \text{by definition of } \doteq && \text{(Definition 12.5 page 203)}
\end{aligned}$$

6. Extraneous proof that (A) $\Longrightarrow$ (4) (this proof is not really necessary here)

   (a) The sequence $\sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2$ is *monotonically increasing* in n.

   (b) By Bessel's inequality (page 160), the sequence is upper bounded by $\|x\|^2$:

$$\sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 \leq \|x\|^2$$

   (c) Because this sequence is both monotonically increasing and bounded in n, it must equal its bound in the limit as n approaches infinity:

$$\lim_{N\to\infty} \sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 = \|x\|^2 \tag{10.1}$$

   (d) If we combine this result with *Bessel's Equality* (Theorem 10.6 page 159) we have

$$\begin{aligned}
\lim_{N\to\infty} \left\| x - \sum_{n=1}^{N} \langle x \mid x_n \rangle x_n \right\|^2 &= \|x\|^2 - \lim_{N\to\infty} \sum_{n=1}^{N} |\langle x \mid x_n \rangle|^2 && \text{by } \textit{Bessel's Equality} \text{ (Theorem 10.6 page 159)}\\
&= \|x\|^2 - \|x\|^2 && \text{by equation (10.1) page 165}\\
&= 0
\end{aligned}$$

⇨

Proposition 10.3 (Fourier expansion). *Let* $\{ x_n \in X | n \in \mathbb{N} \}$ *be a set of vectors in a* HILBERT SPACE *(Definition 8.11 page 138)* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$.

PRP

$$\underbrace{\{x_n\} \text{ is an ORTHONORMAL BASIS for } \boldsymbol{H}}_{(A)} \implies \left\{ \underbrace{x \doteq \sum_{n=1}^{\infty} \alpha_n x_n}_{(1)} \iff \underbrace{\alpha_n = \langle x \mid x_n \rangle}_{(2)} \right\}$$

✎PROOF:

1. Proof that (1) $\implies$ (2): by Lemma 10.2 page 162
2. Proof that (1) $\impliedby$ (2):

$$\begin{aligned}
\left\| x - \sum_{n=1}^{\infty} \alpha_n x_{n \in \mathbb{Z}} \right\|^2 &= \left\| x - \sum_{n=1}^{\infty} \langle x \mid x_n \rangle \, x_{n \in \mathbb{Z}} \right\|^2 && \text{by right hypothesis} \\
&= \|x\|^2 - \sum_{n=1}^{\infty} |\langle x \mid x_n \rangle|^2 && \text{by Bessel's equality} && \text{(Theorem 10.6 page 159)} \\
&= 0 && \text{by } \textit{Parseval's Identity} && \text{(Theorem 10.9 page 164)} \\
&\overset{\text{def}}{\iff} \quad x \doteq \sum_{n=1}^{\infty} \langle x \mid x_n \rangle \, x_n && \text{by definition of } \textit{strong convergence} && \text{(Definition 12.5 page 203)}
\end{aligned}$$

⇰

Proposition 10.4 (Riesz-Fischer Theorem). [26] *Let* $\{ x_n \in X | n \in \mathbb{N} \}$ *be a set of vectors in a* HILBERT SPACE *(Definition 8.11 page 138)* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$.

PRP

$$\underbrace{\{x_n\} \text{ is an ORTHONORMAL BASIS for } \boldsymbol{H}}_{(A)} \implies \left\{ \underbrace{\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty}_{(1)} \iff \underbrace{\exists x \in \boldsymbol{H} \text{ such that } \alpha_n = \langle x \mid x_n \rangle}_{(2)} \right\}$$

✎PROOF:

1. Proof that (1) $\implies$ (2):
   (a) If (1) is true, then let $x \triangleq \sum_{n \in \mathbb{N}} \alpha_n x_n$.
   (b) Then

$$\begin{aligned}
\langle x \mid x_n \rangle &= \left\langle \sum_{m \in \mathbb{N}} \alpha_m x_m \mid x_n \right\rangle && \text{by definition of } x \\
&= \sum_{m \in \mathbb{N}} \alpha_m \langle x_m \mid x_n \rangle && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \mid \triangledown \rangle && \text{(Definition 6.1 page 95)} \\
&= \sum_{m \in \mathbb{N}} \alpha_m \bar{\delta}_{mn} && \text{by (A)} \\
&= \sum_{m \in \mathbb{N}} \alpha_n && \text{by definition of } \bar{\delta} && \text{(Definition 6.3 page 107)}
\end{aligned}$$

[26] Young (2001) page 6

2. Proof that (1) $\Longleftarrow$ (2):

$$\begin{aligned}
\sum_{n\in\mathbb{N}} |\alpha_n|^2 &= \sum_{n\in\mathbb{N}} |\langle x \mid x_n\rangle|^2 && \text{by (2)} \\
&\le \|x\|^2 && \text{by } \textit{Bessel's Inequality} && \text{(Theorem 10.7 page 160)} \\
&\le \infty
\end{aligned}$$

Theorem 10.10. [27]

THM *All* SEPARABLE HILBERT SPACES *are* ISOMORPHIC. *That is,*

$$\left\{\begin{array}{ll} \boldsymbol{X} \text{ is a separable} & \\ \text{Hilbert space} & \text{and} \\ \boldsymbol{Y} \text{ is a separable} & \\ \text{Hilbert space} & \end{array}\right\} \implies \left\{\begin{array}{lllll} \multicolumn{5}{l}{\text{there is a BIJECTIVE operator } \mathbf{M} \in \boldsymbol{Y}^{\boldsymbol{X}} \text{ such that}} \\ \text{(1).} & \boldsymbol{y} = \mathbf{M}\boldsymbol{x} & \forall x\in\boldsymbol{X}, y\in\boldsymbol{Y} & \text{and} \\ \text{(2).} & \|\mathbf{M}\boldsymbol{x}\| = \|\boldsymbol{x}\| & \forall x\in\boldsymbol{X} & \text{and} \\ \text{(3).} & \langle \mathbf{M}\boldsymbol{x} \mid \mathbf{M}\boldsymbol{y}\rangle = \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle & \forall x\in\boldsymbol{X}, y\in\boldsymbol{Y} & \end{array}\right\}$$

✎PROOF:

1. Let $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}), \langle\triangle \mid \nabla\rangle)$ be a *separable Hilbert space* with *orthonormal basis* $\{x_n|\, n\in\mathbb{N}\}$.
   Let $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dotplus, \dot{\times}), \langle\triangle \mid \nabla\rangle)$ be a *separable Hilbert space* with *orthonormal basis* $\{y_n|\, n\in\mathbb{N}\}$.

2. Proof that there exists *bijective* operator $\mathbf{M}$ and its inverse $\mathbf{M}^{-1}$ between $\{x_n\}$ and $\{y_n\}$:

   (a) Let $\mathbf{M}$ be defined such that $y_n \triangleq \mathbf{M}x_n$.

   (b) Thus $\mathbf{M}$ is a *bijection* between $\{x_n\}$ and $\{y_n\}$.

   (c) Because $\mathbf{M}$ is a *bijection* between $\{x_n\}$ and $\{y_n\}$, $\mathbf{M}$ has an inverse operator $\mathbf{M}^{-1}$ between $\{x_n\}$ and $\{y_n\}$ such that $x_n = \mathbf{M}^{-1}y_n$.

3. Proof that $\mathbf{M}$ and $\mathbf{M}^{-1}$ are *bijective* operators between $\boldsymbol{X}$ and $\boldsymbol{Y}$:

   (a) Proof that $\mathbf{M}$ maps $\boldsymbol{X}$ *into* $\boldsymbol{Y}$:

$$\begin{aligned}
x \in \boldsymbol{X} &\iff x \stackrel{*}{=} \sum_{n\in\mathbb{N}} \langle x \mid x_n\rangle x_n && \text{by } \textit{Fourier expansion} && \text{(Theorem 10.9 page 164)} \\
&\implies \exists y \in \boldsymbol{Y} \quad \text{such that} \quad \langle y \mid y_n\rangle = \langle x \mid x_n\rangle && \text{by } \textit{Riesz-Fischer Thm.} && \text{(Proposition 10.4 page 166)} \\
&\implies \\
y &= \sum_{n\in\mathbb{N}} \langle y \mid y_n\rangle y_n && \text{by } \textit{Fourier expansion} && \text{(Theorem 10.9 page 164)} \\
&= \sum_{n\in\mathbb{N}} \langle x \mid x_n\rangle y_n && \text{by } \textit{Riesz-Fischer Thm.} && \text{(Proposition 10.4 page 166)} \\
&= \sum_{n\in\mathbb{N}} \langle x \mid x_n\rangle \mathbf{M}x_n && \text{by definition of } \mathbf{M} && \text{(item (2a) page 167)} \\
&= \mathbf{M}\sum_{n\in\mathbb{N}} \langle x \mid x_n\rangle x_n && \text{by prop. of linear ops.} && \text{(Theorem 14.1 page 214)} \\
&= \mathbf{M}x && \text{by definition of } x
\end{aligned}$$

[27] Young (2001) page 6

(b) Proof that $\mathbf{M}^{-1}$ maps $\boldsymbol{Y}$ *into* $\boldsymbol{X}$:

$$\begin{aligned}
\boldsymbol{y} \in \boldsymbol{Y} &\iff \boldsymbol{y} \stackrel{\star}{=} \sum_{n\in\mathbb{N}} \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle \boldsymbol{y}_n && \text{by \textit{Fourier expansion}} && \text{(Theorem 10.9 page 164)}\\
&\implies \exists \boldsymbol{x} \in \boldsymbol{X} \quad \text{such that} \quad \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle = \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle && \text{by \textit{Riesz-Fischer Thm.}} && \text{(Proposition 10.4 page 166)}\\
&\implies\\
\boldsymbol{x} &= \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n && \text{by \textit{Fourier expansion}} && \text{(Theorem 10.9 page 164)}\\
&= \sum_{n\in\mathbb{N}} \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle \boldsymbol{x}_n && \text{by \textit{Riesz-Fischer Thm.}} && \text{(Proposition 10.4 page 166)}\\
&= \sum_{n\in\mathbb{N}} \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle \mathbf{M}^{-1}\boldsymbol{y}_n && \text{by definition of } \mathbf{M}^{-1} && \text{(item (2c) page 167)}\\
&= \mathbf{M}^{-1} \sum_{n\in\mathbb{N}} \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle \boldsymbol{y}_n && \text{by prop. of \textit{linear ops.}} && \text{(Theorem 14.1 page 214)}\\
&= \mathbf{M}^{-1}\boldsymbol{y} && \text{by definition of } \boldsymbol{y}
\end{aligned}$$

4. Proof for (2):

$$\begin{aligned}
\|\mathbf{M}\boldsymbol{x}\|^2 &= \left\| \mathbf{M} \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 && \text{by \textit{Fourier expansion}} && \text{(Theorem 10.9 page 164)}\\
&= \left\| \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \mathbf{M}\boldsymbol{x}_n \right\|^2 && \text{by property of \textit{linear operators}} && \text{(Theorem 14.1 page 214)}\\
&= \left\| \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{y}_n \right\|^2 && \text{by definition of } \mathbf{M} && \text{(item (2a) page 167)}\\
&= \sum_{n\in\mathbb{N}} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 && \text{by \textit{Parseval's Identity}} && \text{(Proposition 10.4 page 166)}\\
&= \left\| \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 && \text{by \textit{Parseval's Identity}} && \text{(Proposition 10.4 page 166)}\\
&= \|\boldsymbol{x}\|^2 && \text{by \textit{Fourier expansion}} && \text{(Theorem 10.9 page 164)}
\end{aligned}$$

5. Proof for (3): by (2) and Theorem 14.21 page 233

✏

Theorem 10.11. [28] *Let* $\boldsymbol{H}$ *be a* HILBERT SPACE *(Definition 8.11 page 138)*.

THM			
	$\boldsymbol{H}$ *has a* SCHAUDER BASIS	$\iff$	$\boldsymbol{H}$ *is* SEPARABLE

Theorem 10.12. [29] *Let* $\boldsymbol{H}$ *be a* HILBERT SPACE.

THM			
	$\boldsymbol{H}$ *has an* ORTHONORMAL BASIS	$\iff$	$\boldsymbol{H}$ *is* SEPARABLE

[28] Bachman et al. (2000) page 112 ⟨3.4.8⟩, Berberian (1961) page 53 ⟨Theorem II§8.3⟩
[29] Kubrusly (2001) page 357 ⟨Proposition 5.43⟩

10.6 Riesz bases in Hilbert spaces

Definition 10.11. [30] *Let* $\{x_n \in X \mid n\in\mathbb{N}\}$ *be a set of vectors in a* SEPARABLE HILBERT SPACE (Definition 8.11 page 138) $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$.

DEF $\{x_n\}$ *is a* ***Riesz basis*** *for* $\boldsymbol{H}$ *if* $\{x_n\}$ *is* EQUIVALENT (Definition 10.8 page 157) *to some* ORTHONORMAL BASIS (Definition 10.9 page 161) *in* $\boldsymbol{H}$.

Definition 10.12. [31] *Let* $(x_n \in X)_{n\in\mathbb{N}}$ *be a sequence of vectors in a* SEPARABLE HILBERT SPACE (Definition 8.11 page 138)
$\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$.

DEF *The sequence* (x_n) *is a* ***Riesz sequence*** *for* $\boldsymbol{H}$ *if*

$$\exists A, B \in \mathbb{R}^+ \quad \text{such that} \quad A\sum_{n=1}^{\infty}|\alpha_n|^2 \le \left\|\sum_{n=1}^{\infty}\alpha_n x_n\right\|^2 \le B\sum_{n=1}^{\infty}|\alpha_n|^2 \qquad \forall(\alpha_n)\in\ell^2_{\mathbb{F}}.$$

Definition 10.13. *Let* $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ *be an* INNER PRODUCT SPACE (Definition 6.1 page 95).

DEF *The sequences* $(x_n \in X)_{n\in\mathbb{Z}}$ *and* $(y_n \in X)_{n\in\mathbb{Z}}$ *are* ***biorthogonal*** *with respect to each other in* $\boldsymbol{X}$ *if* $\quad \langle x_n \mid y_m \rangle = \bar{\delta}_{nm}$

Lemma 10.3. [32] *Let* $\{x_n \mid n\in\mathbb{N}\}$ *be a sequence in a* HILBERT SPACE $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$. *Let* $\{y_n \mid n\in\mathbb{N}\}$ *be a sequence in a* HILBERT SPACE $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ (Definition 8.11 page 138). *Let*

LEM

$$\left\{\begin{array}{lll}
(i). & \{x_n\} \text{ is TOTAL in } \boldsymbol{X} & \text{and}\\
(ii). & \text{There exists } A>0 \text{ such that } A\sum_{n\in C}|a_n|^2 \le \left\|\sum_{n\in C} a_n x_n\right\|^2 \quad \text{for finite } C & \text{and}\\
(iii). & \text{There exists } B>0 \text{ such that } \left\|\sum_{n=1}^{\infty} b_n y_n\right\|^2 \le B\sum_{n=1}^{\infty}|b_n|^2 \quad \forall(b_n)_{n\in\mathbb{N}}\in\ell^2_{\mathbb{F}} &
\end{array}\right\} \implies$$

$$\left\{\begin{array}{lll}
(1). & \mathbf{R}^\circ \text{ is a linear bounded operator that maps from } \operatorname{span}\{x_n\} \text{ to } \operatorname{span}\{y_n\} & \\
 & \quad \text{where } \mathbf{R}^\circ\sum_{n\in C} c_n x_n \triangleq \sum_{n\in C} c_n y_n, \text{ for some sequence } (c_n) \text{ and finite set } C & \text{and}\\
(2). & \mathbf{R} \text{ has a unique extension to a bounded operator } \mathbf{R} \text{ that maps from } \boldsymbol{X} \text{ to } \boldsymbol{Y} & \text{and}\\
(3). & |||\mathbf{R}^\circ||| \le \frac{B}{A} & \text{and}\\
(4). & |||\mathbf{R}||| \le \frac{B}{A} &
\end{array}\right\}$$

Theorem 10.13. [33] *Let* $\{x_n \in X \mid n\in\mathbb{N}\}$ *be a set of vectors in a* SEPARABLE HILBERT SPACE (Definition 8.11 page 138)
$\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$.

THM

$$\left\{\begin{array}{l}\{x_n\} \text{ is a RIESZ BASIS}\\ \text{for } \boldsymbol{H}\end{array}\right\} \iff \left\{\begin{array}{ll}
(1). \ \{x_n\} \text{ is TOTAL in } \boldsymbol{H} & \text{and}\\
(2). \ \exists A,B\in\mathbb{R}^+ \quad \text{such that} \quad \forall(\alpha_n)\in\ell^2_{\mathbb{F}}, & \\
\quad A\sum_{n=1}^{\infty}|\alpha_n|^2 \le \left\|\sum_{n=1}^{\infty}\alpha_n x_n\right\|^2 \le B\sum_{n=1}^{\infty}|\alpha_n|^2 &
\end{array}\right\}$$

---

[30] Young (2001) page 27 ⟨Definition 1.8.2⟩, Christensen (2003) page 63 ⟨Definition 3.6.1⟩, Heil (2011) page 196 ⟨Definition 7.9⟩

[31] Christensen (2003) pages 66–68 ⟨page 68 and (3.24) on page 66⟩, Wojtaszczyk (1997) page 20 ⟨Definition 2.6⟩

[32] Christensen (2003) pages 65–66 ⟨Lemma 3.6.5⟩

[33] Young (2001) page 27 ⟨Theorem 1.8.9⟩, Christensen (2003) page 66 ⟨Theorem 3.6.6⟩, Heil (2011) pages 197–198 ⟨Theorem 7.13⟩, Christensen (2008) pages 61–62 ⟨Theorem 3.3.7⟩

✎PROOF:

1. Proof for ($\Longrightarrow$) case:
   (a) Proof that *Riesz basis* hypothesis $\Longrightarrow$ (1): all bases for $\boldsymbol{H}$ are *total* in $\boldsymbol{H}$.
   (b) Proof that *Riesz basis* hypothesis $\Longrightarrow$ (2):
      i. Let $(\boldsymbol{u}_n)_{n\in\mathbb{N}}$ be an *orthonormal basis* for $\boldsymbol{H}$.
      ii. Let $\mathbf{R}$ be a *bounded bijective* operator such that $\boldsymbol{x}_n = \mathbf{R}\boldsymbol{u}_n$.
      iii. Proof for upper bound B:

$$
\begin{aligned}
\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 &= \left\|\sum_{n=1}^{\infty}\alpha_n\mathbf{R}\boldsymbol{u}_n\right\|^2 && \text{by definition of } \mathbf{R} && \text{(item (1(b)ii))}\\
&= \left\|\mathbf{R}\sum_{n=1}^{\infty}\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by Theorem 14.1 page 214}\\
&\le |||\mathbf{R}|||^2\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by Theorem 14.6 page 220}\\
&= |||\mathbf{R}|||^2\sum_{n=1}^{\infty}\|\alpha_n\boldsymbol{u}_n\|^2 && \text{by \textit{Pythagorean Theorem}} && \text{(Theorem 6.10 page 108)}\\
&= |||\mathbf{R}|||^2\sum_{n=1}^{\infty}|\alpha|^2\,\|\boldsymbol{u}_n\|^2 && \text{by \textit{homogeneous} property of norms} && \text{(Definition 5.1 page 83)}\\
&= \underbrace{|||\mathbf{R}|||^2}_{B}\sum_{n=1}^{\infty}|\alpha|^2 && \text{by definition of \textit{orthonormality}} && \text{(Definition 6.4 page 107)}
\end{aligned}
$$

      iv. Proof for lower bound A:

$$
\begin{aligned}
\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 &= \frac{|||\mathbf{R}^{-1}|||^2}{|||\mathbf{R}^{-1}|||^2}\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 && \text{because } |||\mathbf{R}^{-1}|||>0 && \text{(Proposition 14.1 page 218)}\\
&\ge \frac{1}{|||\mathbf{R}^{-1}|||^2}\left\|\mathbf{R}^{-1}\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 && \text{by Theorem 14.6 page 220}\\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\left\|\mathbf{R}^{-1}\sum_{n=1}^{\infty}\alpha_n\mathbf{R}\boldsymbol{u}_n\right\|^2 && \text{by definition of } \mathbf{R} && \text{(item (1(b)ii) page 170)}\\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\left\|\mathbf{R}^{-1}\mathbf{R}\sum_{n=1}^{\infty}\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by property of \textit{linear operators}} && \text{(Theorem 14.1 page 214)}\\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by definition of inverse op.} && \text{(Definition 14.2 page 213)}\\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\sum_{n=1}^{\infty}\|\alpha_n\boldsymbol{u}_n\|^2 && \text{by \textit{Pythagorean Theorem}} && \text{(Theorem 6.10 page 108)}\\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\sum_{n=1}^{\infty}|\alpha_n|^2\,\|\boldsymbol{u}_n\|^2 && \text{by } \|\cdot\| \text{ \textit{homogeneous} prop.} && \text{(Definition 5.1 page 83)}\\
&= \underbrace{\frac{1}{|||\mathbf{R}^{-1}|||^2}}_{A}\sum_{n=1}^{\infty}|\alpha_n|^2 && \text{by def. of \textit{orthonormality}} && \text{(Definition 6.4 page 107)}
\end{aligned}
$$

2. Proof for ($\Longrightarrow$) case:

   (a) Let $\{u_n | _{n\in\mathbb{N}}\}$ be an *orthonormal basis* for $\boldsymbol{H}$.

   (b) Using (2) and Lemma 10.3 (page 169), construct an bounded extension operator $\mathbf{R}$ such that $\mathbf{R}u_n = x_n$ for all $n \in \mathbb{N}$.

   (c) Using (2) and Lemma 10.3 (page 169), construct an bounded extension operator $\mathbf{S}$ such that $\mathbf{S}x_n = u_n$ for all $n \in \mathbb{N}$.

   (d) Then, $\mathbf{RV}x = \mathbf{VR}x \implies \mathbf{V} = \mathbf{R}^{-1}$, and so $\mathbf{R}$ is a bounded invertible operator

   (e) and $\{x_n\}$ is a *Riesz sequence*.

⇰

Theorem 10.14. [34] *Let* $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \nabla \rangle)$ *be a* SEPARABLE HILBERT SPACE *(Definition 8.11 page 138).*

THM

$$\left\{\begin{array}{l}(\!(x_n \in \boldsymbol{H})\!)_{n\in\mathbb{Z}} \text{ is a} \\ \text{Riesz basis for } \boldsymbol{H}\end{array}\right\} \implies \left\{\begin{array}{lll} \multicolumn{3}{l}{\text{There exists } (\!(y_n \in \boldsymbol{H})\!)_{n\in\mathbb{Z}} \text{ such that}} \\ (1). & (\!(x_n)\!) \text{ and } (\!(y_n)\!) \text{ are biorthogonal} & \text{and} \\ (2). & (\!(y_n)\!) \text{ is also a Riesz basis for } \boldsymbol{H} & \text{and} \\ (3). & \exists B>A>0 \text{ such that} & \\ & \displaystyle A\sum_{n=1}^{\infty}|a_n|^2 \le \left\|\sum_{n=1}^{\infty} a_n x_n\right\|^2 = \left\|\sum_{n=1}^{\infty} a_n y_n\right\|^2 \le B\sum_{n=1}^{\infty}|a_n|^2 & \\ & \forall (\!(a_n)\!)_{n\in\mathbb{N}}\in\ell^2_{\mathbb{F}} & \end{array}\right\}$$

✎PROOF:

1. Proof for (1):

   (a) Let e_n be the *unit vector* in $\boldsymbol{H}$ such that the nth element of e_n is 1 and all other elements are 0.

   (b) Let $\mathbf{M}$ be an operator on $\boldsymbol{H}$ such that $\mathbf{M}e_n = x_n$.

   (c) Note that $\mathbf{M}$ is *isometric*, and as such $\|\mathbf{M}x\| = \|x\| \quad \forall x\in\boldsymbol{H}$.

   (d) Let $y_n \triangleq (\mathbf{M}^{-1})^*$.

   (e) Then,

$$\begin{aligned}
\langle y_n \,|\, x_m\rangle &= \left\langle (\mathbf{M}^{-1})^* e_n \,|\, \mathbf{M}e_m \right\rangle \\
&= \left\langle e_n \,|\, \mathbf{M}^{-1}\mathbf{M}e_m \right\rangle \\
&= \langle e_n \,|\, e_m\rangle \\
&= \bar{\delta}_{nm} \\
&\implies \{x_n\} \text{ and } \{y_n\} \text{ are biorthogonal} && \text{by Definition 6.4 page 107}
\end{aligned}$$

[34] Wojtaszczyk (1997) page 20 ⟨Lemma 2.7(a)⟩

2. Proof for (3):

$$\begin{aligned}
\left\|\sum_{n\in\mathbb{Z}} \alpha_n \boldsymbol{y}_n\right\| &= \left\|\sum_{n\in\mathbb{Z}} \alpha_n \left(\mathbf{M}^{-1}\right)^* \boldsymbol{e}_n\right\| && \text{by definition of } \boldsymbol{y}_n && \text{(Proposition 1d page 171)}\\
&= \left\|\left(\mathbf{M}^{-1}\right)^* \sum_{n\in\mathbb{Z}} \alpha_n \boldsymbol{e}_n\right\| && \text{by property of \textit{linear ops.}}\\
&= \left\|\sum_{n\in\mathbb{Z}} \alpha_n \boldsymbol{e}_n\right\| && \text{because } \left(\mathbf{M}^{-1}\right)^* \text{ is \textit{isometric}} && \text{(Definition 14.10 page 232)}\\
&= \left\|\mathbf{M}\sum_{n\in\mathbb{Z}} \alpha_n \boldsymbol{e}_n\right\| && \text{because } \mathbf{M} \text{ is \textit{isometric}} && \text{(Definition 14.10 page 232)}\\
&= \left\|\sum_{n\in\mathbb{Z}} \alpha_n \mathbf{M}\boldsymbol{e}_n\right\| && \text{by property of \textit{linear operators}}\\
&= \left\|\sum_{n\in\mathbb{Z}} \alpha_n \boldsymbol{x}_n\right\| && \text{by definition of } \mathbf{M}\\
\implies & \left\{\boldsymbol{y}_n\right\} \text{ is a \textit{Riesz basis}} && \text{by left hypothesis}
\end{aligned}$$

3. Proof for (2): by (3) and definition of *Riesz basis* (Definition 10.11 page 169)

Proposition 10.5. [35] *Let* $\left\{\boldsymbol{x}_n \middle| n\in\mathbb{N}\right\}$ *be a set of vectors in a* HILBERT SPACE *(Definition 8.11 page 138)* $\boldsymbol{H} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla \rangle\right)$.

PRP

$$\left\{\begin{array}{l}\left\{\boldsymbol{x}_n\right\} \text{ is a RIESZ BASIS for } \boldsymbol{H} \text{ with}\\ A\sum_{n=1}^{\infty}|a_n|^2 \le \left\|\sum_{n=1}^{\infty} a_n \boldsymbol{x}_n\right\|^2 \le B\sum_{n=1}^{\infty}|a_n|^2\\ \forall\{a_n\}\in\boldsymbol{\ell}^2_{\mathbb{F}}\end{array}\right\} \implies \left\{\begin{array}{l}\left\{\boldsymbol{x}_n\right\} \text{ is a FRAME for } \boldsymbol{H} \text{ with}\\ \underbrace{\frac{1}{B}\|\boldsymbol{x}\|^2 \le \sum_{n=1}^{\infty}\left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 \le \frac{1}{A}\|\boldsymbol{x}\|^2}_{\text{STABILITY CONDITION}}\\ \forall \boldsymbol{x}\in\boldsymbol{H}\end{array}\right\}$$

✎PROOF:

1. Let $\left\{\boldsymbol{y}_n \middle| n\in\mathbb{N}\right\}$ be a *Riesz basis* that is *biorthogonal* to $\left\{\boldsymbol{x}_n \middle| n\in\mathbb{N}\right\}$ (Theorem 10.14 page 171).

2. Let $\boldsymbol{x} \triangleq \sum_{n=1}^{\infty} a_n \boldsymbol{y}_n$.

3. lemma:

$$\begin{aligned}
\sum_{n=1}^{\infty}\left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 &= \sum_{n=1}^{\infty}\left|\left\langle \sum_{m=1}^{\infty} a_n \boldsymbol{y}_m \,\middle|\, \boldsymbol{x}_n\right\rangle\right|^2 && \text{by definition of } \boldsymbol{x} && \text{(item (2) page 172)}\\
&= \sum_{n=1}^{\infty}\left|\sum_{m=1}^{\infty} a_n \langle \boldsymbol{y}_m \mid \boldsymbol{x}_n\rangle\right|^2 && \text{by \textit{homogeneous} property of } \langle \triangle \mid \nabla \rangle && \text{(Definition 6.1 page 95)}\\
&= \sum_{n=1}^{\infty}\left|\sum_{m=1}^{\infty} a_n \bar{\delta}_{mn}\right|^2 && \text{by definition of \textit{biorthogonal}} && \text{(Definition 10.13 page 169)}
\end{aligned}$$

[35] Igari (1996) page 220 ⟨Lemma 9.8⟩, Wojtaszczyk (1997) pages 20–21 ⟨Lemma 2.7(a)⟩

$$= \sum_{n=1}^{\infty} |a_n|^2 \qquad \text{by definition of } \bar{\delta} \qquad \text{(Definition 6.3 page 107)}$$

4. Then

$$\begin{array}{rlllll}
 & A\sum_{n=1}^{\infty}|a_n|^2 & \le \left\|\sum_{n=1}^{\infty} a_n \boldsymbol{x}_n\right\|^2 & \le B\sum_{n=1}^{\infty}|a_n|^2 & & \text{by definition of } \{\boldsymbol{y}_n\} \text{ (item (1) page 172)}\\
\Longrightarrow & A\sum_{n=1}^{\infty}|a_n|^2 & \le \left\|\sum_{n=1}^{\infty} a_n \boldsymbol{y}_n\right\|^2 & \le B\sum_{n=1}^{\infty}|a_n|^2 & & \text{by definition of } \{\boldsymbol{y}_n\} \text{ (item (1) page 172)}\\
\Longrightarrow & A\sum_{n=1}^{\infty}|a_n|^2 & \le \|\boldsymbol{x}\|^2 & \le B\sum_{n=1}^{\infty}|a_n|^2 & & \text{by definition of } \boldsymbol{x} \text{ (item (2) page 172)}\\
\Longrightarrow & A\sum_{n=1}^{\infty}|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle|^2 & \le \|\boldsymbol{x}\|^2 & \le B\sum_{n=1}^{\infty}|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle|^2 & & \text{by (3) lemma}\\
\Longrightarrow & \frac{1}{B}\|\boldsymbol{x}\|^2 & \le \sum_{n=1}^{\infty}|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle|^2 & \le \frac{1}{A}\|\boldsymbol{x}\|^2 & &
\end{array}$$

✎

Theorem 10.15 (Battle-Lemarié orthogonalization). [36] *Let* $\tilde{\mathsf{f}}(\omega)$ *be the* FOURIER TRANSFORM (*Definition* **??** *page* **??**) *of a function* $\mathsf{f} \in \boldsymbol{L}^2_{\mathbb{R}}$.

THM

$$\left\{\begin{array}{ll} 1. & \{\mathbf{T}^n \mathsf{g} \mid n\in\mathbb{Z}\} \text{ is a RIESZ BASIS for } \boldsymbol{L}^2_{\mathbb{R}} \quad \text{and} \\ 2. & \tilde{\mathsf{f}}(\omega) \triangleq \dfrac{\tilde{\mathsf{g}}(\omega)}{\sqrt{2\pi\sum_{n\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi n)|^2}} \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \{\mathbf{T}^n \mathsf{f} \mid n\in\mathbb{Z}\} \\ \text{is an ORTHONORMAL BASIS for } \boldsymbol{L}^2_{\mathbb{R}} \end{array}\right\}$$

✎PROOF:

1. Proof that $\{\mathbf{T}^n \mathsf{f} \mid n\in\mathbb{Z}\}$ is orthonormal:

$$\begin{aligned}
\tilde{\mathsf{S}}_{\phi\phi}(\omega) &= 2\pi\sum_{n\in\mathbb{Z}}\left|\tilde{\mathsf{f}}(\omega+2\pi n)\right|^2 && \text{by Theorem ?? page ??}\\
&= 2\pi\sum_{n\in\mathbb{Z}}\left|\frac{\tilde{\mathsf{g}}(\omega+2\pi n)}{\sqrt{2\pi\sum_{m\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi(n-m))|^2}}\right|^2 && \text{by left hypothesis}\\
&= \sum_{n\in\mathbb{Z}}\left|\frac{\tilde{\mathsf{g}}(\omega+2\pi n)}{\sqrt{\sum_{m\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi m)|^2}}\right|^2\\
&= \sum_{n\in\mathbb{Z}}\left|\frac{1}{\sqrt{\sum_{m\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi m)|^2}}\right|^2 |\tilde{\mathsf{g}}(\omega+2\pi n)|^2\\
&= \frac{1}{\sum_{m\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi m)|^2}\sum_{n\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi n)|^2
\end{aligned}$$

[36] Wojtaszczyk (1997) page 25 ⟨Remark 2.4⟩, Vidakovic (1999), page 71, Mallat (1989), page 72, Mallat (1999), page 225, Daubechies (1992) page 140 ⟨(5.3.3)⟩

$$= 1$$
$$\implies \{ \mathsf{f}_n | n\in\mathbb{Z} \} \text{ is orthonormal} \qquad \text{by Theorem ?? page ??}$$

2. Proof that $\{ \mathbf{T}^n \mathsf{f} | n\in\mathbb{Z} \}$ is a basis for $\boldsymbol{V}_0$: by Lemma 10.1 page 157.

✏

10.7 Frames in Hilbert spaces

Definition 10.14. [37] *Let* $\{ \boldsymbol{x}_n \in X | n\in\mathbb{N} \}$ *be a set of vectors in a* HILBERT SPACE (Definition 8.11 page 138) $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle | \triangledown \rangle)$.

DEF

The set $\{ \boldsymbol{x}_n \}$ *is a* ***frame*** *for* $\boldsymbol{H}$ *if* (STABILITY CONDITION)

$$\exists A, B \in \mathbb{R}^+ \quad \text{such that} \quad A \|\boldsymbol{x}\|^2 \leq \sum_{n=1}^{\infty} |\langle \boldsymbol{x} | \boldsymbol{x}_n \rangle|^2 \leq B \|\boldsymbol{x}\|^2 \qquad \forall x\in X.$$

The quantities A and B are ***frame bounds****.*

The quantity A′ is the ***optimal lower frame bound*** *if*

$$A' = \sup \{ A \in \mathbb{R}^+ | A \text{ is a lower frame bound} \}.$$

The quantity B′ is the ***optimal upper frame bound*** *if*

$$B' = \inf \{ B \in \mathbb{R}^+ | B \text{ is an upper frame bound} \}.$$

A frame is a ***tight frame*** *if* $A = B$.

A frame is a ***normalized tight frame*** *(or a* ***Parseval frame****) if* $A = B = 1$.

A frame $\{ \boldsymbol{x}_n | n\in\mathbb{N} \}$ *is an* ***exact frame*** *if for some* $m \in \mathbb{Z}$, $\{ \boldsymbol{x}_n | n\in\mathbb{N} \} \setminus \{ \boldsymbol{x}_m \}$ *is* NOT *a frame.*

A frame is a *Parseval frame* (Definition 10.14) if it satisfies *Parseval's Identity* (Theorem 10.9 page 164). All orthonormal bases are Parseval frames (Theorem 10.9 page 164); but not all Parseval frames are orthonormal bases.

Definition 10.15. *Let* $\{ \boldsymbol{x}_n \}$ *be a* ***frame*** (Definition 10.14 page 174) *for the* HILBERT SPACE $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle | \triangledown \rangle)$. *Let* $\mathbf{S}$ *be an* OPERATOR *on* $\boldsymbol{H}$.

DEF

$\mathbf{S}$ *is a* ***frame operator*** *for* $\{ \boldsymbol{x}_n \}$ *if*
$$\mathbf{S}\mathsf{f}(x) = \sum_{n\in\mathbb{Z}} \langle \mathsf{f} | \boldsymbol{x}_n \rangle \boldsymbol{x}_n \qquad \forall \mathsf{f}\in\boldsymbol{H}.$$

Theorem 10.16. [38] *Let* $\mathbf{S}$ *be a* FRAME OPERATOR (Definition 10.15 page 174) *of a* FRAME $\{ \boldsymbol{x}_n \}$ (Definition 10.14 page 174) *for the* HILBERT SPACE $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle | \triangledown \rangle)$.

THM

(1). $\mathbf{S}$ *is* INVERTIBLE. *and*

(2). $$\mathsf{f}(x) = \sum_{n\in\mathbb{Z}} \langle \mathsf{f} | \mathbf{S}^{-1} \boldsymbol{x}_n \rangle \boldsymbol{x}_n = \sum_{n\in\mathbb{Z}} \langle \mathsf{f} | \boldsymbol{x}_n \rangle \mathbf{S}^{-1} \boldsymbol{x}_n \qquad \forall \mathsf{f} \in \boldsymbol{H}$$

Theorem 10.17. [39] *Let* $\{ \boldsymbol{x}_n \in X | n=1,2,\ldots,N \}$ *be a set of vectors in a* HILBERT SPACE (Definition 8.11 page 138) $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle | \triangledown \rangle)$.

THM

$\{ \boldsymbol{x}_n \}$ *is a* FRAME *for* $\operatorname{span}\{ \boldsymbol{x}_n \}$.

[37] Young (2001) pages 154–155, Christensen (2003) page 88 ⟨Definitions 5.1.1, 5.1.2⟩, Heil (2011) pages 204–205 ⟨Definition 8.2⟩, Jørgensen et al. (2008) page 267 ⟨Definition 12.22⟩, Duffin and Schaeffer (1952) page 343, Daubechies et al. (1986), page 1272

[38] Christensen (2008) pages 100–102 ⟨Theorem 5.1.7⟩

[39] Christensen (2003) page 3

✎PROOF:

1. Upper bound: Proof that there exists B such that $\sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 \le B \|\boldsymbol{x}\|^2 \quad \forall \boldsymbol{x} \in \boldsymbol{H}$:

$$\begin{aligned}
\sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 &\le \sum_{n=1}^{N} \langle \boldsymbol{x}_n \mid \boldsymbol{x}_n \rangle \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle && \text{by *Cauchy-Schwarz inequality* (Theorem 6.2 page 96)} \\
&= \underbrace{\left\{ \sum_{n=1}^{N} \|\boldsymbol{x}_n\|^2 \right\}}_{B} \|\boldsymbol{x}\|^2
\end{aligned}$$

2. Lower bound: Proof that there exists A such that $A \|\boldsymbol{x}\|^2 \le \sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 \quad \forall \boldsymbol{x} \in \boldsymbol{H}$:

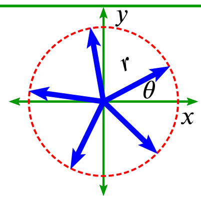
$$\begin{aligned}
\sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 &= \sum_{n=1}^{N} \left| \left\langle \boldsymbol{x}_n \mid \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \right\rangle \right|^2 \|\boldsymbol{x}\|^2 \\
&\ge \underbrace{\left(\inf_{\boldsymbol{y}} \left\{ \sum_{n=1}^{N} |\langle \boldsymbol{x}_n \mid \boldsymbol{y} \rangle|^2 \mid \|\boldsymbol{y}\| = 1 \right\} \right)}_{A} \|\boldsymbol{x}\|^2
\end{aligned}$$

✏

Example 10.1. Let $(\mathbb{R}^2, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle)$ be an inner product space with $\left\langle \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \mid \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right\rangle \triangleq x_1 x_2 + y_1 y_2$. Let $\mathbf{S}$ be the *frame operator* (Definition 10.15 page 174) with *inverse* $\mathbf{S}^{-1}$.

EX

Let $N \in \{3, 4, 5, \dots\}$, $\theta \in \mathbb{R}$, and $r \in \mathbb{R}^+$ $(r > 0)$.
Let $\boldsymbol{x}_n \triangleq r \begin{bmatrix} \cos(\theta + 2n\pi/N) \\ \sin(\theta + 2n\pi/N) \end{bmatrix} \quad \forall n \in \{0, 1, \dots, N-1\}$.
Then, $(\!(\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{N-1})\!)$ is a **tight frame** for $\mathbb{R}^2$ with *frame bound* $A = \frac{Nr^2}{2}$.



Moreover, $\quad \mathbf{S} = \frac{Nr^2}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{S}^{-1} = \frac{2}{Nr^2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad$ and $\quad \boldsymbol{v} = \frac{2}{Nr^2} \sum_{n=0}^{N-1} \langle \boldsymbol{v} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \quad \forall \boldsymbol{v} \in \mathbb{R}^2$.

✎PROOF:

1. Proof that $(\!(\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{N-1})\!)$ is a *tight frame* with *frame bound* $A = \frac{Nr^2}{2}$: Let $\boldsymbol{v} \triangleq (x, y) \in \mathbb{R}^2$.

$$\begin{aligned}
\sum_{n=0}^{N-1} |\langle \boldsymbol{v} \mid \boldsymbol{x}_n \rangle|^2 &\triangleq \sum_{n=0}^{N-1} \left| \boldsymbol{v}^{\mathbf{H}} r \begin{bmatrix} \cos\left(\theta + \frac{2n\pi}{N}\right) \\ \sin\left(\theta + \frac{2n\pi}{N}\right) \end{bmatrix} \right|^2 && \text{by definitions of } \boldsymbol{v} \text{ of } \langle \boldsymbol{y} \mid \boldsymbol{x} \rangle \\
&\triangleq \sum_{n=0}^{N-1} r^2 \left| x\cos\left(\theta + \frac{2n\pi}{N}\right) + y\sin\left(\theta + \frac{2n\pi}{N}\right) \right|^2 && \text{by definition of } \boldsymbol{y}^{\mathbf{H}}\boldsymbol{x} \text{ operation} \\
&= r^2 x^2 \sum_{n=0}^{N-1} \cos^2\left(\theta + \frac{2n\pi}{N}\right) + r^2 y^2 \sum_{n=0}^{N-1} \sin^2\left(\theta + \frac{2n\pi}{N}\right) + r^2 xy \sum_{n=0}^{N-1} \cos\left(\theta + \frac{2n\pi}{N}\right) \sin\left(\theta + \frac{2n\pi}{N}\right) \\
&= r^2 x^2 \frac{N}{2} + r^2 y^2 \frac{N}{2} + r^2 xy0 && \text{by Corollary ?? page ??} \\
&= (x^2 + y^2) \frac{Nr^2}{2} = \left(\frac{Nr^2}{2} \right) \boldsymbol{v}^{\mathbf{H}} \boldsymbol{v} \triangleq \underbrace{\left(\frac{Nr^2}{2} \right)}_{A} \|\boldsymbol{v}\|^2 && \text{by definition of } \|\boldsymbol{v}\|
\end{aligned}$$

2. Proof that $\mathbf{S} = \frac{Nr^2}{2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$:

 (a) Let $\boldsymbol{e}_1 \triangleq \begin{bmatrix}1\\0\end{bmatrix}$ and $\boldsymbol{e}_2 \triangleq \begin{bmatrix}0\\1\end{bmatrix}$.

 (b) lemma: $\mathbf{S}\boldsymbol{e}_1 = \frac{Nr^2}{2}\boldsymbol{e}_1$. Proof:

$$\begin{aligned}\mathbf{S}\boldsymbol{e}_1 &= \sum_{n=0}^{N-1}\langle \boldsymbol{e}_1 \,|\, \boldsymbol{x}_n\rangle \boldsymbol{x}_n\\ &= \sum_{n=0}^{N-1} r\cos\left(\theta+\frac{2n\pi}{N}\right) r\begin{bmatrix}\cos\left(\theta+\frac{2n\pi}{N}\right)\\ \sin\left(\theta+\frac{2n\pi}{N}\right)\end{bmatrix} = r^2\sum_{n=0}^{N-1}\begin{bmatrix}\cos^2\left(\theta+\frac{2n\pi}{N}\right)\\ \cos\left(\theta+\frac{2n\pi}{N}\right)\sin\left(\theta+\frac{2n\pi}{N}\right)\end{bmatrix}\\ &= r^2\begin{bmatrix}N/2\\0\end{bmatrix} = \frac{Nr^2}{2}\boldsymbol{e}_1 \qquad \text{by } \textit{Summation around unit circle} \text{ (Corollary ?? page ??)}\end{aligned}$$

 (c) lemma: $\mathbf{S}\boldsymbol{e}_2 = \frac{Nr^2}{2}\boldsymbol{e}_2$. Proof:

$$\begin{aligned}\mathbf{S}\boldsymbol{e}_2 &= \sum_{n=0}^{N-1}\langle \boldsymbol{e}_2 \,|\, \boldsymbol{x}_n\rangle \boldsymbol{x}_n\\ &= \sum_{n=0}^{N-1} r\sin\left(\theta+\frac{2n\pi}{N}\right) r\begin{bmatrix}\cos\left(\theta+\frac{2n\pi}{N}\right)\\ \sin\left(\theta+\frac{2n\pi}{N}\right)\end{bmatrix} = r^2\sum_{n=0}^{N-1}\begin{bmatrix}\sin\left(\theta+\frac{2n\pi}{N}\right)\cos\left(\theta+\frac{2n\pi}{N}\right)\\ \sin^2\left(\theta+\frac{2n\pi}{N}\right)\end{bmatrix}\\ &= r^2\begin{bmatrix}0\\N/2\end{bmatrix} = \frac{Nr^2}{2}\boldsymbol{e}_2 \qquad \text{by } \textit{Summation around unit circle} \text{ (Corollary ?? page ??)}\end{aligned}$$

 (d) Complete the proof of item (2) using *Eigendecomposition* $\mathbf{S} = \mathbf{Q}\Lambda\mathbf{Q}^{-1}$:

$$\begin{aligned}\mathbf{S}\boldsymbol{e}_1 &= \frac{Nr^2}{2}\boldsymbol{e}_1 && \text{by (2c) lemma}\\ &\implies \boldsymbol{e}_1 \text{ is an } \textit{eigenvector} \text{ of } \mathbf{S} \text{ with } \textit{eigenvalue } \frac{Nr^2}{2}\\ \mathbf{S}\boldsymbol{e}_2 &= \frac{Nr^2}{2}\boldsymbol{e}_2 && \text{by (2c) lemma}\\ &\implies \boldsymbol{e}_2 \text{ is an } \textit{eigenvector} \text{ of } \mathbf{S} \text{ with } \textit{eigenvalue } \frac{Nr^2}{2}\end{aligned}$$

$$\mathbf{S} = \overbrace{\underbrace{\begin{bmatrix}| & |\\ \boldsymbol{e}_1 & \boldsymbol{e}_2\\ | & |\end{bmatrix}}_{\mathbf{Q}}\underbrace{\begin{bmatrix}\frac{Nr^2}{2} & 0\\ 0 & \frac{Nr^2}{2}\end{bmatrix}}_{\Lambda}\underbrace{\begin{bmatrix}| & |\\ \boldsymbol{e}_1 & \boldsymbol{e}_2\\ | & |\end{bmatrix}^{-1}}_{\mathbf{Q}^{-1}}}^{\textit{Eigendecomposition} \text{ of } \mathbf{S}} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}\frac{Nr^2}{2} & 0\\ 0 & \frac{Nr^2}{2}\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}^{-1} = \frac{Nr^2}{2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

3. Proof that $\mathbf{S}^{-1} = \frac{2}{Nr^2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$:

$$\begin{aligned}\mathbf{S}\mathbf{S}^{-1} &= \frac{Nr^2}{2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\frac{2}{Nr^2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \triangleq \mathbf{I}_2 && \text{by item (2)}\\ \mathbf{S}^{-1}\mathbf{S} &= \frac{2}{Nr^2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\frac{Nr^2}{2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \triangleq \mathbf{I}_2 && \text{by item (2)}\end{aligned}$$

4. Proof that $\boldsymbol{v} = \frac{2}{Nr^2}\sum_{n=0}^{N-1}\left(\boldsymbol{v}^{\mathbf{H}}\boldsymbol{x}_n\right)\boldsymbol{x}_n$:

$$\begin{aligned}
\boldsymbol{v} &= \sum_{n=0}^{N-1}\left\langle \boldsymbol{v} \,|\, \mathbf{S}^{-1}\boldsymbol{x}_n \right\rangle \boldsymbol{x}_n = \sum_{n=0}^{N-1}\left\langle \boldsymbol{v} \,|\, \frac{2}{Nr^2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\boldsymbol{x}_n \right\rangle \boldsymbol{x}_n && \text{by item (3)} \\
&= \frac{2}{Nr^2}\sum_{n=0}^{N-1}\langle \boldsymbol{v} \,|\, \boldsymbol{x}_n\rangle \boldsymbol{x}_n = \frac{2}{Nr^2}\sum_{n=0}^{N-1}\left(\boldsymbol{v}^{\mathbf{H}}\boldsymbol{x}_n\right)\boldsymbol{x}_n && \text{by definition of } \langle \boldsymbol{y} \,|\, \boldsymbol{x}\rangle
\end{aligned}$$

✎

Example 10.2 (Peace Frame/Mercedes Frame). [40] Let $\left(\mathbb{R}^2, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \nabla \rangle\right)$ be an inner product space with $\left\langle (x_1, y_1) \,|\, (x_2, y_2) \right\rangle \triangleq x_1 y_1 + x_2 y_2$. Let $\mathbf{S}$ be the *frame operator* (Definition 10.15 page 174) with *inverse* $\mathbf{S}^{-1}$.

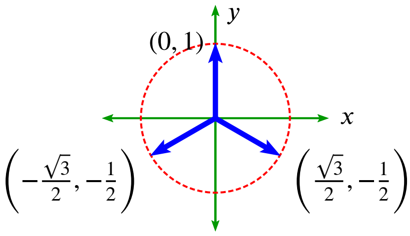
EX

Let $\boldsymbol{x}_1 \triangleq \frac{1}{2}\begin{bmatrix} 0 \\ 2 \end{bmatrix}$, $\boldsymbol{x}_2 \triangleq \frac{1}{2}\begin{bmatrix} -\sqrt{3} \\ -1 \end{bmatrix}$, and $\boldsymbol{x}_3 \triangleq \frac{1}{2}\begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}$.

Then, $(\!(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)\!)$ is a **tight frame** for $\mathbb{R}^2$ with *frame bound* $A = \frac{3}{2}$.

Moreover, $\mathbf{S} = \frac{3}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{S}^{-1} = \frac{2}{3}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,

and $\quad \boldsymbol{v} = \frac{2}{3}\sum_{n=1}^{3}\langle \boldsymbol{v} \,|\, \boldsymbol{x}_n\rangle \boldsymbol{x}_n \triangleq \frac{2}{3}\sum_{n=1}^{3}\left(\boldsymbol{v}^{\mathbf{H}}\boldsymbol{x}_n\right)\boldsymbol{x}_n \qquad \forall \boldsymbol{v} \in \mathbb{R}^2.$



✎PROOF:

1. This frame is simply a special case of the frame presented in Example 10.1 (page 175) with $r = 1$, $N = 3$, and $\theta = \pi/2$.

2. Let's give it a try! Let $\boldsymbol{v} \triangleq \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$$\begin{aligned}
\sum_{n=1}^{3}\left\langle \boldsymbol{v} \,|\, \mathbf{S}^{-1}\boldsymbol{x}_n \right\rangle \boldsymbol{x}_n &= \frac{2}{3}\sum_{n=1}^{3}\left(\boldsymbol{v}^{\mathbf{H}}\boldsymbol{x}_n\right)\boldsymbol{x}_n && \text{by Example 10.1 page 175} \\
&= \left(\boldsymbol{v}^{\mathbf{H}}\boldsymbol{x}_1\right)\boldsymbol{x}_1 + \left(\boldsymbol{v}^{\mathbf{H}}\boldsymbol{x}_2\right)\boldsymbol{x}_2 + \left(\boldsymbol{v}^{\mathbf{H}}\boldsymbol{x}_3\right)\boldsymbol{x}_3 \\
&= \frac{2}{3}\left(\left(\boldsymbol{v}^{\mathbf{H}}\frac{1}{2}\begin{bmatrix} 0 \\ 2 \end{bmatrix}\right)\boldsymbol{x}_1 + \left(\boldsymbol{v}^{\mathbf{H}}\frac{1}{2}\begin{bmatrix} -\sqrt{3} \\ -1 \end{bmatrix}\right)\boldsymbol{x}_2 + \left(\boldsymbol{v}^{\mathbf{H}}\frac{1}{2}\begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}\right)\boldsymbol{x}_3\right) \\
&= \frac{2}{3}\cdot\frac{1}{2}\left(\left(\boldsymbol{v}^{\mathbf{H}}\begin{bmatrix} 0 \\ 2 \end{bmatrix}\right)\boldsymbol{x}_1 + \left(\boldsymbol{v}^{\mathbf{H}}\begin{bmatrix} -\sqrt{3} \\ -1 \end{bmatrix}\right)\boldsymbol{x}_2 + \left(\boldsymbol{v}^{\mathbf{H}}\begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}\right)\boldsymbol{x}_3\right) \\
&= \frac{1}{3}\left((2)\frac{1}{2}\begin{bmatrix} 0 \\ 2 \end{bmatrix} + \left(-\sqrt{3} - 1\right)\frac{1}{2}\begin{bmatrix} -\sqrt{3} \\ -1 \end{bmatrix} + \left(\sqrt{3} - 1\right)\frac{1}{2}\begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}\right) \\
&= \frac{1}{6}\begin{bmatrix} 2(0) & + & (-\sqrt{3} - 1)(-\sqrt{3}) & + & (\sqrt{3} - 1)(\sqrt{3}) \\ 2(2) & + & (-\sqrt{3} - 1)(-1) & + & (\sqrt{3} - 1)(-1) \end{bmatrix} \\
&= \frac{1}{6}\begin{bmatrix} 0 & + & (3 + \sqrt{3}) & + & (3 - \sqrt{3}) \\ 4 & + & (1 + \sqrt{3}) & + & (1 - \sqrt{3}) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \triangleq \boldsymbol{v}
\end{aligned}$$

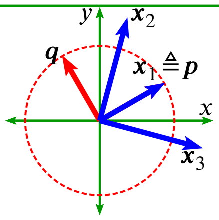
[40] Heil (2011) pages 204–205 ⟨$r = 1$ case⟩, Byrne (2005) page 80 ⟨$r = 1$ case⟩, Han et al. (2007) page 91 ⟨Example 3.9, $r = \sqrt{2/3}$ case⟩

In Example 10.1 (page 175) and Example 10.2 (page 177), the frame operator $\mathbf{S}$ and its inverse $\mathbf{S}^{-1}$ were computed. In general however, it is not always necessary or even possible to compute these, as illustrated in Example 10.3 (next).

Example 10.3. [41] Let $\left(\mathbb{R}^2, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla \rangle\right)$ be an inner product space with $\left\langle (x_1, y_1) \mid (x_2, y_2) \right\rangle \triangleq x_1 y_1 + x_2 y_2$. Let $\mathbf{S}$ be the *frame operator* (Definition 10.15 page 174) with *inverse* $\mathbf{S}^{-1}$.

EX

Let $\boldsymbol{p}$ and $\boldsymbol{q}$ be *orthonormal* vectors in $\boldsymbol{X} \triangleq \operatorname{span}\{\boldsymbol{p}, \boldsymbol{q}\}$.
Let $\boldsymbol{x}_1 \triangleq \boldsymbol{p}$, $\boldsymbol{x}_2 \triangleq \boldsymbol{p}+\boldsymbol{q}$, and $\boldsymbol{x}_3 \triangleq \boldsymbol{p}-\boldsymbol{q}$.
Then, $\{\boldsymbol{x}_1, \boldsymbol{x}_2, , \boldsymbol{x}_3\}$ is a **frame** for $\boldsymbol{X}$ with *frame bounds* $A = 0$ and $B = 5$.



Moreover,
$\mathbf{S}^{-1}\boldsymbol{x}_1 = {}^1\!/\!_3\boldsymbol{p}$ and
$\mathbf{S}^{-1}\boldsymbol{x}_2 = {}^1\!/\!_3\boldsymbol{p} + {}^1\!/\!_2\boldsymbol{q}$ and
$\mathbf{S}^{-1}\boldsymbol{x}_3 = {}^1\!/\!_3\boldsymbol{p} - {}^1\!/\!_2\boldsymbol{q}$.

✎PROOF:

1. Proof that $(\!(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)\!)$ is a *frame* with *frame bounds* $A = 0$ and $B = 5$:

$$\begin{aligned}
\sum_{n=1}^{3} \left|\langle \boldsymbol{v} \mid \boldsymbol{x}_n\rangle\right|^2 &\triangleq |\langle \boldsymbol{v} \mid \boldsymbol{p}\rangle|^2 + |\langle \boldsymbol{v} \mid \boldsymbol{p}+\boldsymbol{q}\rangle|^2 + |\langle \boldsymbol{v} \mid \boldsymbol{p}-\boldsymbol{q}\rangle|^2 && \text{by definitions of } \boldsymbol{x}_1, \boldsymbol{x}_2, \text{ and } \boldsymbol{x}_3\\
&= |\langle \boldsymbol{v} \mid \boldsymbol{p}\rangle|^2 + |\langle \boldsymbol{v} \mid \boldsymbol{p}\rangle + \langle \boldsymbol{v} \mid \boldsymbol{q}\rangle|^2 + |\langle \boldsymbol{v} \mid \boldsymbol{p}\rangle - \langle \boldsymbol{v} \mid \boldsymbol{q}\rangle|^2 && \text{by } \textit{additivity} \text{ of } \langle \triangle \mid \nabla\rangle \text{ (Definition 6.1 page 95)}\\
&= |\langle \boldsymbol{v} \mid \boldsymbol{p}\rangle|^2 + \left(|\langle \boldsymbol{v} \mid \boldsymbol{p}\rangle|^2 + |\langle \boldsymbol{v} \mid \boldsymbol{q}\rangle|^2 + \langle \boldsymbol{v} \mid \boldsymbol{p}\rangle \langle \boldsymbol{v} \mid \boldsymbol{q}\rangle^* + \langle \boldsymbol{v} \mid \boldsymbol{q}\rangle \langle \boldsymbol{v} \mid \boldsymbol{p}\rangle^*\right)\\
&\quad + \left(|\langle \boldsymbol{v} \mid \boldsymbol{p}\rangle|^2 + |\langle \boldsymbol{v} \mid \boldsymbol{q}\rangle|^2 - \langle \boldsymbol{v} \mid \boldsymbol{p}\rangle \langle \boldsymbol{v} \mid \boldsymbol{q}\rangle^* - \langle \boldsymbol{v} \mid \boldsymbol{q}\rangle \langle \boldsymbol{v} \mid \boldsymbol{p}\rangle^*\right)\\
&= 3|\langle \boldsymbol{v} \mid \boldsymbol{p}\rangle|^2 + 2|\langle \boldsymbol{v} \mid \boldsymbol{q}\rangle|^2\\
&\boxed{\leq} 3\,\|\boldsymbol{v}\|\,\|\boldsymbol{p}\| + 2\,\|\boldsymbol{v}\|\,\|\boldsymbol{q}\| && \text{by } \textit{CS Inequality} \text{ (Theorem 6.2 page 96)}\\
&= \|\boldsymbol{v}\|\,(3\,\|\boldsymbol{p}\| + 2\,\|\boldsymbol{q}\|)\\
&= \boxed{5}\,\|\boldsymbol{v}\| && \text{by } \textit{orthonormality} \text{ of } \boldsymbol{p} \text{ and } \boldsymbol{q}
\end{aligned}$$

2. lemma: $\mathbf{S}\boldsymbol{p} = 3\boldsymbol{p}$, $\mathbf{S}\boldsymbol{q} = 2\boldsymbol{q}$, $\mathbf{S}^{-1}\boldsymbol{p} = \frac{1}{3}\boldsymbol{p}$, and $\mathbf{S}^{-1}\boldsymbol{q} = \frac{1}{2}\boldsymbol{q}$. Proof:

$$\begin{aligned}
\mathbf{S}\boldsymbol{p} &\triangleq \sum_{n=1}^{3} \langle \boldsymbol{p} \mid \boldsymbol{x}_n\rangle\, \boldsymbol{x}_n\\
&= \langle \boldsymbol{p} \mid \boldsymbol{p}\rangle\, \boldsymbol{p} + \langle \boldsymbol{p} \mid \boldsymbol{p}+\boldsymbol{q}\rangle\,(\boldsymbol{p}+\boldsymbol{q}) + \langle \boldsymbol{p} \mid \boldsymbol{p}-\boldsymbol{q}\rangle\,(\boldsymbol{p}-\boldsymbol{q})\\
&= (1)\boldsymbol{p} + (1+0)(\boldsymbol{p}+\boldsymbol{q}) + (1-0)(\boldsymbol{p}-\boldsymbol{q})\\
&= 3\boldsymbol{p}\\
\implies \mathbf{S}^{-1}\boldsymbol{p} &= \frac{1}{3}\boldsymbol{p}\\
\mathbf{S}\boldsymbol{q} &\triangleq \sum_{n=1}^{3} \langle \boldsymbol{q} \mid \boldsymbol{x}_n\rangle\, \boldsymbol{x}_n\\
&= \langle \boldsymbol{q} \mid \boldsymbol{p}\rangle\, \boldsymbol{p} + \langle \boldsymbol{q} \mid \boldsymbol{p}+\boldsymbol{q}\rangle\,(\boldsymbol{p}+\boldsymbol{q}) + \langle \boldsymbol{q} \mid \boldsymbol{p}-\boldsymbol{q}\rangle\,(\boldsymbol{p}-\boldsymbol{q})\\
&= (0)\boldsymbol{q} + (0+1)(\boldsymbol{p}+\boldsymbol{q}) + (0-1)(\boldsymbol{p}-\boldsymbol{q})\\
&= 2\boldsymbol{q}\\
\implies \mathbf{S}^{-1}\boldsymbol{q} &= \frac{1}{2}\boldsymbol{q}
\end{aligned}$$

[41] Christensen (2003) pages 7–8 ⟨?⟩

3. Remark: Without knowing $\boldsymbol{p}$ and $\boldsymbol{q}$, from (2) lemma it follows that it is not possible to compute $\mathbf{S}$ or $\mathbf{S}^{-1}$ explicitly.

4. Proof that $\mathbf{S}^{-1}\boldsymbol{x}_1 = \frac{1}{3}\boldsymbol{p}$, $\mathbf{S}^{-1}\boldsymbol{x}_2 = \frac{1}{3}\boldsymbol{p} + \frac{1}{2}\boldsymbol{q}$ and $\mathbf{S}^{-1}\boldsymbol{x}_3 = \frac{1}{3}\boldsymbol{p} - \frac{1}{2}\boldsymbol{q}$:

$$\begin{aligned}
\mathbf{S}^{-1}\boldsymbol{x}_1 &\triangleq \mathbf{S}^{-1}\boldsymbol{p} && \text{by definition of } \boldsymbol{x}_1\\
&= \frac{1}{3}\boldsymbol{p} && \text{by (2) lemma}\\
\mathbf{S}^{-1}\boldsymbol{x}_2 &\triangleq \mathbf{S}^{-1}(\boldsymbol{p}+\boldsymbol{q}) && \text{by definition of } \boldsymbol{x}_2\\
&= \frac{1}{3}\boldsymbol{p} + \frac{1}{2}\boldsymbol{q} && \text{by (2) lemma}\\
\mathbf{S}^{-1}\boldsymbol{x}_3 &\triangleq \mathbf{S}^{-1}(\boldsymbol{p}-\boldsymbol{q}) && \text{by definition of } \boldsymbol{x}_2\\
&= \frac{1}{3}\boldsymbol{p} - \frac{1}{2}\boldsymbol{q} && \text{by (2) lemma}
\end{aligned}$$

5. Check that $\boldsymbol{v} = \sum_n \langle \boldsymbol{v} \,|\, \boldsymbol{x}_n \rangle \boldsymbol{x}_n = \langle \boldsymbol{v} \,|\, \boldsymbol{p} \rangle \boldsymbol{p} + \langle \boldsymbol{v} \,|\, \boldsymbol{q} \rangle \boldsymbol{q}$:

$$\begin{aligned}
\boldsymbol{v} &= \sum_{n=1}^{3} \left\langle \boldsymbol{v} \,|\, \mathbf{S}^{-1}\boldsymbol{x}_n \right\rangle \boldsymbol{x}_n\\
&= \left\langle \boldsymbol{v} \,|\, \mathbf{S}^{-1}\boldsymbol{p} \right\rangle \boldsymbol{p} + \left\langle \boldsymbol{v} \,|\, \mathbf{S}^{-1}(\boldsymbol{p}+\boldsymbol{q}) \right\rangle (\boldsymbol{p}+\boldsymbol{q}) + \left\langle \boldsymbol{v} \,|\, \mathbf{S}^{-1}(\boldsymbol{p}-\boldsymbol{q}) \right\rangle (\boldsymbol{p}-\boldsymbol{q})\\
&= \left\langle \boldsymbol{v} \,|\, \frac{1}{3}\boldsymbol{p} \right\rangle \boldsymbol{p} + \left\langle \boldsymbol{v} \,|\, \frac{1}{3}\boldsymbol{p} + \frac{1}{2}\boldsymbol{q} \right\rangle (\boldsymbol{p}+\boldsymbol{q}) + \left\langle \boldsymbol{v} \,|\, \frac{1}{3}\boldsymbol{p} - \frac{1}{2}\boldsymbol{q}) \right\rangle (\boldsymbol{p}-\boldsymbol{q})\\
&= \left(\frac{1}{3}+\frac{1}{3}+\frac{1}{3}\right) \langle \boldsymbol{v} \,|\, \boldsymbol{p} \rangle \boldsymbol{p} + \left(\frac{1}{3}-\frac{1}{3}\right) \langle \boldsymbol{v} \,|\, \boldsymbol{p} \rangle \boldsymbol{q} + \left(\frac{1}{2}-\frac{1}{2}\right) \langle \boldsymbol{v} \,|\, \boldsymbol{q} \rangle \boldsymbol{p} + \left(\frac{1}{2}+\frac{1}{2}\right) \langle \boldsymbol{v} \,|\, \boldsymbol{q} \rangle \boldsymbol{q}\\
&= \langle \boldsymbol{v} \,|\, \boldsymbol{p} \rangle \boldsymbol{p} + \langle \boldsymbol{v} \,|\, \boldsymbol{q} \rangle \boldsymbol{q}
\end{aligned}$$

CHAPTER 11

FINITE SUMS

❝ *I think that it was Harald Bohr who remarked to me that "all analysts spend half their time hunting through the literature for inequalities which they want to use and cannot prove."* ❞
G.H. Hardy (1877–1947) in his "Presidential Address" to the London Mathematical Society on November 8, 1928, about a remark that he suggested was from Harald Bohr (1887–1951), Danish mathematician pictured to the left. [1]

11.1 Summation

Definition 11.1. [2] *Let $+$ be an addition operator on a tuple $\left(x_n\right)_m^N$.*

DEF

*The **summation** of $\left(x_n\right)$ from index m to index N with respect to $+$ is*

$$\sum_{n=m}^{N} x_n \triangleq \begin{cases} 0 & \text{for } N < m \\ \left(\sum_{n=m}^{N-1} x_n\right) + x_N & \text{for } N \geq m \end{cases}$$

Theorem 11.1 (Generalized associative property). [3] *Let $+$ be an addition operator on a tuple $\left(x_n\right)_m^N$.*

THM

$+$ *is* ASSOCIATIVE $\implies$

$$\underbrace{\sum_{n=m}^{L} x_n + \left(\sum_{n=L+1}^{M} x_n + \sum_{n=M+1}^{N} x_n\right) = \left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \sum_{n=M+1}^{N} x_n \qquad \text{for } m < L < M \leq N}_{\sum_{n=m}^{N} \text{ is ASSOCIATIVE}}$$

✎PROOF:

[1] quote: Hardy (1929), page 64
image: http://www-history.mcs.st-andrews.ac.uk/PictDisplay/Bohr_Harald.html

[2] reference: Berberian (1961) page 8 ⟨Definition I.3.1⟩
"$\sum$" notation: Fourier (1820) page 280

[3] Berberian (1961) pages 9–10 ⟨Theorem I.3.1⟩

1. Proof for $N < m$ case: $\sum_{n=m}^{N} x_n = 0.$

2. Proof for $N = m$ case: $\sum_{n=m}^{m} x_n = \left(\sum_{n=m}^{m-1} x_n\right) + x_m = 0 + x_m = x_m.$

3. Proof for $N = m+1$ case: $\sum_{n=m}^{m+1} x_n = \left(\sum_{n=m}^{m} x_n\right) + x_{m+1} = x_m + x_{m+1}$

4. Proof for $N = m+2$ case:

$$\begin{aligned}
\sum_{n=m}^{m+2} x_n &= \left(\sum_{n=m}^{m+1} x_n\right) + x_{m+2} && \text{by Definition 11.1 page 181} \\
&= \left(x_m + x_{m+1}\right) + x_{m+2} && \text{by item (3)} \\
&= x_m + \left(x_{m+1} + x_{m+2}\right) && \text{by left hypothesis}
\end{aligned}$$

5. Proof that N case $\implies N+1$ case:

$$\begin{aligned}
\sum_{n=m}^{N+1} x_n &= \underbrace{\left(\sum_{n=m}^{N} x_n\right)}_{\textit{associative}} + x_{N+1} && \text{by Definition 11.1 page 181} \\
&= \left(\sum_{n=m}^{L} x_n + \left(\sum_{n=L+1}^{M} x_n + \sum_{n=M+1}^{N} x_n\right)\right) + x_{N+1} \\
&= \left(\left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \sum_{n=M+1}^{N} x_n\right) + x_{N+1} \\
&= \left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \left(\sum_{n=M+1}^{N} x_n + x_{N+1}\right) \\
&= \left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \left(\sum_{n=M+1}^{N+1} x_n\right)
\end{aligned}$$

11.2 Means

11.2.1 Weighted ϕ-means

Definition 11.2. [4]

DEF

The $(\!(\lambda_n)\!)_1^N$ *weighted* ϕ***-mean*** *of a tuple* $(\!(x_n)\!)_1^N$ *is defined as*

$$\mathsf{M}_\phi\left((\!(x_n)\!)\right) \triangleq \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi(x_n)\right)$$

where ϕ *is a* CONTINUOUS *and* STRICTLY MONOTONIC *function in* $\mathbb{R}^{\mathbb{R}^{\vdash}}$ *and* $(\!(\lambda_n)\!)_{n=1}^N$ *is a sequence of weights for which* $\sum_{n=1}^N \lambda_n = 1$.

Lemma 11.1. [5] *Let* $\mathsf{M}_\phi\left((\!(x_n)\!)\right)$ *be the* $(\!(\lambda_n)\!)_1^N$ *weighted* ϕ*-mean of a tuple* $(\!(x_n)\!)_1^N$*. Let the property* CONVEX *be defined as in Definition 9.7 (page 143).*

LEM

$$\begin{array}{lllll}
\phi\psi^{-1} \text{ is CONVEX} & \text{and} & \phi \text{ is INCREASING} & \implies & \mathsf{M}_\phi\left((\!(x_n)\!)\right) \geq \mathsf{M}_\psi\left((\!(x_n)\!)\right)\\
\phi\psi^{-1} \text{ is CONVEX} & \text{and} & \phi \text{ is DECREASING} & \implies & \mathsf{M}_\phi\left((\!(x_n)\!)\right) \leq \mathsf{M}_\psi\left((\!(x_n)\!)\right)\\
\phi\psi^{-1} \text{ is CONCAVE} & \text{and} & \phi \text{ is INCREASING} & \implies & \mathsf{M}_\phi\left((\!(x_n)\!)\right) \leq \mathsf{M}_\psi\left((\!(x_n)\!)\right)\\
\phi\psi^{-1} \text{ is CONCAVE} & \text{and} & \phi \text{ is DECREASING} & \implies & \mathsf{M}_\phi\left((\!(x_n)\!)\right) \geq \mathsf{M}_\psi\left((\!(x_n)\!)\right)
\end{array}$$

✎PROOF:

1. Case where $\phi\psi^{-1}$ is *convex* and ϕ is *increasing*:

$$\begin{aligned}
\mathsf{M}_\phi\left((\!(x_n)\!)\right) &\triangleq \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi(x_n)\right) && \text{by definition of } \mathsf{M}_\phi \text{ (Definition 11.2 page 183)}\\
&= \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi\psi^{-1}\psi(x_n)\right) && \text{by definition of } \psi^{-1}\\
&\geq \phi^{-1}\left(\phi\psi^{-1}\sum_{n=1}^N \lambda_n \psi(x_n)\right) && \text{by \textit{Jensen's Inequality} (Theorem 9.1 page 144)}\\
&= \left(\psi^{-1}\sum_{n=1}^N \lambda_n \psi(x_n)\right) && \text{by definition of } \psi^{-1}\\
&\triangleq \mathsf{M}_\psi\left((\!(x_n)\!)\right) && \text{by definition of } \mathsf{M}_\psi \text{ (Definition 11.2 page 183)}
\end{aligned}$$

2. Case where $\phi\psi^{-1}$ is *convex* and ϕ is *decreasing*:

$$\begin{aligned}
\mathsf{M}_\phi\left((\!(x_n)\!)\right) &\triangleq \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi(x_n)\right) && \text{by definition of } \mathsf{M}_\phi \text{ (Definition 11.2 page 183)}\\
&= \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi\psi^{-1}\psi(x_n)\right) && \text{by definition of } \psi^{-1}\\
&\leq \phi^{-1}\left(\phi\psi^{-1}\sum_{n=1}^N \lambda_n \psi(x_n)\right) && \text{by \textit{Jensen's Inequality} and because } \phi^{-1} \text{ is \textit{decreasing}}
\end{aligned}$$

[4] Bollobás (1999) page 5
[5] Pečarić et al. (1992) page 107, Bollobás (1999) page 5, Hardy et al. (1952) page 75

$$= \left(\psi^{-1}\sum_{n=1}^{N}\lambda_n\psi(x_n)\right) \qquad \text{by definition of } \psi^{-1}$$

$$\triangleq \mathsf{M}_\psi(⦇x_n⦈) \qquad \text{by definition of } \mathsf{M}_\psi \text{ (Definition 11.2 page 183)}$$

✎

One of the most well known inequalities in mathematics is *Minkowski's Inequality* (1910, Theorem 11.5 page 190). In 1946, H.P. Mulholland submitted a result[6] that generalizes Minkowski's Inequality to an equal weighted ϕ-mean. And Milovanović and Milovanovć (1979) generalized this even further to a *weighted* ϕ-mean (Theorem 11.2, next).

Theorem 11.2. [7]

THM

$$\left\{\begin{array}{lll} 1. & \phi \textit{ is } \text{CONVEX} & \textit{and} \\ 2. & \phi \textit{ is } \text{STRICTLY MONOTONIC} & \textit{and} \\ 3. & \phi(0)=0 & \textit{and} \\ 4. & \log\circ\phi\circ\exp \textit{ is } \text{CONVEX} & \end{array}\right\} \implies$$

$$\phi^{-1}\left(\sum_{n=1}^{N}\lambda_n\phi(x_n+y_n)\right) \le \phi^{-1}\left(\sum_{n=1}^{N}\lambda_n\phi(x_n)\right) + \phi^{-1}\left(\sum_{n=1}^{N}\lambda_n\phi(y_n)\right)$$

11.2.2 Power means

Definition 11.3. [8] *Let $\mathsf{M}_{\phi(x;r)}(⦇x_n⦈)$ be the $⦇\lambda_n⦈_1^N$ weighted ϕ-mean of a* NON-NEGATIVE *tuple $⦇x_n⦈_1^N$ (Definition 11.2 page 183).*

DEF

*A mean $\mathsf{M}_{\phi(x;r)}(⦇x_n⦈)$ is a **power mean** with parameter r if $\phi(x)\triangleq x^r$. That is,*

$$\mathsf{M}_{\phi(x;r)}(⦇x_n⦈) = \left(\sum_{n=1}^{N}\lambda_n(x_n)^r\right)^{\frac{1}{r}}$$

Theorem 11.3. [9] *Let $\mathsf{M}_{\phi(x;r)}(⦇x_n⦈)$ be* POWER MEAN *with parameter r of an N-tuple $⦇x_n⦈_1^N$. Let $\mathbb{R}^*$ be the set of extended real numbers $(\mathbb{R}^*\triangleq\mathbb{R}\cup\{-\infty,\ \infty\})$.*[10]

THM

$$\mathsf{M}_{\phi(x;r)}(⦇x_n⦈) \triangleq \left(\sum_{n=1}^{N}\lambda_n(x_n)^r\right)^{\frac{1}{r}} \quad \textit{is } \text{CONTINUOUS} \textit{ and } \text{STRICTLY INCREASING} \textit{ in } \mathbb{R}^*.$$

$$\mathsf{M}_{\phi(x;r)}(⦇x_n⦈) = \begin{cases} \min\limits_{n=1,2,\ldots,N} ⦇x_n⦈ & \textit{for } r=-\infty \\ \prod\limits_{n=1}^{N} x_n^{\lambda_n} & \textit{for } r=0 \\ \max\limits_{n=1,2,\ldots,N} ⦇x_n⦈ & \textit{for } r=+\infty \end{cases}$$

✎PROOF:

[6] Mulholland (1950)
[7] Milovanović and Milovanović (1979), Bullen (2003) page 306 ⟨Theorem 9⟩
[8] Bullen (2003) page 175, Bollobás (1999) page 6
[9] Bullen (2003) pages 175–177 ⟨see also page 203⟩, Bollobás (1999) pages 6–8, Besso (1879), Bienaymé (1840) page 68
[10] Rana (2002) pages 385–388 ⟨Appendix A⟩

1. Proof that $M_{\phi(x;r)}$ is *strictly increasing* in r:

   (a) Let r and s be such that $-\infty < r < s < \infty$.

   (b) Let $\phi_r \triangleq x^r$ and $\phi_s \triangleq x^s$. Then $\phi_r\phi_s^{-1} = x^{\frac{r}{s}}$.

   (c) The composite function $\phi_r\phi_s^{-1}$ is *convex* or *concave* depending on the values of r and s:

	$r < 0$ (ϕ_r decreasing)	$r > 0$ (ϕ_r increasing)
$s < 0$	*convex*	(not possible)
$s > 0$	*convex*	*concave*

   (d) Therefore by Lemma 11.1 (page 183),
$$-\infty < r < s < \infty \implies \mathsf{M}_{\phi(x;r)}(\lparen x_n \rparen) < \mathsf{M}_{\phi(x;s)}(\lparen x_n \rparen).$$

2. Proof that $\mathsf{M}_{\phi(x;r)}$ is continuous in r for $r \in \mathbb{R}\backslash 0$: The sum of continuous functions is continuous. For the cases of $r \in \{-\infty,\ 0,\ \infty\}$, see the items that follow.

3. Lemma: $\mathsf{M}_{\phi(x;-r)}(\lparen x_n \rparen) = \left\{\mathsf{M}_{\phi(x;r)}(\lparen x_n^{-1} \rparen)\right\}^{-1}$. Proof:

$$\begin{aligned}
\left\{\mathsf{M}_{\phi(x;r)}(\lparen x_n^{-1} \rparen)\right\}^{-1} &= \left\{\left(\sum_{n=1}^N \lambda_n \left(x_n^{-1}\right)^r\right)^{\frac{1}{r}}\right\}^{-1} && \text{by definition of } \mathsf{M}_\phi \\
&= \left(\sum_{n=1}^N \lambda_n \left(x_n\right)^{-r}\right)^{\frac{1}{-r}} \\
&= \mathsf{M}_{\phi(x;-r)}(\lparen x_n \rparen) && \text{by definition of } \mathsf{M}_\phi
\end{aligned}$$

4. Proof that $\lim_{r\to\infty} \mathsf{M}_\phi(\lparen x_n \rparen) = \max_{n\in\mathbb{Z}} \lparen x_n \rparen$:

   (a) Let $x_m \triangleq \max_{n\in\mathbb{Z}} \lparen x_n \rparen$

   (b) Note that $\lim_{r\to\infty} \mathsf{M}_\phi \le \max_{n\in\mathbb{Z}} \lparen x_n \rparen$ because

$$\begin{aligned}
\lim_{r\to\infty} \mathsf{M}_\phi(\lparen x_n \rparen) &= \lim_{r\to\infty}\left(\sum_{n=1}^N \lambda_n x_n^r\right)^{\frac{1}{r}} && \text{by definition of } \mathsf{M}_\phi \\
&\le \lim_{r\to\infty}\left(\sum_{n=1}^N \lambda_n x_m^r\right)^{\frac{1}{r}} && \text{by definition of } x_m \text{ in item (4a) and because } \phi(x) \triangleq x^r \text{ and } \phi^{-1} \text{ are both increasing or both decreasing} \\
&= \lim_{r\to\infty}\left(x_m^r \underbrace{\sum_{n=1}^N \lambda_n}_{1}\right)^{\frac{1}{r}} && \text{because } x_m \text{ is a constant} \\
&= \lim_{r\to\infty}\left(x_m^r \cdot 1\right)^{\frac{1}{r}} \\
&= x_m \\
&= \max_{n\in\mathbb{Z}} \lparen x_n \rparen && \text{by definition of } x_m \text{ in item (4a)}
\end{aligned}$$

(c) But also note that $\lim_{r\to\infty} \mathsf{M}_\phi \ge \max_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis$ because

$$\begin{aligned}
\lim_{r\to\infty} \mathsf{M}_\phi\left(\llparenthesis x_n \rrparenthesis\right) &= \lim_{r\to\infty}\left(\sum_{n=1}^{N} \lambda_n x_n^r\right)^{\frac{1}{r}} && \text{by definition of } \mathsf{M}_\phi \\
&\ge \lim_{r\to\infty}\left(w_m x_m^r\right)^{\frac{1}{r}} && \text{by definition of } x_m \text{ in item (4a) and because } \phi(x) \triangleq x^r \text{ and } \phi^{-1} \text{ are both increasing or both decreasing} \\
&= \lim_{r\to\infty} w_m^{\frac{1}{r}} x_m^{\frac{r}{r}} \\
&= x_m \\
&= \max_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis && \text{by definition of } x_m \text{ in item (4a)}
\end{aligned}$$

(d) Combining items (b) and (c) we have $\lim_{r\to\infty} \mathsf{M}_\phi = \max_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis$.

5. Proof that $\lim_{r\to-\infty} \mathsf{M}_\phi\left(\llparenthesis x_n \rrparenthesis\right) = \min_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis$:

$$\begin{aligned}
\lim_{r\to-\infty} \mathsf{M}_{\phi(x;r)}\left(\llparenthesis x_n \rrparenthesis\right) &= \lim_{r\to\infty} \mathsf{M}_{\phi(x;-r)}\left(\llparenthesis x_n \rrparenthesis\right) && \text{by change of variable } r \\
&= \lim_{r\to\infty} \left\{\mathsf{M}_{\phi(x;r)}\left(\llparenthesis x_n^{-1} \rrparenthesis\right)\right\}^{-1} && \text{by Lemma in item (3) page 185} \\
&= \lim_{r\to\infty} \frac{1}{\mathsf{M}_{\phi(x;r)}\left(\llparenthesis x_n^{-1} \rrparenthesis\right)} \\
&= \frac{\lim_{r\to\infty} 1}{\lim_{r\to\infty} \mathsf{M}_{\phi(x;r)}\left(\llparenthesis x_n^{-1} \rrparenthesis\right)} && \text{by property of lim [11]} \\
&= \frac{1}{\max_{n\in\mathbb{Z}} \llparenthesis x_n^{-1} \rrparenthesis} && \text{by item (4)} \\
&= \frac{1}{\left(\min_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis\right)^{-1}} \\
&= \min_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis
\end{aligned}$$

6. Proof that $\lim_{r\to 0} \mathsf{M}_\phi\left(\llparenthesis x_n \rrparenthesis\right) = \prod_{n=1}^{N} x_n^{\lambda_n}$:

$$\begin{aligned}
\lim_{r\to 0} \mathsf{M}_\phi\left(\llparenthesis x_n \rrparenthesis\right) &= \lim_{r\to 0} \exp\left\{\ln\left\{\mathsf{M}_\phi\left(\llparenthesis x_n \rrparenthesis\right)\right\}\right\} \\
&= \lim_{r\to 0} \exp\left\{\ln\left\{\left(\sum_{n=1}^{N} \lambda_n\left(x_n^r\right)\right)^{\frac{1}{r}}\right\}\right\} && \text{by definition of } \mathsf{M}_\phi \\
&= \exp\left.\left\{\frac{\frac{\partial}{\partial r}\ln\left(\sum_{n=1}^{N} \lambda_n\left(x_n^r\right)\right)}{\frac{\partial}{\partial r} r}\right\}\right|_{r=0} && \text{by l'Hôpital's rule [12]}
\end{aligned}$$

[11] Rudin (1976) page 85 ⟨4.4 Theorem⟩

$$
\begin{aligned}
&= \exp\left\{\frac{\sum_{n=1}^{N} \lambda_n \frac{\partial}{\partial r}(x_n^r)}{\sum_{n=1}^{N} \lambda_n (x_n^r)}\right\}_{r=0} \\
&= \exp\left\{\frac{\sum_{n=1}^{N} \lambda_n \frac{\partial}{\partial r}\exp\left(\ln\left(x_n^r\right)\right)}{\sum_{n=1}^{N} \lambda_n}\right\}_{r=0} \\
&= \exp\left\{\frac{\sum_{n=1}^{N} \lambda_n \frac{\partial}{\partial r}\exp\left(r\ln\left(x_n\right)\right)}{1}\right\}_{r=0} \\
&= \exp\left\{\sum_{n=1}^{N} \lambda_n \frac{\partial}{\partial r}\exp\left(r\ln\left(x_n\right)\right)\right\}_{r=0} \\
&= \exp\left\{\sum_{n=1}^{N} \lambda_n \exp\left\{r\ln x_n\right\}\ln\left(x_n\right)\right\}_{r=0} \\
&= \exp\left\{\sum_{n=1}^{N} \lambda_n \ln\left(x_n\right)\right\} \\
&= \exp\left\{\sum_{n=1}^{N} \ln\left(x_n^{\lambda_n}\right)\right\} \\
&= \exp\left\{\ln\prod_{n=1}^{N} x_n^{\lambda_n}\right\} \\
&= \prod_{n=1}^{N} x_n^{\lambda_n}
\end{aligned}
$$

✎

Definition 11.4. *Let* $\langle x_n \rangle_1^N$ *be a tuple. Let* $\langle \lambda_n \rangle_1^N$ *be a tuple of weighting values such that* $\sum_{n=1}^{N} \lambda_n = 1$.

[12] Rudin (1976) page 109 ⟨5.13 Theorem⟩

DEF

*The **harmonic mean** of (x_n) is defined as* $\mu_h \triangleq \left(\sum_{n=1}^{N} \lambda_n \frac{1}{x_n}\right)^{-1}$

*The **geometric mean** of (x_n) is defined as* $\mu_g \triangleq \sum_{n=1}^{N} x_n^{\lambda_n}$

*The **arithmetic mean** of (x_n) is defined as* $\mu_a \triangleq \sum_{n=1}^{N} \lambda_n x_n$

*The **average** of (x_n) is defined as* $\mu_a \triangleq \frac{1}{N}\sum_{n=1}^{N} x_n$

Corollary 11.1. [13] *Let $(x_n)_1^N$ be a tuple. Let $(\lambda_n)_1^N$ be a tuple of weighting values such that $\sum_{n=1}^{N} \lambda_n = 1$.*

COR

$$\min (x_n) \le \underbrace{\left(\sum_{n=1}^{N} \lambda_n \frac{1}{x_n}\right)^{-1}}_{\text{harmonic mean}} \le \underbrace{\prod_{n=1}^{N} x_n^{\lambda_n}}_{\text{geometric mean}} \le \underbrace{\sum_{n=1}^{N} \lambda_n x_n}_{\text{arithmetic mean}} \le \max (x_n)$$

✎PROOF:

1. These five means are all special cases of the *power mean* $\mathsf{M}_{\phi(x:r)}$ (Definition 11.3 page 184):
   - $r = \infty$: $\max (x_n)$
   - $r = 1$: arithmetic mean
   - $r = 0$: geometric mean
   - $r = -1$: harmonic mean
   - $r = -\infty$: $\min (x_n)$
2. The inequalities follow directly from Theorem 11.3 (page 184).
3. Generalized AM-GM inequality: If one is only concerned with the arithmetic mean and geometric mean, their relationship can be established directly using *Jensen's Inequality*:

$$\begin{aligned}
\sum_{n=1}^{N} \lambda_n x_n &= b^{\log_b\left(\sum_{n=1}^{N} \lambda_n x_n\right)} \\
&\ge b^{\left(\sum_{n=1}^{N} \lambda_n \log_b x_n\right)} && \text{by } \textit{Jensen's Inequality} \text{ (Theorem 9.1 page 144)} \\
&= \prod_{n=1}^{N} b^{(\lambda_n \log_b x_n)} \\
&= \prod_{n=1}^{N} b^{(\log_b x_n)\lambda_n} \\
&= \prod_{n=1}^{N} x_n^{\lambda_n}
\end{aligned}$$

⇨

[13] Bullen (2003) page 71, Bollobás (1999) page 5, Cauchy (1821) pages 457–459 ⟨Note II, theorem 17⟩, Jensen (1906) page 183

Lemma 11.2 (Young's Inequality). [14]

LEM

$$xy < \frac{x^p}{p} + \frac{y^q}{q} \quad \textit{with } \tfrac{1}{p} + \tfrac{1}{q} = 1 \quad \forall 1 < p < \infty, \; x, y \geq 0, \quad \textit{but } y \neq x^{p-1}$$

$$xy = \frac{x^p}{p} + \frac{y^q}{q} \quad \textit{with } \tfrac{1}{p} + \tfrac{1}{q} = 1 \quad \forall 1 < p < \infty, \; x, y \geq 0, \quad \textit{and } y = x^{p-1}$$

✎PROOF:

1. Proof that $\frac{1}{p-1} = q - 1$:

$$\begin{aligned}
\frac{1}{p} + \frac{1}{q} = 1 &\iff \frac{q}{q} + \frac{q}{p} = q \\
&\iff q\left(1 - \frac{1}{p}\right) = 1 \\
&\iff q = \frac{1}{1 - \frac{1}{p}} \\
&\iff q = \frac{p}{p-1} \\
&\iff q - 1 = \frac{p}{p-1} - \frac{p-1}{p-1} \\
&\iff q - 1 = \frac{p - (p-1)}{p-1} \\
&\iff q - 1 = \frac{1}{p-1}
\end{aligned}$$

2. Proof that $v = u^{p-1} \iff u = v^{q-1}$:

$$\begin{aligned}
u &= v^{\frac{1}{p-1}} && \text{by left hypothesis} \\
&= v^{q-1} && \text{by item (1)}
\end{aligned}$$

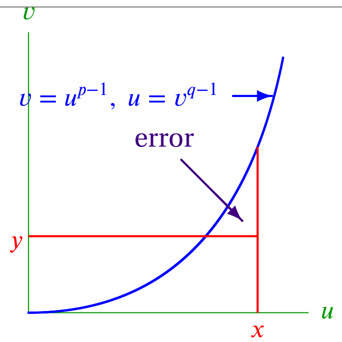
3. Proof that $v = u^{p-1}$ is monotonically increasing in u and $u = v^{q-1}$ is monotonically increasing in v:

$$\begin{aligned}
\frac{\mathrm{d}v}{\mathrm{d}u} &= \frac{\mathrm{d}}{\mathrm{d}u} u^{p-1} \\
&= (p-1)u^{p-2} \\
&> 0 \\
\frac{\mathrm{d}u}{\mathrm{d}v} &= \frac{\mathrm{d}}{\mathrm{d}v} v^{q-1} \\
&= (q-1)v^{q-2} \\
&> 0
\end{aligned}$$

4. Proof that $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$:

[14] Carothers (2000), page 43, Tolsted (1964), page 5, Maligranda (1995), page 257, Hardy et al. (1952) ⟨Theorem 24⟩, Young (1912) page 226

$$
\begin{aligned}
xy &\le \int_0^x u^{p-1}\,\mathrm{d}u + \int_0^y v^{q-1}\,\mathrm{d}v \\
&= \left.\frac{u^p}{p}\right|_0^x + \left.\frac{v^q}{q}\right|_0^y \\
&= \frac{x^p}{p} + \frac{y^q}{q}
\end{aligned}
$$



✏

Theorem 11.4 (Hölder's Inequality). [15] *Let* $(x_n \in \mathbb{C})_1^N$ *and* $(y_n \in \mathbb{C})_1^N$ *be complex* N*-tuples.*

THM

$$
\underbrace{\sum_{n=1}^{N} |x_n y_n|}_{\|\boldsymbol{x}\cdot\boldsymbol{y}\|_1} \le \underbrace{\left(\sum_{n=1}^{N} |x_n|^p\right)^{\frac{1}{p}}}_{\|\boldsymbol{x}\|_p} \underbrace{\left(\sum_{n=1}^{N} |y_n|^q\right)^{\frac{1}{q}}}_{\|\boldsymbol{y}\|_q} \qquad \textit{with} \qquad \frac{1}{p}+\frac{1}{q}=1 \qquad \forall\, 1<p<\infty
$$

✎PROOF: Let $\|x_n\|_p \triangleq \left(\sum_{n=1}^{N} |x_n|^p\right)^{\frac{1}{p}}$.

$$
\begin{aligned}
\sum_{n=1}^{N} |x_n y_n| &= \|(x_n)\|_p \|(y_n)\|_q \sum_{n=1}^{N} \frac{|x_n y_n|}{\|(x_n)\|_p \|(y_n)\|_q} \\
&= \|(x_n)\|_p \|(y_n)\|_q \sum_{n=1}^{N} \frac{|x_n|}{\|(x_n)\|_p} \frac{|y_n|}{\|(y_n)\|_q} \\
&\le \|(x_n)\|_p \|(y_n)\|_q \sum_{n=1}^{N} \left(\frac{1}{p}\frac{|x_n|^p}{\|(x_n)\|_p^p} + \frac{1}{q}\frac{|y_n|^q}{\|(y_n)\|_q^q}\right) && \text{by } \textit{Young's Inequality} \text{ (Lemma 11.2 page 189)} \\
&= \|(x_n)\|_p \|(y_n)\|_q \left(\frac{1}{p}\cdot\frac{\sum |x_n|^p}{\|(x_n)\|_p^p} + \frac{1}{q}\cdot\frac{\sum |y_n|^q}{\|(y_n)\|_q^q}\right) \\
&= \|(x_n)\|_p \|(y_n)\|_q \left(\frac{1}{p}\frac{\|(x_n)\|_p^p}{\|(x_n)\|_p^p} + \frac{1}{q}\frac{\|(y_n)\|_q^q}{\|(y_n)\|_q^q}\right) && \text{by definition of } \|\cdot\| \\
&= \|(x_n)\|_p \|(y_n)\|_q \underbrace{\left(\frac{1}{p}+\frac{1}{q}\right)}_{1} \\
&= \|(x_n)\|_p \|(y_n)\|_q && \text{by } \frac{1}{p}+\frac{1}{q}=1 \text{ constraint}
\end{aligned}
$$

✏

Theorem 11.5 (Minkowski's Inequality for sequences). [16] *Let* $(x_n \in \mathbb{C})_1^N$ *and* $(y_n \in \mathbb{C})_1^N$ *be complex* N*-tuples.*

[15] Bullen (2003) page 178 ⟨2.1⟩, Carothers (2000), page 44, Tolsted (1964), page 6, Maligranda (1995), page 257, Hardy et al. (1952) ⟨Theorem 11⟩, Hölder (1889)

[16] Bullen (2003) page 179, Carothers (2000), page 44, Tolsted (1964), page 7, Maligranda (1995), page 258, Hardy et al. (1952) ⟨Theorem 24⟩, Minkowski (1910) page 115

THM
$$\left(\sum_{n=1}^{N}|x_n+y_n|^p\right)^{\frac{1}{p}} \le \left(\sum_{n=1}^{N}|x_n|^p\right)^{\frac{1}{p}} + \left(\sum_{n=1}^{N}|y_n|^p\right)^{\frac{1}{p}} \qquad \forall\ 1<p<\infty$$

✎PROOF:

1. Define q in terms of p such that $\frac{1}{p}+\frac{1}{q}=1$
2. Proof that $\frac{1}{q}=\frac{p-1}{p}$:
$$\begin{aligned}\frac{1}{p}+\frac{1}{q}=1 &\iff \frac{1}{q}=1-\frac{1}{p}\\ &\iff \frac{1}{q}=\frac{p-1}{p}\end{aligned}$$
3. Define $\|\cdot\|$ as follows:
$$\|\boldsymbol{x}\|_p \triangleq \left(\sum_{n=1}^{N}|x_n|^p\right)^{\frac{1}{p}}$$
4. Proof that $\|x_n+y_n\|_p \le \|x_n\|_p+\|y_n\|_p$:
$$\begin{aligned}
\|x_n+y_n\|_p^p &= \sum_{n=1}^{N}|x_n+y_n|^p && \text{by definition of } \|\cdot\|_p\\
&= \sum_{n=1}^{N}|x_n+y_n||x_n+y_n|^{p-1} && \text{by homogeneous property of } |\cdot|\\
&\le \sum_{n=1}^{N}|x_n||x_n+y_n|^{p-1}+\sum_{n=1}^{N}|y_n||x_n+y_n|^{p-1} && \text{by subadditive property of } |\cdot|\\
&= \sum_{n=1}^{N}\left|x_n(x_n+y_n)^{p-1}\right|+\sum_{n=1}^{N}\left|y_n(x_n+y_n)^{p-1}\right| && \text{by homogeneous property of } |\cdot|\\
&\le \|x_n\|_p\left\|(x_n+y_n)^{p-1}\right\|_q+\|y_n\|_p\left\|(x_n+y_n)^{p-1}\right\|_q && \text{by Hölder's Inequality page 190}\\
&= \left(\|x_n\|_p+\|y_n\|_p\right)\left\|(x_n+y_n)^{p-1}\right\|_q\\
&= \left(\|x_n\|_p+\|y_n\|_p\right)\left(\sum_{n=1}^{N}\left|(x_n+y_n)^{p-1}\right|^q\right)^{\frac{1}{q}} && \text{by definition of } \|\cdot\|_p\\
&= \left(\|x_n\|_p+\|y_n\|_p\right)\left(\sum_{n=1}^{N}\left|(x_n+y_n)^{p-1}\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} && \text{by item (2)}\\
&= \left(\|x_n\|_p+\|y_n\|_p\right)\left(\sum_{n=1}^{N}\left|(x_n+y_n)\right|^p\right)^{\frac{p-1}{p}}\\
&= \left(\|x_n\|_p+\|y_n\|_p\right)\|x_n+y_n\|^{p-1}\\
&\implies \|x_n+y_n\|_p \le \|x_n\|_p+\|y_n\|_p
\end{aligned}$$

✏

❝*Cauchy is the only one occupied with pure mathematics: Poisson, Fourier, Ampere, etc., busy themselves exclusively with magnetism and other physical subjects.*❞
in an 1826 letter written by Niels Henrik Abel [17]

Theorem 11.6 (Cauchy-Schwarz Inequality). [18] *Let $⦇x_n \in \mathbb{C}⦈_1^N$ and $⦇y_n \in \mathbb{C}⦈_1^N$ be complex N-tuples.*

THM

$$\left|\sum_{n=1}^N x_n y_n^*\right|^2 \le \left(\sum_{n=1}^N |x_n|^2\right)\left(\sum_{n=1}^N |y_n|^2\right) \qquad \forall \boldsymbol{x},\boldsymbol{y}\in\boldsymbol{X}$$

$$\left|\sum_{n=1}^N x_n y_n^*\right|^2 = \left(\sum_{n=1}^N |x_n|^2\right)\left(\sum_{n=1}^N |y_n|^2\right) \iff \exists a \in \mathbb{C} \text{ such that } \boldsymbol{y} = a\boldsymbol{x} \qquad \forall \boldsymbol{x},\boldsymbol{y}\in\boldsymbol{X}$$

✎PROOF:

1. The Cauchy-Schwarz inequality for sequences is a special case of the Hölder inequality (Theorem 11.4) for $p = q = 2$.
2. Alternatively, the Cauchy-Schwarz inequality for sequences is a special case of the *Cauchy-Schwarz inequality* in inner-product spaces:
   (a) $\langle x_n \mid y_n \rangle \triangleq \sum_{n=1}^N x_n y_n$ is an inner-product and $\left(⦇x_n⦈, \langle \triangle \mid \triangledown \rangle\right)$ is an inner-product space.
   (b) By the *Cauchy-Schwarz Inequality for inner-product spaces* (Theorem 6.2 page 96),
   $$|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 \le \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \; \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle$$
3. Not only does the Hölder inequality imply the Cauchy-Schwarz inequality, but somewhat surprisingly, the converse is also true: The Cauchy-Schwarz inequality implies the Hölder inequality.[19]

⇨

Proposition 11.1. [20]

PRP

$$(x+y)^p \le 2^p(x^p + y^p) \qquad \forall x, y \ge 0, \; 1 < p < \infty$$

✎PROOF:

$$\begin{aligned}(x+y)^p &\le (2\max\{x,y\})^p \\ &= 2^p(\max\{x,y\})^p \\ &= 2^p(\max\{x^p,y^p\}) \\ &\le 2^p(x^p + y^p)\end{aligned}$$

⇨

[17] quote: Boyer and Merzbach (2011) page 462
image: http://en.wikipedia.org/wiki/File:Augustin-Louis_Cauchy_1901.jpg, public domain
[18] Aliprantis and Burkinshaw (1998), page 278, Scharz (1885), Bouniakowsky (1859), Hardy et al. (1952) page 25 ⟨Theorem 11⟩, Cauchy (1821) page 455 ⟨???⟩
[19] Bullen (2003) pages 183–185 ⟨Theorem 5⟩
[20] Carothers (2000), page 43

11.3 Power Sums

Theorem 11.7 (Geometric Series). [21]

THM
$$\sum_{k=0}^{n-1} r^k = \frac{1-r^n}{1-r} \qquad \forall r \in \mathbb{C}\backslash 0$$

✎PROOF:

$$\begin{aligned}
\sum_{k=0}^{n-1} r^k &= \left(\frac{1}{1-r}\right)\left[(1-r)\sum_{k=0}^{n-1} r^k\right] \\
&= \left(\frac{1}{1-r}\right)\left[\sum_{k=0}^{n-1} r^k - r\sum_{k=0}^{n-1} r^k\right] \\
&= \left(\frac{1}{1-r}\right)\left[\sum_{k=0}^{n-1} r^k - \left(\sum_{k=0}^{n-1} r^k - 1 + r^n\right)\right] \\
&= \left(\frac{1}{1-r}\right)\left[1-r^n\right] \\
&= \frac{1-r^n}{1-r}
\end{aligned}$$

⇨

Lemma 11.3. *Let* $\mathsf{f}(t)$ *be a function.*

LEM
$$\mathsf{S}(x) \triangleq \sum_{n\in\mathbb{Z}} \mathsf{f}(x+n\tau) = \mathsf{S}(x+\tau) \qquad (\mathsf{S}(x) \text{ is PERIODIC with period } \tau)$$

✎PROOF:

$$\begin{aligned}
\mathsf{S}(x+\tau) &\triangleq \sum_{n\in\mathbb{Z}} \mathsf{f}(x+\tau+n\tau) \\
&= \sum_{n\in\mathbb{Z}} \mathsf{f}(x+(n+1)\tau) \\
&= \sum_{m\in\mathbb{Z}} \mathsf{f}(x+m\tau) && \text{where } m \triangleq n+1 \\
&\triangleq \mathsf{S}(x)
\end{aligned}$$

⇨

Proposition 11.2 (Power Sums). [22]

PRP
$$\begin{aligned}
\sum_{k=1}^{n} k &= \frac{n(n+1)}{2} && \forall n\in\mathbb{N} \\
\sum_{k=1}^{n} k^2 &= \frac{n(n+1)(2n+1)}{6} && \forall n\in\mathbb{N} \\
\sum_{k=1}^{n} k^3 &= \frac{n^2(n+1)^2}{4} && \forall n\in\mathbb{N} \\
\sum_{k=1}^{n} k^4 &= \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30} && \forall n\in\mathbb{N}
\end{aligned}$$

[21] Hall and Knight (1894), page 39 ⟨article 55⟩

[22] Amann and Escher (2008) pages 51–57, Menini and Oystaeyen (2004) page 91 ⟨Exercises 5.36–5.39⟩

✎PROOF:

1. Proof that $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$: (proof by induction)

$$\begin{aligned}
\sum_{k=1}^{n=1} k &= 1 \\
&= \frac{1(1+1)}{2} \\
&= \left.\frac{n(n+1)}{2}\right|_{n=1} \\
\sum_{k=1}^{n+1} k &= \left(\sum_{k=1}^{n} k\right) + \left(n+1\right) \\
&= \left(\frac{n(n+1)}{2}\right) + \left(n+1\right) && \text{by left hypothesis} \\
&= \left(n+1\right)\left(\frac{n}{2}+1\right) \\
&= \left(n+1\right)\left(\frac{n+2}{2}\right) \\
&= \frac{(n+1)(n+2)}{2}
\end{aligned}$$

2. Proof that $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$: (proof by induction)

$$\begin{aligned}
\sum_{k=1}^{n=1} k^2 &= 1 \\
&= \frac{1(1+1)(2+1)}{6} \\
&= \left.\frac{n(n+1)(2n+1)}{6}\right|_{n=1} \\
\sum_{k=1}^{n+1} k^2 &= \left(\sum_{k=1}^{n} k^2\right) + \left(n+1\right)^2 \\
&= \left(\frac{n(n+1)(2n+1)}{6}\right) + \left(n+1\right)^2 && \text{by left hypothesis} \\
&= \left(n+1\right)\left(\frac{n(2n+1)+6(n+1)}{6}\right) \\
&= \left(n+1\right)\left(\frac{2n^2+7n+6}{6}\right) \\
&= \left(n+1\right)\left(\frac{(n+2)(2n+3)}{6}\right) \\
&= \frac{(n+1)[(n+1)+1][2(n+1)+1]}{6}
\end{aligned}$$

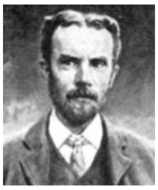
✏

CHAPTER 12

INFINITE SUMS

❝*The divergent series are the invention of the devil, and it is a shame to base on them any demonstration whatsoever. By using them, one may draw any conclusion he pleases and that is why these series have produced so many fallacies and so many paradoxes. . .*❞

Niels Henrik Abel, in a January 16, 1826 letter to Holmboé [1]



❝*The series is divergent; therefore we may be able to do something with it.*❞
Oliver Heaviside (1850–1925) [2]

❝*Some modern appraisals of the cavalier style of 18th-century mathematicians in handling infinite series conveys the impression that these poor men set their brains aside when confronted by them.*❞

Ivor Grattan-Guinness (1990)[3]

12.1 Convergence

An infinite summation $\sum_{n=1}^{\infty} x_n$ is meaningless outside some topological space (e.g. metric space, normed space, etc.). The sum $\sum_{n=1}^{\infty} \boldsymbol{x}_n$ is an abbreviation for $\lim_{N\to\infty} \sum_{n=1}^{N} \boldsymbol{x}_n$ (next definition); and

[1] quote: Kline (1972) page 973 ⟨Chapter 47⟩
image: http://en.wikipedia.org/wiki/File:Niels_Henrik_Abel.jpg, public domain

[2] quote: Kline (1972) page 1096 ⟨Chapter 47⟩
image: http://en.wikipedia.org/wiki/File:Oliver_Heaviside2.jpg, public domain

[3] Grattan-Guinness (1990) page 163

the concept of *limit* (Definition 8.3 page 127) is also itself meaningless outside of a *topological space* (Definition 1.1 page 3).

Definition 12.1. [4] *Let $(X, \mathbf{T})$ be a topological space and* lim *be the limit generated by the topology* $\mathbf{T}$.

DEF

$$\sum_{n=1}^{\infty} x_n \triangleq \sum_{n\in\mathbb{N}} x_n \triangleq \lim_{N\to\infty} \sum_{n=1}^{N} x_n$$

$$\sum_{n=-\infty}^{\infty} x_n \triangleq \sum_{n\in\mathbb{Z}} x_n \triangleq \lim_{N\to\infty} \left(\sum_{n=0}^{N} x_n\right) + \left(\lim_{N\to-\infty} \sum_{n=-1}^{N} x_n\right)$$

In general, the order of summation of an infinite series *does* matter.

Definition 12.2. [5] *Let P be the set of all* PERMUTATIONS *in* $\mathbb{N}^{\mathbb{N}}$.

DEF

A series $\sum_{n=1}^{\infty} x_n$ *is **absolutely convergent** if* $\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty} x_{\mathrm{p}(n)} \quad \forall \mathrm{p} \in P$

*A series is **conditionally convergent** if it is* CONVERGENT *but not* ABSOLUTELY CONVERGENT.

Theorem 12.1 (Riemann Series Theorem). [6] *Let $\mathrm{p}(n)$ be a permutation on $\mathbb{N}$. Let $(a_n)_{n\in\mathbb{Z}}$ be a sequence of real numbers.*

THM

$$\left\{ \sum_{n=1}^{\infty} a_n \text{ is CONDITIONALLY CONVERGENT} \right\} \implies \left\{ \begin{array}{l} \text{For every } x \in \mathbb{R} \text{ there exists } \mathrm{p} \\ \text{such that } \sum_{n=1}^{\infty} a_{\mathrm{p}(n)} = x \\ \text{or such that } \sum_{n=1}^{\infty} a_{\mathrm{p}(n)} \text{ is DIVERGENT} \end{array} \right\}$$

Theorem 12.2. [7]

THM

$$\sum_{n=1}^{\infty} |x_n| < \infty \implies \left\{ \sum_{n=1}^{\infty} x_n \text{ is ABSOLUTELY CONVERGENT.} \right\}$$

Example 12.1 (Logarithmic Series). [8] Consider the sum $\sum_{n=1}^{\infty}(-1)^{(n-1)}\frac{1}{n}$. To which value this sum converges, or whether it even converges at all, depends on the order in which the terms are summed. This is demonstrated by the following series:

[4] Klauder (2010) page 4, Kubrusly (2001) page 43, Bachman and Narici (1966) pages 3–4

[5] Kadets and Kadets (1997) page 5 ⟨Theorem 1.1.1 (Riemann's theorem)⟩, Bromwich (1908), page 64 ⟨IV. Absolute Convergence.⟩, Szász and Barlaz (1952), page 2

[6] Kadets and Kadets (1997) page 5 ⟨Theorem 1.1.1 (Riemann's theorem)⟩, Bromwich (1908), page 68 ⟨Article 28. Riemann's Theorem⟩

[7] Kadets and Kadets (1997) page 5 ⟨Theorem 1.1.1 (Riemann's theorem)⟩

[8] Bromwich (1908), pages 51–52 ⟨Article 21 Example 1⟩, Hall and Knight (1894), page 191 ⟨article 223⟩, Jolley (1961) pages 14–15 ⟨item (71)⟩, Sloane (2014) ⟨http://oeis.org/A002939⟩⟨2n(2n − 1)⟩, Graham et al. (1994) page 99 ⟨n.w. diagonal of *spiral function*⟩, Many many thanks to Po-Ning Chen (Chinese: ???, pinyin: Chén Bó Níng) for his consultation regarding this series.

EX

If the series is added in the given order, the result is ln 2:

$$\sum_{n=1}^{\infty}(-1)^{(n-1)}\frac{1}{n} \triangleq \lim_{N\to\infty}\sum_{n=1}^{N}(-1)^{(n-1)}\frac{1}{n} = \ln 2$$

But if the order is changed, the sum can be any real value:

Let x be any real value (even an irrational one such as π or $\sqrt{2}$).

$$\text{Let} \quad \mathsf{p}(N) \triangleq \begin{cases} \text{next unused odd value} & \text{if } \sum_{n=1}^{N-1}(-1)^{(\mathsf{p}(n)-1)}\frac{1}{\mathsf{p}(n)} \le x \\ \text{next unused even value} & \text{if } \sum_{n=1}^{N-1}(-1)^{(\mathsf{p}(n)-1)}\frac{1}{\mathsf{p}(n)} > x \end{cases}$$

$$\sum_{n=1}^{\infty}(-1)^{(\mathsf{p}(n)-1)}\frac{1}{\mathsf{p}(n)} \triangleq \lim_{N\to\infty}\sum_{n=1}^{N}(-1)^{(\mathsf{p}(n)-1)}\frac{1}{\mathsf{p}(n)} = x$$

The series can even be summed in such a way that it does not converge at all:

Let $\mathsf{q}(n)$ be a permutation that partitions the natural numbers into odd and even values such that $\left(\!\left(x_{\mathsf{q}(n)} \right)\!\right) = (\!(1, 3, 5, \ldots, 2, 4, 6, \ldots)\!)$.

$$\sum_{n=1}^{\infty}(-1)^{(\mathsf{q}(n)-1)}\frac{1}{\mathsf{q}(n)} = \underbrace{\sum_{n=1}^{\infty}\frac{1}{2n-1}}_{\infty} - \underbrace{\sum_{n=1}^{\infty}\frac{1}{2n}}_{\infty} \implies \text{diverges}$$

✎PROOF:

1. Proof that $\sum_{n=1}^{\infty}(-1)^{(n-1)}\frac{1}{n} = \ln 2$ using polynomial expansion:[9]

   (a) Lemma: Proof that $\frac{1}{1+x} = \sum_{k=0}^{2n-1}(-1)^k x^k + \frac{x^{2n}}{1+x}$:

$$\begin{aligned}(1+x)\left(\sum_{k=0}^{2n-1}(-1)^k x^k + \frac{x^{2n}}{1+x}\right) &= \sum_{k=0}^{2n-1}(-1)^k x^k + \sum_{k=0}^{2n-1}(-1)^k x^{k+1} + x^{2n}\\ &= 1 + \sum_{k=1}^{2n-1}(-1)^k x^k - \sum_{k=1}^{2n-1}(-1)^k x^k + x^{2n}\\ &= 1 + x^{2n}\end{aligned}$$

   (b) Lemma: Proof that $\lim_{n\to\infty}\int_0^1 \frac{x^{2n}}{1+x}\,\mathrm{d}x = 0$:

$$\begin{aligned}0 &< \lim_{n\to\infty}\int_0^1 \frac{x^{2n}}{1+x}\,\mathrm{d}x\\ &< \lim_{n\to\infty}\int_0^1 x^{2n}\,\mathrm{d}x\\ &= \lim_{n\to\infty}\left.\frac{x^{2n+1}}{2n+1}\right|_0^1\\ &= \lim_{n\to\infty}\frac{1}{2n+1}\\ &= 0\end{aligned}$$

9 Bromwich (1908), pages 51–52 ⟨Article 21 Example 1⟩

(c) Proof that sum = ln 2:

$$\begin{aligned}
\ln 2 &= \ln 2 - \ln 1 \\
&= \int_1^2 \frac{1}{x}\,\mathrm{d}x \\
&= \int_0^1 \frac{1}{x+1}\,\mathrm{d}x \\
&= \lim_{n\to\infty} \int_0^1 \left\{ \sum_{k=0}^{2n-1} (-1)^k x^k + \frac{x^{2n}}{1+x} \right\} \mathrm{d}x && \text{by item (1a)} \\
&= \sum_{k=0}^{2n-1} (-1)^k \int_0^1 x^k\,\mathrm{d}x + \lim_{n\to\infty} \int_0^1 \frac{x^{2n}}{1+x}\,\mathrm{d}x \\
&= \lim_{n\to\infty} \sum_{k=0}^{2n-1} (-1)^k \frac{1}{k} + 0 && \text{by item (1b)} \\
&= \lim_{n\to\infty} \sum_{k=1}^{2n} (-1)^{k-1} \frac{1}{k} \\
&= \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}
\end{aligned}$$

2. Proof that $\sum_{n=1}^{\infty} (-1)^{(n-1)} \frac{1}{n} = \ln 2$ using Taylor expansion:[10]

(a) Lemma: Proof that $\ln(x+1) = \sum_{n=1}^{\infty} (-1)^{(n+1)} \frac{1}{n} x^n$:

$$\begin{aligned}
\ln(x+1) &= \sum_{n=0}^{\infty} \frac{\left[\mathbf{D}^n \ln(x+1)\right](0)}{n!} x^n && \text{by \textit{Taylor series expansion}} \\
&= \frac{\ln(1+0)}{0!}x^0 + \frac{\frac{1}{1+0}}{1!}x^1 - \frac{\frac{1}{(1+0)^2}}{2!}x^2 + \frac{\frac{2}{(1+0)^3}}{3!}x^3 - \frac{\frac{6}{(1+0)^4}}{4!}x^4 + \frac{\frac{24}{(1+0)^5}}{5!}x^5 - \frac{\frac{120}{(1+0)^6}}{6!}x^6 + \cdots \\
&= 0 + x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \frac{1}{5}x^5 - \frac{1}{6}x^6 + \cdots \\
&= \sum_{n=1}^{\infty} (-1)^{(n+1)} \frac{1}{n} x^n
\end{aligned}$$

(b) Proof that $\sum_{n=1}^{\infty} (-1)^{(n-1)} \frac{1}{n} = \ln 2$:

$$\begin{aligned}
&= \underbrace{\left(1 - \frac{1}{2}\right)}_{\frac{1}{2}} + \underbrace{\left(\frac{1}{3} - \frac{1}{4}\right)}_{\frac{1}{12}} + \underbrace{\left(\frac{1}{5} - \frac{1}{6}\right)}_{\frac{1}{30}} + \underbrace{\left(\frac{1}{7} - \frac{1}{8}\right)}_{\frac{1}{56}} + \ldots \\
&= \left.\sum_{n=1}^{\infty} (-1)^{(n+1)} \frac{1}{n} x^n \right|_{x=1} \\
&= \ln 2 && \text{by Lemma in item 1} \\
&\approx 0.693147
\end{aligned}$$

[10] Hall and Knight (1894), page 191 ⟨article 223⟩

3. Proof that $\sum_{n=1}^{\infty}(-1)^{(\mathsf{p}(n)-1)}\frac{1}{\mathsf{p}(n)} = x$: If the partial sum is less than x, positive values are added. If the partial sum is greater than x, negative values are added. The limit is x.

4. Proof that $\sum_{n=1}^{\infty}(-1)^{(\mathsf{q}(n)-1)}\frac{1}{\mathsf{q}(n)} = \infty - \infty$:

$$\begin{aligned}
\sum_{n=1}^{\infty}(-1)^{(\mathsf{q}(n)-1)}\frac{1}{\mathsf{q}(n)} &= \underbrace{\sum_{n=1}^{\infty}(-1)^{(2n-1-1)}\frac{1}{2n-1}}_{\text{odd indices}} + \underbrace{\sum_{n=1}^{\infty}(-1)^{(2n-1)}\frac{1}{2n}}_{\text{even indices}} \\
&= \underbrace{\sum_{n=1}^{\infty}\frac{1}{2n-1}}_{\infty} - \underbrace{\sum_{n=1}^{\infty}\frac{1}{2n}}_{\infty} \\
&\to \text{(diverges)}
\end{aligned}$$

✏

Divergent series could even result in decisions that may be considered extremely irrational, as demonstrated by *St. Petersberg Paradox* (next).

Example 12.2 (St. Petersburg Paradox). [11] There is a lottery with a prize pot of \$1. A coin is tossed. If the coin is a tail, the money in the lottery is doubled (\$2, \$4, \$8, \$16, ...). If the coin is a head, you win the money and the game is finished.

How much money would you be willing to play this game? The answer to this question for some people may depend on the expected value of how much money would be won. But the expected value of the amount of money you would win is

$$\tfrac{1}{2}\times\$1 + \tfrac{1}{4}\times\$2 + \tfrac{1}{8}\times\$4 + \tfrac{1}{16}\times\$8 + \ldots = \tfrac{1}{2}+\tfrac{1}{2}+\tfrac{1}{2}+\tfrac{1}{2}+\ldots = \infty$$

Since the expected value of the win is infinity, you should be willing to pay any finite amount of money to play this game (even trillions of dollars). But yet common sense would tell most people that this would be an unwise investment.

12.2 Multiplication

Theorem 12.3. [12] *Let* $(\!(x_n)\!)_1^N$ *and* $(\!(y_n)\!)_1^N$ *be sequences over a ring* $(\mathbb{X}, +, \times)$.

THM

$$\left(\sum_{n=0}^{p} x_n\right)\left(\sum_{m=0}^{q} y_m\right) = \sum_{n=0}^{p+q}\underbrace{\left(\sum_{k=\max(0,n-q)}^{\min(n,p)} x_k y_{n-k}\right)}_{\textit{Cauchy product}}$$

✎PROOF:

[11] Székely (1986) pages 27–28, Bernoulli (1783) pages 31–32 ⟨§17⟩, de Montmort (1713) page 402 ⟨1713 letter from Nicolas Bernoulli⟩

[12] Apostol (1975) page 204 ⟨Definition 8.45⟩

1.
$$\begin{aligned}
\left(\sum_{n=0}^{p} x_n\right)\left(\sum_{m=0}^{q} y_m\right) &= \sum_{n=0}^{p}\sum_{m=0}^{q} x_n y_m z^{n+m} \\
&= \sum_{n=0}^{p}\sum_{k=n}^{q+n} x_n y_{k-n} && k=n+m \qquad m=k-n \\
&\vdots \\
&= \sum_{n=0}^{p+q}\left(\sum_{k=0}^{n} x_k y_{n-k}\right)
\end{aligned}$$

2. Perhaps the easiest way to see the relationship is by illustration with a matrix of product terms:

	y_0	y_1	y_2	y_3	$\cdots$	y_q
x_0	x_0y_0	x_0y_1	x_0y_2	x_0y_3	$\cdots$	x_0y_q
x_1	x_1y_0	x_1y_1	x_1y_2	x_1y_3	$\cdots$	x_1y_q
x_2	x_2y_0	x_2y_1	x_2y_2	x_2y_3	$\cdots$	x_2y_q
x_3	x_3y_0	x_3y_1	x_3y_2	x_3y_3	$\cdots$	x_3y_q
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_p	x_py_0	x_py_1	x_py_2	x_py_3	$\cdots$	x_py_q

(a) The expression $\sum_{n=0}^{p}\sum_{m=0}^{q} x_n y_m z^{n+m}$ is equivalent to adding *horizontally* from left to right, from the first row to the last.

(b) If we switched the order of summation to $\sum_{m=0}^{q}\sum_{n=0}^{p} x_n y_m z^{n+m}$, then it would be equivalent to adding *vertically* from top to bottom, from the first column to the last.

(c) However the final result expression $\sum_{n=0}^{p+q}\left(\sum_{k=0}^{n} x_k y_{n-k}\right)$ is equivalent to adding *diagonally* starting from the upper left corner and proceding to the lower right.

(d) Upper limit on inner summation: Looking at the x_k terms, we see that there are two constraints on k:
$$\left.\begin{array}{ccc} k & \le & n \\ k & \le & p \end{array}\right\} \implies k \le \min(n,p)$$

(e) Lower limit on inner summation: Looking at the x_k terms, we see that there are two constraints on k:
$$\left.\begin{array}{ccc} k & \ge & 0 \\ k & \ge & n-q \end{array}\right\} \implies k \ge \max(0,n-q)$$

Corollary 12.1. *Let* $(\!|x_n \in \mathbb{C}|\!)$ *and* $(\!|y_n \in \mathbb{C}|\!)$.

COR
$$\left(\sum_{n=0}^{\infty} x_n\right)\left(\sum_{m=0}^{\infty} y_m\right) = \underbrace{\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} x_k y_{n-k}\right)}_{\textit{Cauchy product}}$$

PROOF:
$$\begin{aligned}
\left(\sum_{n=0}^{p=\infty} x_n\right)\left(\sum_{m=0}^{q=\infty} y_m\right) &= \sum_{n=0}^{\infty}\left[\sum_{k=\max(0,n-\infty)}^{\min(n,\infty)} x_k y_{n-k}\right] && \text{by Theorem 12.3 page 199} \\
&= \sum_{n=0}^{\infty}\left((\sum_{k=0}^{n} x_k y_{n-k}\right)
\end{aligned}$$

Theorem 12.4. [13] *Let* $X \triangleq \sum_{n=0}^{\infty} x_n$, $Y \triangleq \sum_{n=0}^{\infty} y_n$, *and* $Z \triangleq \left(\sum_{n=0}^{\infty} \sum_{k=0}^{n} x_k y_{n-k} \right)$.

THM
$$\left\{ \begin{array}{ll} X \text{ is ABSOLUTELY CONVERGENT} & \text{and} \\ Y \text{ is ABSOLUTELY CONVERGENT} & \end{array} \right\} \implies \left\{ \begin{array}{ll} Z \text{ is ABSOLUTELY CONVERGENT} & \text{and} \\ Z = XY. & \end{array} \right\}$$

Theorem 12.5. [14] *Let* $X \triangleq \sum_{n=0}^{\infty} x_n$, $Y \triangleq \sum_{n=0}^{\infty} y_n$, *and* $Z \triangleq \left(\sum_{n=0}^{\infty} \sum_{k=0}^{n} x_k y_{n-k} \right)$.

THM
$$\left\{ \begin{array}{lll} 1. & X \text{ is ABSOLUTELY CONVERGENT} & \text{and} \\ 2. & Y \text{ is CONVERGENT} & \end{array} \right\} \implies \left\{ \begin{array}{lll} 1. & Z \text{ is CONVERGENT} & \text{and} \\ 2. & Z = XY & \end{array} \right\}$$

Theorem 12.6. [15] *Let* $X \triangleq \sum_{n=0}^{\infty} x_n$, $Y \triangleq \sum_{n=0}^{\infty} y_n$, *and* $Z \triangleq \left(\sum_{n=0}^{\infty} \sum_{k=0}^{n} x_k y_{n-k} \right)$.

THM
$$\left\{ \begin{array}{lll} 1. & X \text{ is CONVERGENT} & \text{and} \\ 2. & Y \text{ is CONVERGENT} & \text{and} \\ 3. & Z \text{ is CONVERGENT} & \end{array} \right\} \implies \{ Z = XY \}$$

12.3 Summability

Cauchy and Abel, the 19th century champions of rigour in analysis, firmly rejected any and all divergent sums. However in more recent times, certain classes of divergent sums have been found to be extremely useful. Often such sums are ones that are said to be *summable*.

Definition 12.3. [16]

DEF

The series $\sum_{n=0}^{\infty} x_n$ *is* ***summable by the*** k***-th arithmetic mean of Cesàro to limit*** x, *or* ***summable*** (C, k) ***to the limit*** x, *if*

$$\lim_{n\to\infty} \frac{S_n^k}{A_n^k} = x \quad \text{for } n \in \mathbb{W} \text{ and where}$$

$$S_n^k \triangleq \begin{cases} \sum_{m=0}^{n} x_m & \text{for } k = 0 \\ \sum_{m=0}^{n} S_m^{k-1} & \text{for } k = 1, 2, 3, \ldots \end{cases} \quad \text{and} \quad A_n^k \triangleq \begin{cases} 1 & \text{for } k = 0 \\ \sum_{m=0}^{n} A_m^{k-1} & \text{for } k = 1, 2, 3, \ldots \end{cases}$$

Proposition 12.1. [17]

[13] Hardy (1949), pages 227–228 ⟨Theorem 160⟩, Bromwich (1908), page 66 ⟨Article 27.⟩, Cauchy (1821) pages 147–148 ⟨6.^{e} Théorème⟩

[14] Hardy (1949), page 228 ⟨Theorem 161⟩, Bromwich (1908), pages 85–86 ⟨Article 35.⟩, Mertens (1875)

[15] Hardy (1949), page 228 ⟨Theorem 162⟩, Abel (1826)

[16] Zygmund (2002) pages 75–76, Hardy (1949), page 96 ⟨5.4 Cesàro means⟩, Whittaker and Watson (1920), pages 155–156 ⟨8.43, 8.431⟩, Cesàro (1890)

[17] Zygmund (2002) pages 75–76, Thomson et al. (2008) page 129 ⟨Definition 3.54⟩, Szász and Barlaz (1952), page 13

PRP

$$\sum_{n=0}^{\infty} x_n \quad \textit{is summable } (C,0) \textit{ to the limit } x \textit{ if } \quad \lim_{N\to\infty} \sum_{n=0}^{N} x_n = x \qquad \textit{(normal convergence)}$$

$$\sum_{n=0}^{\infty} x_n \quad \textit{is summable } (C,1) \textit{ to the limit } x \textit{ if } \quad \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} s_n = x \qquad \textit{(arthimetic mean)}$$

$$\textit{where } s_n \triangleq \sum_{m=0}^{n} x_m$$

Definition 12.4. [18]

DEF

The series $\sum_{n=0}^{\infty} a_n$ *is* ***summable by Euler's method to limit*** a *if*

$$\lim_{x\to 1-0} \sum_{n=0}^{\infty} a_n x^n = a$$

Example 12.3. [19]

EX

The series $\sum_{n=0}^{\infty} (-1)^n = 1 - 1 + 1 - 1 + 1 - 1 + \cdots$ is *divergent*
(it is *not* summable $(C,0)$),
but yet it *is* summable $(C,1)$ to the limit $\frac{1}{2}$.
It is also summable by Euler's method to the limit $\frac{1}{2}$.

✎PROOF:

1. Proof for Cesàro summability:

   (a) Note that the sequence of partial sums s_n is $s_0 = 1$, $s_1 = 0$, $s_2 = 1$, $s_3 = 0$, $s_4 = 1$, ... That is
   $$s_n = \begin{cases} 1 & \text{for } n \text{ even} \\ 0 & \text{for } n \text{ odd} \end{cases}$$

   (b) Then

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} s_k &= \lim_{n\to\infty} \frac{1}{2n+1} \sum_{k=0}^{2n} s_k \\
&= \lim_{n\to\infty} \frac{1}{2n+1} \left(\underbrace{\sum_{k=0}^{n} s_{2k}}_{\text{even terms}} + \underbrace{\sum_{k=0}^{n-1} s_{2k+1}}_{\text{odd terms}} \right) \\
&= \lim_{n\to\infty} \frac{1}{2n+1} \left(\sum_{k=0}^{n} 1 + \sum_{k=0}^{n-1} 0 \right) \\
&= \lim_{n\to\infty} \frac{n+1}{2n+1} \\
&= \frac{1}{2}
\end{aligned}$$

[18] Whittaker and Watson (1920), page 155 ⟨8.42⟩
[19] Thomson et al. (2008) page 130 ⟨Example 3.56⟩, Whittaker and Watson (1920), page 155 ⟨8.42⟩

2. Proof for Euler summability:

$$\begin{aligned}
\lim_{x\to 1-0}\sum_{n=0}^{\infty}(-1)^n &= \lim_{x\to 1-0}\lim_{n\to\infty}\left(\frac{1}{1+x} = \sum_{k=0}^{2n-1}(-1)^k x^k + \frac{x^{2n}}{1+x}\right) \\
&= \lim_{x\to 1-0}\frac{1}{1+x} && \text{by item (1a) (page 197) of Example 12.1} \\
&= \frac{1}{2}
\end{aligned}$$

✎

12.4 Convergence in Banach spaces

The properties of *strong convergence* and *weak convergence* are defined on *sequences* (Definition 8.6 page 135). An infinite sum $\sum_{n=1}^{\infty} \boldsymbol{x}_n$ in a Banach space is the limit of a sequence of partial sums $\left(\!\left(\sum_{n=1}^{N} \boldsymbol{x}_n\right)\!\right)$, so the properties of strong and weak convergence apply to infinite sums as well. Definition 12.5 (next) assigns special equality symbols for these sums.

Definition 12.5. *Let* $\boldsymbol{B} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *be a Banach space.*

DEF

The expression $\boldsymbol{x} \stackrel{\star}{=} \sum_{n=1}^{\infty} \boldsymbol{x}_n$ *denotes that the sum* ***converges strongly*** *to* $\boldsymbol{x}$.

The expression $\boldsymbol{x} \doteq \sum_{n=1}^{\infty} \boldsymbol{x}_n$ *denotes that the sum* ***converges weakly*** *to* $\boldsymbol{x}$.

12.5 Convergence tests for real sequences

Theorem 12.7 (*comparison test*). [20]

THM

$$\left\{\begin{array}{lll} 1. & \sum_{n=1}^{\infty}(\!(y_n)\!) \text{ CONVERGES} & \text{and} \\ 2. & x_n \le y_n & \forall n\in\mathbb{N} \end{array}\right\} \implies \sum_{n=1}^{\infty}(\!(x_n)\!) \text{ CONVERGES}$$

[20] Bonar et al. (2006) page 26 ⟨Theorem 1.53 (Limit Comparison Test Strengthened)⟩, Heinbockel (2010) page 152 ⟨Comparison Tests⟩

Part IV

Structures between Spaces

CHAPTER 13

LINEAR FUNCTIONALS

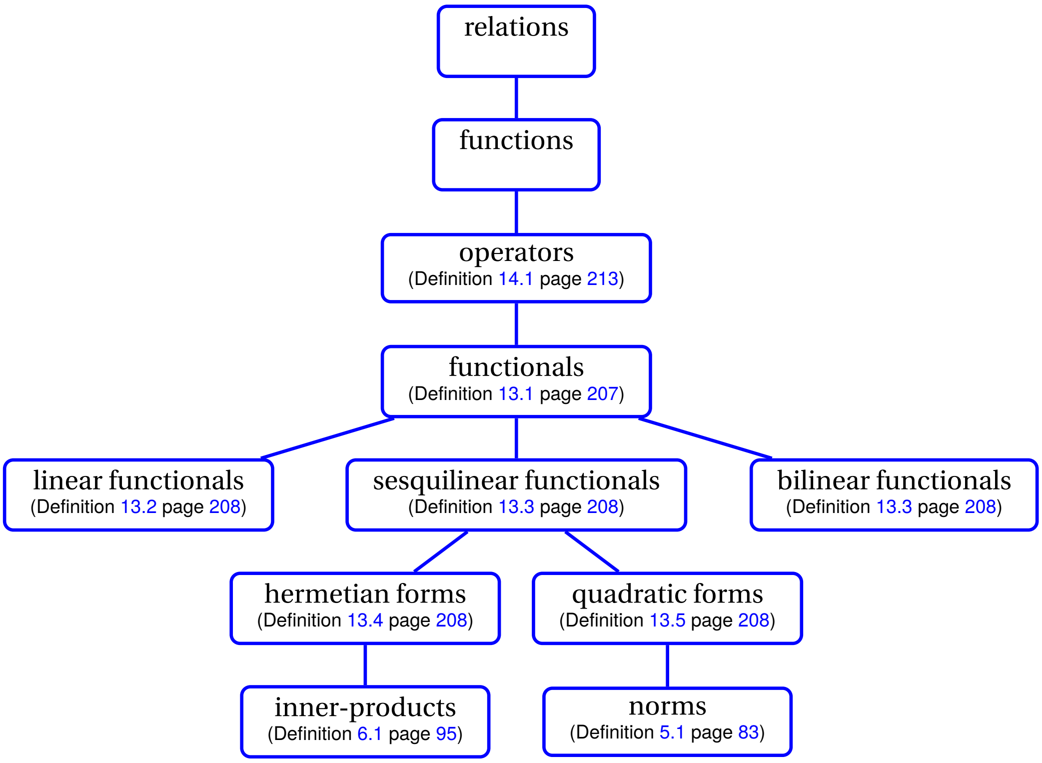


Figure 13.1: Relations

13.1 Definitions

Definition 13.1. [1] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE *(Definition 3.1 page 67). Let* $\mathbb{F}^X$ *be the set of all functions from the set* X *into the* FIELD *(Definition F.5 page 346)* $\mathbb{F}$*.*

DEF *A function* f *is a* ***functional*** *on* $\boldsymbol{\Omega}$ *if* f *is in* $\mathbb{F}^X$*.*

[1] Bachman and Narici (1966) page 5, Michel and Herget (1993) pages 109–114 ⟨Definitions 3.5.1, 3.6.1⟩

Definition 13.2. [2] *Let* $\mathsf{f} \in \mathbb{F}^X$ *be a* FUNCTIONAL *on a* LINEAR SPACE $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$.

DEF

f *is* ***linear*** *if*	$\mathsf{f}(\alpha\boldsymbol{x} + \beta\boldsymbol{y})$	$=$	$\alpha\mathsf{f}(\boldsymbol{x}) + \beta\mathsf{f}(\boldsymbol{y})$	$\forall \boldsymbol{x},\boldsymbol{y} \in X, \forall \alpha,\beta \in \mathbb{F}$.
f *is a* ***conjugate linear*** *if*	$\mathsf{f}(\alpha\boldsymbol{x} + \beta\boldsymbol{y})$	$=$	$\overline{\alpha}\mathsf{f}(\boldsymbol{x}) + \overline{\beta}\mathsf{f}(\boldsymbol{y})$	$\forall \boldsymbol{x},\boldsymbol{y} \in X, \forall \alpha,\beta \in \mathbb{F}$.
f *is* ***subaddtive*** *if*	$\mathsf{f}(\boldsymbol{x} + \boldsymbol{y})$	$\leq$	$\mathsf{f}(\boldsymbol{x}) + \mathsf{f}(\boldsymbol{y})$	$\forall \boldsymbol{x},\boldsymbol{y} \in X$.

Definition 13.3. [3] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE *(Definition 3.1 page 67).*

DEF

A function $\mathsf{b} \in \mathbb{F}^{X \times X}$ *is a* ***bilinear functional*** *or* ***bilinear form*** *on* $\mathbb{F}^{X \times X}$ *if*

1. $\mathsf{b}(\alpha\boldsymbol{x} + \beta\boldsymbol{y}, \boldsymbol{u}) = \alpha\mathsf{b}(\boldsymbol{x}, \boldsymbol{u}) + \beta\mathsf{b}(\boldsymbol{y}, \boldsymbol{u}) \quad \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{u} \in X, \alpha,\beta \in \mathbb{F}$ (LINEAR *in first variable) and*
2. $\mathsf{b}(\boldsymbol{u}, \alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha\mathsf{b}(\boldsymbol{u}, \boldsymbol{x}) + \beta\mathsf{b}(\boldsymbol{u}, \boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{u} \in X, \alpha,\beta \in \mathbb{F}$ (LINEAR *in second variable).*

A function $\mathsf{s} \in \mathbb{F}^{X \times X}$ *is a* ***sesquilinear functional*** *or* ***sesquilinear form*** *on* $\mathbb{F}^{X \times X}$ *if*[4]

1. $\mathsf{b}(\alpha\boldsymbol{x} + \beta\boldsymbol{y}, \boldsymbol{u}) = \alpha\mathsf{b}(\boldsymbol{u}, \boldsymbol{x}) + \beta\mathsf{b}(\boldsymbol{u}, \boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{u} \in X, \alpha,\beta \in \mathbb{F}$ (LINEAR *in first variable) and*
2. $\mathsf{b}(\boldsymbol{u}, \alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \overline{\alpha}\mathsf{b}(\boldsymbol{u}, \boldsymbol{x}) + \overline{\beta}\mathsf{b}(\boldsymbol{u}, \boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{u} \in X, \alpha,\beta \in \mathbb{F}$ (CONJUGATE LINEAR *in second variable).*

Definition 13.4. [5] *Let* $\phi \in \mathbb{F}^{X \times X}$ *be a* BILINEAR FUNCTIONAL *or a* SESQUILINEAR FUNCTIONAL *(Definition 13.3 page 208) on a* LINEAR SPACE $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$.

DEF

1. ϕ *is* ***symmetric*** *if* $\quad \phi(\boldsymbol{x}, \boldsymbol{y}) = \underline{\phi(\boldsymbol{y}, \boldsymbol{x})} \quad \forall \boldsymbol{x},\boldsymbol{y} \in X$.
2. ϕ *is* ***Hermitian symmetric*** *if* $\quad \phi(\boldsymbol{x}, \boldsymbol{y}) = \overline{\phi(\boldsymbol{y}, \boldsymbol{x})} \quad \forall \boldsymbol{x},\boldsymbol{y} \in X$.
3. ϕ *is* ***nonnegative*** *if* $\quad \phi(\boldsymbol{x}, \boldsymbol{y}) \geq 0 \quad \forall \boldsymbol{x},\boldsymbol{y} \in X$.
4. ϕ *is* ***positive*** *if* $\quad \phi(\boldsymbol{x}, \boldsymbol{x}) > 0 \quad \forall \boldsymbol{x} \in X \setminus 0$.
5. *If* ϕ *is* HERMITIAN SYMMETRIC, *then* ϕ *is a* ***hermitian form***.

Definition 13.5. [6] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE *(Definition 3.1 page 67).*

DEF

A function $\mathsf{q} \in \mathbb{F}^X$ *is the* ***quadratic form*** *induced (or generated) by* $\mathsf{s} \in \mathbb{F}^{X \times X}$ *if*

1. s *is a* SESQUILINEAR FUNCTIONAL *and*
2. $\mathsf{q}(\boldsymbol{x}) \triangleq \mathsf{s}(\boldsymbol{x}, \boldsymbol{x}) \quad \forall \boldsymbol{x} \in X$

13.2 Basic results

Lemma 13.1. *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE *(Definition 3.1 page 67).*

LEM

$$\mathsf{q} \in \mathbb{F}^X \text{ is a QUADRATIC FORM} \implies \left\{ \begin{array}{lrcll} 1. & \mathsf{q}(-\boldsymbol{x}) & = & \mathsf{q}(\boldsymbol{x}) & \forall \boldsymbol{x} \in X \quad \text{and} \\ 2. & \mathsf{q}(i\boldsymbol{x}) & = & \mathsf{q}(\boldsymbol{x}) & \forall \boldsymbol{x} \in X. \end{array} \right\}$$

[2] Michel and Herget (1993) pages 109–114 ⟨Definitions 3.5.1, 3.6.1⟩, Nikol'skiĭ (1992) pages 109–110 ⟨subadditive property⟩

[3] Kubrusly (2001) page 312, Brown and Pearcy (1995) page 58

[4] The prefix *sesqui–* is Latin for "one and a half". Reference: merriamwebsterdictionary ⟨`http://www.merriam-webster.com/dictionary/sesqui-`⟩

[5] Kubrusly (2001) page 312, Brown and Pearcy (1995) page 58, Michel and Herget (1993) page 115 ⟨Definition 3.6.10⟩, Debnath and Mikusiński (2005) page 152 ⟨Theorem 4.3.5⟩, Grove (2002) page 85 ⟨item 4⟩

[6] Kubrusly (2001) page 312, Michel and Herget (1993) page 115 ⟨Definition 3.6.11⟩, Debnath and Mikusiński (2005) page 152 ⟨Theorem 4.3.6⟩

✎PROOF: Let $\mathsf{s} \in \mathbb{F}^{X \times X}$ be the *sesquilinear functional* in $\mathbb{F}^{X \times X}$ that induces q.

$$\begin{aligned}
\mathsf{q}(\boldsymbol{x}) &= \mathsf{s}(\boldsymbol{x}, \boldsymbol{x}) \\
&= (-1)(-1)\mathsf{s}(\boldsymbol{x}, \boldsymbol{x}) \\
&= \mathsf{s}((-1)\boldsymbol{x}, \overline{(-1)}\boldsymbol{x}) \\
&= \mathsf{s}(-\boldsymbol{x}, -\boldsymbol{x}) \\
&= \mathsf{q}(-\boldsymbol{x}) \\
\mathsf{q}(\boldsymbol{x}) &= \mathsf{s}(\boldsymbol{x}, \boldsymbol{x}) \\
&= (i)(-i)\mathsf{s}(\boldsymbol{x}, \boldsymbol{x}) \\
&= \mathsf{s}(i\boldsymbol{x}, \overline{(-i)}\boldsymbol{x}) \\
&= \mathsf{s}(i\boldsymbol{x}, i\boldsymbol{x}) \\
&= \mathsf{q}(i\boldsymbol{x})
\end{aligned}$$

✏

Theorem 13.1. [7] *Let* $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *be a* LINEAR SPACE (Definition 3.1 page 67). *Let* $\mathsf{q} \in \mathbb{F}^{\boldsymbol{X}}$ *be the* QUADRATIC FORM (Definition 13.5 page 208) *induced by a function* $\mathsf{s} \in \mathbb{F}^{X \times X}$.

THM

s *is a* SESQUILINEAR FUNCTIONAL $\implies$

$$2\mathsf{s}(\boldsymbol{x}, \boldsymbol{y}) + 2\mathsf{s}(\boldsymbol{y}, \boldsymbol{x}) \quad = \quad \mathsf{q}(\boldsymbol{x} + \boldsymbol{y}) - \mathsf{q}(\boldsymbol{x} - \boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in X$$

✎PROOF:

$$\begin{aligned}
\mathsf{q}(\boldsymbol{x} + \boldsymbol{y}) - \mathsf{q}(\boldsymbol{x} - \boldsymbol{y}) &= \mathsf{s}(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y}) - \mathsf{s}(\boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}) \\
&= \underbrace{\{\mathsf{s}(\boldsymbol{x}, \boldsymbol{x}) + \mathsf{s}(\boldsymbol{x}, \boldsymbol{y}) + \mathsf{s}(\boldsymbol{y}, \boldsymbol{x}) + \mathsf{s}(\boldsymbol{y}, \boldsymbol{y})\}}_{\mathsf{s}(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y})} - \underbrace{\{\mathsf{s}(\boldsymbol{x}, \boldsymbol{x}) - \mathsf{s}(\boldsymbol{x}, \boldsymbol{y}) - \mathsf{s}(\boldsymbol{y}, \boldsymbol{x}) + \mathsf{s}(\boldsymbol{y}, \boldsymbol{y})\}}_{\mathsf{s}(\boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y})} \\
&= \{\mathsf{s}(\boldsymbol{x}, \boldsymbol{y}) + \mathsf{s}(\boldsymbol{y}, \boldsymbol{x})\} - \{-\mathsf{s}(\boldsymbol{x}, \boldsymbol{y}) - \mathsf{s}(\boldsymbol{y}, \boldsymbol{x})\} \\
&= 2\mathsf{s}(\boldsymbol{x}, \boldsymbol{y}) + 2\mathsf{s}(\boldsymbol{y}, \boldsymbol{x})
\end{aligned}$$

✏

Theorem 13.2 (polarization identities). [8] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *be a* LINEAR SPACE (Definition 3.1 page 67). *Let* $\mathsf{q} \in \mathbb{F}^{\boldsymbol{X}}$ *be the* QUADRATIC FORM (Definition 13.5 page 208) *induced by a function* $\mathsf{s} \in \mathbb{F}^{X \times X}$.

THM

1. s *is a* SESQUILINEAR FUNCTIONAL $\implies$
$$4\mathsf{s}(\boldsymbol{x}, \boldsymbol{y}) \quad = \quad \mathsf{q}(\boldsymbol{x} + \boldsymbol{y}) - \mathsf{q}(\boldsymbol{x} - \boldsymbol{y}) + i\mathsf{q}(\boldsymbol{x} + i\boldsymbol{y}) - i\mathsf{q}(\boldsymbol{x} - i\boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in X$$
2. s *is a* HERMITIAN SYMMETRIC *sesquilinear functional and* $\mathbb{F} = \mathbb{R}$ $\implies$
$$4\mathsf{s}(\boldsymbol{x}, \boldsymbol{y}) \quad = \quad \mathsf{q}(\boldsymbol{x} + \boldsymbol{y}) - \mathsf{q}(\boldsymbol{x} - \boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in X$$

✎PROOF:

[7] Michel and Herget (1993) pages 115–116 ⟨Theorem 3.6.12⟩

[8] Brown and Pearcy (1995) pages 62–63 ⟨Problem P(ii)⟩, Halmos (1998b) pages 13–14 ⟨Theorem 1⟩, Michel and Herget (1993) page 116 ⟨Theorem 3.6.13⟩, Debnath and Mikusiński (2005) pages 152–153 ⟨Theorem 4.3.7⟩

1. Proof that $4\mathsf{s}(\boldsymbol{x}, \boldsymbol{y}) = \mathsf{q}(\boldsymbol{x}+\boldsymbol{y}) - \mathsf{q}(\boldsymbol{x}-\boldsymbol{y}) + i\mathsf{q}(\boldsymbol{x}+i\boldsymbol{y}) - i\mathsf{q}(\boldsymbol{x}-i\boldsymbol{y})$:

$$\begin{aligned}
&\mathsf{q}(\boldsymbol{x}+\boldsymbol{y}) - \mathsf{q}(\boldsymbol{x}-\boldsymbol{y}) + i\mathsf{q}(\boldsymbol{x}+i\boldsymbol{y}) - i\mathsf{q}(\boldsymbol{x}-i\boldsymbol{y})\\
&= \mathsf{s}(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}) - \mathsf{s}(\boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y}) + i\mathsf{s}(\boldsymbol{x}+i\boldsymbol{y}, \boldsymbol{x}+i\boldsymbol{y}) - i\mathsf{s}(\boldsymbol{x}-i\boldsymbol{y}, \boldsymbol{x}-i\boldsymbol{y})\\
&= \underbrace{\{\mathsf{s}(\boldsymbol{x},\boldsymbol{x}) + \mathsf{s}(\boldsymbol{x},\boldsymbol{y}) + \mathsf{s}(\boldsymbol{y},\boldsymbol{x}) + \mathsf{s}(\boldsymbol{y},\boldsymbol{y})\}}_{\mathsf{s}(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{x}+\boldsymbol{y})} - \underbrace{\{\mathsf{s}(\boldsymbol{x},\boldsymbol{x}) - \mathsf{s}(\boldsymbol{x},\boldsymbol{y}) - \mathsf{s}(\boldsymbol{y},\boldsymbol{x}) + \mathsf{s}(\boldsymbol{y},\boldsymbol{y})\}}_{\mathsf{s}(\boldsymbol{x}-\boldsymbol{y},\boldsymbol{x}-\boldsymbol{y})}\\
&\quad + i\underbrace{\left\{\mathsf{s}(\boldsymbol{x},\boldsymbol{x}) - i\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) + i\mathsf{s}(\boldsymbol{y},\boldsymbol{x}) - i^2\mathsf{s}(\boldsymbol{y},\boldsymbol{y})\right\}}_{\mathsf{s}(\boldsymbol{x}+i\boldsymbol{y},\boldsymbol{x}+i\boldsymbol{y})} - i\underbrace{\left\{\mathsf{s}(\boldsymbol{x},\boldsymbol{x}) + i\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) - i\mathsf{s}(\boldsymbol{y},\boldsymbol{x}) - i^2\mathsf{s}(\boldsymbol{y},\boldsymbol{y})\right\}}_{\mathsf{s}(\boldsymbol{x}-i\boldsymbol{y},\boldsymbol{x}-i\boldsymbol{y})}\\
&= \{\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) + \mathsf{s}(\boldsymbol{y},\boldsymbol{x})\} - \{-\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) - \mathsf{s}(\boldsymbol{y},\boldsymbol{x})\} + i\{-i\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) + i\mathsf{s}(\boldsymbol{y},\boldsymbol{x})\} - i\{+i\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) - i\mathsf{s}(\boldsymbol{y},\boldsymbol{x})\}\\
&= \{\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) + \mathsf{s}(\boldsymbol{y},\boldsymbol{x})\} + \{\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) + \mathsf{s}(\boldsymbol{y},\boldsymbol{x})\} + \{\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) - \mathsf{s}(\boldsymbol{y},\boldsymbol{x})\} + \{\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) - \mathsf{s}(\boldsymbol{y},\boldsymbol{x})\}\\
&= 4\mathsf{s}(\boldsymbol{x},\boldsymbol{y})
\end{aligned}$$

2. Proof that s is a *Hermitian symmetric* sesquilinear functional and $\mathbb{F} = \mathbb{R} \implies 4\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) = \mathsf{q}(\boldsymbol{x}+\boldsymbol{y}) - \mathsf{q}(\boldsymbol{x}-\boldsymbol{y})$:

$$\begin{aligned}
\mathsf{q}(\boldsymbol{x}+\boldsymbol{y}) - \mathsf{q}(\boldsymbol{x}-\boldsymbol{y}) &= 2\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) + 2\mathsf{s}(\boldsymbol{y},\boldsymbol{x}) && \text{by Theorem 13.1 page 209}\\
&= 2\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) + 2\overline{\mathsf{s}(\boldsymbol{x},\boldsymbol{y})} && \text{by } \textit{symmetric} \text{ hypothesis}\\
&= 2\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) + 2\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) && \text{by } \textit{real} \text{ hypothesis}\\
&= 4\mathsf{s}(\boldsymbol{x},\boldsymbol{y})
\end{aligned}$$

⇨

Theorem 13.3. [9] *Let $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ be a* LINEAR SPACE. *Let* $\mathsf{q}_1 \in \mathbb{F}^{\boldsymbol{X}}$ *be the* QUADRATIC FORM (Definition 13.5 page 208) *induced by a sesquilinear functional* (Definition 13.3 page 208) $\mathsf{s}_1 \in \mathbb{F}^{X\times X}$ *and* $\mathsf{q}_2 \in \mathbb{F}^{\boldsymbol{X}}$ *be the* QUADRATIC FORM *induced by a sesquilinear functional* $\mathsf{s}_2 \in \mathbb{F}^{X\times X}$.

THM
$$\left\{\mathsf{q}_1(\boldsymbol{x}) = \mathsf{q}_2(\boldsymbol{x}) \quad \forall \boldsymbol{x}\in X\right\} \iff \left\{\mathsf{s}_1(\boldsymbol{x},\boldsymbol{y}) = \mathsf{s}_2(\boldsymbol{x},\boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y}\in X\right\}$$

✎PROOF:

1. Proof that $\mathsf{q}_1(\boldsymbol{x}) = \mathsf{q}_2(\boldsymbol{x}) \implies \mathsf{s}_1(\boldsymbol{x},\boldsymbol{y}) = \mathsf{s}_2(\boldsymbol{x},\boldsymbol{y})$:

$$\begin{aligned}
\mathsf{s}_1(\boldsymbol{x},\boldsymbol{y}) &= \mathsf{q}_1(\boldsymbol{x}+\boldsymbol{y}) - \mathsf{q}_1(\boldsymbol{x}-\boldsymbol{y}) + i\mathsf{q}_1(\boldsymbol{x}+i\boldsymbol{y}) - i\mathsf{q}_1(\boldsymbol{x}-i\boldsymbol{y}) && \text{by } \textit{polarization identity} \text{ (Theorem 13.2 page 209)}\\
&= \mathsf{q}_2(\boldsymbol{x}+\boldsymbol{y}) - \mathsf{q}_2(\boldsymbol{x}-\boldsymbol{y}) + i\mathsf{q}_2(\boldsymbol{x}+i\boldsymbol{y}) - i\mathsf{q}_2(\boldsymbol{x}-i\boldsymbol{y}) && \text{by left hypothesis}\\
&= \mathsf{s}_2(\boldsymbol{x},\boldsymbol{y})
\end{aligned}$$

2. Proof that $\mathsf{q}_1(\boldsymbol{x}) = \mathsf{q}_2(\boldsymbol{x}) \implies \mathsf{s}_1(\boldsymbol{x},\boldsymbol{y}) = \mathsf{s}_2(\boldsymbol{x},\boldsymbol{y})$:

$$\begin{aligned}
\mathsf{q}_1(\boldsymbol{x}) &= \mathsf{s}_1(\boldsymbol{x},\boldsymbol{x}) && \text{by Definition 13.5 page 208}\\
&= \mathsf{s}_2(\boldsymbol{x},\boldsymbol{x}) && \text{by right hypothesis}\\
&= \mathsf{q}_2(\boldsymbol{x}) && \text{by Definition 13.5 page 208}
\end{aligned}$$

⇨

Theorem 13.4. [10] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE. *Let* $\mathsf{q} \in \mathbb{F}^{\boldsymbol{X}}$ *be the* QUADRATIC FORM (Definition 13.5 page 208) *induced by a* SESQUILINEAR FUNCTIONAL (Definition 13.3 page 208) $\mathsf{s} \in \mathbb{F}^{X\times X}$.

THM
$$\underbrace{\mathsf{q} \in \mathbb{R}}_{\mathsf{q}\ \textit{is}\ \text{REAL}} \iff \underbrace{\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) = \overline{\mathsf{s}(\boldsymbol{y},\boldsymbol{x})}}_{\mathsf{s}\ \textit{is}\ \text{HERMITIAN SYMMETRIC}}$$

[9] Michel and Herget (1993) page 117 ⟨Theorem 3.6.15⟩, Debnath and Mikusiński (2005) page 153 ⟨Corollary 4.3.8⟩

[10] Kubrusly (2001) page 312, Michel and Herget (1993) pages 115–116 ⟨Theorem 3.6.12⟩, Debnath and Mikusiński (2005) page 153 ⟨Theorem 4.3.9⟩

✎PROOF:

1. Proof that q is *real* $\implies$ s is *hermitian*:

$$\begin{aligned}
\overline{4\mathsf{s}(\boldsymbol{y},\boldsymbol{x})} &= \overline{\mathsf{q}(\boldsymbol{y}+\boldsymbol{x})} - \overline{\mathsf{q}(\boldsymbol{y}-\boldsymbol{x})} + \overline{i\mathsf{q}(\boldsymbol{y}+i\boldsymbol{x})} - \overline{i\mathsf{q}(\boldsymbol{y}-i\boldsymbol{x})} && \text{by \textit{polarization id.} (Theorem 13.2 page 209)}\\
&= \mathsf{q}(\boldsymbol{y}+\boldsymbol{x}) - \mathsf{q}(\boldsymbol{y}-\boldsymbol{x}) - i\mathsf{q}(\boldsymbol{y}+i\boldsymbol{x}) + i\mathsf{q}(\boldsymbol{y}-i\boldsymbol{x}) && \text{by left hypothesis}\\
&= \mathsf{q}(\boldsymbol{y}+\boldsymbol{x}) - \mathsf{q}(-\boldsymbol{y}+\boldsymbol{x}) - i\mathsf{q}(-i\boldsymbol{y}-i^2\boldsymbol{x}) + i\mathsf{q}(i\boldsymbol{y}-i^2\boldsymbol{x}) && \text{by Lemma 13.1 page 208}\\
&= \mathsf{q}(\boldsymbol{x}+\boldsymbol{y}) - \mathsf{q}(\boldsymbol{x}-\boldsymbol{y}) - i\mathsf{q}(\boldsymbol{x}-i\boldsymbol{y}) + i\mathsf{q}(\boldsymbol{x}+i\boldsymbol{y})\\
&= 4\mathsf{s}(\boldsymbol{x},\boldsymbol{y}) && \text{by \textit{polarization id.} (Theorem 13.2 page 209)}
\end{aligned}$$

2. Proof that q is *real* $\impliedby$ s is *Hermitian symmetric*:

$$\begin{aligned}
\overline{\mathsf{q}(\boldsymbol{x})} &= \overline{\mathsf{s}(\boldsymbol{x},\boldsymbol{x})} && \text{by definition of } \mathsf{q} \text{ (Definition 13.5 page 208)}\\
&= \overline{\overline{\mathsf{s}(\boldsymbol{x},\boldsymbol{x})}} && \text{by right hypothesis}\\
&= \mathsf{s}(\boldsymbol{x},\boldsymbol{x})\\
&= \mathsf{q}(\boldsymbol{x}) && \text{by definition of } \mathsf{q} \text{ (Definition 13.5 page 208)}\\
&\implies \quad \mathsf{q}\in\mathbb{R}
\end{aligned}$$

✏

Theorem 13.5 (Riesz Representation Theorem). [11] *Let* $\boldsymbol{H} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla \rangle\right)$ *be a* HILBERT SPACE (Definition 8.11 page 138) *and* f *a* BOUNDED LINEAR FUNCTIONAL *in* $\mathcal{B}(\boldsymbol{H},\mathbb{F})$.

THM: *For every* $\mathsf{f} \in \mathcal{B}(\boldsymbol{X},\mathbb{F})$, *there exists a* UNIQUE $\boldsymbol{y}$ *such that*

$$\mathsf{f}(\boldsymbol{x}) = \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle \quad \forall x\in X;$$

Moreover $$|||\mathsf{f}||| = \|\boldsymbol{y}\|$$

✎PROOF:

1. Proof that $\boldsymbol{y}$ is unique:

$$\begin{aligned}
&& \mathsf{f}(\boldsymbol{x}) &= \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = \langle \boldsymbol{x} \mid \boldsymbol{z} \rangle\\
&\implies & 0 &= \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle - \langle \boldsymbol{x} \mid \boldsymbol{z} \rangle\\
&& &= \langle \boldsymbol{x} \mid \boldsymbol{y}-\boldsymbol{z} \rangle && \text{by Definition 6.1 page 95}\\
&\implies & \boldsymbol{y} &= \boldsymbol{z} && \text{by Theorem 6.1}
\end{aligned}$$

2. Proof that $\exists \boldsymbol{y}$ such that $\mathsf{f}(\boldsymbol{x}) = \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$:

[11] Yosida (1980) page 90, Schechter (2002) pages 29–30 ⟨Theorem 2.1⟩, Kubrusly (2001) page 374 ⟨Proposition 5.62⟩

3. Proof that $\|\mathsf{f}\| = \|\boldsymbol{y}\|$:

$$\begin{aligned}
\|\mathsf{f}\| &\triangleq \sup_{\|\boldsymbol{x}\|\le 1} |\mathsf{f}(\boldsymbol{x})| \\
&\ge \left|\mathsf{f}\left(\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|}\right)\right| \\
&= \left\langle \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \,\middle|\, \boldsymbol{y} \right\rangle \\
&= \frac{1}{\|\boldsymbol{y}\|} \langle \boldsymbol{y} \,|\, \boldsymbol{y} \rangle \\
&= \|\boldsymbol{y}\| \\
\|\mathsf{f}\| &\triangleq \sup_{\|\boldsymbol{x}\|\le 1} |\mathsf{f}(\boldsymbol{x})| \\
&= \sup_{\|\boldsymbol{x}\|\le 1} |\langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle| \\
&\le \sup_{\|\boldsymbol{x}\|\le 1} \|\boldsymbol{x}\| \; \|\boldsymbol{y}\| \\
&= \|\boldsymbol{y}\|
\end{aligned}$$

✏

CHAPTER 14

OPERATORS ON LINEAR SPACES

❝*And I am not afraid to say that there is a way to advance algebra as far beyond what Vieta and Descartes have left us as Vieta and Descartes carried it beyond the ancients.... we need still another analysis which is distinctly geometrical or linear, and which will express situation directly as algebra expresses magnitude directly.*❞

Gottfried Leibniz (1646–1716), German mathematician, in a September 8, 1679 letter to Christian Huygens. [1]

14.1 Operators on linear spaces

14.1.1 Operator Algebra

An operator is simply a function that maps from a linear space to another linear space (or to the same linear space).

Definition 14.1. [2]

DEF

A function $\mathbf{A}$ *in* $\boldsymbol{Y}^{\boldsymbol{X}}$ *is an* ***operator*** *in* $\boldsymbol{Y}^{\boldsymbol{X}}$ *if* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *are both* LINEAR SPACES *(Definition 3.1 page 67)*.

Two operators $\mathbf{A}$ and $\mathbf{B}$ in $\boldsymbol{Y}^{\boldsymbol{X}}$ are **equal** if $\mathbf{A}x = \mathbf{B}x$ for all $x \in \boldsymbol{X}$. The inverse relation of an operator $\mathbf{A}$ in $\boldsymbol{Y}^{\boldsymbol{X}}$ always exists as a *relation* in $2^{\boldsymbol{X}\boldsymbol{Y}}$, but may not always be a *function* (may not always be an operator) in $\boldsymbol{Y}^{\boldsymbol{X}}$.

The operator $\mathbf{I} \in \boldsymbol{X}^{\boldsymbol{X}}$ is the *identity* operator if $\mathbf{I}x = \mathbf{I}$ for all $x \in \boldsymbol{X}$.

Definition 14.2. [3] *Let* $\boldsymbol{X}^{\boldsymbol{X}}$ *be the set of all operators with from a* LINEAR SPACE $\boldsymbol{X}$ *to* $\boldsymbol{X}$. *Let* $\mathbf{I}$ *be an operator in* $\boldsymbol{X}^{\boldsymbol{X}}$. *Let* ①($\boldsymbol{X}$) *be the* IDENTITY ELEMENT *in* $\boldsymbol{X}^{\boldsymbol{X}}$.

[1] quote: Leibniz (1679) pages 248–249
image: http://en.wikipedia.org/wiki/File:Gottfried_Wilhelm_von_Leibniz.jpg, public domain
[2] Heil (2011) page 42
[3] Michel and Herget (1993) page 411

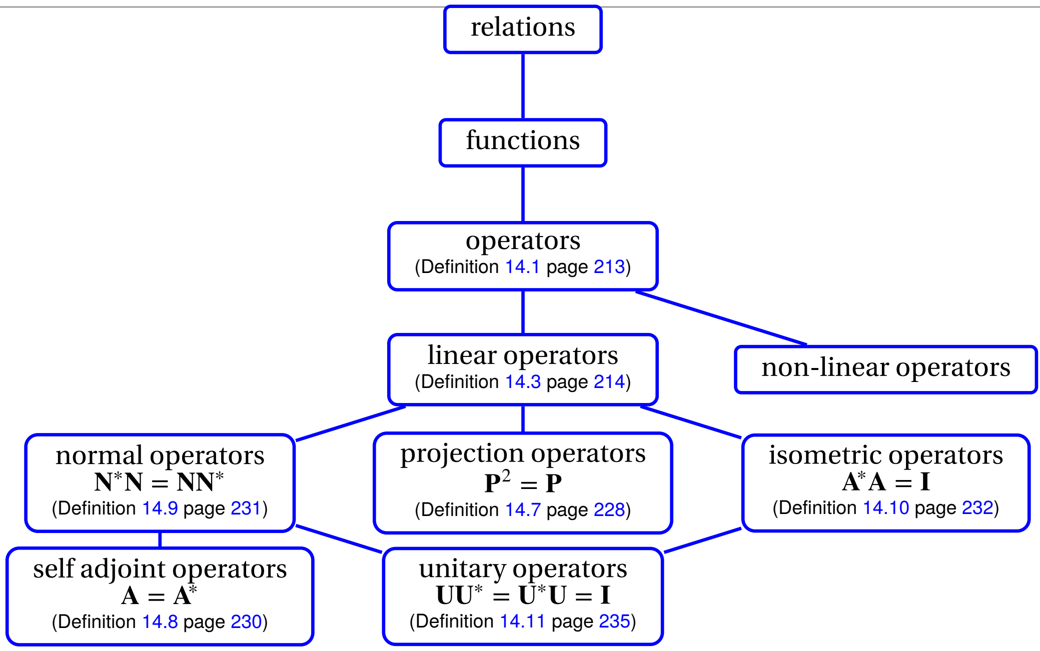


Figure 14.1: Some operator types

DEF: $\mathbf{I}$ *is the* ***identity operator*** *in* $\boldsymbol{X}^{\boldsymbol{X}}$ *if* $\quad \mathbf{I} = \mathbb{1}(\boldsymbol{X})$.

14.1.2 Linear operators

Definition 14.3. [4] *Let* $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *and* $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be linear spaces.*

DEF: *An operator* $\mathbf{L} \in \boldsymbol{Y}^{\boldsymbol{X}}$ *is* ***linear*** *if*

1. $\mathbf{L}(\boldsymbol{x}+\boldsymbol{y}) = \mathbf{L}\boldsymbol{x} + \mathbf{L}\boldsymbol{y} \quad \forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{X}$ (ADDITIVE) *and*
2. $\mathbf{L}(\alpha\boldsymbol{x}) = \alpha\mathbf{L}\boldsymbol{x} \quad \forall \boldsymbol{x} \in \boldsymbol{X}, \ \forall \alpha \in \mathbb{F}$ (HOMOGENEOUS).

The set of all linear operators from $\boldsymbol{X}$ *to* $\boldsymbol{Y}$ *is denoted* $\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$ *such that*

$$\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}) \triangleq \left\{\mathbf{L} \in \boldsymbol{Y}^{\boldsymbol{X}} | \mathbf{L} \text{ is linear}\right\}.$$

Theorem 14.1. [5] *Let* $\mathbf{L}$ *be an operator from a linear space* $\boldsymbol{X}$ *to a linear space* $\boldsymbol{Y}$, *both over a field* $\mathbb{F}$.

THM:

$$\mathbf{L} \text{ is LINEAR} \implies \left\{\begin{array}{llll}
1. & \mathbf{L}\mathbb{0} = \mathbb{0} & & \text{and} \\
2. & \mathbf{L}(-\boldsymbol{x}) = -(\mathbf{L}\boldsymbol{x}) & \forall \boldsymbol{x} \in \boldsymbol{X} & \text{and} \\
3. & \mathbf{L}(\boldsymbol{x}-\boldsymbol{y}) = \mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y} & \forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{X} & \text{and} \\
4. & \mathbf{L}\left(\sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n\right) = \sum_{n=1}^{N} \alpha_n (\mathbf{L}\boldsymbol{x}_n) & \boldsymbol{x}_n \in \boldsymbol{X}, \alpha_n \in \mathbb{F} &
\end{array}\right.$$

✎PROOF:

[4] Kubrusly (2001) page 55, Aliprantis and Burkinshaw (1998) page 224, Hilbert et al. (1927) page 6, Stone (1932) page 33

[5] Berberian (1961) page 79 ⟨Theorem IV.1.1⟩

1. Proof that $\mathbf{L}\mathbb{0} = \mathbb{0}$:

$$\begin{aligned}
\mathbf{L}\mathbb{0} &= \mathbf{L}(0\cdot\mathbb{0}) && \text{by additive identity property (Theorem 3.1 page 69)}\\
&= 0\cdot(\mathbf{L}\mathbb{0}) && \text{by \textit{homogeneous} property of } \mathbf{L} \text{ (Definition 14.3 page 214)}\\
&= \mathbb{0} && \text{by additive identity property (Theorem 3.1 page 69)}
\end{aligned}$$

2. Proof that $\mathbf{L}(-\boldsymbol{x}) = -(\mathbf{L}\boldsymbol{x})$:

$$\begin{aligned}
\mathbf{L}(-\boldsymbol{x}) &= \mathbf{L}(-1\cdot\boldsymbol{x}) && \text{by additive inverse property (Theorem 3.2 page 70)}\\
&= -1\cdot(\mathbf{L}\boldsymbol{x}) && \text{by \textit{homogeneous} property of } \mathbf{L} \text{ (Definition 14.3 page 214)}\\
&= -(\mathbf{L}\boldsymbol{x}) && \text{by additive inverse property (Theorem 3.2 page 70)}
\end{aligned}$$

3. Proof that $\mathbf{L}(\boldsymbol{x}-\boldsymbol{y}) = \mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y}$:

$$\begin{aligned}
\mathbf{L}(\boldsymbol{x}-\boldsymbol{y}) &= \mathbf{L}(\boldsymbol{x}+(-\boldsymbol{y})) && \text{by additive inverse property (Theorem 3.2 page 70)}\\
&= \mathbf{L}(\boldsymbol{x})+\mathbf{L}(-\boldsymbol{y}) && \text{by \textit{linearity} property of } \mathbf{L} \text{ (Definition 14.3 page 214)}\\
&= \mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y} && \text{by 2.}
\end{aligned}$$

4. Proof that $\mathbf{L}\left(\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right) = \sum_{n=1}^{N}\alpha_n\left(\mathbf{L}\boldsymbol{x}_n\right)$:

   (a) Proof for $N=1$:

$$\begin{aligned}
\mathbf{L}\left(\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right) &= \mathbf{L}\left(\alpha_1\boldsymbol{x}_1\right) && \text{by } N=1 \text{ hypothesis}\\
&= \alpha_1\left(\mathbf{L}\boldsymbol{x}_1\right) && \text{by \textit{homogeneous} property of Definition 14.3 page 214}
\end{aligned}$$

   (b) Proof that N case $\implies$ $N+1$ case:

$$\begin{aligned}
\mathbf{L}\left(\sum_{n=1}^{N+1}\alpha_n\boldsymbol{x}_n\right) &= \mathbf{L}\left(\alpha_{N+1}\boldsymbol{x}_{N+1} + \sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right)\\
&= \mathbf{L}\left(\alpha_{N+1}\boldsymbol{x}_{N+1}\right) + \mathbf{L}\left(\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right) && \text{by \textit{linearity} property of Definition 14.3 page 214}\\
&= \alpha_{N+1}\mathbf{L}\left(\boldsymbol{x}_{N+1}\right) + \sum_{n=1}^{N}\mathbf{L}\left(\alpha_n\boldsymbol{x}_n\right) && \text{by left } N+1 \text{ hypothesis}\\
&= \sum_{n=1}^{N+1}\mathbf{L}\left(\alpha_n\boldsymbol{x}_n\right)
\end{aligned}$$

✎

Theorem 14.2. [6] *Let $\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$ be the set of all linear operators from a linear space $\boldsymbol{X}$ to a linear space $\boldsymbol{Y}$. Let $\mathcal{N}(\mathbf{L})$ be the* NULL SPACE *of an operator $\mathbf{L}$ in $\boldsymbol{Y}^{\boldsymbol{X}}$ and $\mathcal{I}(\mathbf{L})$ the* IMAGE SET *of $\mathbf{L}$ in $\boldsymbol{Y}^{\boldsymbol{X}}$.*

THM			
$\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	*is a linear space*		*(space of linear transforms)*
$\mathcal{N}(\mathbf{L})$	*is a linear subspace of $\boldsymbol{X}$*	$\forall\mathbf{L}\in\boldsymbol{Y}^{\boldsymbol{X}}$	
$\mathcal{I}(\mathbf{L})$	*is a linear subspace of $\boldsymbol{Y}$*	$\forall\mathbf{L}\in\boldsymbol{Y}^{\boldsymbol{X}}$	

✎PROOF:

[6] Michel and Herget (1993) pages 98–104, Berberian (1961) pages 80–85 ⟨Theorem IV.1.4 and Theorem IV.3.1⟩

1. Proof that $\mathcal{N}(\mathbf{L})$ is a linear subspace of $\boldsymbol{X}$:
   (a) $\mathbb{0} \in \mathcal{N}(\mathbf{L}) \implies \mathcal{N}(\mathbf{L}) \neq \varnothing$
   (b) $\mathcal{N}(\mathbf{L}) \triangleq \{\boldsymbol{x} \in \boldsymbol{X} | \mathbf{L}\boldsymbol{x} = \mathbb{0}\} \subseteq \boldsymbol{X}$
   (c) $\boldsymbol{x}+\boldsymbol{y} \in \mathcal{N}(\mathbf{L}) \implies \mathbb{0} = \mathbf{L}(\boldsymbol{x}+\boldsymbol{y}) = \mathbf{L}(\boldsymbol{y}+\boldsymbol{x}) \implies \boldsymbol{y}+\boldsymbol{x} \in \mathcal{N}(\mathbf{L})$
   (d) $\alpha \in \mathbb{F},\ \boldsymbol{x} \in \boldsymbol{X} \implies \mathbb{0} = \mathbf{L}\boldsymbol{x} \implies \mathbb{0} = \alpha\mathbf{L}\boldsymbol{x} \implies \mathbb{0} = \mathbf{L}(\alpha\boldsymbol{x}) \implies \alpha\boldsymbol{x} \in \mathcal{N}(\mathbf{L})$
2. Proof that $\mathcal{I}(\mathbf{L})$ is a linear subspace of $\boldsymbol{Y}$:
   (a) $\mathbb{0} \in \mathcal{I}(\mathbf{L}) \implies \mathcal{I}(\mathbf{L}) \neq \varnothing$
   (b) $\mathcal{I}(\mathbf{L}) \triangleq \{\boldsymbol{y} \in \boldsymbol{Y} | \exists \boldsymbol{x} \in \boldsymbol{X} \text{ such that } \boldsymbol{y} = \mathbf{L}\boldsymbol{x}\} \subseteq \boldsymbol{Y}$
   (c) $\boldsymbol{x}+\boldsymbol{y} \in \mathcal{I}(\mathbf{L}) \implies \exists \boldsymbol{v} \in \boldsymbol{X} \text{ such that } \mathbf{L}\boldsymbol{v} = \boldsymbol{x}+\boldsymbol{y} = \boldsymbol{y}+\boldsymbol{x} \implies \boldsymbol{y}+\boldsymbol{x} \in \mathcal{I}(\mathbf{L})$
   (d) $\alpha \in \mathbb{F},\ \boldsymbol{x} \in \mathcal{I}(\mathbf{L}) \implies \exists \boldsymbol{x} \in \boldsymbol{X} \text{ such that } \boldsymbol{y} = \mathbf{L}\boldsymbol{x} \implies \alpha\boldsymbol{y} = \alpha\mathbf{L}\boldsymbol{x} = \mathbf{L}(\alpha\boldsymbol{x}) \implies \alpha\boldsymbol{x} \in \mathcal{I}(\mathbf{L})$

✎

Example 14.1. [7] Let $C([a:b],\mathbb{R})$ be the set of all *continuous* functions from the closed real interval $[a:b]$ to $\mathbb{R}$.

EX $C([a:b],\mathbb{R})$ is a linear space.

Theorem 14.3. [8] *Let $\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$ be the set of linear operators from a linear space $\boldsymbol{X}$ to a linear space $\boldsymbol{Y}$. Let $\mathcal{N}(\mathbf{L})$ be the* NULL SPACE *of a linear operator* $\mathbf{L} \in \mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$.

THM

$\mathbf{L}\boldsymbol{x} = \mathbf{L}\boldsymbol{y}$	$\iff$	$\boldsymbol{x}-\boldsymbol{y} \in \mathcal{N}(\mathbf{L})$
$\mathbf{L}$ *is* INJECTIVE	$\iff$	$\mathcal{N}(\mathbf{L}) = \{\mathbb{0}\}$

✎PROOF:

1. Proof that $\mathbf{L}\boldsymbol{x} = \mathbf{L}\boldsymbol{y} \implies \boldsymbol{x}-\boldsymbol{y} \in \mathcal{N}(\mathbf{L})$:

$$\begin{aligned} \mathbf{L}(\boldsymbol{x}-\boldsymbol{y}) &= \mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y} && \text{by Theorem 14.1 page 214} \\ &= \mathbb{0} && \text{by left hypothesis} \\ &\implies \boldsymbol{x}-\boldsymbol{y} \in \mathcal{N}(\mathbf{L}) && \text{by definition of } \textit{null space} \end{aligned}$$

2. Proof that $\mathbf{L}\boldsymbol{x} = \mathbf{L}\boldsymbol{y} \impliedby \boldsymbol{x}-\boldsymbol{y} \in \mathcal{N}(\mathbf{L})$:

$$\begin{aligned} \mathbf{L}\boldsymbol{y} &= \mathbf{L}\boldsymbol{y} + \mathbb{0} && \text{by definition of linear space (Definition 3.1 page 67)} \\ &= \mathbf{L}\boldsymbol{y} + \mathbf{L}(\boldsymbol{x}-\boldsymbol{y}) && \text{by right hypothesis} \\ &= \mathbf{L}\boldsymbol{y} + (\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y}) && \text{by Theorem 14.1 page 214} \\ &= (\mathbf{L}\boldsymbol{y} - \mathbf{L}\boldsymbol{y}) + \mathbf{L}\boldsymbol{x} && \text{by \textit{associative} and \textit{commutative} properties (Definition 3.1 page 67)} \\ &= \mathbf{L}\boldsymbol{x} \end{aligned}$$

3. Proof that $\mathbf{L}$ is *injective* $\iff \mathcal{N}(\mathbf{L}) = \{\mathbb{0}\}$:

$$\begin{aligned} \mathbf{L} \text{ is \textit{injective}} &\iff \{(\mathbf{L}\boldsymbol{x} = \mathbf{L}\boldsymbol{y} \iff \boldsymbol{x} = \boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y} \in X\} \\ &\iff \left\{\left[\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y} = \mathbb{0} \iff (\boldsymbol{x}-\boldsymbol{y}) = \mathbb{0}\right] \quad \forall \boldsymbol{x},\boldsymbol{y} \in X\right\} \\ &\iff \left\{\left[\mathbf{L}(\boldsymbol{x}-\boldsymbol{y}) = \mathbb{0} \iff (\boldsymbol{x}-\boldsymbol{y}) = \mathbb{0}\right] \quad \forall \boldsymbol{x},\boldsymbol{y} \in X\right\} \\ &\iff \mathcal{N}(\mathbf{L}) = \{\mathbb{0}\} \end{aligned}$$

[7] Eidelman et al. (2004) page 3
[8] Berberian (1961) page 88 ⟨Theorem IV.1.4⟩

Theorem 14.4. [9] *Let $\boldsymbol{W}$, $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ be linear spaces over a field $\mathbb{F}$.*

1.	$\mathbf{L}(\mathbf{MN})$	$=$	$(\mathbf{LM})\mathbf{N}$	$\forall \mathbf{L}\in\mathcal{L}(\boldsymbol{Z},\boldsymbol{W}), \mathbf{M}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{N}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(ASSOCIATIVE)
2.	$\mathbf{L}(\mathbf{M} \dotplus \mathbf{N})$	$=$	$(\mathbf{LM}) \dotplus (\mathbf{LN})$	$\forall \mathbf{L}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{M}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), \mathbf{N}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(LEFT DISTRIBUTIVE)
3.	$(\mathbf{L} \dotplus \mathbf{M})\mathbf{N}$	$=$	$(\mathbf{LN}) \dotplus (\mathbf{MN})$	$\forall \mathbf{L}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{M}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{N}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(RIGHT DISTRIBUTIVE)
4.	$\alpha(\mathbf{LM})$	$=$	$(\alpha\mathbf{L})\mathbf{M} = \mathbf{L}(\alpha\mathbf{M})$	$\forall \mathbf{L}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{M}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), \alpha\in\mathbb{F}$	(HOMOGENEOUS)

✎PROOF:

1. Proof that $\mathbf{L}(\mathbf{MN}) = (\mathbf{LM})\mathbf{N}$: Follows directly from property of *associative* operators.

2. Proof that $\mathbf{L}(\mathbf{M} \dotplus \mathbf{N}) = (\mathbf{LM}) \dotplus (\mathbf{LN})$:

$$\begin{aligned}
[\mathbf{L}(\mathbf{M} \dotplus \mathbf{N})]\boldsymbol{x} &= \mathbf{L}[(\mathbf{M} \dotplus \mathbf{N})\boldsymbol{x}] \\
&= \mathbf{L}[(\mathbf{M}\boldsymbol{x}) \dotplus (\mathbf{N}\boldsymbol{x})] \\
&= [\mathbf{L}(\mathbf{M}\boldsymbol{x})] \dotplus [\mathbf{L}(\mathbf{N}\boldsymbol{x})] && \text{by } \textit{additive} \text{ property Definition 14.3 page 214} \\
&= [(\mathbf{LM})\boldsymbol{x}] \dotplus [(\mathbf{LN})\boldsymbol{x}]
\end{aligned}$$

3. Proof that $(\mathbf{L} \dotplus \mathbf{M})\mathbf{N} = (\mathbf{LN}) \dotplus (\mathbf{MN})$: Follows directly from property of *associative* operators.

4. Proof that $\alpha(\mathbf{LM}) = (\alpha\mathbf{L})\mathbf{M}$: Follows directly from *associative* property of linear operators.

5. Proof that $\alpha(\mathbf{LM}) = \mathbf{L}(\alpha\mathbf{M})$:

$$\begin{aligned}
[\alpha(\mathbf{LM})]\boldsymbol{x} &= \alpha[(\mathbf{LM})\boldsymbol{x}] \\
&= \mathbf{L}[\alpha(\mathbf{M}\boldsymbol{x})] && \text{by } \textit{homogeneous} \text{ property Definition 14.3 page 214} \\
&= \mathbf{L}[(\alpha\mathbf{M})\boldsymbol{x}] \\
&= [\mathbf{L}(\alpha\mathbf{M})]\boldsymbol{x}
\end{aligned}$$

Theorem 14.5 (Fundamental theorem of linear equations). *Michel and Herget (1993) page 99 Let $\boldsymbol{Y}^{\boldsymbol{X}}$ be the set of all operators from a linear space $\boldsymbol{X}$ to a linear space $\boldsymbol{Y}$. Let $\mathcal{N}(\mathbf{L})$ be the* NULL SPACE *of an operator $\mathbf{L}$ in $\boldsymbol{Y}^{\boldsymbol{X}}$ and $\mathcal{I}(\mathbf{L})$ the* IMAGE SET *of $\mathbf{L}$ in $\boldsymbol{Y}^{\boldsymbol{X}}$ (Definition ?? page ??).*

$$\dim \mathcal{I}(\mathbf{L}) + \dim \mathcal{N}(\mathbf{L}) = \dim \boldsymbol{X} \qquad \forall \mathbf{L}\in\boldsymbol{Y}^{\boldsymbol{X}}$$

✎PROOF: Let $\{\psi_k | k = 1, 2, \ldots, p\}$ be a basis for $\boldsymbol{X}$ constructed such that $\{\psi_{p-n+1}, \psi_{p-n+2}, \ldots, \psi_p\}$ is a basis for

[9] Berberian (1961) page 88 ⟨Theorem IV.5.1⟩

$\mathcal{N}(\mathbf{L})$.

Let $p \triangleq \dim \boldsymbol{X}$.
Let $n \triangleq \dim \mathcal{N}(\mathbf{L})$.

$$\begin{aligned}
\dim \mathcal{I}(\mathbf{L}) &= \dim \{ \boldsymbol{y} \in \boldsymbol{Y} | \exists \boldsymbol{x} \in \boldsymbol{X} \quad \text{such that} \quad \boldsymbol{y} = \mathbf{L}\boldsymbol{x} \} \\
&= \dim \left\{ \boldsymbol{y} \in \boldsymbol{Y} | \exists (\alpha_1, \alpha_2, \ldots, \alpha_p) \quad \text{such that} \quad \boldsymbol{y} = \mathbf{L} \sum_{k=1}^{p} \alpha_k \psi_k \right\} \\
&= \dim \left\{ \boldsymbol{y} \in \boldsymbol{Y} | \exists (\alpha_1, \alpha_2, \ldots, \alpha_p) \quad \text{such that} \quad \boldsymbol{y} = \sum_{k=1}^{p} \alpha_k \mathbf{L} \psi_k \right\} \\
&= \dim \left\{ \boldsymbol{y} \in \boldsymbol{Y} | \exists (\alpha_1, \alpha_2, \ldots, \alpha_p) \quad \text{such that} \quad \boldsymbol{y} = \sum_{k=1}^{p-n} \alpha_k \mathbf{L} \psi_k + \sum_{k=1}^{n} \alpha_k \mathbf{L} \psi_k \right\} \\
&= \dim \left\{ \boldsymbol{y} \in \boldsymbol{Y} | \exists (\alpha_1, \alpha_2, \ldots, \alpha_p) \quad \text{such that} \quad \boldsymbol{y} = \sum_{k=1}^{p-n} \alpha_k \mathbf{L} \psi_k + \mathbb{0} \right\} \\
&= p - n \\
&= \dim \boldsymbol{X} - \dim \mathcal{N}(\mathbf{L})
\end{aligned}$$

Note: This "proof" may be missing some necessary detail. ✎

14.2 Operators on Normed linear spaces

14.2.1 Operator norm

Definition 14.4. [10] *Let $\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$ be the space of linear operators over normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$.* [11]

DEF

*The **operator norm** $|||\cdot|||$ is defined as*

$$|||\mathbf{A}||| \triangleq \sup_{x \in X} \{ \|\mathbf{A}\boldsymbol{x}\| \mid \|\boldsymbol{x}\| \le 1 \} \qquad \forall \mathbf{A} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$$

*The pair $(\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}), |||\cdot|||)$ is the **normed space of linear operators** on $(\boldsymbol{X}, \boldsymbol{Y})$.*

Proposition 14.1 (next) shows that the functional defined in Definition 14.4 (previous) is a *norm*.[12]

Proposition 14.1. [13] *Let $(\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}), |||\cdot|||)$ be the normed space of linear operators over the normed linear spaces $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ and $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$.*

PRP

*The functional $|||\cdot|||$ is a **norm** on $\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$. In particular,*

1. $|||\mathbf{A}||| \ge 0 \qquad \forall \mathbf{A} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$ (NON-NEGATIVE) *and*
2. $|||\mathbf{A}||| = 0 \iff \mathbf{A} \doteq \mathbb{0} \qquad \forall \mathbf{A} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$ (NONDEGENERATE) *and*
3. $|||\alpha\mathbf{A}||| = |\alpha|\, |||\mathbf{A}||| \qquad \forall \mathbf{A} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}), \alpha \in \mathbb{F}$ (HOMOGENEOUS) *and*
4. $|||\mathbf{A} \dot{+} \mathbf{B}||| \le |||\mathbf{A}||| + |||\mathbf{B}||| \qquad \forall \mathbf{A} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$ (SUBADDITIVE).

*Moreover, $(\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}), |||\cdot|||)$ is a **normed linear space**.*

---

[10] Rudin (1991) page 92, Aliprantis and Burkinshaw (1998) page 225
[11] The operator norm notation $|||\cdot|||$ is introduced (as a Matrix norm) in Horn and Johnson (1990), page 290
[12] *norm* $\|\cdot\|$: Definition 5.1 (page 83)
[13] Rudin (1991) page 93

✎PROOF:

1. Proof that $|||\mathbf{A}||| > 0$ for $\mathbf{A} \neq \mathbb{0}$:

$$\begin{aligned}
|||\mathbf{A}||| &\triangleq \sup_{x \in \boldsymbol{X}} \{\|\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition 14.4 page 218)}\\
&> 0
\end{aligned}$$

2. Proof that $|||\mathbf{A}||| = 0$ for $\mathbf{A} \doteq \mathbb{0}$:

$$\begin{aligned}
|||\mathbf{A}||| &\triangleq \sup_{x \in \boldsymbol{X}} \{\|\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition 14.4 page 218)}\\
&= \sup_{x \in \boldsymbol{X}} \{\|\mathbb{0}x\| \mid \|x\| \le 1\}\\
&= 0
\end{aligned}$$

3. Proof that $|||\alpha\mathbf{A}||| = |\alpha|\ |||\mathbf{A}|||$:

$$\begin{aligned}
|||\alpha\mathbf{A}||| &\triangleq \sup_{x \in \boldsymbol{X}} \{\|\alpha\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition 14.4 page 218)}\\
&= \sup_{x \in \boldsymbol{X}} \{|\alpha|\ \|\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition 14.4 page 218)}\\
&= |\alpha| \sup_{x \in \boldsymbol{X}} \{\|\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of sup}\\
&= |\alpha|\ \ |||\mathbf{A}||| && \text{by definition of } |||\cdot||| \text{ (Definition 14.4 page 218)}
\end{aligned}$$

4. Proof that $|||\mathbf{A} \mathring{+} \mathbf{B}||| \le |||\mathbf{A}||| + |||\mathbf{B}|||$:

$$\begin{aligned}
|||\mathbf{A} \mathring{+} \mathbf{B}||| &\triangleq \sup_{x \in \boldsymbol{X}} \left\{\left\|(\mathbf{A} \mathring{+} \mathbf{B})x\right\| \mid \|x\| \le 1\right\} && \text{by definition of } |||\cdot||| \text{ (Definition 14.4 page 218)}\\
&= \sup_{x \in \boldsymbol{X}} \{\|\mathbf{A}x + \mathbf{B}x\| \mid \|x\| \le 1\}\\
&\le \sup_{x \in \boldsymbol{X}} \{\|\mathbf{A}x\| + \|\mathbf{B}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition 14.4 page 218)}\\
&\le \sup_{x \in \boldsymbol{X}} \{\|\mathbf{A}x\| \mid \|x\| \le 1\} + \sup_{x \in \boldsymbol{X}} \{\|\mathbf{B}x\| \mid \|x\| \le 1\}\\
&\triangleq |||\mathbf{A}||| + |||\mathbf{B}||| && \text{by definition of } |||\cdot||| \text{ (Definition 14.4 page 218)}
\end{aligned}$$

✏

Lemma 14.1. *Let* $(\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}), |||\cdot|||)$ *be the normed space of linear operators over normed linear spaces* $\boldsymbol{X} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *and* $\boldsymbol{Y} \triangleq \left(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$.

LEM

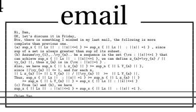
$$|||\mathbf{L}||| = \sup_{x} \{\|\mathbf{L}x\| \mid \|x\| = 1\} \qquad \forall x \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$$

✎PROOF: [14]

1. Proof that $\sup_x \{\|\mathbf{L}x\| \mid \|x\| \le 1\} \ge \sup_x \{\|\mathbf{L}x\| \mid \|x\| = 1\}$:

$$\sup_{x} \{\|\mathbf{L}x\| \mid \|x\| \le 1\} \ge \sup_{x} \{\|\mathbf{L}x\| \mid \|x\| = 1\} \qquad \text{because } A \subseteq B \implies \sup A \le \sup B$$

[14] email



Many many thanks to former NCTU Ph.D. student Chien Yao (Chinese: 姚建; PinYin: Yáo Jiàn) for his brilliant help with this proof. (If you are viewing this text as a pdf file, zoom in on the figure to the left to see text from Chien Yao's 2007 April 16 email.)

2. Let the subset $Y \subsetneq X$ be defined as

$$Y \triangleq \left\{ y \in X \,\middle|\, \begin{array}{ll} 1. & \|\mathbf{L}y\| = \sup_{x \in X} \{\|\mathbf{L}x\| \mid \|x\| \le 1\} \text{ and} \\ 2. & 0 < \|y\| \le 1 \end{array} \right\}$$

3. Proof that $\sup_x \{\|\mathbf{L}x\| \mid \|x\| \le 1\} \le \sup_x \{\|\mathbf{L}x\| \mid \|x\| = 1\}$:

$$\begin{aligned}
\sup_{x \in X} \{\|\mathbf{L}x\| \mid \|x\| \le 1\} &= \|\mathbf{L}y\| && \text{by definition of set } Y \\
&= \frac{\|y\|}{\|y\|} \|\mathbf{L}y\| \\
&= \|y\| \left\| \frac{1}{\|y\|} \mathbf{L}y \right\| && \text{by homogeneous property (page 83)} \\
&= \|y\| \left\| \mathbf{L} \frac{y}{\|y\|} \right\| && \text{by homogeneous property (page 214)} \\
&\le \|y\| \sup_{y \in Y} \left\{ \left\| \mathbf{L} \frac{y}{\|y\|} \right\| \right\} && \text{by definition of supremum} \\
&= \|y\| \sup_{y \in Y} \left\{ \left\| \mathbf{L} \frac{y}{\|y\|} \right\| \,\middle|\, \left\| \frac{y}{\|y\|} \right\| = 1 \right\} && \text{because } \left\| \frac{y}{\|y\|} \right\| = 1 \text{ for all } y \in Y \\
&\le \sup_{y \in Y} \left\{ \left\| \mathbf{L} \frac{y}{\|y\|} \right\| \,\middle|\, \left\| \frac{y}{\|y\|} \right\| = 1 \right\} && \text{because } 0 < \|y\| \le 1 \\
&\le \sup_{x \in X} \{\|\mathbf{L}x\| \mid \|x\| = 1\} && \text{because } \frac{y}{\|y\|} \in X \qquad \forall y \in Y
\end{aligned}$$

4. By (1) and (3),

$$\sup_{x \in X} \{\|\mathbf{L}x\| \mid \|x\| \le 1\} = \sup_{x \in X} \{\|\mathbf{L}x\| \mid \|x\| = 1\}$$

⇨

Proposition 14.2. [15] *Let* $\mathbf{I}$ *be the identity operator in the normed space of linear operators* $(\mathcal{L}(\boldsymbol{X}, \boldsymbol{X}), |||\cdot|||)$.

PRP: $|||\mathbf{I}||| = 1$

✎PROOF:

$$\begin{aligned}
|||\mathbf{I}||| &\triangleq \sup \{\|\mathbf{I}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition 14.4 page 218)} \\
&= \sup \{\|x\| \mid \|x\| \le 1\} && \text{by definition of } \mathbf{I} \text{ (Definition 14.2 page 213)} \\
&= 1
\end{aligned}$$

⇨

Theorem 14.6. [16] *Let* $(\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}), |||\cdot|||)$ *be the normed space of linear operators over normed linear spaces* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$.

THM:

$$\begin{array}{rcll}
\|\mathbf{L}x\| & \le & |||\mathbf{L}|||\ \|x\| & \forall \mathbf{L} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}),\ x \in \boldsymbol{X} \\
|||\mathbf{KL}||| & \le & |||\mathbf{K}|||\ |||\mathbf{L}||| & \forall \mathbf{K}, \mathbf{L} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})
\end{array}$$

[15] Michel and Herget (1993) page 410
[16] Rudin (1991) page 103, Aliprantis and Burkinshaw (1998) page 225

✎PROOF:

1. Proof that $\|\mathbf{L}x\| \le |||\mathbf{L}|||\ \|x\|$:

$$\begin{aligned}
\|\mathbf{L}x\| &= \frac{\|x\|}{\|x\|}\|\mathbf{L}x\| \\
&= \|x\| \left\| \frac{1}{\|x\|}\mathbf{L}x \right\| && \text{by property of norms} \\
&= \|x\| \left\| \mathbf{L}\frac{x}{\|x\|} \right\| && \text{by property of linear operators} \\
&\triangleq \|x\|\ \|\mathbf{L}y\| && \text{where } y \triangleq \frac{x}{\|x\|} \\
&\le \|x\| \sup_{y} \|\mathbf{L}y\| && \text{by definition of supremum} \\
&= \|x\| \sup_{y} \{\|\mathbf{L}y\| \mid \|y\| = 1\} && \text{because } \|y\| = \left\|\frac{x}{\|x\|}\right\| = \frac{\|x\|}{\|x\|} = 1 \\
&\triangleq \|x\|\ \|\mathbf{L}\| && \text{by definition of operator norm}
\end{aligned}$$

2. Proof that $|||\mathbf{KL}||| \le |||\mathbf{K}|||\ |||\mathbf{L}|||$:

$$\begin{aligned}
|||\mathbf{KL}||| &\triangleq \sup_{x\in\boldsymbol{X}} \{\|(\mathbf{KL})x\| \mid \|x\| \le 1\} && \text{by Definition 14.4 page 218 } (|||\cdot|||) \\
&= \sup_{x\in\boldsymbol{X}} \{\|\mathbf{K}(\mathbf{L}x)\| \mid \|x\| \le 1\} \\
&\le \sup_{x\in\boldsymbol{X}} \{|||\mathbf{K}|||\ \|\mathbf{L}x\| \mid \|x\| \le 1\} && \text{by 1.} \\
&\le \sup_{x\in\boldsymbol{X}} \{|||\mathbf{K}|||\ |||\mathbf{L}|||\ \|x\| \mid \|x\| \le 1\} && \text{by 1.} \\
&= \sup_{x\in\boldsymbol{X}} \{|||\mathbf{K}|||\ |||\mathbf{L}|||\ 1 \mid \|x\| \le 1\} && \text{by definition of sup} \\
&= |||\mathbf{K}|||\ |||\mathbf{L}||| && \text{by definition of sup}
\end{aligned}$$

⇨

14.2.2 Bounded linear operators

Definition 14.5. [17] *Let $(\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), |||\cdot|||)$ be a normed space of linear operators.*

DEF

*An operator $\mathbf{B}$ is **bounded** if $|||\mathbf{B}||| < \infty$.*
*The quantity $\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$ is the set of all **bounded linear operators** on $(\boldsymbol{X},\boldsymbol{Y})$ such that*
$$\mathcal{B}(\boldsymbol{X},\boldsymbol{Y}) \triangleq \{\mathbf{L} \in \mathcal{L}(\boldsymbol{X},\boldsymbol{Y}) \mid |||\mathbf{L}||| < \infty\}.$$

Theorem 14.7. [18] *Let $(\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), |||\cdot|||)$ be the set of linear operators over normed linear spaces $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ and $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$.*

THM

The following conditions are all EQUIVALENT:

1. $\mathbf{L}$ *is continuous at* A SINGLE POINT $x_0 \in X$ $\quad \forall \mathbf{L}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}) \quad \Longleftrightarrow$
2. $\mathbf{L}$ *is* CONTINUOUS *(at every point $x \in X$)* $\quad \forall \mathbf{L}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}) \quad \Longleftrightarrow$
3. $|||\mathbf{L}||| < \infty$ ($\mathbf{L}$ *is* BOUNDED) $\quad \forall \mathbf{L}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}) \quad \Longleftrightarrow$
4. $\exists M \in \mathbb{R}$ such that $\|\mathbf{L}x\| \le M\|x\|$ $\quad \forall \mathbf{L}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), x\in X$

[17] Rudin (1991) pages 92–93
[18] Aliprantis and Burkinshaw (1998) page 227

✎PROOF:

1. Proof that $1 \implies 2$:

$$\begin{aligned}
\epsilon &> \left\|\mathbf{L}x - \mathbf{L}x_0\right\| && \text{by hypothesis 1}\\
&= \left\|\mathbf{L}(x - x_0)\right\| && \text{by linearity (Definition 14.3 page 214)}\\
&= \left\|\mathbf{L}(x + y - x_0 - y)\right\| && \\
&= \left\|\mathbf{L}(x + y) - \mathbf{L}(x_0 + y)\right\| && \text{by linearity (Definition 14.3 page 214)}\\
\implies & \mathbf{L} \text{ is continuous at point } x + y && \\
\implies & \mathbf{L} \text{ is continuous at every point in } X && \text{(hypothesis 2)}
\end{aligned}$$

2. Proof that $2 \implies 1$: obvious:

3. Proof that $4 \implies 2$:[19]

$$\begin{aligned}
|||\mathbf{L}x||| \le M\,\|x\| &\implies |||\mathbf{L}(x-y)||| \le M\,\|x-y\| && \text{by hypothesis 4}\\
&\implies |||\mathbf{L}x - \mathbf{L}y||| \le M\,\|x-y\| && \text{by linearity of } \mathbf{L} \text{ (Definition 14.3 page 214)}\\
&\implies |||\mathbf{L}x - \mathbf{L}y||| \le \epsilon \text{ whenever } M\,\|x-y\| < \epsilon && \\
&\implies |||\mathbf{L}x - \mathbf{L}y||| \le \epsilon \text{ whenever } \|x-y\| < \frac{\epsilon}{M} && \text{(hypothesis 2)}
\end{aligned}$$

4. Proof that $3 \implies 4$:

$$\begin{aligned}
\|\mathbf{L}x\| &\le \underbrace{|||\mathbf{L}|||}_{M}\ \|x\| && \text{by Theorem 14.6 page 220}\\
&= M\ \|x\| && \text{where } M \triangleq |||\mathbf{L}||| < \infty \text{ (by hypothesis 1)}
\end{aligned}$$

5. Proof that $1 \implies 3$:[20]

$$\begin{aligned}
|||\mathbf{L}||| = \infty &\implies \{\|\mathbf{L}x\| \mid \|x\| \le 1\} = \infty\\
&\implies \exists\,(x_n) \ \text{ such that } \ \left\|x_n\right\| = 1 \text{ and } |||\mathbf{L}||| = \left\{\left\|\mathbf{L}x_n\right\| \mid \left\|x_n\right\| \le 1\right\} = \infty\\
&\implies \left\|x_n\right\| = 1 \text{ and } \infty = |||\mathbf{L}||| = \left\|\mathbf{L}x_n\right\|\\
&\implies \left\|x_n\right\| = 1 \text{ and } \left\|\mathbf{L}x_n\right\| \ge n\\
&\implies \frac{1}{n}\left\|x_n\right\| = \frac{1}{n} \text{ and } \frac{1}{n}\left\|\mathbf{L}x_n\right\| \ge 1\\
&\implies \left\|\frac{x_n}{n}\right\| = \frac{1}{n} \text{ and } \left\|\mathbf{L}\frac{x_n}{n}\right\| \ge 1\\
&\implies \lim_{n\to\infty}\left\|\frac{x_n}{n}\right\| = 0 \text{ and } \lim_{n\to\infty}\left\|\mathbf{L}\frac{x_n}{n}\right\| \ge 1\\
&\implies \mathbf{L} \text{ is not continuous at } 0
\end{aligned}$$

But by hypothesis, $\mathbf{L}$ *is* continuous. So the statement $|||\mathbf{L}||| = \infty$ must be *false* and thus $|||\mathbf{L}||| < \infty$ ($\mathbf{L}$ is *bounded*).

✏

[19] Bollobás (1999), page 29
[20] Aliprantis and Burkinshaw (1998), page 227

14.2.3 Adjoints on normed linear spaces

Definition 14.6. *Let $\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. Let $\boldsymbol{X}^*$ be the* TOPOLOGICAL DUAL SPACE (Definition 4.4 page 80) *of $\boldsymbol{X}$.*

DEF: $\mathbf{B}^*$ *is the **adjoint** of an operator* $\mathbf{B}\in\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$ *if*

$$\mathsf{f}(\mathbf{B}\boldsymbol{x}) = \left[\mathbf{B}^*\mathsf{f}\right](\boldsymbol{x}) \qquad \forall \mathsf{f}\in\boldsymbol{X}^*,\ \boldsymbol{x}\in\boldsymbol{X}$$

Theorem 14.8. [21] *Let $\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$.*

THM:
$$\begin{array}{rcll}
(\mathbf{A}\mathring{+}\mathbf{B})^* &=& \mathbf{A}^*\mathring{+}\mathbf{B}^* & \forall\mathbf{A},\mathbf{B}\in\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})\\
(\lambda\mathbf{A})^* &=& \lambda\mathbf{A}^* & \forall\mathbf{A},\mathbf{B}\in\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})\\
(\mathbf{A}\mathbf{B})^* &=& \mathbf{B}^*\mathbf{A}^* & \forall\mathbf{A},\mathbf{B}\in\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})
\end{array}$$

✎PROOF:

$$\begin{aligned}
\left[\mathbf{A}\mathring{+}\mathbf{B}\right]^*\mathsf{f}(\boldsymbol{x}) &= \mathsf{f}\left(\left[\mathbf{A}\mathring{+}\mathbf{B}\right]\boldsymbol{x}\right) && \text{by definition of adjoint (Definition 14.6 page 223)}\\
&= \mathsf{f}(\mathbf{A}\boldsymbol{x}+\mathbf{B}\boldsymbol{x}) && \text{by definition of linear operators (Definition 14.3 page 214)}\\
&= \mathsf{f}(\mathbf{A}\boldsymbol{x})+\mathsf{f}(\mathbf{B}\boldsymbol{x}) && \text{by definition of \textit{linear functional} (Definition 13.1 page 207)}\\
&= \mathbf{A}^*\mathsf{f}(\boldsymbol{x})+\mathbf{B}^*\mathsf{f}(\boldsymbol{x}) && \text{by definition of \textit{adjoint} (Definition 14.6 page 223)}\\
&= \left[\mathbf{A}^*+\mathbf{B}^*\right]\mathsf{f}(\boldsymbol{x}) && \text{by definition of \textit{linear functional} (Definition 13.1 page 207)}
\end{aligned}$$

$$\begin{aligned}
[\lambda\mathbf{A}]^*\mathsf{f}(\boldsymbol{x}) &= \mathsf{f}([\lambda\mathbf{A}]\boldsymbol{x}) && \text{by definition of \textit{adjoint} (Definition 14.6 page 223)}\\
&= \lambda\mathsf{f}(\mathbf{A}\boldsymbol{x}) && \text{by definition of \textit{linear functional} (Definition 13.1 page 207)}\\
&= \left[\lambda\mathbf{A}^*\right]\mathsf{f}(\boldsymbol{x}) && \text{by definition of \textit{adjoint} (Definition 14.6 page 223)}
\end{aligned}$$

$$\begin{aligned}
[\mathbf{A}\mathbf{B}]^*\mathsf{f}(\boldsymbol{x}) &= \mathsf{f}([\mathbf{A}\mathbf{B}]\boldsymbol{x}) && \text{by definition of \textit{adjoint} (Definition 14.6 page 223)}\\
&= \mathsf{f}(\mathbf{A}[\mathbf{B}\boldsymbol{x}]) && \text{by definition of \textit{linear operators} (Definition 14.3 page 214)}\\
&= \left[\mathbf{A}^*\mathsf{f}\right](\mathbf{B}\boldsymbol{x}) && \text{by definition of \textit{adjoint} (Definition 14.6 page 223)}\\
&= \mathbf{B}^*\left[\mathbf{A}^*\mathsf{f}\right](\boldsymbol{x}) && \text{by definition of \textit{adjoint} (Definition 14.6 page 223)}\\
&= \left[\mathbf{B}^*\mathbf{A}^*\right]\mathsf{f}(\boldsymbol{x}) && \text{by definition of \textit{adjoint} (Definition 14.6 page 223)}
\end{aligned}$$

✏

Theorem 14.9. [22] *Let $\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. Let $\mathbf{B}^*$ be the adjoint of an operator $\mathbf{B}$.*

THM:
$$|||\mathbf{B}||| = |||\mathbf{B}^*||| \qquad \forall\mathbf{B}\in\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$$

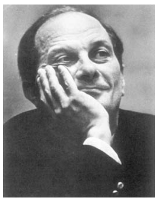
✎PROOF:

$$\begin{aligned}
|||\mathbf{B}||| &\triangleq \sup\{\|\mathbf{B}\boldsymbol{x}\| \mid \|\boldsymbol{x}\|\le 1\} && \text{by Definition 14.4 page 218}\\
&\overset{?}{=} \sup\left\{|\mathsf{g}(\mathbf{B}\boldsymbol{x};\boldsymbol{y}^*)|\,|\,\|\boldsymbol{x}\|\le 1,\ \left\|\boldsymbol{y}^*\right\|\le 1\right\}\\
&= \sup\left\{|\mathsf{f}(\boldsymbol{x};\mathbf{B}^*\boldsymbol{y}^*)|\,|\,\|\boldsymbol{x}\|\le 1,\ \left\|\boldsymbol{y}^*\right\|\le 1\right\}\\
&\triangleq \sup\left\{\left\|\mathbf{B}^*\boldsymbol{y}^*\right\| \mid \|\boldsymbol{x}\|\le 1,\ \left\|\boldsymbol{y}^*\right\|\le 1\right\}\\
&= \sup\left\{\left\|\mathbf{B}^*\boldsymbol{y}^*\right\| \mid \left\|\boldsymbol{y}^*\right\|\le 1\right\}\\
&\triangleq |||\mathbf{B}^*||| && \text{by Definition 14.4 page 218}
\end{aligned}$$

✏

[21] Bollobás (1999), page 156
[22] Rudin (1991) page 98

14.2.4 More properties



"Beginning with the third year of studies, most of my mathematical work was really started in conversations with Mazur and Banach. And according to Banach some of my own contributions were characterized by a certain "strangeness" in the formulation of problems and in the outline of possible proofs. As he told me once some years later, he was surprised how often these "strange" approaches really worked."

Stanislaus M. Ulam (1909–1984), Polish mathematician [23]

Theorem 14.10 (Mazur-Ulam theorem). [24] *Let $\phi \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$ be a function on normed linear spaces $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}})$ and $(\boldsymbol{Y}, \|\cdot\|_{\boldsymbol{Y}})$. Let $\mathbf{I} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{X})$ be the identity operator on $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}})$.*

THM

$$\left.\begin{array}{ll} 1. & \underbrace{\phi^{-1}\phi = \phi\phi^{-1} = \mathbf{I}}_{\textit{bijective}} \qquad \textit{and} \\ 2. & \underbrace{\|\phi\boldsymbol{x} - \phi\boldsymbol{y}\|_{\boldsymbol{Y}} = \|\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{X}} \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}}_{\textit{isometric}} \end{array}\right\} \implies \underbrace{\phi([1-\lambda]\boldsymbol{x} + \lambda\boldsymbol{y}) = [1-\lambda]\phi\boldsymbol{x} + \lambda\phi\boldsymbol{y} \forall \lambda \in \mathbb{R}}_{\textit{affine}}$$

✎PROOF: Proof not yet complete.

1. Let ψ be the *reflection* of $\boldsymbol{z}$ in $\boldsymbol{X}$ such that $\psi\boldsymbol{x} = 2\boldsymbol{z} - \boldsymbol{x}$

   (a) $\|\psi\boldsymbol{x} - \boldsymbol{z}\| = \|\boldsymbol{x} - \boldsymbol{z}\|$

2. Let $\lambda \triangleq \sup_{\mathrm{g}} \{\|\mathrm{g}\boldsymbol{z} - \boldsymbol{z}\|\}$

3. Proof that $\mathrm{g} \in W \implies \mathrm{g}^{-1} \in W$:

   Let $\hat{\boldsymbol{x}} \triangleq \mathrm{g}^{-1}\boldsymbol{x}$ and $\hat{\boldsymbol{y}} \triangleq \mathrm{g}^{-1}\boldsymbol{y}$.

$$\begin{aligned} \left\|\mathrm{g}^{-1}\boldsymbol{x} - \mathrm{g}^{-1}\boldsymbol{y}\right\| &= \|\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}\| && \text{by definition of } \hat{\boldsymbol{x}} \text{ and } \hat{\boldsymbol{y}} \\ &= \|\mathrm{g}\hat{\boldsymbol{x}} - \mathrm{g}\hat{\boldsymbol{y}}\| && \text{by left hypothesis} \\ &= \left\|\mathrm{g}\mathrm{g}^{-1}\boldsymbol{x} - \mathrm{g}\mathrm{g}^{-1}\boldsymbol{y}\right\| && \text{by definition of } \hat{\boldsymbol{x}} \text{ and } \hat{\boldsymbol{y}} \\ &= \|\boldsymbol{x} - \boldsymbol{y}\| && \text{by definition of } \mathrm{g}^{-1} \end{aligned}$$

[23] quote: Ulam (1991), page 33
image: http://www-history.mcs.st-andrews.ac.uk/Biographies/Ulam.html

[24] Oikhberg and Rosenthal (2007), page 598, Väisälä (2003), page 634, Giles (2000), page 11, Dunford and Schwartz (1957), page 91, Mazur and Ulam (1932)

4. Proof that $gz = z$:

$$\begin{aligned}
2\lambda &= 2\sup\{\|gz - z\|\} && \text{by definition of } \lambda \text{ item (2)}\\
&\le 2\|gz - z\| && \text{by definition of sup}\\
&= \|2z - 2gz\|\\
&= \|\psi gz - gz\| && \text{by definition of } \psi \text{ item (1)}\\
&= \left\|g^{-1}\psi gz - g^{-1}gz\right\| && \text{by item (3)}\\
&= \left\|g^{-1}\psi gz - z\right\| && \text{by definition of } g^{-1}\\
&= \left\|\psi g^{-1}\psi gz - z\right\|\\
&= \left\|g^{*}z - z\right\|\\
&\le \lambda && \text{by definition of } \lambda \text{ item (2)}\\
&\implies 2\lambda \le \lambda\\
&\implies \lambda = 0\\
&\implies gz = z
\end{aligned}$$

5. Proof that $\phi\left(\frac{1}{2}x + \frac{1}{2}y\right) = \frac{1}{2}\phi x + \frac{1}{2}\phi y$:

$$\begin{aligned}
\phi\left(\frac{1}{2}x + \frac{1}{2}y\right) &=\\
&= \frac{1}{2}\phi x + \frac{1}{2}\phi y
\end{aligned}$$

6. Proof that $\phi([1-\lambda]x + \lambda y) = [1-\lambda]\phi x + \lambda\phi y$:

$$\begin{aligned}
\phi([1-\lambda]x + \lambda y) &=\\
&= [1-\lambda]\phi x + \lambda\phi y
\end{aligned}$$

Theorem 14.11 (Neumann Expansion Theorem). [25] *Let* $\mathbf{A} \in \boldsymbol{X}^{\boldsymbol{X}}$ *be an operator on a linear space* $\boldsymbol{X}$. *Let* $\mathbf{A}^0 \triangleq \mathbf{I}$.

THM
$$\left.\begin{array}{ll}
1. & \mathbf{A} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{X}) \quad (\mathbf{A} \text{ is bounded})\\
2. & |||\mathbf{A}||| < 1
\end{array}\right\} \implies \left\{\begin{array}{llcl}
1. & (\mathbf{I} - \mathbf{A})^{-1} & & \textit{exists}\\
2. & \left|\left|\left|(\mathbf{I} - \mathbf{A})^{-1}\right|\right|\right| & \le & \dfrac{1}{1 - |||\mathbf{A}|||}\\
3. & (\mathbf{I} - \mathbf{A})^{-1} & = & \displaystyle\sum_{n=0}^{\infty} \mathbf{A}^n\\
& \multicolumn{3}{l}{\textit{with uniform convergence}}
\end{array}\right.$$

14.3 Operators on Inner product spaces

14.3.1 General Results

Theorem 14.12. [26] *Let* $\mathbf{A}, \mathbf{B} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{X})$ *be bounded linear operators on an inner product space* $\boldsymbol{X} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla \rangle\right)$.

THM
$$\begin{array}{lclll}
\langle \mathbf{B}x \mid x \rangle = 0 & \forall x \in X & \iff & \mathbf{B}x = \mathbb{0} & \forall x \in X\\
\langle \mathbf{A}x \mid x \rangle = \langle \mathbf{B}x \mid x \rangle & \forall x \in X & \iff & \mathbf{A} = \mathbf{B} &
\end{array}$$

[25] Michel and Herget (1993) page 415
[26] Rudin (1991) page 310 ⟨Theorem 12.7, Corollary⟩

✎PROOF:

1. Proof that $\langle \mathbf{B}x \mid x \rangle = 0 \implies \mathbf{B}x = \mathbb{0}$:

$$\begin{aligned}
0 &= \langle \mathbf{B}(x+\mathbf{B}x) \mid (x+\mathbf{B}x)\rangle + i\,\langle \mathbf{B}(x+i\mathbf{B}x) \mid (x+i\mathbf{B}x)\rangle && \text{by left hypothesis}\\
&= \left\{\left\langle \mathbf{B}x+\mathbf{B}^2x) \mid x+\mathbf{B}x\right\rangle\right\} + i\left\{\left\langle \mathbf{B}x+i\mathbf{B}^2x) \mid x+i\mathbf{B}x\right\rangle\right\} && \text{by Definition 14.3 page 214}\\
&= \left\{\langle \mathbf{B}x \mid x\rangle + \langle \mathbf{B}x \mid \mathbf{B}x\rangle + \left\langle \mathbf{B}^2x \mid x\right\rangle + \left\langle \mathbf{B}^2x \mid \mathbf{B}x\right\rangle\right\} && \text{by Definition 6.1 page 95}\\
&\quad + i\left\{\langle \mathbf{B}x \mid x\rangle - i\,\langle \mathbf{B}x \mid \mathbf{B}x\rangle + i\left\langle \mathbf{B}^2x \mid x\right\rangle - i^2\left\langle \mathbf{B}^2x \mid \mathbf{B}x\right\rangle\right\}\\
&= \left\{0 + \langle \mathbf{B}x \mid \mathbf{B}x\rangle + \left\langle \mathbf{B}^2x \mid x\right\rangle + 0\right\} + i\left\{0 - i\,\langle \mathbf{B}x \mid \mathbf{B}x\rangle + i\left\langle \mathbf{B}^2x \mid x\right\rangle - i^2 0\right\} && \text{by left hypothesis}\\
&= \left\{\langle \mathbf{B}x \mid \mathbf{B}x\rangle + \left\langle \mathbf{B}^2x \mid x\right\rangle\right\} + \left\{\langle \mathbf{B}x \mid \mathbf{B}x\rangle - \left\langle \mathbf{B}^2x \mid x\right\rangle\right\}\\
&= 2\,\langle \mathbf{B}x \mid \mathbf{B}x\rangle\\
&= 2\,\|\mathbf{B}x\|^2\\
&\implies \mathbf{B}x = \mathbb{0} && \text{by Definition 5.1 page 83}
\end{aligned}$$

2. Proof that $\langle \mathbf{B}x \mid x \rangle = 0 \impliedby \mathbf{B}x = \mathbb{0}$: by property of inner products (Theorem 6.1 page 95).

3. Proof that $\langle \mathbf{A}x \mid x \rangle = \langle \mathbf{B}x \mid x \rangle \implies \mathbf{A} \doteq \mathbf{B}$:

$$\begin{aligned}
0 &= \langle \mathbf{A}x \mid x\rangle - \langle \mathbf{B}x \mid x\rangle && \text{by left hypothesis}\\
&= \langle \mathbf{A}x - \mathbf{B}x \mid x\rangle && \text{by } \textit{additivity} \text{ property of } \langle \triangle \mid \triangledown\rangle \text{ (Definition 6.1 page 95)}\\
&= \langle (\mathbf{A}-\mathbf{B})x \mid x\rangle && \text{by definition of operator addition}\\
\implies (\mathbf{A}-\mathbf{B})x &= \mathbb{0} && \text{by item 1}\\
\implies \mathbf{A} &= \mathbf{B} && \text{by definition of operator subtraction}
\end{aligned}$$

4. Proof that $\langle \mathbf{A}x \mid x \rangle = \langle \mathbf{B}x \mid x \rangle \impliedby \mathbf{A} \doteq \mathbf{B}$:

$$\langle \mathbf{A}x \mid x\rangle = \langle \mathbf{B}x \mid x\rangle \qquad \text{by } \mathbf{A} \doteq \mathbf{B} \text{ hypothesis}$$

✏

14.3.2 Operator adjoint

A fundamental concept of operators on inner product spaces is the *operator adjoint* (Proposition 14.3 page 226). The adjoint of an operator is a kind of generalization of the conjugate of a complex number in that

- Both are *star-algebras* (Theorem 14.13 page 227).
- Both support decomposition into "real" and "imaginary" parts (Theorem 16.3 page 260).

Structurally, the operator adjoint provides a convenient symmetric relationship between the *range space* and *null space* of an operator (Theorem 14.14 page 228).

Proposition 14.3. [27] *Let $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$ be the space of bounded linear operators on a Hilbert space $\boldsymbol{H}$.*[28]

PRP *An operator $\mathbf{B}^*$ is the* ADJOINT *of $\mathbf{B} \in \mathcal{B}(\boldsymbol{H},\boldsymbol{H})$ if*
$$\langle \mathbf{B}x \mid y\rangle = \langle x \mid \mathbf{B}^* y\rangle \qquad \forall x, y \in \boldsymbol{H}.$$

[27] Michel and Herget (1993) page 220, Rudin (1991) page 311, Giles (2000), page 182, von Neumann (1929) page 49, Stone (1932) page 41

[28] *bounded operator*: Definition 14.5 (page 221); *adjoint*: Definition 14.6 (page 223)

✎PROOF:

1. For fixed y, $\mathsf{f}(\boldsymbol{x}) \triangleq \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$ is a *functional* in $\mathbb{F}^{\boldsymbol{X}}$ (Definition 13.1 page 207).
2. $\mathbf{B}^*$ is the *adjoint* of $\mathbf{B}$ because

$$\begin{aligned}
\langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{y} \rangle &\triangleq \mathsf{f}(\mathbf{B}\boldsymbol{x}) && \text{by Definition 13.1 (page 207)}\\
&\triangleq \mathbf{B}^*\mathsf{f}(\boldsymbol{x}) && \text{by Definition 14.6 (page 223)}\\
&= \langle \boldsymbol{x} \mid \mathbf{B}^*\boldsymbol{y} \rangle && \text{by Definition 13.1 (page 207)}
\end{aligned}$$

✏

Example 14.2 (Matrix algebra: $\mathbf{A}^* = \mathbf{A}^H$). In matrix algebra,

EX
- The inner product operation $\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$ is represented by $\boldsymbol{y}^H\boldsymbol{x}$.
- The linear operator is represented as a matrix $\boldsymbol{A}$.
- The operation of $\boldsymbol{A}$ on vector $\boldsymbol{x}$ is represented as $\boldsymbol{A}\boldsymbol{x}$.
- The adjoint of matrix $\boldsymbol{A}$ is the Hermitian matrix $\boldsymbol{A}^H$.

✎PROOF:

$$\langle \boldsymbol{A}\boldsymbol{x} \mid \boldsymbol{y} \rangle \triangleq \boldsymbol{y}^H\boldsymbol{A}\boldsymbol{x} = [(\boldsymbol{A}\boldsymbol{x})^H\boldsymbol{y}]^H = [\boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{y}]^H = (\boldsymbol{A}^H\boldsymbol{y})^H\boldsymbol{x} \triangleq \left\langle \boldsymbol{x} \mid \boldsymbol{A}^H\boldsymbol{y} \right\rangle$$

✏

Structures that satisfy the four conditions of the next theorem are known as **-algebras* ("*star-algebras*", Definition 16.3 page 258). Other structures which are *-algebras include the *field of complex numbers* $\mathbb{C}$ and any *ring of complex square $n \times n$ matrices*.[29]

Theorem 14.13 (operator star-algebra). [30] *Let $\boldsymbol{H}$ be a Hilbert space with operators $\mathbf{A}, \mathbf{B} \in \mathcal{B}(\boldsymbol{H}, \boldsymbol{H})$ and with adjoints $\mathbf{A}^*, \mathbf{B}^* \in \mathcal{B}(\boldsymbol{H}, \boldsymbol{H})$. Let $\bar{\alpha}$ be the complex conjugate of some $\alpha \in \mathbb{C}$.*

THM
The pair $(\boldsymbol{H}, *)$ is a $*$-ALGEBRA (STAR-ALGEBRA). In particular,

1.	$(\mathbf{A} \dotplus \mathbf{B})^*$	$=$	$\mathbf{A}^* + \mathbf{B}^*$	$\forall \mathbf{A},\mathbf{B} \in \boldsymbol{H}$	(DISTRIBUTIVE)	*and*
2.	$(\alpha\mathbf{A})^*$	$=$	$\bar{\alpha}\mathbf{A}^*$	$\forall \mathbf{A},\mathbf{B} \in \boldsymbol{H}$	(CONJUGATE LINEAR)	*and*
3.	$(\mathbf{A}\mathbf{B})^*$	$=$	$\mathbf{B}^*\mathbf{A}^*$	$\forall \mathbf{A},\mathbf{B} \in \boldsymbol{H}$	(ANTIAUTOMORPHIC)	*and*
4.	$\mathbf{A}^{**}$	$=$	$\mathbf{A}$	$\forall \mathbf{A},\mathbf{B} \in \boldsymbol{H}$	(INVOLUTARY)	

✎PROOF:

$$\begin{aligned}
\langle \boldsymbol{x} \mid (\mathbf{A} \dotplus \mathbf{B})^*\boldsymbol{y} \rangle &= \langle (\mathbf{A} \dotplus \mathbf{B})\boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by definition of adjoint} && \text{(Proposition 14.3 page 226)}\\
&= \langle \mathbf{A}\boldsymbol{x} \mid \boldsymbol{y} \rangle + \langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by definition of inner product} && \text{(Definition 6.1 page 95)}\\
&= \langle \boldsymbol{x} \mid \mathbf{A}^*\boldsymbol{y} \rangle + \langle \boldsymbol{x} \mid \mathbf{B}^*\boldsymbol{y} \rangle && \text{by definition of operator addition}\\
&= \langle \boldsymbol{x} \mid \mathbf{A}^*\boldsymbol{y} + \mathbf{B}^*\boldsymbol{y} \rangle && \text{by definition of inner product} && \text{(Definition 6.1 page 95)}\\
&= \langle \boldsymbol{x} \mid (\mathbf{A}^* + \mathbf{B}^*)\boldsymbol{y} \rangle && \text{by definition of operator addition}
\end{aligned}$$

$$\begin{aligned}
\langle \boldsymbol{x} \mid (\alpha\mathbf{A})^*\boldsymbol{y} \rangle &= \langle (\alpha\mathbf{A})\boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by definition of adjoint} && \text{(Proposition 14.3 page 226)}\\
&= \langle \alpha(\mathbf{A}\boldsymbol{x}) \mid \boldsymbol{y} \rangle && \text{by definition of scalar multiplication}\\
&= \alpha \langle \mathbf{A}\boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by definition of inner product} && \text{(Definition 6.1 page 95)}\\
&= \alpha \langle \boldsymbol{x} \mid \mathbf{A}^*\boldsymbol{y} \rangle && \text{by definition of adjoint} && \text{(Proposition 14.3 page 226)}
\end{aligned}$$

[29] Sakai (1998) page 1

[30] Halmos (1998a), pages 39–40, Rudin (1991) page 311

$$
\begin{aligned}
&= \langle x \mid \alpha^* \mathbf{A}^* y \rangle && \text{by definition of inner product} && \text{(Definition 6.1 page 95)}\\
\langle x \mid (\mathbf{AB})^* y \rangle &= \langle (\mathbf{AB})x \mid y \rangle && \text{by definition of adjoint} && \text{(Proposition 14.3 page 226)}\\
&= \langle \mathbf{A}(\mathbf{B}x) \mid y \rangle && \text{by definition of operator multiplication}\\
&= \langle (\mathbf{B}x) \mid \mathbf{A}^* y \rangle && \text{by definition of adjoint} && \text{(Proposition 14.3 page 226)}\\
&= \langle x \mid \mathbf{B}^* \mathbf{A}^* y \rangle && \text{by definition of adjoint} && \text{(Proposition 14.3 page 226)}\\
\langle x \mid \mathbf{A}^{**} y \rangle &= \langle \mathbf{A}^* x \mid y \rangle && \text{by definition of adjoint} && \text{(Proposition 14.3 page 226)}\\
&= \langle y \mid \mathbf{A}^* x \rangle^* && \text{by definition of inner product} && \text{(Definition 6.1 page 95)}\\
&= \langle \mathbf{A} y \mid x \rangle^* && \text{by definition of adjoint} && \text{(Proposition 14.3 page 226)}\\
&= \langle x \mid \mathbf{A} y \rangle && \text{by definition of inner product} && \text{(Definition 6.1 page 95)}
\end{aligned}
$$

⇰

Theorem 14.14. [31] *Let $\boldsymbol{Y}^{\boldsymbol{X}}$ be the set of all operators from a linear space $\boldsymbol{X}$ to a linear space $\boldsymbol{Y}$. Let $\mathcal{N}(\mathbf{L})$ be the* NULL SPACE *of an operator $\mathbf{L}$ in $\boldsymbol{Y}^{\boldsymbol{X}}$ and $\mathcal{I}(\mathbf{L})$ the* IMAGE SET *of $\mathbf{L}$ in $\boldsymbol{Y}^{\boldsymbol{X}}$.*

THM

$$\mathcal{N}(\mathbf{A}) = \mathcal{I}(\mathbf{A}^*)^{\perp}$$
$$\mathcal{N}(\mathbf{A}^*) = \mathcal{I}(\mathbf{A})^{\perp}$$

✎PROOF:

$$
\begin{aligned}
\mathcal{I}(\mathbf{A}^*)^{\perp} &= \left\{ y \in \boldsymbol{H} \mid \langle y \mid u \rangle = 0 \quad \forall u \in \mathcal{I}(\mathbf{A}^*) \right\}\\
&= \left\{ y \in \boldsymbol{H} \mid \langle y \mid \mathbf{A}^* x \rangle = 0 \quad \forall x \in \boldsymbol{H} \right\}\\
&= \{ y \in \boldsymbol{H} \mid \langle \mathbf{A} y \mid x \rangle = 0 \quad \forall x \in \boldsymbol{H} \} && \text{by definition of } \mathbf{A}^* && \text{(Proposition 14.3 page 226)}\\
&= \{ y \in \boldsymbol{H} \mid \mathbf{A} y = 0 \}\\
&= \mathcal{N}(\mathbf{A}) && \text{by definition of } \mathcal{N}(\mathbf{A})
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{I}(\mathbf{A})^{\perp} &= \{ y \in \boldsymbol{H} \mid \langle y \mid u \rangle = 0 \quad \forall u \in \mathcal{I}(\mathbf{A}) \}\\
&= \{ y \in \boldsymbol{H} \mid \langle y \mid \mathbf{A} x \rangle = 0 \quad \forall x \in \boldsymbol{H} \} && \text{by definition of } \mathcal{I}\\
&= \left\{ y \in \boldsymbol{H} \mid \langle \mathbf{A}^* y \mid x \rangle = 0 \quad \forall x \in \boldsymbol{H} \right\} && \text{by definition of } \mathbf{A}^* && \text{(Proposition 14.3 page 226)}\\
&= \left\{ y \in \boldsymbol{H} \mid \mathbf{A}^* y = 0 \right\}\\
&= \mathcal{N}(\mathbf{A}^*) && \text{by definition of } \mathcal{N}(\mathbf{A})
\end{aligned}
$$

⇰

14.4 Special Classes of Operators

14.4.1 Projection operators

Definition 14.7. [32] *Let $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. Let $\mathbf{P}$ be a bounded linear operator in $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$.*

DEF

$\mathbf{P}$ *is a **projection** operator if* $\quad \mathbf{P}^2 = \mathbf{P}$.

[31] Rudin (1991) page 312

[32] Rudin (1991) page 133 ⟨5.15 Projections⟩, Kubrusly (2001) page 70, Bachman and Narici (1966) page 6, Halmos (1958) page 73 ⟨§41. Projections⟩

Theorem 14.15. [33] *Let $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. Let $\mathbf{P}$ be a bounded linear operator in $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ with* NULL SPACE $\mathcal{N}(\mathbf{P})$ *and* IMAGE SET $\mathcal{I}(\mathbf{P})$.

THM

$$\left.\begin{array}{lrcll} 1. & \mathbf{P}^2 &=& \mathbf{P} & (\mathbf{P} \text{ is a projection operator}) \quad \text{and} \\ 2. & \boldsymbol{\Omega} &=& \boldsymbol{X} \dotplus \boldsymbol{Y} & (\boldsymbol{Y} \text{ compliments } \boldsymbol{X} \text{ in } \boldsymbol{\Omega}) \quad \text{and} \\ 3. & \mathbf{P}\boldsymbol{\Omega} &=& \boldsymbol{X} & (\mathbf{P} \text{ projects onto } \boldsymbol{X}) \end{array}\right\} \implies \left\{\begin{array}{lrcll} 1. & \mathcal{I}(\mathbf{P}) &=& \boldsymbol{X} & \text{and} \\ 2. & \mathcal{N}(\mathbf{P}) &=& \boldsymbol{Y} & \text{and} \\ 3. & \boldsymbol{\Omega} &=& \mathcal{I}(\mathbf{P}) \dotplus \mathcal{N}(\mathbf{P}) & \end{array}\right.$$

✎PROOF:

$$\begin{aligned} \mathcal{I}(\mathbf{P}) &= \mathbf{P}\boldsymbol{\Omega} \\ &= \mathbf{P}(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2) \\ &= \mathbf{P}\boldsymbol{\Omega}_1 + \mathbf{P}\boldsymbol{\Omega}_2 \\ &= \boldsymbol{\Omega}_1 + \{\mathbb{0}\} \\ &= \boldsymbol{\Omega}_1 \end{aligned}$$

$$\begin{aligned} \mathcal{N}(\mathbf{P}) &= \{x \in \boldsymbol{\Omega} | \mathbf{P}x = \mathbb{0}\} \\ &= \left\{x \in (\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2) | \mathbf{P}x = \mathbb{0}\right\} \\ &= \left\{x \in \boldsymbol{\Omega}_1 | \mathbf{P}x = \mathbb{0}\right\} + \left\{x \in \boldsymbol{\Omega}_2 | \mathbf{P}x = \mathbb{0}\right\} \\ &= \{\mathbb{0}\} + \boldsymbol{\Omega}_2 \\ &= \boldsymbol{\Omega}_2 \end{aligned}$$

✏

Theorem 14.16. [34] *Let $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. Let $\mathbf{P}$ be a bounded linear operator in $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$.*

THM

$$\underbrace{\mathbf{P}^2 = \mathbf{P}}_{\mathbf{P} \text{ is a projection operator}} \iff \underbrace{(\mathbf{I} - \mathbf{P})^2 = (\mathbf{I} - \mathbf{P})}_{(\mathbf{I} - \mathbf{P}) \text{ is a projection operator}}$$

✎PROOF:

Proof that $\mathbf{P}^2 = \mathbf{P} \implies (\mathbf{I} - \mathbf{P})^2 = (\mathbf{I} - \mathbf{P})$:

$$\begin{aligned} (\mathbf{I} - \mathbf{P})^2 &= (\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{P}) \\ &= \mathbf{I}(\mathbf{I} - \mathbf{P}) + (-\mathbf{P})(\mathbf{I} - \mathbf{P}) \\ &= \mathbf{I} - \mathbf{P} - \mathbf{P}\mathbf{I} + \mathbf{P}^2 \\ &= \mathbf{I} - \mathbf{P} - \mathbf{P} + \mathbf{P} && \text{by left hypothesis} \\ &= \mathbf{I} - \mathbf{P} \end{aligned}$$

Proof that $\mathbf{P}^2 = \mathbf{P} \impliedby (\mathbf{I} - \mathbf{P})^2 = (\mathbf{I} - \mathbf{P})$:

$$\begin{aligned} \mathbf{P}^2 &= \underbrace{\mathbf{I} - \mathbf{P} - \mathbf{P} + \mathbf{P}^2}_{(\mathbf{I} - \mathbf{P})^2} - (\mathbf{I} - \mathbf{P} - \mathbf{P}) \\ &= (\mathbf{I} - \mathbf{P})^2 - (\mathbf{I} - \mathbf{P} - \mathbf{P}) \\ &= (\mathbf{I} - \mathbf{P}) - (\mathbf{I} - \mathbf{P} - \mathbf{P}) && \text{by right hypothesis} \\ &= \mathbf{P} \end{aligned}$$

✏

[33] Michel and Herget (1993) pages 120–121

[34] Michel and Herget (1993) page 121

14.4.2 Self Adjoint Operators

Definition 14.8. [35] *Let* $\mathbf{B} \in \mathcal{B}(\boldsymbol{H}, \boldsymbol{H})$ *be a bounded operator with adjoint* $\mathbf{B}^*$ *on a Hilbert space* $\boldsymbol{H}$.

DEF: *The operator* $\mathbf{B}$ *is said to be* ***self-adjoint*** *or* ***hermitian*** *if* $\mathbf{B} \stackrel{\circ}{=} \mathbf{B}^*$.

Example 14.3 (Autocorrelation operator). Let $\mathsf{x}(t)$ be a random process with autocorrelation

$$\mathsf{R}_{\mathsf{xx}}(t,u) \triangleq \underbrace{\mathsf{E}[\mathsf{x}(t)\mathsf{x}^*(u)]}_{\text{expectation}}.$$

Let an autocorrelation operator $\mathbf{R}$ be defined as $[\mathbf{R}\mathsf{f}](t) \triangleq \int_{\mathbb{R}} \underbrace{\mathsf{R}_{\mathsf{xx}}(t,u)}_{\text{kernel}} \mathsf{f}(u)\,\mathsf{d}u$.

EX: $\mathbf{R} = \mathbf{R}^*$ (The autocorrelation operator $\mathbf{R}$ is *self-adjoint*)

✎PROOF:

1. First note that the *autocorrelation kernel* $\mathsf{R}_{\mathsf{xx}}(t,u)$ is *hermitian symmetric*:

$$\begin{aligned}\mathsf{R}_{\mathsf{xx}}(t,u) &\triangleq \mathsf{E}\mathsf{x}(t)\mathsf{x}^*(u) = [\mathsf{E}\mathsf{x}^*(t)\mathsf{x}(u)]^* = [\mathsf{E}\mathsf{x}(u)\mathsf{x}^*(t)]^* \\ &= \mathsf{R}^*_{\mathsf{xx}}(u,t)\end{aligned}$$

2. Proof that the *autocorrelation operator* $\mathbf{R}$ is *self-adjoint*:

$$\begin{aligned}
\langle \mathbf{R}\mathsf{f} \,|\, \mathsf{g} \rangle &= \left\langle \int_{u\in\mathbb{R}} \mathsf{R}_{\mathsf{xx}}(t,u)\mathsf{f}(u)\,\mathsf{d}u \,\middle|\, \mathsf{g}(t) \right\rangle && \text{by definition of } \mathbf{R} \\
&= \int_{u\in\mathbb{R}} \mathsf{f}(u) \langle \mathsf{R}_{\mathsf{xx}}(t,u) \,|\, \mathsf{g}(t) \rangle \,\mathsf{d}u \\
&= \int_{u\in\mathbb{R}} \mathsf{f}(u) \int_t \mathsf{R}_{\mathsf{xx}}(t,u)\mathsf{g}^*(t)\,\mathsf{d}t\,\mathsf{d}u \\
&= \int_{u\in\mathbb{R}} \mathsf{f}(u) \int_t \mathsf{R}^*_{\mathsf{xx}}(u,t)\mathsf{g}^*(t)\,\mathsf{d}t\,\mathsf{d}u && \text{by 1.} \\
&= \int_{u\in\mathbb{R}} \mathsf{f}(u) \left[\int_t \mathsf{R}_{\mathsf{xx}}(u,t)\mathsf{g}(t)\,\mathsf{d}t\right]^* \mathsf{d}u \\
&= \int_{u\in\mathbb{R}} \mathsf{f}(u) \left[\mathbf{R}\mathsf{g}\right]^* \mathsf{d}u && \text{by definition of } \mathbf{R} \\
&= \langle \mathsf{f} \,|\, \mathbf{R}\mathsf{g} \rangle
\end{aligned}$$

✇

Theorem 14.17. [36] *Let* $\mathbf{S} : \boldsymbol{H} \to \boldsymbol{H}$ *be an operator over a Hilbert space* $\boldsymbol{H}$ *with eigenvalues* $\{\lambda_n\}$ *and eigenfunctions* $\{\psi_n\}$ *such that* $\mathbf{S}\psi_n = \lambda_n\psi_n$ *and let* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \,|\, \boldsymbol{x} \rangle}$.

THM: $\underbrace{\mathbf{S} = \mathbf{S}^*}_{\mathbf{S}\text{ is SELF ADJOINT}} \implies \begin{cases} 1.\ \langle \mathbf{S}\boldsymbol{x} \,|\, \boldsymbol{x} \rangle \in \mathbb{R} & \text{(the hermitian quadratic form of } \mathbf{S} \text{ is real)} \\ 2.\ \lambda_n \in \mathbb{R} & \text{(eigenvalues of } \mathbf{S} \text{ are real)} \\ 3.\ \lambda_n \neq \lambda_m \implies \langle \psi_n \,|\, \psi_m \rangle = 0 & \text{(eigenfunctions associated with distinct eigenvalues are orthogonal)} \end{cases}$

✎PROOF:

[35] Historical works regarding self-adjoint operators: von Neumann (1929) page 49, "linearer Operator R selbstadjungiert oder Hermitesch", Stone (1932) page 50 ⟨"self-adjoint transformations"⟩

[36] Lax (2002), pages 315–316, Keener (1988), pages 114–119

1. Proof that $\mathbf{S} = \mathbf{S}^* \implies \langle \mathbf{S}\boldsymbol{x} \,|\, \boldsymbol{x} \rangle \in \mathbb{R}$:

$$\begin{aligned}
\langle \boldsymbol{x} \,|\, \mathbf{S}\boldsymbol{x} \rangle &= \langle \mathbf{S}\boldsymbol{x} \,|\, \boldsymbol{x} \rangle && \text{by left hypothesis} \\
&= \langle \boldsymbol{x} \,|\, \mathbf{S}\boldsymbol{x} \rangle^* && \text{by definition of } \langle \triangle \,|\, \nabla \rangle \text{ Definition 6.1 page 95}
\end{aligned}$$

2. Proof that $\mathbf{S} = \mathbf{S}^* \implies \lambda_n \in \mathbb{R}$:

$$\begin{aligned}
\lambda_n \|\psi_n\|^2 &= \lambda_n \langle \psi_n \,|\, \psi_n \rangle && \text{by definition} \\
&= \langle \lambda_n \psi_n \,|\, \psi_n \rangle && \text{by definition of } \langle \triangle \,|\, \nabla \rangle \text{ Definition 6.1 page 95} \\
&= \langle \mathbf{S}\psi_n \,|\, \psi_n \rangle && \text{by definition of eigenpairs} \\
&= \langle \psi_n \,|\, \mathbf{S}\psi_n \rangle && \text{by left hypothesis} \\
&= \langle \psi_n \,|\, \lambda_n \psi_n \rangle && \text{by definition of eigenpairs} \\
&= \lambda_n^* \langle \psi_n \,|\, \psi_n \rangle && \text{by definition of } \langle \triangle \,|\, \nabla \rangle \text{ Definition 6.1 page 95} \\
&= \lambda_n^* \|\psi_n\|^2 && \text{by definition}
\end{aligned}$$

3. Proof that $\mathbf{S} = \mathbf{S}^* \implies [\lambda_n \neq \lambda_m \implies \langle \psi_n \,|\, \psi_m \rangle = 0]$:

$$\begin{aligned}
\lambda_n \langle \psi_n \,|\, \psi_m \rangle &= \langle \lambda_n \psi_n \,|\, \psi_m \rangle && \text{by definition of } \langle \triangle \,|\, \nabla \rangle \text{ Definition 6.1 page 95} \\
&= \langle \mathbf{S}\psi_n \,|\, \psi_m \rangle && \text{by definition of eigenpairs} \\
&= \langle \psi_n \,|\, \mathbf{S}\psi_m \rangle && \text{by left hypothesis} \\
&= \langle \psi_n \,|\, \lambda_m \psi_m \rangle && \text{by definition of eigenpairs} \\
&= \lambda_m^* \langle \psi_n \,|\, \psi_m \rangle && \text{by definition of } \langle \triangle \,|\, \nabla \rangle \text{ Definition 6.1 page 95} \\
&= \lambda_m \langle \psi_n \,|\, \psi_m \rangle && \text{because } \lambda_m \text{ is real}
\end{aligned}$$

This implies for $\lambda_n \neq \lambda_m \neq 0$, $\langle \psi_n \,|\, \psi_m \rangle = 0$.

⇨

14.4.3 Normal Operators

Definition 14.9. [37] *Let $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. Let $\mathbf{N}^*$ be the adjoint of an operator $\mathbf{N} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$.*

DEF: $\mathbf{N}$ *is **normal** if* $\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^*$.

Theorem 14.18. [38] *Let $\mathcal{B}(\boldsymbol{H}, \boldsymbol{H})$ be the space of bounded linear operators on a Hilbert space $\boldsymbol{H}$. Let $\mathcal{N}(\mathbf{N})$ be the* NULL SPACE *of an operator $\mathbf{N}$ in $\mathcal{B}(\boldsymbol{H}, \boldsymbol{H})$ and $\mathcal{I}(\mathbf{N})$ the* IMAGE SET *of $\mathbf{N}$ in $\mathcal{B}(\boldsymbol{H}, \boldsymbol{H})$.*

THM: $$\underbrace{\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^*}_{\mathbf{N} \text{ is normal}} \iff \|\mathbf{N}^*\boldsymbol{x}\| = \|\mathbf{N}\boldsymbol{x}\| \qquad \forall \boldsymbol{x} \in \boldsymbol{H}$$

✎PROOF:

[37] Rudin (1991) page 312, Michel and Herget (1993) page 431, Dieudonné (1969), page 167, Frobenius (1878), Frobenius (1968) page 391
[38] Rudin (1991) pages 312–313

1. Proof that $\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^* \implies \|\mathbf{N}^*\boldsymbol{x}\| = \|\mathbf{N}\boldsymbol{x}\|$:

$$\begin{aligned}
\|\mathbf{N}\boldsymbol{x}\|^2 &= \langle \mathbf{N}\boldsymbol{x} \mid \mathbf{N}\boldsymbol{x}\rangle && \text{by definition}\\
&= \langle \boldsymbol{x} \mid \mathbf{N}^*\mathbf{N}\boldsymbol{x}\rangle && \text{by Proposition 14.3 page 226 (definition of } \mathbf{N}^*)\\
&= \langle \boldsymbol{x} \mid \mathbf{N}\mathbf{N}^*\boldsymbol{x}\rangle && \text{by left hypothesis (}\mathbf{N}\text{ is normal)}\\
&= \langle \mathbf{N}\boldsymbol{x} \mid \mathbf{N}^*\boldsymbol{x}\rangle && \text{by Proposition 14.3 page 226 (definition of } \mathbf{N}^*)\\
&= \|\mathbf{N}^*\boldsymbol{x}\|^2 && \text{by definition}
\end{aligned}$$

2. Proof that $\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^* \impliedby \|\mathbf{N}^*\boldsymbol{x}\| = \|\mathbf{N}\boldsymbol{x}\|$:

$$\begin{aligned}
\langle \mathbf{N}^*\mathbf{N}\boldsymbol{x} \mid \boldsymbol{x}\rangle &= \langle \mathbf{N}\boldsymbol{x} \mid \mathbf{N}^{**}\boldsymbol{x}\rangle && \text{by Proposition 14.3 page 226 (definition of } \mathbf{N}^*)\\
&= \langle \mathbf{N}\boldsymbol{x} \mid \mathbf{N}\boldsymbol{x}\rangle && \text{by Theorem 14.13 page 227 (property of adjoint)}\\
&= \|\mathbf{N}\boldsymbol{x}\|^2 && \text{by definition}\\
&= \|\mathbf{N}^*\boldsymbol{x}\|^2 && \text{by right hypothesis } (\|\mathbf{N}^*\boldsymbol{x}\| = \|\mathbf{N}\boldsymbol{x}\|)\\
&= \langle \mathbf{N}^*\boldsymbol{x} \mid \mathbf{N}^*\boldsymbol{x}\rangle && \text{by definition}\\
&= \langle \mathbf{N}\mathbf{N}^*\boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by Proposition 14.3 page 226 (definition of } \mathbf{N}^*)
\end{aligned}$$

✏

Theorem 14.19. [39] *Let $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$ be the space of bounded linear operators on a Hilbert space $\boldsymbol{H}$. Let $\mathcal{N}(\mathbf{N})$ be the* NULL SPACE *of an operator* $\mathbf{N}$ *in $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$ and $\mathcal{I}(\mathbf{N})$ the* IMAGE SET *of* $\mathbf{N}$ *in $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$.*

THM
$$\underbrace{\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^*}_{\mathbf{N}\text{ is normal}} \implies \underbrace{\mathcal{N}(\mathbf{N}^*) = \mathcal{N}(\mathbf{N})}_{\mathbf{N}\text{ and }\mathbf{N}^*\text{ have the same null space}}$$

✎PROOF:

$$\begin{aligned}
\mathcal{N}(\mathbf{N}^*) &= \left\{\boldsymbol{x} \middle| \mathbf{N}^*\boldsymbol{x} = 0 \quad \forall \boldsymbol{x}\in\boldsymbol{X}\right\} && \text{(definition of } \mathcal{N})\\
&= \left\{\boldsymbol{x} \middle| \|\mathbf{N}^*\boldsymbol{x}\| = 0 \quad \forall \boldsymbol{x}\in\boldsymbol{X}\right\} && \text{by definition of } \|\cdot\| \text{ (Definition 5.1 page 83)}\\
&= \left\{\boldsymbol{x} \middle| \|\mathbf{N}\boldsymbol{x}\| = 0 \quad \forall \boldsymbol{x}\in\boldsymbol{X}\right\}\\
&= \left\{\boldsymbol{x} \middle| \mathbf{N}\boldsymbol{x} = 0 \quad \forall \boldsymbol{x}\in\boldsymbol{X}\right\} && \text{by definition of } \|\cdot\| \text{ (Definition 5.1 page 83)}\\
&= \mathcal{N}(\mathbf{N}) && \text{(definition of } \mathcal{N})
\end{aligned}$$

✏

14.4.4 Isometric operators

An operator on a pair of normed linear spaces is *isometric* (next definition) if it is an *isometry*.

Definition 14.10. *Let* $(X, +, \cdot, (\mathbb{F}, \dotplus, \dot\times), \|\cdot\|)$ *and* $(Y, +, \cdot, (\mathbb{F}, \dotplus, \dot\times), \|\cdot\|)$ *be* NORMED LINEAR SPACES *(Definition 5.1 page 83).*

DEF
An operator $\mathbf{M}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$ *is **isometric** if*
$$\|\mathbf{M}\boldsymbol{x}\| = \|\boldsymbol{x}\| \qquad \forall x\in X.$$

Theorem 14.20. [40] *Let* $(X, +, \cdot, (\mathbb{F}, \dotplus, \dot\times), \|\cdot\|)$ *and* $(Y, +, \cdot, (\mathbb{F}, \dotplus, \dot\times), \|\cdot\|)$ *be* NORMED LINEAR SPACES. *Let* $\mathbf{M}$ *be a linear operator in* $\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$.

THM
$$\underbrace{\|\mathbf{M}\boldsymbol{x}\| = \|\boldsymbol{x}\| \quad \forall x\in X}_{\text{isometric in length}} \iff \underbrace{\|\mathbf{M}\boldsymbol{x} - \mathbf{M}\boldsymbol{y}\| = \|\boldsymbol{x}-\boldsymbol{y}\| \quad \forall x,y\in X}_{\text{isometric in distance}}$$

[39] Rudin (1991) pages 312–313
[40] Kubrusly (2001) page 239 ⟨Proposition 4.37⟩, Berberian (1961) page 27 ⟨Theorem IV.7.5⟩

✎PROOF:

1. Proof that $\|\mathbf{M}x\| = \|x\| \implies \|\mathbf{M}x - \mathbf{M}y\| = \|x - y\|$:

$$\begin{aligned}
\|\mathbf{M}x - \mathbf{M}y\| &= \|\mathbf{M}(x-y)\| && \text{by definition of linear operators (Definition 14.3 page 214)}\\
&= \|\mathbf{M}u\| && \text{let } u \triangleq x - y\\
&= \|x - y\| && \text{by left hypothesis}
\end{aligned}$$

2. Proof that $\|\mathbf{M}x\| = \|x\| \impliedby \|\mathbf{M}x - \mathbf{M}y\| = \|x - y\|$:

$$\begin{aligned}
\|\mathbf{M}x\| &= \|\mathbf{M}(x - \mathbb{0})\|\\
&= \|\mathbf{M}x - \mathbf{M}\mathbb{0})\| && \text{by definition of linear operators (Definition 14.3 page 214)}\\
&= \|x - \mathbb{0}\| && \text{by right hypothesis}\\
&= \|x\|
\end{aligned}$$

✏

Isometric operators have already been defined (Definition 14.10 page 232) in the more general normed linear spaces, while Theorem 14.20 (page 232) demonstrated that in a normed linear space $\boldsymbol{X}$, $\|\mathbf{M}x\| = \|x\| \iff \|\mathbf{M}x - \mathbf{M}y\| = \|x - y\|$ for all $x, y \in \boldsymbol{X}$. Here in the more specialized inner product spaces, Theorem 14.21 (next) demonstrates two additional equivalent properties.

Theorem 14.21. [41] *Let $\mathcal{B}(\boldsymbol{X}, \boldsymbol{X})$ be the space of bounded linear operators on a normed linear space $\boldsymbol{X} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$. Let $\mathbf{N}$ be a bounded linear operator in $\mathcal{L}(\boldsymbol{X}, \boldsymbol{X})$, and $\mathbf{I}$ the identity operator in $\mathcal{L}(\boldsymbol{X}, \boldsymbol{X})$. Let $\|x\| \triangleq \sqrt{\langle x \mid x\rangle}$.*

THM

*The following conditions are all **equivalent**:*

$$\begin{array}{llcll l}
1. & \mathbf{M}^*\mathbf{M} & = & \mathbf{I} & & & \iff\\
2. & \langle \mathbf{M}x \mid \mathbf{M}y\rangle & = & \langle x \mid y\rangle & \forall x,y \in X & (\mathbf{M} \text{ is surjective}) & \iff\\
3. & \|\mathbf{M}x - \mathbf{M}y\| & = & \|x - y\| & \forall x,y \in X & (\text{isometric in distance}) & \iff\\
4. & \|\mathbf{M}x\| & = & \|x\| & \forall x \in X & (\text{isometric in length}) &
\end{array}$$

✎PROOF:

1. Proof that (1) $\implies$ (2):

$$\begin{aligned}
\langle \mathbf{M}x \mid \mathbf{M}y\rangle &= \langle x \mid \mathbf{M}^*\mathbf{M}y\rangle && \text{by Proposition 14.3 page 226 (definition of adjoint)}\\
&= \langle x \mid \mathbf{I}y\rangle && \text{by (1)}\\
&= \langle x \mid y\rangle && \text{by Definition 14.2 page 213 (definition of } \mathbf{I})
\end{aligned}$$

2. Proof that (2) $\implies$ (4):

$$\begin{aligned}
\|\mathbf{M}x\| &= \sqrt{\langle \mathbf{M}x \mid \mathbf{M}x\rangle} && \text{by definition of } \|\cdot\|\\
&= \sqrt{\langle x \mid x\rangle} && \text{by right hypothesis}\\
&= \|x\| && \text{by definition of } \|\cdot\|
\end{aligned}$$

3. Proof that (2) $\impliedby$ (4):

$$\begin{aligned}
4\langle \mathbf{M}x \mid \mathbf{M}y\rangle &= \|\mathbf{M}x + \mathbf{M}y\|^2 - \|\mathbf{M}x - \mathbf{M}y\|^2 + i\,\|\mathbf{M}x + i\mathbf{M}y\|^2 - i\,\|\mathbf{M}x - i\mathbf{M}y\|^2 && \text{by polarization id.}\\
&= \|\mathbf{M}(x+y)\|^2 - \|\mathbf{M}(x-y)\|^2 + i\,\|\mathbf{M}(x+iy)\|^2 - i\,\|\mathbf{M}(x-iy)\|^2 && \text{by Definition 14.3}\\
&= \|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2 && \text{by left hypothesis}
\end{aligned}$$

[41] Michel and Herget (1993) page 432 ⟨Theorem 7.5.8⟩, Kubrusly (2001) page 391 ⟨Proposition 5.72⟩

4. Proof that (3) $\iff$ (4): by Theorem 14.20 page 232

5. Proof that (4) $\implies$ (1):

$$\begin{aligned}
\langle \mathbf{M}^*\mathbf{M}x \mid x \rangle &= \langle \mathbf{M}x \mid \mathbf{M}^{**}x \rangle && \text{by Proposition 14.3 page 226 (definition of adjoint)}\\
&= \langle \mathbf{M}x \mid \mathbf{M}x \rangle && \text{by Theorem 14.13 page 227 (property of adjoint)}\\
&= \|\mathbf{M}x\|^2 && \text{by definition}\\
&= \|x\|^2 && \text{by left hypothesis with } y = 0\\
&= \langle x \mid x \rangle && \text{by definition}\\
&= \langle \mathbf{I}x \mid x \rangle && \text{by Definition 14.2 page 213 (definition of } \mathbf{I})\\
\implies \quad \mathbf{M}^*\mathbf{M} &= \mathbf{I} && \forall x \in X
\end{aligned}$$

Theorem 14.22. [42] *Let $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. Let $\mathbf{M}$ be a bounded linear operator in $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$, and $\mathbf{I}$ the identity operator in $\mathcal{L}(\boldsymbol{X}, \boldsymbol{X})$. Let Λ be the set of eigenvalues of $\mathbf{M}$. Let $\|x\| \triangleq \sqrt{\langle x \mid x \rangle}$.*

THM

$$\underbrace{\mathbf{M}^*\mathbf{M} = \mathbf{I}}_{\mathbf{M} \text{ is isometric}} \implies \left\{ \begin{array}{rcll} |||\mathbf{M}||| &=& 1 & \text{(UNIT LENGTH)} \quad \textit{and} \\ |\lambda| &=& 1 & \forall \lambda \in \Lambda \end{array} \right.$$

✎PROOF:

1. Proof that $\mathbf{M}^*\mathbf{M} = \mathbf{I} \implies |||\mathbf{M}||| = 1$:

$$\begin{aligned}
|||\mathbf{M}||| &= \sup_{x \in X} \{\|\mathbf{M}x\| \mid \|x\| = 1\} && \text{by Definition 14.4 page 218}\\
&= \sup_{x \in X} \{\|x\| \mid \|x\| = 1\} && \text{by Theorem 14.21 page 233}\\
&= \sup_{x \in X} \{1\}\\
&= 1
\end{aligned}$$

2. Proof that $|\lambda| = 1$: Let (x, λ) be an eigenvector-eigenvalue pair.

$$\begin{aligned}
1 &= \frac{1}{\|x\|} \|x\|\\
&= \frac{1}{\|x\|} \|\mathbf{M}x\| && \text{by Theorem 14.21 page 233}\\
&= \frac{1}{\|x\|} \|\lambda x\| && \text{by definition of } \lambda\\
&= \frac{1}{\|x\|} |\lambda| \ \|x\| && \text{by homogeneous property of } \|\cdot\|\\
&= |\lambda|
\end{aligned}$$

Example 14.4 (One sided shift operator). [43] Let $\boldsymbol{X}$ be the set of all sequences with range $\mathbb{W}$ (0, 1, 2, ...) and shift operators defined as

1. $\mathbf{S}_r\,(\!(x_0, x_1, x_2, \dots)\!) \triangleq (\!(0, x_0, x_1, x_2, \dots)\!)$ (right shift operator)
2. $\mathbf{S}_l\,(\!(x_0, x_1, x_2, \dots)\!) \triangleq (\!(x_1, x_2, x_3, \dots)\!)$ (left shift operator)

[42] Michel and Herget (1993) page 432
[43] Michel and Herget (1993) page 441

EX
1. $\mathbf{S}_r$ is an isometric operator.
2. $\mathbf{S}_r^* = \mathbf{S}_l$

✎PROOF:

1. Proof that $\mathbf{S}_r^* = \mathbf{S}_l$:

$$\begin{aligned}
\langle \mathbf{S}_r \left(\!\left(x_0, x_1, x_2, \ldots \right)\!\right) \mid \left(\!\left(y_0, y_1, y_2, \ldots \right)\!\right) \rangle &= \langle \left(\!\left(0, x_0, x_1, x_2, \ldots \right)\!\right) \mid \left(\!\left(y_0, y_1, y_2, \ldots \right)\!\right) \rangle \\
&= \sum_{n=1}^{\infty} x_{n-1}\ y_n^* \\
&= \sum_{n=0}^{\infty} x_n\ y_{n+1}^* \\
&= \sum_{n=0}^{\infty} x_n\ y_{n+1}^* \\
&= \langle \left(\!\left(x_0, x_1, x_2, \ldots \right)\!\right) \mid \left(\!\left(y_1, y_2, y_3, \ldots \right)\!\right) \rangle \\
&= \left\langle \left(\!\left(x_0, x_1, x_2, \ldots \right)\!\right) \mid \underbrace{\mathbf{S}_l}_{\mathbf{S}_r^*} \left(\!\left(y_0, y_1, y_2, \ldots \right)\!\right) \right\rangle
\end{aligned}$$

2. Proof that $\mathbf{S}_r$ is isometric ($\mathbf{S}_r^*\mathbf{S}_r = \mathbf{I}$):

$$\begin{aligned}
\mathbf{S}_r^*\mathbf{S}_r &= \mathbf{S}_l\mathbf{S}_r && \text{by 1.} \\
&= \mathbf{I}
\end{aligned}$$

✏

14.4.5 Unitary operators

Definition 14.11. [44] *Let $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. Let $\mathbf{U}$ be a bounded linear operator in $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$, and $\boldsymbol{I}$ the identity operator in $\mathcal{B}(\boldsymbol{X}, \boldsymbol{X})$.*

DEF
*The operator $\mathbf{U}$ is **unitary** if* $\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I}$.

Proposition 14.4. *Let $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ be the space of bounded linear operators on normed linear spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. Let $\mathbf{U}$ and $\mathbf{V}$ be bounded linear operators in $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$.*

PRP
$$\left.\begin{array}{l}\mathbf{U} \textit{ is } \text{UNITARY} \quad \textit{and} \\ \mathbf{V} \textit{ is } \text{UNITARY}\end{array}\right\} \implies (\mathbf{UV}) \textit{ is } \text{UNITARY}.$$

[44] Rudin (1991) page 312, Michel and Herget (1993) page 431, Autonne (1901) page 209, Autonne (1902), Schur (1909), Steen (1973)

✎PROOF:

$$\begin{aligned}
(\mathbf{UV})(\mathbf{UV})^* &= (\mathbf{UV})\left(\mathbf{V}^*\mathbf{U}^*\right) && \text{by Theorem 14.8 page 223}\\
&= \mathbf{U}\left(\mathbf{VV}^*\right)\mathbf{U}^* && \text{by associative property}\\
&= \mathbf{UIU}^* && \text{by definition of unitary operators—Definition 14.11 page 235}\\
&= \mathbf{I} && \text{by definition of unitary operators—Definition 14.11 page 235}
\end{aligned}$$

$$\begin{aligned}
(\mathbf{UV})^*(\mathbf{UV}) &= \left(\mathbf{V}^*\mathbf{U}^*\right)(\mathbf{UV}) && \text{by Theorem 14.8 page 223}\\
&= \mathbf{V}^*\left(\mathbf{U}^*\mathbf{U}\right)\mathbf{V} && \text{by associative property}\\
&= \mathbf{V}^*\mathbf{IV} && \text{by definition of unitary operators—Definition 14.11 page 235}\\
&= \mathbf{I} && \text{by definition of unitary operators—Definition 14.11 page 235}
\end{aligned}$$

⇨

Theorem 14.23. [45] *Let $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$ be the space of bounded linear operators on a Hilbert space $\boldsymbol{H}$. Let $\mathbf{U}$ be a bounded linear operator in $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$, and $\mathcal{I}(\mathbf{U})$ the* IMAGE SET *of $\mathbf{U}$.*

THM

*The following conditions are **equivalent**:*

1. $\mathbf{UU}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}$ *(unitary)* $\iff$
2. $\langle \mathbf{U}\boldsymbol{x} \mid \mathbf{U}\boldsymbol{y}\rangle = \langle \mathbf{U}^*\boldsymbol{x} \mid \mathbf{U}^*\boldsymbol{y}\rangle = \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle$ *and* $\mathcal{I}(\mathbf{U}) = X$ *(surjective)* $\iff$
3. $\|\mathbf{U}\boldsymbol{x} - \mathbf{U}\boldsymbol{y}\| = \left\|\mathbf{U}^*\boldsymbol{x} - \mathbf{U}^*\boldsymbol{y}\right\| = \|\boldsymbol{x}-\boldsymbol{y}\|$ *and* $\mathcal{I}(\mathbf{U}) = X$ *(isometric in distance)* $\iff$
4. $\|\mathbf{U}\boldsymbol{x}\| = \|\boldsymbol{x}\|$ *and* $\mathcal{I}(\mathbf{U}) = X$ *(isometric in length)*

✎PROOF:

1. Proof that (1) $\implies$ (2):
   (a) $\langle \mathbf{U}\boldsymbol{x} \mid \mathbf{U}\boldsymbol{y}\rangle = \langle \mathbf{U}^*\boldsymbol{x} \mid \mathbf{U}^*\boldsymbol{y}\rangle = \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle$ by Theorem 14.21 (page 233).
   (b) Proof that $\mathcal{I}(\mathbf{U}) = X$:

$$\begin{aligned}
X &\supseteq \mathcal{I}(\mathbf{U}) && \text{because } \mathbf{U} \in X^X\\
&\supseteq \mathcal{I}(\mathbf{UU}^*)\\
&= \mathcal{I}(\mathbf{I}) && \text{by left hypothesis } (\mathbf{U}^*\mathbf{U} = \mathbf{UU}^* = \mathbf{I})\\
&= X && \text{by Definition 14.2 page 213 (definition of } \mathbf{I})
\end{aligned}$$

2. Proof that (2) $\iff$ (3) $\iff$ (4): by Theorem 14.21 page 233.
3. Proof that (3) $\implies$ (1):
   (a) Proof that $\|\mathbf{U}\boldsymbol{x} - \mathbf{U}\boldsymbol{y}\| = \|\boldsymbol{x}-\boldsymbol{y}\| \implies \mathbf{U}^*\mathbf{U} = \mathbf{I}$: by Theorem 14.21 page 233
   (b) Proof that $\left\|\mathbf{U}^*\boldsymbol{x} - \mathbf{U}^*\boldsymbol{y}\right\| = \|\boldsymbol{x}-\boldsymbol{y}\| \implies \mathbf{UU}^* = \mathbf{I}$:

$$\begin{aligned}
\left\|\mathbf{U}^*\boldsymbol{x} - \mathbf{U}^*\boldsymbol{y}\right\| = \|\boldsymbol{x}-\boldsymbol{y}\| \implies\quad & \mathbf{U}^{**}\mathbf{U}^* = \mathbf{I} && \text{by Theorem 14.21 page 233}\\
& \mathbf{UU}^* = \mathbf{I} && \text{by Theorem 14.13 page 227}
\end{aligned}$$

⇨

Theorem 14.24. *Let $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$ be the space of bounded linear operators on a Hilbert space $\boldsymbol{H}$. Let $\mathbf{U}$ be a bounded linear operator in $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$, $\mathcal{N}(\mathbf{U})$ the* NULL SPACE *of $\mathbf{U}$, and $\mathcal{I}(\mathbf{U})$ the* IMAGE SET *of $\mathbf{U}$.*

THM

$$\underbrace{\mathbf{UU}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}}_{\mathbf{U}\text{ is unitary}} \implies \left\{\begin{array}{lllll}
\mathbf{U}^{-1} & = & \mathbf{U}^* & & \\
\mathcal{I}(\mathbf{U}) & = & \mathcal{I}(\mathbf{U}^*) & = & X\\
\mathcal{N}(\mathbf{U}) & = & \mathcal{N}(\mathbf{U}^*) & = & \{\mathbb{0}\}\\
\|\!|\mathbf{U}|\!\| & = & \|\!|\mathbf{U}^*|\!\| & = & 1 \quad (\text{UNIT LENGTH})
\end{array}\right.$$

[45] Rudin (1991) pages 313–314 ⟨Theorem 12.13⟩, Knapp (2005) page 45 ⟨Proposition 2.6⟩

✎PROOF:

1. Note that $\mathbf{U}$, $\mathbf{U}^*$, and $\mathbf{U}^{-1}$ are all both **isometric** and **normal**:

$$\begin{array}{lclcll}
\mathbf{U}^*\mathbf{U} & & & = & \mathbf{I} & \implies \mathbf{U} \text{ is isometric}\\
\mathbf{U}\mathbf{U}^* & = & \mathbf{U}^*\mathbf{U} & = & \mathbf{I} & \implies \mathbf{U}^* \text{ is isometric}\\
\mathbf{U}^{-1} & = & \mathbf{U}^* & & & \implies \mathbf{U}^{-1} \text{ is isometric}\\
\\
\mathbf{U}^*\mathbf{U} & = & \mathbf{U}\mathbf{U}^* & = & & \implies \mathbf{U} \text{ is normal}\\
\mathbf{U}\mathbf{U}^* & = & \mathbf{U}^*\mathbf{U} & = & & \implies \mathbf{U}^* \text{ is normal}\\
\mathbf{U}^{-1} & = & \mathbf{U}^* & & & \implies \mathbf{U}^{-1} \text{ is normal}
\end{array}$$

2. Proof that $\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I} \implies \mathcal{I}(\mathbf{U}) = \mathcal{I}(\mathbf{U}^*) = \boldsymbol{H}$: by Theorem 14.23 page 236.

3. Proof that $\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I} \implies \mathcal{N}(\mathbf{U}) = \mathcal{N}(\mathbf{U}^*) = \mathcal{N}(\mathbf{U}^{-1})$:

$$\begin{aligned}
\mathcal{N}(\mathbf{U}^*) = \mathcal{N}(\mathbf{U}) & && \text{because } \mathbf{U} \text{ and } \mathbf{U}^* \text{ are both normal and by Theorem 14.18 page 231}\\
&= \mathcal{I}(\mathbf{U})^\perp && \text{by Theorem 14.14 page 228}\\
&= X^\perp && \text{by above result}\\
&= \{\mathbb{0}\} && \text{by Proposition 7.6 page 117}
\end{aligned}$$

4. Proof that $\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I} \implies \|\mathbf{U}^{-1}\| = \|\mathbf{U}^*\| = \|\mathbf{U}\| = 1$:
Because $\mathbf{U}$, $\mathbf{U}^*$, and $\mathbf{U}^{-1}$ are all isometric and by Theorem 14.22 page 234.

✏

Example 14.5. Examples of *Fredholm integral operators* include

$$\begin{array}{llrclrcl}
1. & \text{Fourier Transform} & [\tilde{\mathbf{F}}\mathrm{x}](f) & = & \int_t \mathrm{x}(t)e^{-i2\pi ft}\,\mathrm{d}t & \kappa(t,f) & = & e^{-i2\pi ft}\\
2. & \text{Inverse Fourier Transform} & [\tilde{\mathbf{F}}^{-1}\tilde{\mathrm{x}}](t) & = & \int_f \tilde{\mathrm{x}}(f)e^{i2\pi ft}\,\mathrm{d}f & \kappa(f,t) & = & e^{i2\pi ft}\\
3. & \text{Laplace operator} & [\mathbf{L}\mathrm{x}](s) & = & \int_t \mathrm{x}(t)e^{-st}\,\mathrm{d}t & \kappa(t,s) & = & e^{-st}
\end{array}$$

Example 14.6 (Translation operator). Let $\boldsymbol{X} = \boldsymbol{L}^2_{\mathbb{R}}$ and $\mathbf{T} \in \boldsymbol{X}^{\boldsymbol{X}}$ be defined as

$$\mathbf{T}\mathrm{f}(x) \triangleq \mathrm{f}(x-1) \qquad \forall \mathrm{f} \in \boldsymbol{L}^2_{\mathbb{R}} \qquad \text{(translation operator)}$$

EX

$$\begin{array}{lrcll}
1. & \mathbf{T}^{-1}\mathrm{f}(x) & = & \mathrm{f}(x+1) \quad \forall \mathrm{f} \in \boldsymbol{L}^2_{\mathbb{R}} & \text{(inverse translation operator)}\\
2. & \mathbf{T}^* & = & \mathbf{T}^{-1} & (\mathbf{T} \text{ is invertible})\\
3. & \mathbf{T}^*\mathbf{T} & = & \mathbf{T}\mathbf{T}^* = \mathbf{I} & (\mathbf{T} \text{ is unitary})
\end{array}$$

✎PROOF:

1. Proof that $\mathbf{T}^{-1}\mathrm{f}(x) = \mathrm{f}(x+1)$:

$$\mathbf{T}^{-1}\mathbf{T} = \mathbf{I}$$
$$\mathbf{T}\mathbf{T}^{-1} = \mathbf{I}$$

2. Proof that $\mathbf{T}$ is unitary:

$$\begin{aligned}
\langle \mathbf{T}\mathrm{f}(x) \,|\, \mathrm{g}(x)\rangle &= \langle \mathrm{f}(x-1) \,|\, \mathrm{g}(x)\rangle && \text{by definition of } \mathbf{T}\\
&= \int_x \mathrm{f}(x-1)\mathrm{g}^*(x)\,\mathrm{d}x\\
&= \int_x \mathrm{f}(x)\mathrm{g}^*(x+1)\,\mathrm{d}x\\
&= \langle \mathrm{f}(x) \,|\, \mathrm{g}(x+1)\rangle\\
&= \Big\langle \mathrm{f}(x) \,\Big|\, \underbrace{\mathbf{T}^{-1}}_{\mathbf{T}^*}\mathrm{g}(x) \Big\rangle && \text{by 1.}
\end{aligned}$$

Example 14.7 (Dilation operator). Let $\boldsymbol{X} = \boldsymbol{L}^2_{\mathbb{R}}$ and $\mathbf{T} \in \boldsymbol{X}^{\boldsymbol{X}}$ be defined as

$$\mathbf{D}\mathsf{f}(x) \triangleq \sqrt{2}\mathsf{f}(2x) \qquad \forall \mathsf{f} \in \boldsymbol{L}^2_{\mathbb{R}} \qquad \text{(dilation operator)}$$

EX

1. $\mathbf{D}^{-1}\mathsf{f}(x) = \frac{1}{\sqrt{2}}\mathsf{f}\left(\frac{1}{2}x\right) \quad \forall \mathsf{f} \in \boldsymbol{L}^2_{\mathbb{R}}$ (inverse dilation operator)
2. $\mathbf{D}^* = \mathbf{D}^{-1}$ ($\mathbf{D}$ is invertible)
3. $\mathbf{D}^*\mathbf{D} = \mathbf{D}\mathbf{D}^* = \mathbf{I}$ ($\mathbf{D}$ is unitary)

✎PROOF:

1. Proof that $\mathbf{D}^{-1}\mathsf{f}(x) = \frac{1}{\sqrt{2}}\mathsf{f}\left(\frac{1}{2}x\right)$:

$$\mathbf{D}^{-1}\mathbf{D} = \mathbf{I}$$
$$\mathbf{D}\mathbf{D}^{-1} = \mathbf{I}$$

2. Proof that $\mathbf{D}$ is unitary:

$$\begin{aligned}
\langle \mathbf{D}\mathsf{f}(x) \,|\, \mathsf{g}(x)\rangle &= \left\langle \sqrt{2}\mathsf{f}(2x) \,|\, \mathsf{g}(x)\right\rangle && \text{by definition of } \mathbf{D}\\
&= \int_x \sqrt{2}\mathsf{f}(2x)\mathsf{g}^*(x)\,\mathrm{d}x\\
&= \int_{u\in\mathbb{R}} \sqrt{2}\mathsf{f}(u)\mathsf{g}^*\left(\frac{1}{2}u\right)\frac{1}{2}\,\mathrm{d}u && \text{let } u \triangleq 2x \implies \mathrm{d}x = \frac{1}{2}\,\mathrm{d}u\\
&= \int_{u\in\mathbb{R}} \mathsf{f}(u)\left[\frac{1}{\sqrt{2}}\mathsf{g}\left(\frac{1}{2}u\right)\right]^*\,\mathrm{d}u\\
&= \left\langle \mathsf{f}(x) \,|\, \frac{1}{\sqrt{2}}\mathsf{g}\left(\frac{1}{2}x\right)\right\rangle\\
&= \left\langle \mathsf{f}(x) \,|\, \underbrace{\mathbf{D}^{-1}}_{\mathbf{D}^*}\mathsf{g}(x)\right\rangle && \text{by 1.}
\end{aligned}$$

Example 14.8 (Delay operator). Let $\boldsymbol{X}$ be the set of all sequences and $\mathbf{D} \in \boldsymbol{X}^{\boldsymbol{X}}$ be a delay operator.

EX The delay operator $\mathbf{D}\,(\!(x_n)\!)_{n\in\mathbb{Z}} \triangleq (\!(x_{n-1})\!)_{n\in\mathbb{Z}}$ is unitary.

✎PROOF: The inverse $\mathbf{D}^{-1}$ of the delay operator $\mathbf{D}$ is

$$\mathbf{D}^{-1}\,(\!(x_n)\!)_{n\in\mathbb{Z}} \triangleq (\!(x_{n+1})\!)_{n\in\mathbb{Z}}.$$

$$\begin{aligned}
\langle \mathbf{D}\,(\!(x_n)\!) \,|\, (\!(y_n)\!)\rangle &= \langle (\!(x_{n-1})\!) \,|\, (\!(y_n)\!)\rangle && \text{by definition of } \mathbf{D}\\
&= \sum_n x_{n-1}\, y_n^*\\
&= \sum_n x_n\, y_{n+1}^*\\
&= \left\langle (\!(x_n)\!) \,|\, (\!(y_{n+1})\!)\right\rangle\\
&= \left\langle (\!(x_n)\!) \,|\, \underbrace{\mathbf{D}^{-1}}_{\mathbf{D}^*}(\!(y_n)\!)\right\rangle
\end{aligned}$$

Therefore, $\mathbf{D}^* = \mathbf{D}^{-1}$. This implies that $\mathbf{D}\mathbf{D}^* = \mathbf{D}^*\mathbf{D} = \mathbf{I}$ which implies that $\mathbf{D}$ is unitary.

Example 14.9 (Fourier transform). Let $\tilde{\mathbf{F}}$ be the *Fourier Transform* and $\tilde{\mathbf{F}}^{-1}$ the *inverse Fourier Transform* operator

$$[\tilde{\mathbf{F}}\mathrm{x}](f) \triangleq \int_t \mathrm{x}(t)\underbrace{e^{-i2\pi ft}}_{\kappa(t,f)}\,\mathrm{d}t \qquad\qquad \left[\tilde{\mathbf{F}}^{-1}\tilde{\mathrm{x}}\right](t) \triangleq \int_f \tilde{\mathrm{x}}(f)\underbrace{e^{i2\pi ft}}_{\kappa^*(t,f)}\,\mathrm{d}f.$$

EX $\tilde{\mathbf{F}}^* = \tilde{\mathbf{F}}^{-1}$ (the Fourier Transform operator $\tilde{\mathbf{F}}$ is unitary)

✎PROOF:

$$\begin{aligned}
\langle \tilde{\mathbf{F}}\mathrm{x} \mid \tilde{\mathrm{y}} \rangle &= \left\langle \int_t \mathrm{x}(t)e^{-i2\pi ft}\,\mathrm{d}t \mid \tilde{\mathrm{y}}(f) \right\rangle \\
&= \int_t \mathrm{x}(t)\left\langle e^{-i2\pi ft} \mid \tilde{\mathrm{y}}(f) \right\rangle\,\mathrm{d}t \\
&= \int_t \mathrm{x}(t)\int_f e^{-i2\pi ft}\tilde{\mathrm{y}}^*(f)\,\mathrm{d}f\,\mathrm{d}t \\
&= \int_t \mathrm{x}(t)\left[\int_f e^{i2\pi ft}\tilde{\mathrm{y}}(f)\,\mathrm{d}f\right]^*\mathrm{d}t \\
&= \left\langle \mathrm{x}(t) \mid \int_f \tilde{\mathrm{y}}(f)e^{i2\pi ft}\,\mathrm{d}f \right\rangle \\
&= \left\langle \mathrm{x} \mid \underbrace{\tilde{\mathbf{F}}^{-1}}_{\tilde{\mathbf{F}}^*}\tilde{\mathrm{y}} \right\rangle
\end{aligned}$$

This implies that $\tilde{\mathbf{F}}$ is unitary ($\tilde{\mathbf{F}}^* = \tilde{\mathbf{F}}^{-1}$). ⇨

Example 14.10 (Rotation matrix). [46] Let the rotation matrix $\mathbf{R}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be defined as

$$\mathbf{R}_\theta \triangleq \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

EX
1. $\mathbf{R}^{-1}{}_\theta = \mathbf{R}_{-\theta}$
2. $\mathbf{R}^*{}_\theta = \mathbf{R}^{-1}{}_\theta$ ($\mathbf{R}$ is unitary)

✎PROOF:

$$\begin{aligned}
\mathbf{R}^* &= \mathbf{R}^H \\
&= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^H && \text{by definition of } \mathbf{R} \\
&= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} && \text{by definition of Hermetian transpose operator } H \\
&= \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} && \text{by Theorem ?? page ??} \\
&= \mathbf{R}_{-\theta} && \text{by definition of } \mathbf{R} \\
&= \mathbf{R}^{-1} && \text{by 1.}
\end{aligned}$$

⇨

[46] Noble and Daniel (1988), page 311

14.5 Operator order

Definition 14.12. [47] *Let* $\mathbf{P} \in \boldsymbol{Y}^{\boldsymbol{X}}$ *be an operator.*

DEF

$\mathbf{P}$ *is **positive** if* $\langle \mathbf{P}\boldsymbol{x} \mid \boldsymbol{x} \rangle \geq 0 \; \forall \boldsymbol{x} \in \boldsymbol{X}$.
This condition is denoted $\mathbf{P} \geq 0$.

Theorem 14.25. [48]

THM

$$\underbrace{\mathbf{P} \geq 0 \text{ and } \mathbf{Q} \geq 0}_{\mathbf{P} \text{ and } \mathbf{Q} \text{ are both positive}} \implies \left\{\begin{array}{llll} (\mathbf{P}+\mathbf{Q}) & \geq 0 & & ((\mathbf{P}+\mathbf{Q}) \text{ is positive}) \\ \mathbf{A}^*\mathbf{P}\mathbf{A} & \geq 0 & \forall \mathbf{A} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{X}) & (\mathbf{A}^*\mathbf{P}\mathbf{A} \text{ is positive}) \\ \mathbf{A}^*\mathbf{A} & \geq 0 & \forall \mathbf{A} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{X}) & (\mathbf{A}^*\mathbf{A} \text{ is positive}) \end{array}\right.$$

✎PROOF:

$$\begin{aligned}
\langle (\mathbf{P}+\mathbf{Q})\boldsymbol{x} \mid \boldsymbol{x} \rangle &= \langle \mathbf{P}\boldsymbol{x} \mid \boldsymbol{x} \rangle + \langle \mathbf{Q}\boldsymbol{x} \mid \boldsymbol{x} \rangle && \text{by additive property of } \langle \triangle \mid \nabla \rangle \text{ (Definition 6.1 page 95)} \\
&\geq \langle \mathbf{P}\boldsymbol{x} \mid \boldsymbol{x} \rangle && \text{by left hypothesis} \\
&\geq 0 && \text{by left hypothesis} \\
\langle \mathbf{A}^*\mathbf{P}\mathbf{A}\boldsymbol{x} \mid \boldsymbol{x} \rangle &= \langle \mathbf{P}\mathbf{A}\boldsymbol{x} \mid \mathbf{A}\boldsymbol{x} \rangle && \text{by definition of adjoint (Proposition 14.3 page 226)} \\
&= \langle \mathbf{P}\boldsymbol{y} \mid \boldsymbol{y} \rangle && \text{where } \boldsymbol{y} \triangleq \mathbf{A}\boldsymbol{x} \\
&\geq 0 && \text{by left hypothesis} \\
\langle \mathbf{I}\boldsymbol{x} \mid \boldsymbol{x} \rangle &= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle && \text{by definition of } \mathbf{I} \text{ (Definition 14.2 page 213)} \\
&\geq 0 && \text{by non-negative property of } \langle \triangle \mid \nabla \rangle \text{ (Definition 6.1 page 95)} \\
&\implies \mathbf{I} \text{ is positive} \\
\langle \mathbf{A}^*\mathbf{A}\boldsymbol{x} \mid \boldsymbol{x} \rangle &= \langle \mathbf{A}^*\mathbf{I}\mathbf{A}\boldsymbol{x} \mid \boldsymbol{x} \rangle && \text{by definition of } \mathbf{I} \text{ (Definition 14.2 page 213)} \\
&\geq 0 && \text{by two previous results}
\end{aligned}$$

✏

Definition 14.13. [49] *Let* $\mathbf{A}, \mathbf{B} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ *be bounded operators.*

DEF

$$\underbrace{\mathbf{A} \geq \mathbf{B}}_{\text{"}\mathbf{A} \text{ is greater than or equal to } \mathbf{B}\text{"}} \iff \underbrace{(\mathbf{A}-\mathbf{B}) \geq 0}_{\text{"}(\mathbf{A}-\mathbf{B}) \text{ is positive"}}$$

[47] Michel and Herget (1993) page 429 ⟨Definition 7.4.12⟩
[48] Michel and Herget (1993) page 429
[49] Michel and Herget (1993) page 429

Part V

Structure of Spaces

CHAPTER 15

ORTHOCOMPLEMENTED LATTICES

Orthocomplemented lattices (Definition 15.1 page 244) are a kind of generalization of *Boolean algebras*. The relationship between lattices of several types, including orthocomplemented and Boolean lattices, is stated in Theorem 15.7 (page 255) and illustrated in Figure 15.1 (page 243).

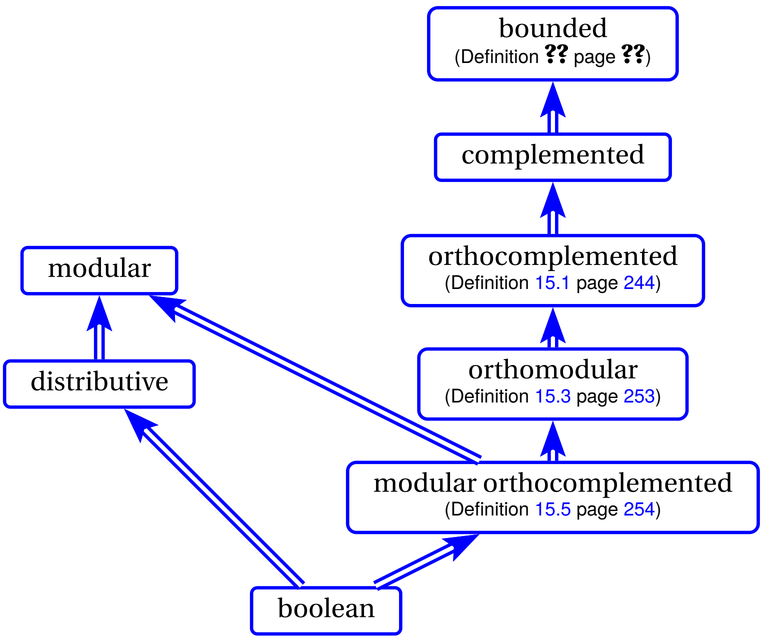


Figure 15.1: lattice of orthocomplemented lattices

15.1 Orthocomplemented Lattices

15.1.1 Definition

Definition 15.1. [1] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ *be a* BOUNDED LATTICE *(Definition* **??** *page* **??***).*

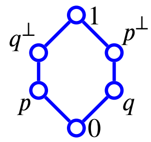
DEF

An element $x^\perp \in X$ *is an* ***orthocomplement*** *of an element* $x \in X$ *if*

1. $x^{\perp\perp} = x$ (INVOLUTORY) *and*
2. $x \wedge x^\perp = 0$ (NON-CONTRADICTION) *and*
3. $x \leq y \implies y^\perp \leq x^\perp \quad \forall y \in X$ (ANTITONE).

The LATTICE $\mathbf{L}$ *is* ***orthocomplemented*** *($\mathbf{L}$ is an* ***orthocomplemented lattice****) if every element x in X has an* ORTHOCOMPLEMENT $x^\perp$ *in* X.

Definition 15.2. [2]



DEF

The $\mathbf{O_6}$ ***lattice*** *is the ordered set* $\left(\{0, p, q, p^\perp, q^\perp, 1\}, \leq\right)$ *with cover relation* $\lessdot = \left\{(0,p), (0,q), (p,q^\perp), (q,p^\perp), (p^\perp,1), (q^\perp,1)\right\}$.

The O_6 *lattice is illustrated by the Hasse diagram to the right.*

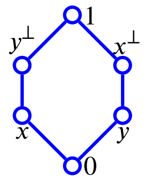
Example 15.1. [3]

EX

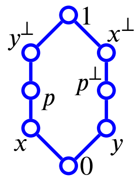
The O_6 **lattice** (Definition 15.2 page 244) is an **orthocomplemented lattice** (Definition 15.1 page 244).

Example 15.2. [4]There are a total of 10 **orthocomplemented lattices** with 8 elements or less. These 10, along with 3 other orthocomplemented lattices with 10 elements, are illustrated next:

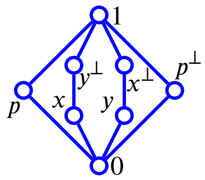
Lattices that are **orthocomplemented** but *non-orthomodular* and hence also *not modular orthocomplemented* and *non-Boolean*:



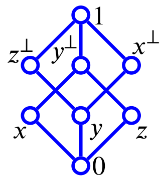
1. O_6 *lattice*



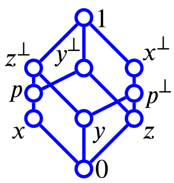
2. O_8 *lattice*



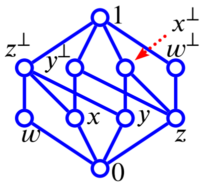
3.



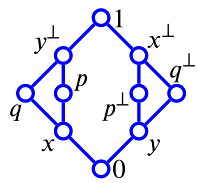
4.



5.



6.



7.

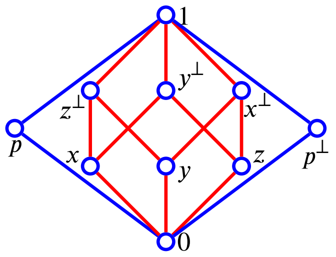
Lattices that are **orthocomplemented** and **orthomodular** but *not modular orthocomplemented* and hence also *non-Boolean*:

[1] Stern (1999) page 11, Beran (1985) page 28, Kalmbach (1983) page 16, Gudder (1988) page 76, Loomis (1955) page 3, Birkhoff and Neumann (1936) page 830 ⟨L71–L73⟩

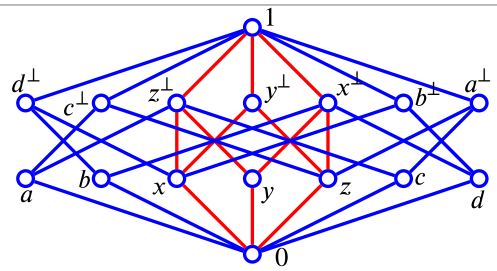
[2] Kalmbach (1983) page 22, Holland (1970), page 50, Beran (1985) page 33, Stern (1999) page 12, The O_6 *lattice* is also called the **Benzene ring** or the **hexagon**.

[3] Holland (1963), page 50

[4] Beran (1985) pages 33–42, Maeda (1966) page 250, Kalmbach (1983) page 24 ⟨Figure 3.2⟩, Stern (1999) page 12, Holland (1970), page 50

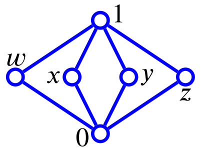


8.

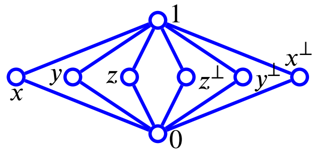


9.

Lattices that are **orthocomplemented**, **orthomodular**, and **modular orthocomplemented** but *non-Boolean*:



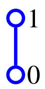
10. M_4 *lattice*



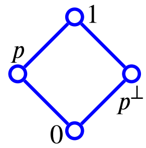
11. M_6 *lattice*

Lattices that are **orthocomplemented**, **orthomodular**, **modular orthocomplemented** and **Boolean**:

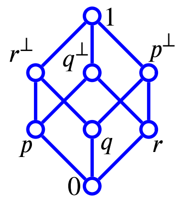
12. L_1 *lattice*



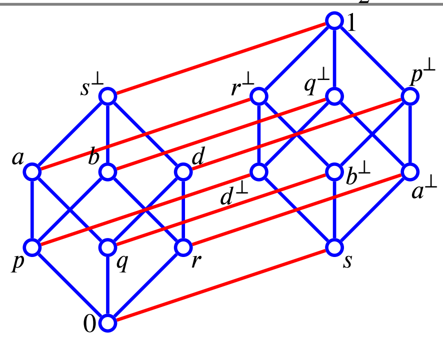
13. L_2 *lattice*



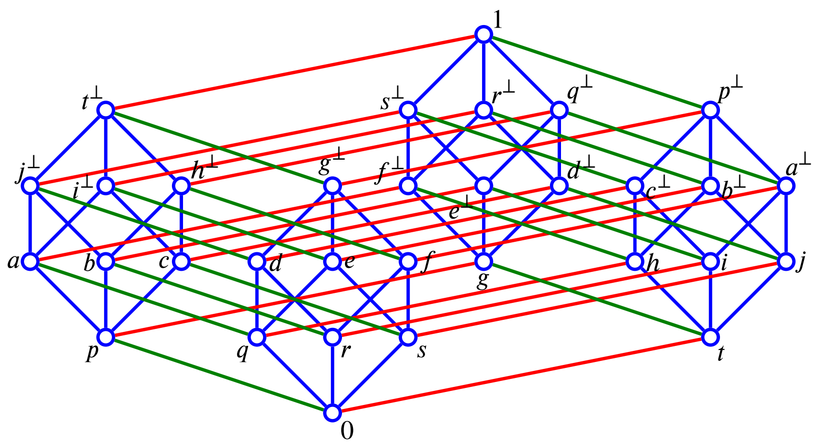
14. L_2^2 *lattice*



15. L_2^3 *lattice*



16. L_2^4 *lattice*



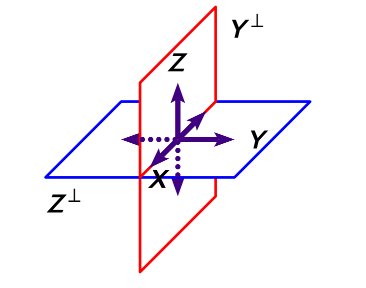
17. L_2^5 *lattice*

Example 15.3.

EX

The structure $\left(2^{\mathbb{R}^N}, +, \cap, \varnothing, \boldsymbol{H}\,;\subseteq\right)$ is an **orthocomplemented lattice** where

- $\mathbb{R}^N$ is an **Euclidean space** with dimension N
- $2^{\mathbb{R}^N}$ is the set of all subspaces of $\mathbb{R}^N$
- $\boldsymbol{V}+\boldsymbol{W}$ is the *Minkowski sum* of subspaces $\boldsymbol{V}$ and $\boldsymbol{W}$
- $\boldsymbol{V}\cap\boldsymbol{W}$ is the *intersection* of subspaces $\boldsymbol{V}$ and $\boldsymbol{W}$



Example 15.4.

EX

The structure $\left(2^{\boldsymbol{H}}, \oplus, \cap, \varnothing, \boldsymbol{H}\,;\subseteq\right)$ is an **orthocomplemented lattice** where

- $\boldsymbol{H}$ is a **Hilbert space**
- $2^{\boldsymbol{H}}$ is the set of all closed subspaces of $\boldsymbol{H}$
- $\boldsymbol{X}+\boldsymbol{Y}$ is the *Minkowski sum* of subspaces $\boldsymbol{X}$ and $\boldsymbol{Y}$
- $\boldsymbol{X}\oplus\boldsymbol{Y}\triangleq(\boldsymbol{X}+\boldsymbol{Y})^{-}$ is the *closure* of $\boldsymbol{X}+\boldsymbol{Y}$
- $\boldsymbol{X}\cap\boldsymbol{Y}$ is the *intersection* of subspaces $\boldsymbol{X}$ and $\boldsymbol{Y}$

15.1.2 Properties

Theorem 15.1. [5] *Let* $\boldsymbol{L}\triangleq(X, \vee, \wedge, 0, 1\,;\leq)$ *be a* BOUNDED LATTICE.

THM

$$\left\{\begin{array}{l}\boldsymbol{L}\text{ is}\\ \text{ORTHOCOMPLEMENTED}\\ \text{(Definition 15.1 page 244)}\end{array}\right\} \implies \left\{\begin{array}{lrcll} (1). & 0^{\perp} &=& 1 & & \text{(BOUNDARY CONDITION) and}\\ (2). & 1^{\perp} &=& 0 & & \text{(BOUNDARY CONDITION) and}\\ (3). & (x\vee y)^{\perp} &=& x^{\perp}\wedge y^{\perp} & \forall x,y\in X & \text{(DISJUNCTIVE DE MORGAN) and}\\ (4). & (x\wedge y)^{\perp} &=& x^{\perp}\vee y^{\perp} & \forall x,y\in X & \text{(CONJUNCTIVE DE MORGAN) and}\\ (5). & x\vee x^{\perp} &=& 1 & \forall x\in X & \text{(EXCLUDED MIDDLE).}\end{array}\right.$$

✎PROOF: Let $x^{\perp}\triangleq\neg x$, where $\neg$ is an *ortho negation* function (Definition D.3 page 322). Then, this theorem follows directly from Theorem D.5 (page 326). ✑

Corollary 15.1. *Let* $\boldsymbol{L}\triangleq(X, \vee, \wedge, 0, 1\,;\leq)$ *be a* BOUNDED LATTICE *(Definition* **??** *page* **??***).*

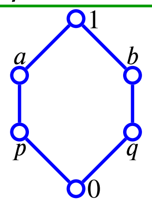
COR

$$\left\{\begin{array}{l}\boldsymbol{L}\textit{ is orthocomplemented}\\ \text{(Definition 15.1 page 244)}\end{array}\right\} \implies \left\{\boldsymbol{L}\textit{ is complemented}\right\}$$

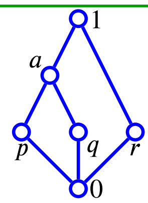
✎PROOF: This follows directly from the definition of *orthocomplemented lattices* (Definition 15.1 page 244) and *complemented lattices*. ✑

Example 15.5.

EX



The O_6 *lattice* (Definition 15.2 page 244) illustrated to the left is both **orthocomplemented** (Definition 15.1 page 244) and **multiply complemented**. The lattice illustrated to the right is **multiply complemented**, but is **non-orthocomplemented**.



✎PROOF:

1. Proof that O_6 *lattice* is multiply complemented: b and q are both *complements* of p.

[5] Beran (1985) pages 30–31, Birkhoff and Neumann (1936) page 830 ⟨L74⟩, Cohen (1989) page 37 ⟨3B.13. Theorem⟩

2. Proof that the right side lattice is multiply complemented: a, p, and q are all *complements* of r.

⇒

Lemma 15.1 (next) is useful in proving that *de Morgan*'s laws (Theorem A.6 page 278) hold in orthocomplemented lattices (Theorem 15.1 page 246) and in proving the characterization of Theorem 15.2 (page 248).

Lemma 15.1. [6] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge, 0, 1\,;\, \leq)$ *be an* ORTHOCOMPLEMENTED LATTICE *(Definition 15.1 page 244).*

LEM

$$\underbrace{x \leq y \implies y^\perp \leq x^\perp}_{\text{ANTITONE}} \iff \underbrace{\left\{\begin{array}{lcll} (x \vee y)^\perp &=& x^\perp \wedge y^\perp & {\scriptstyle x,y \in X} \quad \textit{and} \\ (x \wedge y)^\perp &=& x^\perp \vee y^\perp & {\scriptstyle x,y \in X} \end{array}\right.}_{\text{DE MORGAN}}$$

✎PROOF: This follows directly from Lemma D.2 (page 324). ⇒

Lemma 15.2. *Let* $\mathbf{L} \triangleq (X, \vee, \wedge, 0, 1\,;\, \leq)$ *be an* ORTHOCOMPLEMENTED LATTICE *(Definition 15.1 page 244).*

LEM

The set $\{0, x, x^\perp\}$ *is* DISTRIBUTIVE *(Definition ?? page ??) for all* $x \in X$.

✎PROOF:

$0 \wedge (x \vee x^\perp) = 0$	by *lower bounded* property	(Proposition ?? page ??)
$= 0 \vee 0$	by *join identity*	(Proposition ?? page ??)
$= (0 \wedge x) \vee (0 \wedge x^\perp)$	by *lower bounded* property	(Proposition ?? page ??)
$0 \wedge (x^\perp \vee x) = 0$	by *lower bounded* property	(Proposition ?? page ??)
$= 0 \vee 0$	by *join identity*	(Proposition ?? page ??)
$= (0 \wedge x^\perp) \vee (0 \wedge x)$	by *lower bounded* property	(Proposition ?? page ??)
$x \wedge (x^\perp \vee 0) = x \wedge x^\perp$	by *join identity*	(Proposition ?? page ??)
$= 0$	by *non-contradiction* property	(Definition 15.1 page 244)
$= 0 \vee 0$	by *join identity*	(Proposition ?? page ??)
$= (x \wedge x^\perp) \vee 0$	by *non-contradiction* property	(Definition 15.1 page 244)
$= (x \wedge x^\perp) \vee (x \wedge 0)$	by *lower bounded* property	(Proposition ?? page ??)
$x \wedge (0 \vee x^\perp) = x \wedge (x^\perp \vee 0)$	by *commutative* property of *lattices*	(Theorem C.3 page 306)
$= (x \wedge x^\perp) \vee (x \wedge 0)$	by previous result	
$= (x \wedge 0) \vee (x \wedge x^\perp)$	by *commutative* property of *lattices*	(Theorem C.3 page 306)
$x^\perp \wedge (x \vee 0) = (x^\perp \wedge x) \vee (x^\perp \wedge 0)$	by $x \wedge (x^\perp \vee 0)$ result	
$x^\perp \wedge (0 \vee x) = (x^\perp \wedge 0) \vee (x^\perp \wedge x)$	by $x \wedge (0 \vee x^\perp)$ result	

⇒

[6] Beran (1985) pages 30–31, Fáy (1967) ⟨cf Beran 1985 page 30⟩, Nakano and Romberger (1971) ⟨cf Beran 1985⟩

15.1.3 Characterization

Theorem 15.2. [7] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge, 0, 1 ; \leq)$ *be an algebraic structure.*

THM $\left.\begin{array}{l}\mathbf{L} \textit{ is an} \\ \textit{orthocomplemented lattice}\end{array}\right\} \iff \left\{\begin{array}{llcll} 1. & (z^\perp \wedge y^\perp)^\perp \vee x & = & (x \vee y) \vee z & \forall x,y,z \in X \quad \textit{and} \\ 2. & x \wedge (x \vee y) & = & x & \forall x,y \in X \quad \textit{and} \\ 3. & x \vee (y \wedge y^\perp) & = & x & \forall x,y \in X. \end{array}\right.$

✎PROOF:

1. Proof that orthocomplemented lattice $\implies$ 3 properties:

$$\begin{aligned}
(z^\perp \wedge y^\perp)^\perp \vee x &= \left[(z^\perp)^\perp \vee (y^\perp)^\perp\right] \vee x && \text{by } \textit{de Morgan} \text{ property (Theorem 15.1 page 246)}\\
&= (z \vee y) \vee x && \text{by } \textit{involutory} \text{ property (Definition 15.1 page 244)}\\
&= x \vee (z \vee y) && \text{by } \textit{commutative} \text{ property (Theorem C.3 page 306)}\\
&= x \vee (y \vee z) && \text{by } \textit{commutative} \text{ property (Theorem C.3 page 306)}\\
&= (x \vee y) \vee z && \text{by } \textit{associative} \text{ property (Theorem C.3 page 306)}
\end{aligned}$$

$$x \wedge (x \vee y) = x \qquad \text{by } \textit{absorptive} \text{ property (Theorem C.3 page 306)}$$

$$\begin{aligned}
x \vee (y \wedge y^\perp) &= x \vee 0 && \text{by } \textit{complemented} \text{ property (Definition 15.1 page 244)}\\
&= x
\end{aligned}$$

2. Proof that orthocomplemented lattice $\impliedby$ 3 properties:

   (a) Proof that $\mathbf{L}$ is *meet-idempotent*:

$$\begin{aligned}
x \wedge x &= x \wedge \left[x \vee (y \wedge y^\perp)\right] && \text{by (3)}\\
&= x \wedge \left[x \vee (y \wedge y^\perp)\right] && \text{by (3)}\\
&= x && \text{by (2)}
\end{aligned}$$

   (b) Define $0 \triangleq xx^\perp$ for some $x \in X$. Proof that 0 is the *greatest lower bound* of $\mathbf{L}$: The element 0 is the greatest lower bound if and only if $xx^\perp = yy^\perp \quad \forall x, y \in X \ldots$

      i. Proof that $(xx^\perp)^{\perp\perp} = (xx^\perp) \quad \forall x \in X$:

$$\begin{aligned}
(xx^\perp)^{\perp\perp} &= (xx^\perp)^{\perp\perp} + (xx^\perp) && \text{by (3)}\\
&= \left[(xx^\perp)^\perp (xx^\perp)^\perp\right]^\perp + (xx^\perp) && \text{by item (2a)}\\
&= \left[(xx^\perp) + (xx^\perp)\right] + (xx^\perp) && \text{by (1)}\\
&= \left[(xx^\perp)\right] + (xx^\perp) && \text{by (3)}\\
&= (xx^\perp) && \text{by (3)}
\end{aligned}$$

      ii. Proof that $a = (xx^\perp) + a \quad \forall a, x \in X$:

$$\begin{aligned}
a &= a + (xx^\perp) && \text{by (3)}\\
&= \left[a + (xx^\perp)\right] + (xx^\perp) && \text{by (3)}\\
&= \left[(xx^\perp)^\perp (xx^\perp)^\perp\right]^\perp + a && \text{by (1)}\\
&= \left[(xx^\perp)^\perp\right]^\perp + a && \text{by item (2a)}\\
&= (xx^\perp) + a && \text{by item (2(b)i)}
\end{aligned}$$

[7] Beran (1985) pages 31–33, Beran (1976) pages 251–252

iii. Proof that $(xx^\perp) = (yy^\perp) \quad \forall x, y \in X$:

$$\begin{aligned}
(xx^\perp) &= (xx^\perp) + (yy^\perp) && \text{by (3)}\\
&= (yy^\perp) && \text{by item (2(b)ii)}
\end{aligned}$$

(c) Proof that $x + 0 = 0 + x = x \quad \forall x \in X$ (*join identity*):

$$\begin{aligned}
x + 0 &= x + (yy^\perp) && \text{by item (2(b)iii)}\\
&= x && \text{by (3)}\\
0 + x &= (uu^\perp) + x && \text{by item (2(b)iii)}\\
&= x && \text{by item (2(b)ii)}
\end{aligned}$$

(d) Proof that $x + y = (y^\perp x^\perp)^\perp \quad \forall x, y \in X$:

$$\begin{aligned}
(y^\perp x^\perp)^\perp &= (y^\perp x^\perp)^\perp + 0 && \text{by item (2c)}\\
&= (0 + x) + y && \text{by (1)}\\
&= x + y && \text{by item (2c)}
\end{aligned}$$

(e) Proof that $x + x = x^{\perp\perp} \quad \forall x \in X$:

$$\begin{aligned}
x + x &= (x^\perp x^\perp)^\perp && \text{by item (2d)}\\
&= (x^\perp)^\perp && \text{by item (2a)}
\end{aligned}$$

(f) Proof that $x + y = y + x \quad \forall x, y \in X$ (*join-commutative*):

$$\begin{aligned}
x + y &= (x + 0) + y && \text{by item (2c)}\\
&= (0^\perp x^\perp)^\perp + y && \text{by item (2d)}\\
&= (y + x) + 0 && \text{by (1)}\\
&= y + x && \text{by item (2c)}
\end{aligned}$$

(g) Proof that $(x + y) + z = x + (y + z) \quad \forall x, y, z \in X$ (*join-associative*):

$$\begin{aligned}
(x + y) + z &= (z^\perp y^\perp)^\perp + x && \text{by (1)}\\
&= (y + z) + x && \text{by item (2d)}\\
&= x + (y + z) && \text{by item (2f)}
\end{aligned}$$

(h) Proof that $x^{\perp\perp} = x \quad \forall x \in X$ (*involutory*):

$$\begin{aligned}
x^{\perp\perp} &= \left(x^\perp\right)\perp && \text{by definition of } x^{\perp\perp}\\
&= \left[x^\perp\left(x^\perp + x\right)\right]\perp && \text{by (2)}\\
&= \left[x^\perp\left(x^\perp x^{\perp\perp}\right)^\perp\right]\perp && \text{by item (2d)}\\
&= \left(x^\perp x^{\perp\perp}\right) + x && \text{by item (2d)}\\
&= (0) + x && \text{by item (2b)}\\
&= x && \text{by item (2c)}
\end{aligned}$$

(i) Proof of *de Morgan*'s laws:

$$\begin{aligned}
(x + y)^\perp &= (y + x)^\perp && \text{by item (2g)}\\
&= \left[\left(x^\perp y^\perp\right)^\perp\right]^\perp && \text{by item (2d)}\\
&= x^\perp y^\perp && \text{by item (2h)}
\end{aligned}$$

$$\begin{aligned}
(xy)^\perp &= \left(x^{\perp\perp} y^{\perp\perp}\right)^\perp && \text{by item (2h)}\\
&= y^\perp + x^\perp && \text{by item (2d)}\\
&= x^\perp + y^\perp && \text{by item (2g)}
\end{aligned}$$

(j) Proof that $(xy)z = x(yz) \quad \forall x, y, z \in X \quad$ (*meet-commutative*):

$$\begin{aligned}
xy &= (xy)^{\perp\perp} && \text{by item (2h)}\\
&= \left(x^{\perp} + y^{\perp}\right)^{\perp} && \text{by item (2i)}\\
&= \left(y^{\perp} + x^{\perp}\right)^{\perp} && \text{by item (2g)}\\
&= y^{\perp\perp}x^{\perp\perp} && \text{by item (2i)}\\
&= yx && \text{by item (2i)}
\end{aligned}$$

(k) Proof that $(xy)z = x(yz) \quad \forall x, y, z \in X \quad$ (*meet-associative*):

$$\begin{aligned}
(xy)z &= [(xy)z]^{\perp} \perp && \text{by item (2h)}\\
&= \left[(xy)^{\perp} + z^{\perp}\right]^{\perp} && \text{by item (2i)}\\
&= \left[\left(x^{\perp} + y^{\perp}\right) + z^{\perp}\right]^{\perp} && \text{by item (2i)}\\
&= \left[x^{\perp} + \left(y^{\perp} + z^{\perp}\right)\right]^{\perp} && \text{by item (2g)}\\
&= x^{\perp\perp}\left(y^{\perp} + z^{\perp}\right)^{\perp} && \text{by item (2i)}\\
&= x^{\perp\perp}\left(y^{\perp\perp}z^{\perp\perp}\right) && \text{by item (2i)}\\
&= x(yz) && \text{by item (2h)}
\end{aligned}$$

(l) Proof that $x + (xz) = x$ (*join-meet-absorptive*):

$$\begin{aligned}
x \vee (xz) &= [x + (xz)]^{\perp\perp} && \text{by item (2h)}\\
&= \left[x^{\perp}(xz)^{\perp}\right]^{\perp} && \text{by item (2i)}\\
&= \left[x^{\perp}\left(x^{\perp} + z^{\perp}\right)\right]^{\perp} && \text{by item (2i)}\\
&= \left[x^{\perp}\right]^{\perp} && \text{by (2)}\\
&= x && \text{by item (2h)}
\end{aligned}$$

(m) Because $\boldsymbol{L}$ is *commutative* (item (2f) and item (2j)), *associative* (item (2g) and item (2k)), and *absorptive* ((2) and item (2l)), and by Theorem C.8 (page 314), $\boldsymbol{L}$ is a *lattice*.

(n) Define $1 \triangleq x + x^{\perp}$ for some $x \in X$. Proof that 1 is the *least upper bound* of $\boldsymbol{L}$: The element 1 is the least upper bound if and only if $x + x^{\perp} = y + y^{\perp} \quad \forall x, y \in X\ldots$

$$\begin{aligned}
1 &= \left(x + x^{\perp}\right) && \text{by definition of 1}\\
&= \left(x + x^{\perp}\right)^{\perp\perp} && \text{by item (2h)}\\
&= \left(x^{\perp}x\right)^{\perp} && \text{by item (2h)}\\
&= \left(xx^{\perp}\right)^{\perp} && \text{by item (2j)}\\
&= \left(yy^{\perp}\right)^{\perp} && \text{by item (2(b)iii)}\\
&= y^{\perp} + y^{\perp\perp} && \text{by item (2i)}\\
&= y^{\perp} + y && \text{by item (2h)}\\
&= y + y^{\perp} && \text{by item (2f)}
\end{aligned}$$

(o) Proof that $\boldsymbol{L}$ is *antitone*: by Theorem D.4 (page 326).

(p) Proof that $\boldsymbol{L}$ is *complemented*: by item (2(b)iii) and item (2n).

(q) Because $\boldsymbol{L}$ is a *bounded* (item (2b) and item (2n)) lattice (item (2m)), and because $\boldsymbol{L}$ is *complemented* (item (2p)), is *involutory* (item (2h)), and is *antitone* (item (2o)), and by Definition 15.1 (page 244), $\boldsymbol{L}$ is an *orthocomplemented lattice*.

15.1.4 Restrictions resulting in Boolean algebras

Proposition 15.1. [8] *Let* $\mathbf{L} = (X, \vee, \wedge, 0, 1; \leq)$ *be a* LATTICE *(Definition C.3 page 305).*

PRP: $\left\{\begin{array}{ll} 1. & \mathbf{L} \text{ is \textbf{orthocomplemented}} \text{ (Definition 15.1 page 244) and} \\ 2. & \mathbf{L} \text{ is \textbf{distributive}} \end{array}\right\} \implies \left\{ \mathbf{L} \text{ is \textbf{Boolean}} \right\}$

✎PROOF: To be a *Boolean algebra*, $\mathbf{L}$ must satisfy the 8 requirements of *boolean algebras*:

1. Proof for *commutative* properties: These are true for *all* lattices (Definition C.3 page 305).
2. Proof for *join-distributive* property: by hypothesis (2).
3. Proof for *meet-distributive* property: by *join-distributive* property and the *Principle of duality* (Theorem C.4 page 307) for lattices.
4. Proof for *identity* properties: because $\mathbf{L}$ is a *bounded lattice* and by definitions of 1 (*least upper bound*), 0 (*greatest lower bound*), $\vee$, and $\wedge$.
5. Proof for *complemented* properties: by hypothesis (1) and definition of *orthocomplemented lattices* (Definition 15.1 page 244).

⇒

Proposition 15.2. *Let* $\mathbf{L} = (X, \vee, \wedge, 0, 1; \leq)$ *be a* LATTICE *(Definition C.3 page 305).*

PRP: $\left\{\begin{array}{ll} 1. & \mathbf{L} \text{ is \textbf{orthocomplemented}} \text{ (Definition 15.1 page 244) and} \\ 2. & \text{Every } x \in \mathbf{L} \text{ is in the \textbf{center} of } \mathbf{L} \text{ (Definition E.4 page 340)} \end{array}\right\} \iff \left\{ \mathbf{L} \text{ is \textbf{Boolean}} \right\}$

✎PROOF:

1. Proof that (1,2) $\implies$ *Boolean*: $\mathbf{L}$ is *Boolean* because it satisfies *Huntington's Fourth Set*, as demonstrated by the following ...
   (a) Proof that $x \vee x = x$ (*idempotent*): $\mathbf{L}$ is a *lattice* (by definition of $\mathbf{L}$), and all lattices are *idempotent* (Definition C.3 page 305).
   (b) Proof that $x \vee y = y \vee x$ (*commutative*): $\mathbf{L}$ is a *lattice* (by definition of $\mathbf{L}$), and all lattices are *commutative* (Definition C.3 page 305).
   (c) Proof that $(x \vee y) \vee z = x \vee (y \vee z)$ (*associative*): $\mathbf{L}$ is a *lattice* (by definition of $\mathbf{L}$), and all lattices are *associative* (Definition C.3 page 305).
   (d) Proof that $(x^{\perp} \vee y^{\perp})^{\perp} \vee (x^{\perp} \vee y)^{\perp} = x$ (*Huntington's axiom*):

$$\begin{aligned}
(x^{\perp} \vee y^{\perp})^{\perp} \vee (x^{\perp} \vee y)^{\perp} &= (x^{\perp}\perp \wedge y^{\perp}\perp) \vee (x^{\perp}\perp \wedge y^{\perp}) && \text{by \textit{de Morgan} property (Theorem 15.1 page 246)} \\
&= (x \wedge y) \vee (x \wedge y^{\perp}) && \text{by \textit{involution} property (Definition 15.1 page 244)} \\
&= x && \text{by definition of \textit{center} (Definition E.4 page 340)}
\end{aligned}$$

2. Proof that (1) $\impliedby$ *Boolean*:
   (a) Proof that $x \vee x^{\perp} = 1$: by definition of *Boolean algebras*.
   (b) Proof that $x \wedge x^{\perp} = 0$: by definition of *Boolean algebras*.

[8] Kalmbach (1983) page 22

(c) Proof that $x^{\perp\perp} = x$: by *involutory* property of *Boolean algebra*.

(d) Proof that $x \le y \implies y^{\perp} \le x^{\perp}$:

$$
\begin{aligned}
y^{\perp} \le x^{\perp} \iff y^{\perp} &= y^{\perp} \wedge x^{\perp} && \text{by Lemma C.1 page 307}\\
\iff y^{\perp\perp} &= \left(y^{\perp} \wedge x^{\perp}\right)^{\perp} \\
\iff y^{\perp\perp} &= y^{\perp\perp} \vee x^{\perp\perp} && \text{by \textit{de Morgan} property}\\
\iff y &= y \vee x && \text{by \textit{involutory} property}\\
\iff y &= y && \text{by } x \le y \text{ hypothesis}
\end{aligned}
$$

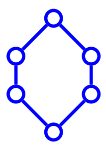
3. Proof that (2) $\Longleftarrow$ *Boolean*: for all $x, y \in \boldsymbol{L}$

$$
\begin{aligned}
(x \wedge y) \vee \left(x \wedge y^{\perp}\right) &= [(x \wedge y) \vee x] \wedge \left[(x \wedge y) \vee y^{\perp}\right] && \text{by \textit{distributive} property}\\
&= x \wedge \left[(x \wedge y) \vee y^{\perp}\right] && \text{by \textit{absorptive} property}\\
&= x \wedge \left[\left(x \vee y^{\perp}\right) \wedge \left(y \vee y^{\perp}\right)\right] && \text{by \textit{distributive} property}\\
&= x \wedge \left(x \vee y^{\perp}\right) \wedge 1 && \text{by \textit{complement} property}\\
&= x && \text{by \textit{absorptive} property}\\
&\implies x \circledC y \quad \forall x, y \in \boldsymbol{L} && \text{by Definition E.2 page 337}\\
&\implies x \text{ is in the \textit{center} of } \boldsymbol{L} \text{ for all } x \in \boldsymbol{L} && \text{by Definition E.4 page 340}
\end{aligned}
$$

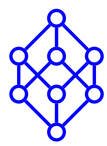
✎

Example 15.6.

EX



The O_6 *lattice* (Definition 15.2 page 244) illustrated to the left is **orthocomplemented** (Definition 15.1 page 244) but **non-join-distributive**, and hence *non-Boolean*. The lattice illustrated to the right is **orthocomplemented** *and* **distributive** and hence also **Boolean** (Proposition 15.1 page 251). Alternatively, the right side lattice is **orthocomplemented** *and* every element is in the *center*, and hence also **Boolean** (Proposition 15.2 page 251).



✎PROOF:

1. Proof that the O_6 *lattice* is *non-join-distributive*:

$$
\begin{aligned}
x \vee (x^{\perp} \wedge z^{\perp}) &= x \vee 0\\
&= x\\
&\ne z^{\perp}\\
&= 1 \wedge z^{\perp}\\
&= (x \vee x^{\perp}) \wedge (x \vee z^{\perp})
\end{aligned}
$$

2. Proof that the O_6 *lattice* is also *non-meet-distributive*:

$$
\begin{aligned}
z^{\perp} \wedge (x \vee z) &= z^{\perp} \wedge 1\\
&= z^{\perp}\\
&\ne x\\
&= x \vee 1\\
&= (z^{\perp} \wedge x) \vee (z^{\perp} \wedge z)
\end{aligned}
$$

✎

15.2 Orthomodular lattices

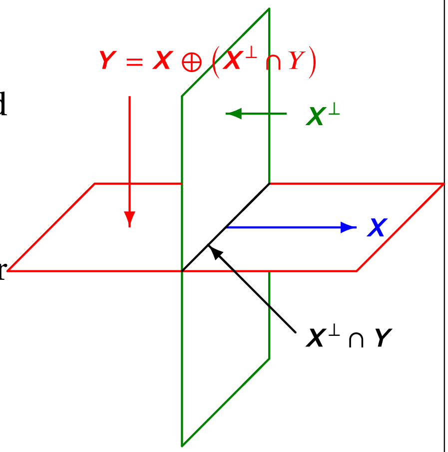
15.2.1 Properties

Definition 15.3. [9] *Let* $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ *be an algebraic structure.*

DEF
$\boldsymbol{L}$ *is an* ***orthomodular lattice*** *if*
1. $\boldsymbol{L}$ *is an* ORTHOCOMPLEMENTED LATTICE *and*
2. $x \leq y \implies x \vee (x^\perp \wedge y) = y \quad \forall x,y \in X$ (ORTHOMODULAR IDENTITY)

Example 15.7.

EX
The O_6 *lattice* (Definition 15.2 page 244) is *orthocomplemented*, but *non-orthomodular* (and hence, *non-modular* and *non-Boolean*).



Example 15.8. [10]Let $\boldsymbol{H}$ be a Hilbert space and $2^{\boldsymbol{H}}$ the set of closed linear subspaces of $\boldsymbol{H}$.

EX
$(2^{\boldsymbol{H}}, \oplus, \cap, \varnothing, \boldsymbol{H}; \subseteq)$
is an orthomodular lattice.

This concept is illustrated to the right where $\boldsymbol{X}, \boldsymbol{Y} \in 2^{\boldsymbol{H}}$ are linear subspaces of the linear space $\boldsymbol{H}$ and

$$\boldsymbol{X} \subseteq \boldsymbol{Y} \implies \boldsymbol{Y} = \boldsymbol{X} \oplus (\boldsymbol{X}^\perp \cap Y).$$

Theorem 15.3. [11] *Let* $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ *be a lattice.*

THM
$$\left.\begin{array}{ll} 1. & \boldsymbol{L} \text{ is ORTHOMODULAR and} \\ 2. & y \copyright x \text{ and } z \copyright x \end{array}\right\} \implies (x, y, z) \in Ⓓ$$

15.2.2 Characterizations

Theorem 15.4. [12] *Let* $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ *be an* ORTHOCOMPLEMENTED LATTICE *(Definition 15.1 page 244). Let* Ⓜ *and* Ⓜ* *be the modularity relation and dual modularity relation, respectively,* $\perp$ *the orthogonality relation (Definition E.1 page 335), and* © *the commutes relation (Definition E.2 page 337).*

THM
The following statements are EQUIVALENT*:*

	1.	$\boldsymbol{L}$ *is* ORTHOMODULAR
$\iff$	2.	$x \leq y$ *and* $y \wedge x^\perp = 0 \implies x = y$
$\iff$	3.	$\boldsymbol{L}$ *does* NOT *contain the* O_6 *lattice*
$\iff$	4.	$x \copyright y \iff y \copyright x$ (© *is* SYMMETRIC)
$\iff$	5.	$x Ⓜ x^\perp \quad \forall x \in X$
$\iff$	6.	$x Ⓜ^* x^\perp \quad \forall x \in X$
$\iff$	7.	$x \vee [x^\perp \wedge (x \vee y)] = x \vee y \quad \forall x, y \in X$
$\iff$	8.	$x \leq y \implies \exists p \in X$ such that $x \perp p$ *and* $x \vee p = y$

[9] Kalmbach (1983) page 22, Lidl and Pilz (1998) page 90, Husimi (1937)
[10] Iturrioz (1985) pages 56–57
[11] Kalmbach (1983) page 25, Holland (1963) pages 69–70 ⟨THEOREM 3⟩, Foulis (1962) page 68 ⟨THEOREM 5⟩
[12] Kalmbach (1983) page 22, Stern (1999) page 12, Nakamura (1957), Holland (1963), Foulis (1962), Maeda and Maeda (1970), page 132 ⟨Theorem 29.13⟩

✎PROOF:

1. Proof that *orthomodular* $\iff$ *symmetric*: by Proposition E.3 (page 338).

⇒

15.2.3 Restrictions resulting in Boolean algebras

Theorem 15.5. [13] *Let* $\mathbf{L} = (X, \vee, \wedge, 0, 1; \leq)$ *be an algebraic structure.*

THM

$$\left\{\begin{array}{ll} \mathbf{L} \text{ is an } \textbf{orthomodular lattice} & \text{and} \\ \underbrace{\left(x \wedge y^{\perp}\right)^{\perp} = y \vee \left(x^{\perp} \wedge y^{\perp}\right)}_{\text{ELKAN'S LAW}} & \forall x, y \in X \end{array}\right\} \implies \left\{\begin{array}{l} \mathbf{L} \text{ is a} \\ \textbf{Boolean algebra} \end{array}\right\}$$

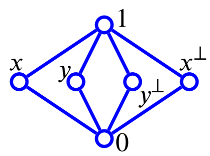
Definition 15.4. [14]

DEF

The $\mathbf{MO_2}$ ***lattice*** *is the ordered set* $\left(\{0, x, y, x^{\perp}, y^{\perp}, 1\}, \leq\right)$ *with cover relation*

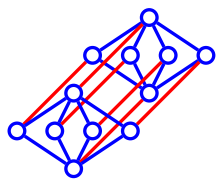
$$\lessdot = \left\{(0,x), (0,y), \left(0,x^{\perp}\right), \left(0,y^{\perp}\right), (x,1), (y,1), \left(x^{\perp},1\right), \left(y^{\perp},1\right)\right\}$$

This lattice is also called the ***Chinese lantern****.*



MO_2

L_2



$MO_2 \times L_2$

Theorem 15.6. [15] *Let* $\mathbf{M} = (X, \vee, \wedge, 0, 1; \leq)$ *be an* ORTHOMODULAR *lattice.*

THM

$$\left\{\begin{array}{l} \mathbf{M} \text{ is} \\ \text{BOOLEAN} \end{array}\right\} \iff \left\{\begin{array}{lll} 1. & \mathbf{M} \text{ does not contain the } MO_2 \text{ lattice (Definition 15.4 page 254)} & \text{and} \\ 2. & \mathbf{M} \text{ does not contain the } MO_2 \times L_2 \text{ lattice.} & \end{array}\right\}$$

15.3 Modular orthocomplemented lattices

Definition 15.5. *Let* $\mathbf{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ *be a* BOUNDED LATTICE *(Definition ?? page ??).*

DEF

$\mathbf{L}$ *is a* ***modular orthocomplemeted lattice*** *if*

1. $\mathbf{L}$ *is* ***orthocomplemented*** (Definition 15.1 page 244) *and*
2. $\mathbf{L}$ *is* ***modular***

[13] Renedo et al. (2003) page 72

[14] Iturrioz (1985) page 57, Davey and Priestley (2002) pages 18–19 ⟨1.25 Products⟩

[15] Iturrioz (1985) page 57, Carrega (1982) ⟨cf Iturrioz 1985 page 57⟩

15.4 Relationships between orthocomplemented lattices

Theorem 15.7. [16] *Let* L *be a lattice.*

THM
$$\left\{\begin{array}{l}\mathsf{L} \textit{ is} \\ \text{BOOLEAN}\end{array}\right\} \implies \left\{\begin{array}{l}\mathsf{L} \textit{ is} \\ \text{MODULAR} \\ \text{ORTHOCOM-} \\ \text{PLEMENTED}\end{array}\right\} \implies \left\{\begin{array}{l}\mathsf{L} \textit{ is } \text{ORTHO-} \\ \text{MODULAR}\end{array}\right\} \implies \left\{\begin{array}{l}\mathsf{L} \textit{ is} \\ \text{ORTHOCOM-} \\ \text{PLEMENTED}\end{array}\right\}$$

Remark 15.1. [17] Lattice number 8 in Example 15.2 (page 244) was originally introduced by Dilworth as a counterexample to *Husimi's conjecture* (1937). Kalmbach(1983) points out that this lattice was the first example of a *finite orthomodular* lattice.

[16] Kalmbach (1983) page 32 ⟨20.⟩, Iturrioz (1985) page 57
[17] Dilworth (1940), Dilworth (1990), Kalmbach (1983) page 9

CHAPTER 16

NORMED ALGEBRAS

16.1 Algebras

All *linear spaces* (Definition 3.1 page 67) are equipped with an operation by which vectors in the spaces can be added together. Linear spaces also have an operation that allows a scalar and a vector to be "multiplied" together. But linear spaces in general have no operation that allows two vectors to be multiplied together. A linear space together with such an operator is an **algebra**.[1]

There are many many possible algebras—many more than one can shake a stick at, as indicated by Michiel Hazewinkel in his book, *Handbook of Algebras*: "Algebra, as we know it today (2005), consists of many different ideas, concepts and results. A reasonable estimate of the number of these different items would be somewhere between 50,000 and 200,000. Many of these have been named and many more could (and perhaps should) have a "name" or other convenient designation."[2]

Definition 16.1. [3] *Let* $\boldsymbol{A}$ *be an* ALGEBRA.

DEF *An algebra* $\boldsymbol{A}$ *is* ***unital*** *if* $\exists u \in \boldsymbol{A}$ such that $ux = xu = x \qquad \forall x \in \boldsymbol{A}$

Definition 16.2. [4] *Let* $\boldsymbol{A}$ *be an* UNITAL ALGEBRA *(Definition 16.1 page 257) with unit* e.

DEF

The ***spectrum*** *of* $x \in \boldsymbol{A}$ *is* $\sigma(x) \triangleq \{\lambda \in \mathbb{C} | \lambda e - x \text{ is not invertible}\}$.

The ***resolvent*** *of* $x \in \boldsymbol{A}$ *is* $\rho_x(\lambda) \triangleq (\lambda e - x)^{-1} \qquad \forall \lambda \notin \sigma(x)$.

The ***spectral radius*** *of* $x \in \boldsymbol{A}$ *is* $r(x) \triangleq \sup\{|\lambda| | \lambda \in \sigma(x)\}$.

---

[1] Fuchs (1995) page 2
[2] Hazewinkel (2000) page v
[3] Folland (1995) page 1
[4] Folland (1995) pages 3–4

16.2 Star-Algebras

Definition 16.3. [5] *Let $\mathbf{A}$ be an* ALGEBRA.

DEF

The pair $(\mathbf{A}, *)$ is a ***$*$-algebra***, *or* ***star-algebra***, *if*

1.	$(x+y)^* = x^* + y^*$	$\forall x,y \in \mathbf{A}$	(DISTRIBUTIVE)	and
2.	$(\alpha x)^* = \bar{\alpha} x^*$	$\forall x \in \mathbf{A}, \alpha \in \mathbb{C}$	(CONJUGATE LINEAR)	and
3.	$(xy)^* = y^* x^*$	$\forall x,y \in \mathbf{A}$	(ANTIAUTOMORPHIC)	and
4.	$x^{**} = x$	$\forall x \in \mathbf{A}$	(INVOLUTORY)	

The operator $*$ is called an ***involution*** *on the algebra $\mathbf{A}$.*

Proposition 16.1. [6] *Let $(\mathbf{A}, *)$ be an* UNITAL $*$-ALGEBRA.

PRP

$$x \text{ is invertible} \implies \left\{ \begin{array}{lll} 1. & x^* \text{ is INVERTIBLE} & \forall x \in \mathbf{A} \quad \text{and} \\ 2. & (x^*)^{-1} = (x^{-1})^* & \forall x \in \mathbf{A} \end{array} \right.$$

✎PROOF: Let e be the unit element of $(\mathbf{A}, *)$.

1. Proof that $e^* = e$:

$$\begin{aligned}
x\,e^* &= (x\,e^*)^{**} && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= (x^*\,e^{**})^* && \text{by } \textit{antiautomorphic} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= (x^*\,e)^* && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= (x^*)^* && \text{by definition of } e\\
&= x && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
e^*\,x &= (e^*\,x)^{**} && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= (e^{**}\,x^*)^* && \text{by } \textit{antiautomorphic} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= (e\,x^*)^* && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= (x^*)^* && \text{by definition of } e\\
&= x && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition 16.3 page 258)}
\end{aligned}$$

2. Proof that $(x^*)^{-1} = (x^{-1})^*$:

$$\begin{aligned}
(x^{-1})^*\,(x^*) &= \left[x\,(x^{-1})\right]^* && \text{by } \textit{antiautomorphic} \text{ and } \textit{involution} \text{ properties of } * && \text{(Definition 16.3 page 258)}\\
&= e^*\\
&= e && \text{by item (1) page 258}\\
(x^*)\,(x^{-1})^* &= \left[x^{-1}\,x\right]^* && \text{by } \textit{antiautomorphic} \text{ and } \textit{involution} \text{ properties of } * && \text{(Definition 16.3 page 258)}\\
&= e^*\\
&= e && \text{by item (1) page 258}
\end{aligned}$$

⇨

Definition 16.4. [7] *Let $(\mathbf{A}, \|\cdot\|)$ be a* $*$-ALGEBRA *(Definition 16.3 page 258).*

DEF

- *An element $x \in \mathbf{A}$ is* ***hermitian*** *or* ***self-adjoint*** *if $x^* = x$.*
- *An element $x \in \mathbf{A}$ is* ***normal*** *if $xx^* = x^*x$.*
- *An element $x \in \mathbf{A}$ is a* ***projection*** *if $xx = x$ (INVOLUTORY) and $x^* = x$ (HERMITIAN).*

[5] Rickart (1960), page 178, Gelfand and Naimark (1964), page 241
[6] Folland (1995) page 5
[7] Rickart (1960), page 178, Gelfand and Naimark (1964), page 242

Theorem 16.1. [8] *Let* $(\boldsymbol{A}, \|\cdot\|)$ *be a* $*$-ALGEBRA *(Definition 16.3 page 258).*

T H M

$$\underbrace{x = x^* \text{ and } y = y^*}_{x \text{ and } y \text{ are hermitian}} \implies \left\{\begin{array}{lcll} x+y &=& (x+y)^* & (x+y \text{ is self adjoint}) \\ x^* &=& (x^*)^* & (x^* \text{ is self adjoint}) \\ \multicolumn{4}{l}{\underbrace{xy = (xy)^*}_{(xy) \text{ is hermitian}} \iff \underbrace{xy = yx}_{\text{commutative}}} \end{array}\right.$$

✎PROOF:

$$\begin{aligned}(x+y)^* &= x^* + y^* && \text{by \textit{distributive} property of } * && \text{(Definition 16.3 page 258)}\\ &= x + y && \text{by left hypothesis}\end{aligned}$$

$$(x^*)^* = x \qquad \text{by \textit{involutory} property of } * \qquad \text{(Definition 16.3 page 258)}$$

Proof that $xy = (xy)^* \implies xy = yx$

$$\begin{aligned}xy &= (xy)^* && \text{by left hypothesis}\\ &= y^*x^* && \text{by \textit{antiautomorphic} property of } * && \text{(Definition 16.3 page 258)}\\ &= yx && \text{by left hypothesis}\end{aligned}$$

Proof that $xy = (xy)^* \impliedby xy = yx$

$$\begin{aligned}(xy)^* &= (yx)^* && \text{by left hypothesis}\\ &= x^*y^* && \text{by \textit{antiautomorphic} property of } * && \text{(Definition 16.3 page 258)}\\ &= xy && \text{by left hypothesis}\end{aligned}$$

✏

Definition 16.5 (Hermitian components). [9] *Let* $(\boldsymbol{A}, \|\cdot\|)$ *be a* $*$-ALGEBRA *(Definition 16.3 page 258).*

D E F

*The **real part** of x is defined as* $\mathbf{R}_{\mathrm{e}}x \triangleq \frac{1}{2}\left(x + x^*\right)$

*The **imaginary part** of x is defined as* $\mathbf{I}_{\mathrm{m}}x \triangleq \frac{1}{2i}\left(x - x^*\right)$

Theorem 16.2. [10] *Let* $(\boldsymbol{A}, *)$ *be a* $*$-ALGEBRA *(Definition 16.3 page 258).*

T H M

$$\begin{array}{rcllr} \Re x &=& (\Re x)^* & \forall x \in \boldsymbol{A} & (\Re x \text{ is hermitian})\\ \Im x &=& (\Im x)^* & \forall x \in \boldsymbol{A} & (\Im x \text{ is hermitian})\end{array}$$

✎PROOF:

$$\begin{aligned}(\Re x)^* &= \left(\frac{1}{2}(x + x^*)\right)^* && \text{by definition of } \Re && \text{(Definition 16.5 page 259)}\\ &= \frac{1}{2}(x^* + x^{**}) && \text{by \textit{distributive} property of } * && \text{(Definition 16.3 page 258)}\\ &= \frac{1}{2}(x^* + x) && \text{by \textit{involutory} property of } * && \text{(Definition 16.3 page 258)}\\ &= \Re x && \text{by definition of } \Re && \text{(Definition 16.5 page 259)}\end{aligned}$$

$$(\Im x)^* = \left(\frac{1}{2i}(x - x^*)\right)^* \qquad \text{by definition of } \Im \qquad \text{(Definition 16.5 page 259)}$$

[8] Michel and Herget (1993) page 429
[9] Michel and Herget (1993) page 430, Rickart (1960), page 179, Gelfand and Naimark (1964), page 242
[10] Michel and Herget (1993) page 430, Halmos (1998a) page 42

$$\begin{aligned}
&= \frac{1}{2i}\left(x^* - x^{**}\right) && \text{by } \textit{distributive} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= \frac{1}{2i}\left(x^* - x\right) && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
&= \Im x && \text{by definition of } \Im && \text{(Definition 16.5 page 259)}
\end{aligned}$$

⇰

Theorem 16.3 (Hermitian representation). [11] *Let* $(\boldsymbol{A}, *)$ *be a* $*$-ALGEBRA *(Definition 16.3 page 258).*

THM $a = x + iy \iff x = \Re a$ *and* $y = \Im a$

✎PROOF:

Proof that $a = x + iy \implies x = \Re a$ and $y = \Im a$:

$$\begin{aligned}
& & a &= x + iy && \text{by left hypothesis}\\
&\implies & a^* &= (x + iy)^* && \text{by definition of } \textit{adjoint} && \text{(Definition 16.4 page 258)}\\
& & &= x^* - iy^* && \text{by } \textit{distributive} \text{ property of } * && \text{(Definition 16.3 page 258)}\\
& & &= x - iy && \text{by Theorem 16.2 page 259}\\
&\implies & x &= a - iy && \text{by solving for } x \text{ in } a = x + iy \text{ equation}\\
& & x &= a^* + iy && \text{by solving for } x \text{ in } a^* = x - iy \text{ equation}\\
&\implies & x + x &= a + a^* && \text{by adding previous 2 equations}\\
&\implies & 2x &= a + a^* && \text{by solving for } x \text{ in previous equation}\\
&\implies & x &= \frac{1}{2}\left(a + a^*\right)\\
& & &= \Re a && \text{by definition of } \Re && \text{(Definition 16.5 page 259)}\\
\\
& & iy &= a - x && \text{by solving for } iy \text{ in } a = x + iy \text{ equation}\\
& & iy &= -a^* + x && \text{by solving for } iy \text{ in } a = x + iy \text{ equation}\\
&\implies & iy + iy &= a - a^* && \text{by adding previous 2 equations}\\
&\implies & y &= \frac{1}{2i}\left(a - a^*\right) && \text{by solving for } iy \text{ in previous equations}\\
& & &= \Im a && \text{by definition of } \Im && \text{(Definition 16.5 page 259)}
\end{aligned}$$

Proof that $a = x + iy \impliedby x = \Re a$ and $y = \Im a$:

$$\begin{aligned}
x + iy &= \Re a + i\,\Im a && \text{by right hypothesis}\\
&= \underbrace{\frac{1}{2}\left(a + a^*\right)}_{\Re a} + i\underbrace{\frac{1}{2i}\left(a - a^*\right)}_{\Im a} && \text{by definition of } \Re \text{ and } \Im && \text{(Definition 16.5 page 259)}\\
&= \left(\frac{1}{2}a + \frac{1}{2}a\right) + \left(\cancelto{0}{\frac{1}{2}a^* - \frac{1}{2}a^*}\right)\\
&= a
\end{aligned}$$

⇰

[11] Michel and Herget (1993) page 430, Rickart (1960), page 179, Gelfand and Neumark (1943b), page 7

16.3 Normed Algebras

Definition 16.6. [12] *Let $\boldsymbol{A}$ be an algebra.*

DEF

*The pair $(\boldsymbol{A}, \|\cdot\|)$ is a **normed algebra** if*

$$\|xy\| \le \|x\|\,\|y\| \qquad \forall x, y \in \boldsymbol{A} \qquad \text{(multiplicative condition)}$$

*A normed algebra $(\boldsymbol{A}, \|\cdot\|)$ is a **Banach algebra** if $(\boldsymbol{A}, \|\cdot\|)$ is also a Banach space.*

Proposition 16.2.

PRP

$(\boldsymbol{A}, \|\cdot\|)$ *is a normed algebra* $\implies$ *multiplication is **continuous** in* $(\boldsymbol{A}, \|\cdot\|)$

✎PROOF:

1. Define $\mathsf{f}(x) \triangleq zx$. That is, the function f represents multiplication of x times some arbitrary value z.
2. Let $\delta \triangleq \|x - y\|$ and $\epsilon \triangleq \|\mathsf{f}(x) - \mathsf{f}(y)\|$.
3. To prove that multiplication (f) is *continuous* with respect to the metric generated by $\|\cdot\|$, we have to show that we can always make ϵ arbitrarily small for some $\delta > 0$.
4. And here is the proof that multiplication is indeed continuous in $(\boldsymbol{A}, \|\cdot\|)$:

$$\begin{aligned}
\|\mathsf{f}(x) - \mathsf{f}(y)\| &\triangleq \|zx - zy\| && \text{by definition of } \mathsf{f} && \text{(item (1) page 261)}\\
&= \|z(x-y)\| \\
&\le \|z\|\,\|x-y\| && \text{by definition of } \textit{normed algebra} && \text{(Definition 16.6 page 261)}\\
&\triangleq \|z\|\,\delta && \text{by definition of } \delta && \text{(item (2) page 261)}\\
&\le \epsilon && \text{for some value of } \delta > 0
\end{aligned}$$

✏

Theorem 16.4 (Gelfand-Mazur Theorem). [13] *Let $\mathbb{C}$ be the field of complex numbers.*

THM

$$\left.\begin{array}{l} (\boldsymbol{A}, \|\cdot\|) \textit{ is a Banach algebra} \\ \textit{every nonzero } x \in \boldsymbol{A} \textit{ is invertible} \end{array}\right\} \implies \boldsymbol{A} \equiv \mathbb{C} \quad (\boldsymbol{A} \textit{ is isomorphic to } \mathbb{C})$$

16.4 C* Algebras

Definition 16.7. [14]

DEF

*The triple $(\boldsymbol{A}, \|\cdot\|, *)$ is a C^* **algebra** if*

1. $(\boldsymbol{A}, \|\cdot\|)$ *is a Banach algebra* and
2. $(\boldsymbol{A}, *)$ *is a $*$-algebra* and
3. $\|x^* x\| = \|x\|^2 \quad \forall x \in \boldsymbol{A}$.

*A C^* **algebra** $(\boldsymbol{A}, \|\cdot\|, *)$ is also called a **C star algebra**.*

[12] Rickart (1960), page 2, Berberian (1961) page 103 ⟨Theorem IV.9.2⟩
[13] Folland (1995) page 4, Mazur (1938) ⟨(statement)⟩, Gelfand (1941) ⟨(proof)⟩
[14] Folland (1995) page 1, Gelfand and Naimark (1964), page 241, Gelfand and Neumark (1943a), Gelfand and Neumark (1943b)

Theorem 16.5. [15] *Let $\boldsymbol{A}$ be an algebra.*

THM: $(\boldsymbol{A}, \|\cdot\|, *)$ *is a* C^* ***algebra*** $\implies$ $\|x^*\| = \|x\|$

✎PROOF:

$$
\begin{aligned}
\|x\| &= \frac{1}{\|x\|}\,\|x\|^2 \\
&= \frac{1}{\|x\|}\,\|x^* x\| && \text{by definition of } C^*\text{-}\textit{algebra} && \text{(Definition 16.7 page 261)} \\
&\le \frac{1}{\|x\|}\,\|x^*\|\,\|x\| && \text{by definition of } \textit{normed algebra} && \text{(Definition 16.6 page 261)} \\
&= \|x^*\| \\
\|x^*\| &\le \|x^{**}\| && \text{by previous result} \\
&= \|x\| && \text{by } \textit{involution} \text{ property of } * && \text{(Definition 16.3 page 258)}
\end{aligned}
$$

⇰

[15] Folland (1995) page 1, Gelfand and Neumark (1943b), page 4, Gelfand and Neumark (1943a)

Part VI

Appendices

APPENDIX A

SET STRUCTURES

A.1 General set structures

Similar to the definition of a *relation* on a set X as being any subset of the *Cartesian product* $X \times X$, a *set structure* on a set X is simply any subset of the *power set* 2^X (next) of the set X.

Definition A.1.

DEF *The **power set** 2^X on a set X is defined as*

$$2^X \triangleq \{A | A \subseteq X\} \qquad \text{(the set of all subsets of } X\text{)}$$

Definition A.2. [1] *Let 2^X be the* POWER SET *(Definition A.1 page 265) of a set X.*

DEF *A set $\mathcal{S}(X)$ is a **set structure** on X if* $\quad \mathcal{S}(X) \subseteq 2^X$.
A SET STRUCTURE $\mathcal{Q}(X)$ *is a **paving** on X if* $\quad \varnothing \in \mathcal{Q}(X)$.

Definition A.3. [2] *Let $\mathcal{Q}(X)$ be a* PAVING *(Definition A.2 page 265) on a set X. Let Y be a set containing the element* 0.

DEF *A function $\mathsf{m} \in Y^{\mathcal{Q}(X)}$ is a **set function** if*

$$\mathsf{m}(\varnothing) = 0.$$

A.2 Operations on the power set

A.2.1 Standard operations

Definition A.4. [3] *Let 2^X be a set. Let $|X|$ be a function in the function space* $[0 : +\infty]^X$.

DEF $|X|$ *is the **cardinality** or **order** of X if*

$$|X| \triangleq \begin{cases} \text{number of elements in } X & \text{if } X \text{ is FINITE} \\ +\infty & \text{otherwise} \end{cases}$$

[1] Molchanov (2005) page 389, Pap (1995) page 7, Hahn and Rosenthal (1948) page 254

[2] Pap (1995) page 8 ⟨Definition 2.3: extended real-valued set function⟩, Halmos (1950) page 30 ⟨§7. MEASURE ON RINGS⟩, Hahn and Rosenthal (1948), Choquet (1954)

[3] Tao (2011) page 12 ⟨Example 3.6⟩, Tao (2010) page 7 ⟨Example 1.1.14⟩

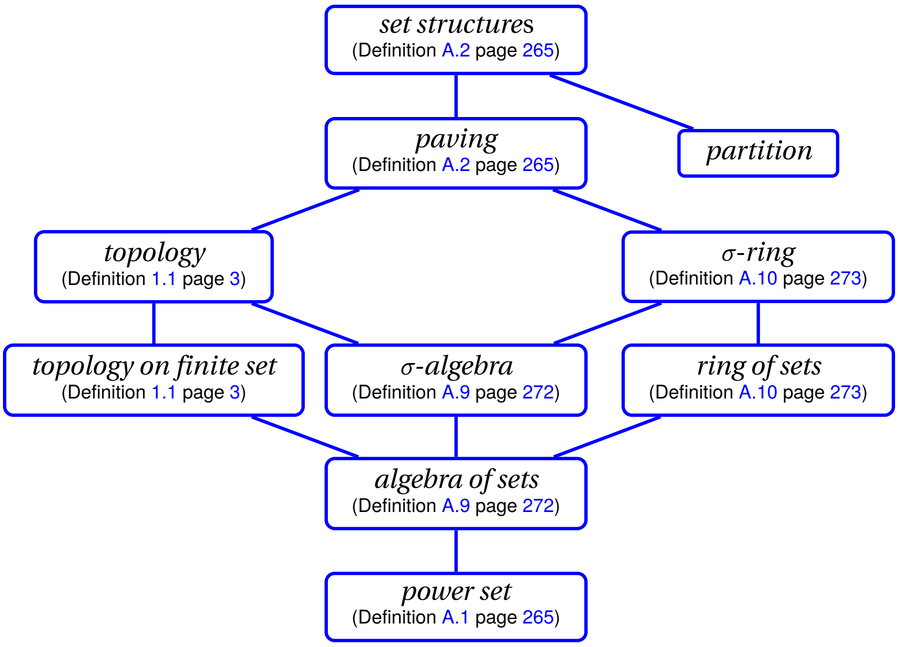


Figure A.1: some standard set structures

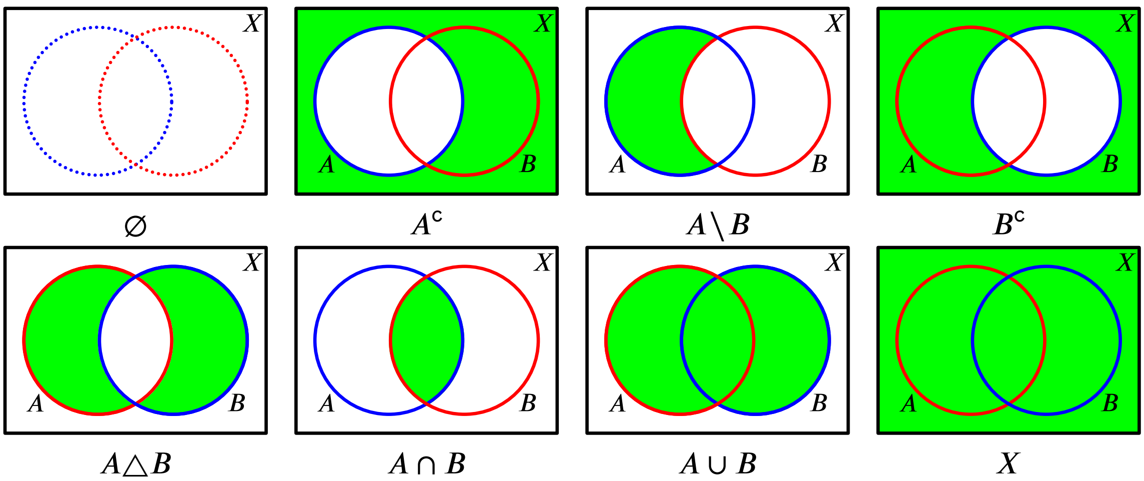


Figure A.2: Venn diagrams for standard set operations (Definition A.5 page 266)

Definition A.5 (next) introduces seven standard set operations: two *nullary* operations, one *unary* operation, and four *binary operations*.

Definition A.5. [4] *Let 2^X be the* POWER SET *(Definition A.1 page 265) on a set X. Let $\neg$ represent the* LOGICAL NOT *operation, $\vee$ represent the* LOGICAL OR *operation, $\wedge$ represent the* LOGICAL AND *operation, and $\oplus$ represent the* LOGICAL EXCLUSIVE-OR *operation.*

DEF	*name/symbol*		*arity*	*definition*	*domain*
	emptyset	$\varnothing$	0	$\varnothing \triangleq \{x \in X \mid x \neq x\}$	
	universal set	X	0	$X \triangleq \{x \in X \mid x = x\}$	
	complement	^{c}	1	$A^c \triangleq \{x \in X \mid \neg(x \in A)\}$	$\forall A \in 2^X$
	union	$\cup$	2	$A \cup B \triangleq \{x \in X \mid (x \in A) \vee (x \in B)\}$	$\forall A, B \in 2^X$
	intersection	$\cap$	2	$A \cap B \triangleq \{x \in X \mid (x \in A) \wedge (x \in B)\}$	$\forall A, B \in 2^X$
	difference	$\setminus$	2	$A \setminus B \triangleq \{x \in X \mid (x \in A) \wedge \neg(x \in B)\}$	$\forall A, B \in 2^X$
	symmetric difference	$\triangle$	2	$A \triangle B \triangleq \{x \in X \mid (x \in A) \oplus (x \in B)\}$	$\forall A, B \in 2^X$

[4] Aliprantis and Burkinshaw (1998) pages 2–4

With regards to the standard seven set operations only, Theorem A.1 (next) expresses each of the set operations in terms of pairs of other operations.

Theorem A.1.

THM

$$\begin{array}{rclclclcl}
X &=& \varnothing^c \\
\varnothing &=& X^c = \left(A \cup A^c\right)^c &=& A \cap A^c &=& A \setminus A &=& A \triangle A \\
X &=& A \cup A^c &=& \left(A \cap A^c\right)^c \\
A^c &=& X \setminus A &=& X \triangle A \\
A \cup B &=& \left(A^c \cap B^c\right)^c &=& (A \triangle B) \triangle (A \cap B) &=& (A \setminus B) \triangle B \\
A \cap B &=& \left(A^c \cup B^c\right)^c &=& (A \cup B) \triangle A \triangle B &=& A \setminus (A \setminus B) \\
A \setminus B &=& \left(A^c \cup B\right)^c &=& A \cap B^c &=& (A \cup B) \triangle B &=& (A \triangle B) \cap A \\
A \triangle B &=& \left[\left(A^c \cup B\right)^c\right] \cup \left[\left(A \cup B^c\right)^c\right] & & &=& \left[\left(A^c \cap B^c\right)^c\right] \cap (A \cap B)^c \\
&=& (A \setminus B) \cup (B \setminus A)
\end{array}$$

Proposition A.1. *Let X be a set and 2^X the power set of X. Let $\boldsymbol{R} \subseteq X$ such that $\boldsymbol{R}$ is closed with respect to the set symmetric difference operator $\triangle$.*

PRP

$(\boldsymbol{R}, \triangle)$ *is a* GROUP. *In particular,*

1. $\varnothing \triangle A = A \triangle \varnothing = A$ $\quad \forall A \in \boldsymbol{R}$ $\quad$ *($\varnothing$ is the* IDENTITY *element)*
2. $A \triangle A = \varnothing$ $\quad \forall A \in \boldsymbol{R}$ $\quad$ *(A is the* INVERSE *of A)*
3. $A \triangle (B \triangle C) = (A \triangle B) \triangle C$ $\quad \forall A, B, C \in \boldsymbol{R}$ $\quad$ *(*ASSOCIATIVE*)*

✎PROOF: The definition of a group is given by Definition F.1 (page 345).

Proof that $\varnothing$ is the *identity* element:

1a. Proof that $\varnothing \in \boldsymbol{R}$:

$$\begin{aligned}
\varnothing &= A \triangle A && \triangle \text{ closed with respect to } \boldsymbol{R} \\
&\in \boldsymbol{R}
\end{aligned}$$

1b. Proof that $\varnothing \triangle A = A$:

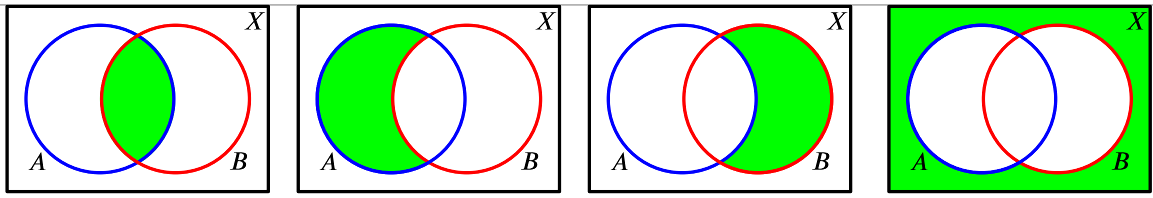
$$\begin{aligned}
\varnothing \triangle A &= \{x \in X | (x \in \varnothing) \oplus (x \in A)\} && \text{by definition of } \triangle \text{ page 266} \\
&= \{x \in X | (x \in \{x \in X | x \neq x\}) \oplus (x \in A)\} && \text{by definition of } \triangle \text{ page 266} \\
&= \{x \in X | 0 \oplus (x \in A)\} \\
&= \{x \in X | (x \in A)\} && \text{by definition of } \oplus \\
&= A
\end{aligned}$$

1c. Proof that $A \triangle \varnothing = A$:

$$\begin{aligned}
A \triangle \varnothing &= \{x \in X | (x \in A) \oplus (x \in \varnothing)\} && \text{by definition of } \triangle \text{ page 266} \\
&= \{x \in X | (x \in A) \oplus (x \in \{x \in X | x \neq x\})\} && \text{by definition of } \triangle \text{ page 266} \\
&= \{x \in X | (x \in A) \oplus 0\} \\
&= \{x \in X | (x \in A)\} && \text{by definition of } \oplus \\
&= A
\end{aligned}$$

2. Proof that $A \triangle A$:

$$\begin{aligned}
A \triangle A &= \{x \in X | (x \in A) \oplus (x \in A)\} && \text{by definition of } \triangle \text{ page 266} \\
&= \{x \in X | 0\} && \text{by definition of } \triangle \text{ page 266} \\
&= \varnothing && \text{by definition of } \triangle \text{ page 266}
\end{aligned}$$



$A \cap B$	$A \cap B^c$	$A^c \cap B$	$A^c \cap B^c$
(in both A and B)	(in A but not in B)	(not in A but in B)	(neither in A nor B)

Figure A.3: The partition of a set X into 4 regions by subsets A and B

3. Proof that $A\triangle(B\triangle C) = (A\triangle B)\triangle C$:

$$\begin{aligned}
A\triangle(B\triangle C) &= \{x \in X | (x \in A) \oplus [x \in (B\triangle C)]\} && \text{by definition of } \triangle \text{ page 266} \\
&= \{x \in X | (x \in A) \oplus [(x \in B) \oplus (x \in C)]\} && \text{by definition of } \triangle \text{ page 266} \\
&= \{x \in X | [(x \in A) \oplus (x \in B)] \oplus (x \in C)\} \\
&= (A\triangle B)\triangle C
\end{aligned}$$

A.2.2 Non-standard operations

Two subsets A and B of a set X that are intersecting but yet one is not contained in the other, partition the set X into four regions, as illustrated in Figure A.3 (page 268). Because there are four regions, the number of ways we can select one or more of them is $2^4 = 16$. Therefore, a binary operator on sets A and B can likewise result in one of $2^4 = 16$ possibilities. Definition A.6 (page 268) presents 7 set operations. Therefore, there should be an additional $16-7 = 9$ operations. Definition A.6 (next definition) attempts to define these additional operations. Some definitions are adapted from logic. But in general these definitions are non-standard definitions with respect to set theory. The 16 set operations under the inclusion relation $\subseteq$ form a lattice; this lattice is illustrated by a *Hasse diagram* in Figure A.4 (page 269).

Definition A.6. [5] *Let 2^X be the power set on a set X. For any sets $A, B \in 2^X$, let $AB \triangleq (A \cap B)$.*

DEF

name/symbol		*arity*	*definition*	*domain*
empty set	$\oslash$	2	$A \oslash B \triangleq \varnothing$	$\forall A,B \in 2^X$
rejection	$\downarrow$	2	$A \downarrow B \triangleq A^c B^c$	$\forall A,B \in 2^X$
inhibit x	$\ominus$	2	$A \ominus B \triangleq A^c B$	$\forall A,B \in 2^X$
complement x	c_x	2	$A\, c_x\, B \triangleq A^c B \cup A^c B^c$	$\forall A,B \in 2^X$
difference	$\setminus$	2	$A \setminus B \triangleq AB^c$	$\forall A,B \in 2^X$
complement y	c_y	2	$A\, c_y\, B \triangleq AB^c \cup A^c B^c$	$\forall A,B \in 2^X$
symmetric difference	$\triangle$	2	$A \triangle B \triangleq AB^c \cup A^c B$	$\forall A,B \in 2^X$
Sheffer stroke	$\mid$	2	$A \mid B \triangleq AB^c \cup A^c B \cup A^c B^c$	$\forall A,B \in 2^X$
intersection	$\cap$	2	$A \cap B \triangleq AB$	$\forall A,B \in 2^X$
equivalence	$\Leftrightarrow$	2	$A \Leftrightarrow B \triangleq AB \cup A^c B^c$	$\forall A,B \in 2^X$
projection y	$\Vdash$	2	$A \Vdash B \triangleq AB \cup A^c B$	$\forall A,B \in 2^X$
implication	$\Rightarrow$	2	$A \Rightarrow B \triangleq AB \cup A^c B \cup A^c B^c$	$\forall A,B \in 2^X$
projection x	$\dashv\mid$	2	$A \dashv\mid B \triangleq AB \cup AB^c$	$\forall A,B \in 2^X$
adjunction	$\div$	2	$A \div B \triangleq AB \cup AB^c \cup A^c B^c$	$\forall A,B \in 2^X$
union	$\cup$	2	$A \cup B \triangleq AB \cup AB^c \cup A^c B$	$\forall A,B \in 2^X$
universal set	$\otimes$	2	$A \otimes B \triangleq AB \cup AB^c \cup A^c B \cup A^c B^c$	$\forall A,B \in 2^X$

[5] standard ops: Aliprantis and Burkinshaw (1998) pages 2–4

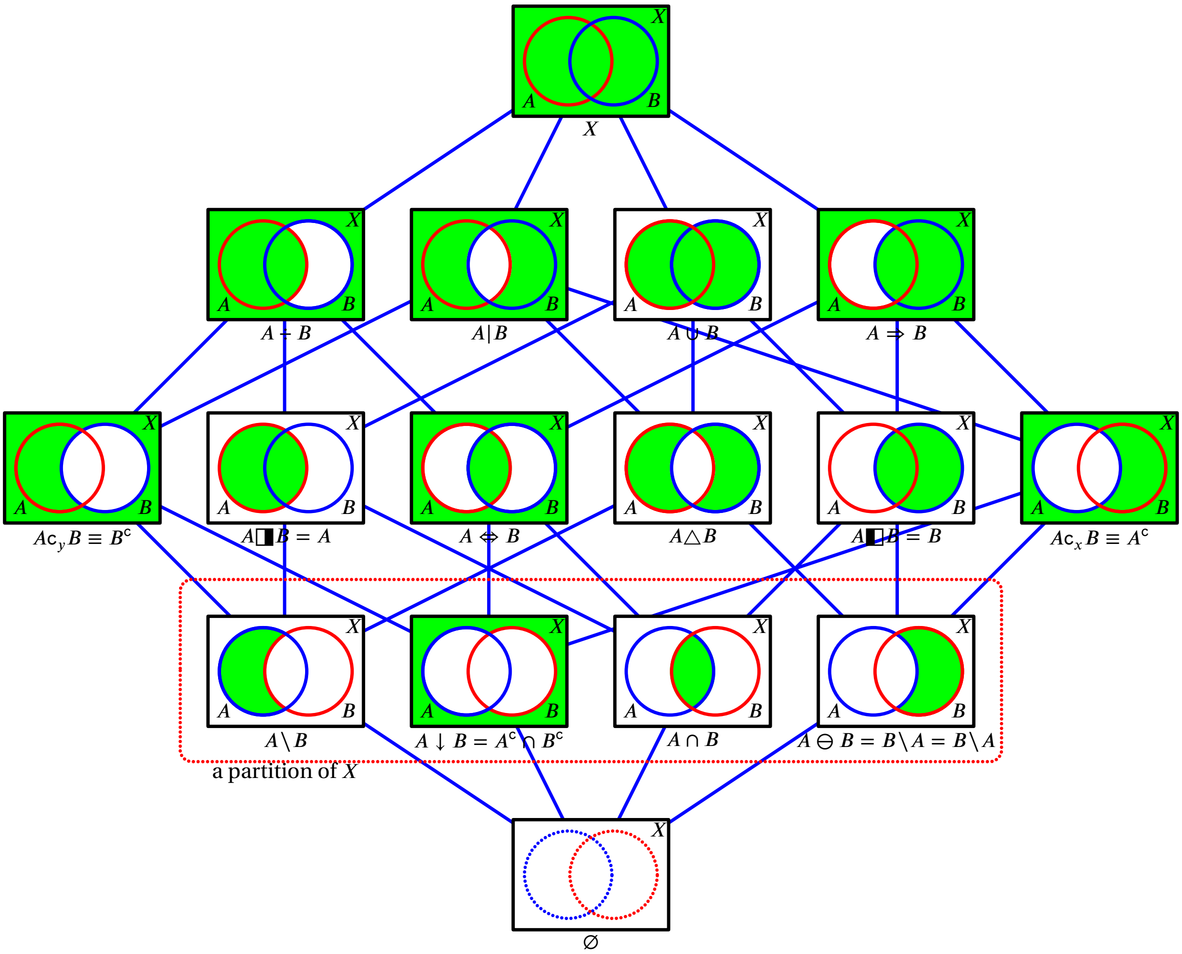


Figure A.4: lattice of set operations

A.2.3 Generated operations

Definition A.5 (page 266) defines set operations in terms of logical operations. However, it is also possible to express set operations in terms of two or more other set operations. When all the set operations can be expressed in terms of a set of operations, then that set of operations is *functionally complete* (next definition).

Definition A.7. [6] *Let S be a set structure.*

DEF

A set of operations Φ is ***functionally complete*** *in S if*

$\cup, \cap, \mathsf{c}, \varnothing$, *and* X

can all be expressed in terms of elements of Φ.

Example A.1. Here are some examples of *functionally complete* sets:

EX

🙛	$\{\downarrow\}$	(*rejection*)
🙛	$\{\mid\}$	(*Sheffer stroke*)
🙛	$\{\div, \varnothing\}$	(*adjunction* and $\varnothing$)
🙛	$\{\setminus, X\}$	(*set difference* and X)
🙛	$\{\cup, \mathsf{c}\}$	(*union* and *complement*)
🙛	$\{\cap, \mathsf{c}\}$	(*intersection* and *complement*)
🙛	$\{\triangle, \cap, X\}$	(*symmetric difference*, *intersection*, and X)
🙛	$\{\triangle, \cup, X\}$	(*symmetric difference*, *union*, and X)
🙛	$\{\triangle, \setminus, \mathsf{c}\}$	(*symmetric difference*, *set difference*, and *complement*)

A.2.4 Set multiplication

The *Cartesian product* operation $\times$ (next definition) is a kind of *set multiplication* operation.

Definition A.8. [7] *Let X and Y be sets, and let (x, y) be an* ORDERED PAIR.

DEF

The ***Cartesian product*** *$X \times Y$ of X and Y is*

$$X \times Y \triangleq \{(x, y) \mid (x \in X) \text{ and } (y \in Y)\}$$

Theorem A.2 (next theorem) demonstrates how this set operation interacts with certain other set operations. The Cartesian product is of critical importance in general because, for example, relations and functions are subsets of Cartesian products.

Theorem A.2. [8] *Let X, Y, Z be sets.*

THM

$$\begin{array}{rcll}
X \times (Y \cup Z) &=& (X \times Y) \cup (X \times Z) & (\times \text{ distributes over } \cup) \\
X \times (Y \cap Z) &=& (X \times Y) \cap (X \times Z) & (\times \text{ distributes over } \cap) \\
X \times (Y \setminus Z) &=& (X \times Y) \setminus (X \times Z) & (\times \text{ distributes over } \setminus) \\
(X \times Y) \cap (Y \times X) &=& (X \cap Y) \times (Y \cap X) & \\
(X \times X) \cap (Y \times Y) &=& (X \cap Y) \times (X \cap Y) &
\end{array}$$

[6] Whitesitt (1995) page 69

[7] Halmos (1960) page 24
G. Frege, 2007 August 25, http://groups.google.com/group/sci.logic/msg/3b3294f5ac3a76f0

[8] Menini and Oystaeyen (2004), page 50, Halmos (1960) page 25

✎PROOF:

$$
\begin{aligned}
X\times(Y\cup Z) &= \{(a,b)\,|(a\in X)\wedge(b\in Y\cup Z)\} \\
&= \{(a,b)\,|(a\in X)\wedge[(b\in Y)\vee(b\cup Z)]\} && \text{by Definition A.5} \\
&= \{(a,b)\,|\,[(a\in X)\wedge(b\in Y)]\vee[(a\in X)\wedge(b\in Z)]\} \\
&= \underbrace{\{(a,b)\,|\,[(a\in X)\wedge(b\in Y)]\}}_{X\times Y}\cup\underbrace{\{(a,b)\,|\,[(a\in X)\wedge(b\in Z)]\}}_{X\times Z} && \text{by Definition A.5} \\
&= (X\times Y)\cup(X\times Z)
\end{aligned}
$$

$$
\begin{aligned}
X\times(Y\cap Z) &= \{(a,b)\,|(a\in X)\wedge(b\in Y\cap Z)\} \\
&= \{(a,b)\,|(a\in X)\wedge[(b\in Y)\wedge(b\cup Z)]\} && \text{by Definition A.5} \\
&= \{(a,b)\,|\,[(a\in X)\wedge(b\in Y)]\wedge[(a\in X)\wedge(b\in Z)]\} \\
&= \underbrace{\{(a,b)\,|\,[(a\in X)\wedge(b\in Y)]\}}_{X\times Y}\cap\underbrace{\{(a,b)\,|\,[(a\in X)\wedge(b\in Z)]\}}_{X\times Z} && \text{by Definition A.5} \\
&= (X\times Y)\cap(X\times Z)
\end{aligned}
$$

$$
\begin{aligned}
X\times(Y\backslash Z) &= \{(a,b)\,|(a\in X)\wedge(b\in Y\backslash Z)\} \\
&= \left\{(a,b)\,|(a\in X)\wedge(b\in Y\cap Z^{c})\right\} \\
&= \left\{(a,b)\,|(a\in X)\wedge\left[(b\in Y)\wedge(b\in Z^{c})\right]\right\} && \text{by Definition A.5} \\
&= \left\{(a,b)\,|\,[(a\in X)\wedge(b\in Y)]\wedge\left[(a\in X)\wedge(b\in Z^{c})\right]\right\} \\
&= \underbrace{\{(a,b)\,|\,[(a\in X)\wedge(b\in Y)]\}}_{X\times Y}\cap\underbrace{\left\{(a,b)\,|\,\left[(a\in X)\wedge(b\in Z^{c})\right]\right\}}_{X\times Z^{c}} && \text{by Definition A.5} \\
&= (X\times Y)\cap(X\times Z^{c}) \\
&\neq (X\times Y)\backslash(X\times Z)
\end{aligned}
$$

$$
\begin{aligned}
(X\times Y)\cap(Y\times X) &= \{(a,b)\,|(a\in X)\wedge(b\in Y)\}\cap\{(a,b)\,|(a\in Y)\wedge(b\in X)\} \\
&= \{(a,b)\,|[(a\in X)\wedge(b\in Y)]\wedge[(a\in Y)\wedge(b\in X)]\} && \text{by Definition A.5} \\
&= \{(a,b)\,|[(a\in X)\wedge(a\in Y)]\wedge[(b\in Y)\wedge(b\in X)]\} \\
&= \{(a,b)\,|(a\in X\cap Y)\wedge(b\in Y\cap X)\} \\
&= (X\cap Y)\times(Y\cap X)
\end{aligned}
$$

$$
\begin{aligned}
(X\times X)\cap(Y\times Y) &= \{(a,b)\,|(a\in X)\wedge(b\in X)\}\cap\{(a,b)\,|(a\in Y)\wedge(b\in Y)\} \\
&= \{(a,b)\,|[(a\in X)\wedge(b\in X)]\ \wedge\ [(a\in Y)\wedge(b\in Y)]\} && \text{by Definition A.5} \\
&= \{(a,b)\,|[(a\in X)\wedge(a\in Y)]\ \wedge\ [(b\in X)\wedge(b\in Y)]\} \\
&= \{(a,b)\,|(a\in X\cap Y)\ \wedge\ (b\in X\cap Y)\} \\
&= (X\cap Y)\times(X\cap Y)
\end{aligned}
$$

⇨

A.3 Standard set structures

Set structures are typically designed to satisfy some special properties— such as being closed with respect to certain set operations. Examples of commonly occurring set structures include

A.3.1 Topologies

See CHAPTER 1 (page 3)

A.3.2 Algebras of sets

Definition A.9. [9] *Let X be a set with* POWER SET 2^X (Definition A.1 page 265).

DEF

*$\mathbf{A} \subseteq 2^X$ is an **algebra of sets** on X if*

1. $A \in \mathbf{A} \implies A^c \in \mathbf{A}$ (closed under complement operation) and
2. $A, B \in \mathbf{A} \implies A \cap B \in \mathbf{A}$ (closed under $\cap$)

The set of all algebra of sets on a set X is denoted $\mathcal{A}(X)$ such that

$$\mathcal{A}(X) \triangleq \left\{ \mathbf{A} \subseteq 2^X \,|\, \mathbf{A} \text{ is an algebra of sets} \right\}.$$

An ALGEBRA OF SETS $\mathbf{A}$ *on X is a **σ-algebra** on X if*

3. $\left\{ A_n \middle| n \in \mathbb{Z} \right\} \subseteq \mathbf{A} \implies \bigcup_{n\in\mathbb{Z}} A_n \in \mathbf{A}$ (closed under countable union operations).

On every set X with at least 2 elements, there are always two particular algebras of sets: the *smallest algebra* and the *largest algebra*, as demonstrated by Example A.2 (next).

Example A.2. [10] Let $\mathcal{A}(X)$ be the set of *algebras of sets* (Definition A.9 page 272) on a set X and 2^X the *power set* (Definition A.1 page 265) on X.

EX

$$\begin{array}{lcll} \{\varnothing,\ X\} & \in & \mathcal{A}(X) & \text{(smallest algebra)} \\ 2^X & \in & \mathcal{A}(X) & \text{(largest algebra)} \end{array}$$

Isomorphically, all *algebras of sets* are *boolean algebras* and all boolean algebras are algebras of sets (next theorem).

Theorem A.3 (Stone Representation Theorem). [11] *Let $\mathbf{L} \triangleq (X, \vee, \wedge\,;\, \leq)$ be a* LATTICE.

THM

$$\mathbf{L} \text{ is BOOLEAN} \iff \left\{ \begin{array}{l} \mathbf{L} \text{ is isomorphic to } (\mathbf{A}, \cup, \cap, \varnothing, X\,;\, \subseteq) \\ \text{for some ALGEBRA OF SETS (Definition A.9 page 272) } \mathbf{A} \end{array} \right\}$$

✎PROOF:

1. Proof that *algebra of sets* $\implies$ *Boolean algebra*:
   (a) Proof that $\mathbf{S}$ is closed under $\cup$ and $\cap$: by hypothesis.

[9] Aliprantis and Burkinshaw (1998) page 95, Aliprantis and Burkinshaw (1998) page 151, Halmos (1950) page 21, Hausdorff (1937) page 91

[10] Stroock (1999) page 33, Aliprantis and Burkinshaw (1998) pages 95–96

[11] Levy (2002) page 257, Grätzer (2003) page 85, Joshi (1989) page 224, Saliĭ (1988) page 32 ⟨"Stone's Theorem"⟩, Stone (1936a)

(b) By item (1b) and by Theorem A.4 (page 277), $\mathbf{L}$ is a *distributive* lattice.

(c) By item (1b) and properties of *lattices* (Theorem C.3 page 306), $\mathbf{L}$ is *idempotent*, *commutative*, *associative*, and *absorptive*.

(d) Proof that $\mathbf{L}$ has *identity*:

$$\begin{aligned}
A \cup \varnothing &= \{x \in X | (x \in A) \vee (x \in \varnothing)\} && \text{by definition of } \cup \text{ Definition A.5 page 266}\\
&= \{x \in X | x \in A\} && \text{by definition of } \varnothing \text{ Definition A.5 page 266}\\
&= A\\
A \cap X &= \{x \in X | (x \in A) \wedge (x \in X)\} && \text{by definition of } \cap \text{ Definition A.5 page 266}\\
&= \{x \in X | x \in A\} && \text{by definition of } \varnothing \text{ Definition A.5 page 266}\\
&= A
\end{aligned}$$

(e) Proof that $\mathbf{L}$ is *complemented*: by hypothesis.

(f) Because $\mathbf{L}$ is *commutative* (item (1c) page 273), *distributive* (item (1b) page 273), has *identity* (item (1d) page 273), and is *complemented* (item (1e) page 273), and by the definition of *Boolean algebras*, $\mathbf{L}$ is a *Boolean algebra*.

2. Proof that *Boolean algebra* $\implies$ *algebra of sets*: not included at this time.

⇰

A.3.3 Rings of sets

A *ring of sets* (next definition) is a family of subsets that is closed under an "addition-like" set union operator $\cup$ and "subtraction-like" set difference operator $\setminus$. Using these two operations, it is not difficult to show that a ring of sets is also closed under a "multiplication-like" set intersection operator $\cap$. Because of this, a ring of sets behaves like an *algebraic ring*. Note however that a ring of sets is not necessarily a *topology* (Definition 1.1 page 3) because it does not necessarily include X itself.

Definition A.10. [12] *Let X be a set with* POWER SET 2^X *(Definition A.1 page 265).*

DEF

*$\mathbf{R} \subseteq 2^X$ is a **ring of sets** on X if*

1. $A, B \in \mathbf{R} \implies A \cup B$ *(closed under $\cup$)* *and*
2. $A, B \in \mathbf{R} \implies A \setminus B \in \mathbf{R}$ *(closed under $\setminus$)*

The set of all rings of sets on a set X is denoted $\mathcal{R}(X)$ such that

$$\mathcal{R}(X) \triangleq \left\{\mathbf{R} \subseteq 2^X | \mathbf{R} \text{ is a ring of sets}\right\}.$$

A RING OF SETS *$\mathbf{R}$ on X is a **σ-ring** on X if*

3. $\left\{A_n \middle| n \in \mathbb{Z}\right\} \subseteq \mathbf{R} \implies \bigcup_{n\in\mathbb{Z}} A_n \in \mathbf{R}$ *(closed under countable union operations).*

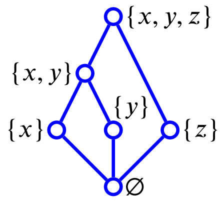
Example A.3. Table A.1 (page 274) lists some *rings of sets* on a finite set X.

Example A.4. Let $X \triangleq \{x, y, z\}$ be a set and $\mathbf{R}$ be the family of sets

$\mathbf{R} \triangleq \{\varnothing,\ X,\ \{x\},\ \{y\},\ \{z\},\ \{x, y\}\}$.

Note that $(\mathbf{R}, \subseteq, \cup, \cap)$ is a lattice as illustrated in the figure to the right. However, $\mathbf{R}$ is *not* a ring of sets on X because, for example,

$\{x, y, z\} \setminus \{x\} = \{y, z\} \notin \mathbf{R}$.



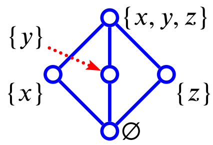
[12] Berezansky et al. (1996) page 4, Halmos (1950) page 19, Hausdorff (1937) page 90

rings $\mathcal{R}(X)$ on a set X
$\mathcal{R}(\varnothing) = \left\{ \boldsymbol{R}_1 = \{\varnothing\} \right\}$
$\mathcal{R}(\{x\}) = \left\{ \begin{array}{lllll} \boldsymbol{R}_1 & = \{ & \varnothing, & & \} \\ \boldsymbol{R}_2 & = \{ & \varnothing, & \{x\} & \} \end{array} \right\}$
$\mathcal{R}(\{x,y\}) = \left\{ \begin{array}{lllllll} \boldsymbol{R}_1 & = \{ & \varnothing, & & & & \} \\ \boldsymbol{R}_2 & = \{ & \varnothing, & \{x\}, & & & \} \\ \boldsymbol{R}_3 & = \{ & \varnothing, & & \{y\}, & & \} \\ \boldsymbol{R}_4 & = \{ & \varnothing, & & & \{x,y\} & \} \\ \boldsymbol{R}_5 & = \{ & \varnothing, & \{x\}, & \{y\}, & \{x,y\} & \} \end{array} \right\}$
$\mathcal{R}(\{x,y,z\}) = \left\{ \begin{array}{lllllllllll} \boldsymbol{R}_1 & = \{ & \varnothing, & & & & & & & & \} \\ \boldsymbol{R}_2 & = \{ & \varnothing, & \{x\}, & & & & & & & \} \\ \boldsymbol{R}_3 & = \{ & \varnothing, & & \{y\}, & & & & & & \} \\ \boldsymbol{R}_4 & = \{ & \varnothing, & & & \{z\}, & & & & & \} \\ \boldsymbol{R}_5 & = \{ & \varnothing, & & & & \{x,y\}, & & & & \} \\ \boldsymbol{R}_6 & = \{ & \varnothing, & & & & & \{x,z\}, & & & \} \\ \boldsymbol{R}_7 & = \{ & \varnothing, & & & & & & \{y,z\}, & & \} \\ \boldsymbol{R}_8 & = \{ & \varnothing, & \{x\}, & \{y\}, & & \{x,y\}, & & & & \} \\ \boldsymbol{R}_9 & = \{ & \varnothing, & \{x\}, & & \{z\}, & & \{x,z\}, & & & \} \\ \boldsymbol{R}_{10} & = \{ & \varnothing, & & \{y\}, & \{z\}, & & & \{y,z\}, & & \} \\ \boldsymbol{R}_{11} & = \{ & \varnothing, & & & & & & & X & \} \\ \boldsymbol{R}_{12} & = \{ & \varnothing, & \{x\}, & & & & & \{y,z\}, & X & \} \\ \boldsymbol{R}_{13} & = \{ & \varnothing, & & \{y\}, & & & \{x,z\}, & & X & \} \\ \boldsymbol{R}_{14} & = \{ & \varnothing, & & & \{z\}, & \{x,y\}, & & & X & \} \\ \boldsymbol{R}_{15} & = \{ & \varnothing, & \{x\}, & \{y\}, & \{z\}, & \{x,y\}, & \{x,z\}, & \{y,z\}, & X & \} \end{array} \right\}$

Table A.1: some *rings of sets* on a finite set X (Example A.3 page 273)

Example A.5. Let $X \triangleq \{x, y, z\}$ be a set and $\boldsymbol{R}$ be the family of sets $\boldsymbol{R} \triangleq \{\varnothing, X, \{x\}, \{y\}, \{z\}\}$. Note that $(\boldsymbol{T}, \subseteq) \cup \cap$ is a lattice as illustrated in the figure to the right. However, $\boldsymbol{R}$ is *not* a ring of sets on X because, for example,

$\{x, y, z\} \setminus \{x\} = \{y, z\} \notin \boldsymbol{R}$.



Proposition A.2. [13] *Let $\mathcal{R}(X)$ be the set of* RINGS OF SETS (Definition A.10 page 273) *on a set X.*

PRP
$$\left\{\begin{array}{l}\boldsymbol{R}_1 \text{ and } \boldsymbol{R}_2 \\ \text{are } \textbf{rings of sets}\end{array}\right\} \implies \left\{\begin{array}{l}(\boldsymbol{R}_1 \cap \boldsymbol{R}_2) \\ \text{is a } \textbf{ring of sets}\end{array}\right\}$$

A.3.4 Partitions

The following definition is a special case of *partition* defined on lattices.

Definition A.11. [14]

DEF

A SET STRUCTURE $\left\{P_n \in 2^X \,\middle|\, n=1,2,\ldots,N\right\}$ *is a* ***partition*** *of the set X if*

1.	$P_n \neq \varnothing$	$\forall n \in \{1,2,\ldots,N\}$	NON-EMPTY	*and*
2.	$P_n \cap P_m = \varnothing$	$\forall n \neq m$	MUTUALLY EXCLUSIVE	*and*
3.	$\bigcup_{n \in \mathbb{Z}} P_n = X$			

Example A.6. Let $A, B \subseteq X$, as illustrated in Figure A.3 (page 268). There are a total of 15 partitions of X induced by A and B (Proposition ?? page ??). Here are 5 of these partitions:

EX
1. $\{X\}$ (1 region)
2. $\{A, A^c\}$ (2 regions)
3. $\{A \cup B, A^c \cap B^c\}$ (2 regions)
4. $\{A \cap B, A \triangle B, A^c \cap B^c\}$ (3 regions)
5. $\{A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c\}$ (4 regions) [see also Figure A.3 page 268 and Figure A.4 page 269]

Proposition A.3. [15] *Let $\mathcal{P}(X)$ be the set of partitions on a set X.*

PRP

The relation $\trianglelefteq \in 2^{\mathbb{PP}}$ defined as

$$\boldsymbol{P} \trianglelefteq \boldsymbol{Q} \quad \overset{\text{def}}{\iff} \quad \forall B \in \boldsymbol{Q}, \exists A \in \boldsymbol{P} \quad \text{such that} \quad B \subseteq A$$

is an ordering relation on $\mathcal{P}(X)$.

Example A.7. Table A.2 (page 276) lists some partitions $\boldsymbol{P}(X)$ on a finite set X.

[13] Kolmogorov and Fomin (1975) page 32, Bartle (2001) page 318
[14] Munkres (2000), page 23, Rota (1964), page 498, Halmos (1950) page 31
[15] Roman (2008) page 111, Comtet (1974) page 220, Grätzer (2007), page 697

partitions $\mathcal{P}(X)$ on a set X		
$\mathcal{P}(\varnothing)$	$=$	$\{\ \boldsymbol{P}_1 = \varnothing\ \}$
$\mathcal{P}(\{x\})$	$=$	$\{\ \boldsymbol{P}_1 = \{\ \{x\}\ \}\ \}$
$\mathcal{P}(\{x,y\})$	$=$	$\boldsymbol{P}_1 = \{\ \{x\},\ \{y\},\ \}$
		$\boldsymbol{P}_2 = \{\ \{x,y\}\ \}$
$\mathcal{P}(\{x,y,z\})$	$=$	$\boldsymbol{P}_1 = \{\ \{x,y,z\}\ \}$
		$\boldsymbol{P}_2 = \{\ \{x\},\ \{y,z\},\ \}$
		$\boldsymbol{P}_3 = \{\ \{y\},\ \{x,z\},\ \}$
		$\boldsymbol{P}_4 = \{\ \{z\},\ \{x,y\}\ \}$
		$\boldsymbol{P}_5 = \{\ \{x\},\ \{y\},\ \{z\}\ \}$
$\mathcal{P}(\{w,x,y,z\})$	$=$	$\boldsymbol{P}_1 = \{\ X\ \}$
		$\boldsymbol{P}_2 = \{\ \{w\},\ \{x,y,z\}\ \}$
		$\boldsymbol{P}_3 = \{\ \{x\},\ \{w,y,z\}\ \}$
		$\boldsymbol{P}_4 = \{\ \{y\},\ \{w,x,z\}\ \}$
		$\boldsymbol{P}_5 = \{\ \{z\},\ \{w,x,y\}\ \}$
		$\boldsymbol{P}_6 = \{\ \{w,x\},\{y,z\}\ \}$
		$\boldsymbol{P}_7 = \{\ \{w,y\},\{x,z\}\ \}$
		$\boldsymbol{P}_8 = \{\ \{w,z\},\{x,y\},\ \}$
		$\boldsymbol{P}_9 = \{\ \{w\},\{x\}\ \ \{y,z\}\ \}$
		$\boldsymbol{P}_{10} = \{\ \{w\},\{y\},\ \{x,z\}\ \}$
		$\boldsymbol{P}_{11} = \{\ \{w\},\{z\},\ \{x,y\}\ \}$
		$\boldsymbol{P}_{12} = \{\ \{x\},\{y\},\ \{w,z\}\ \}$
		$\boldsymbol{P}_{13} = \{\ \{x\},\{z\},\ \{w,y\}\ \}$
		$\boldsymbol{P}_{14} = \{\ \{y\},\{z\},\ \{w,x\}\ \}$
		$\boldsymbol{P}_{15} = \{\ \{w\},\{x\},\{y\},\{z\},\ \}$

Table A.2: some partitions $\boldsymbol{P}(X)$ on a finite set X (Example A.7 page 275)

A.4 Operations on set structures

Proposition A.4.

PRP

closed under	*partition*	*ring of sets*	*algebra of sets*	*topology*
$\varnothing$		✓	✓	✓
X	✓		✓	✓
c			✓	
$\cup$		✓	✓	✓
$\cap$		✓	✓	✓
$\triangle$		✓	✓	
$\setminus$		✓	✓	

✎PROOF:

1. Proof for closure in a *topology*: Definition 1.1 (page 3)
2. Proof for closure in a *ring of sets*: Definition A.10 (page 273) and Theorem A.6 (page 279)
3. Proof for closure in an *algebra of sets*: Definition A.9 (page 272) and Theorem A.5 (page 277)

⇒

Theorem A.4. *Let T be a* SET STRUCTURE (Definition A.2 page 265) *on a set X.*

THM

T *is a* ***topology*** $\implies \forall A, B, C \in T$

property with emphasis on $\cup$	dual property with emphasis on $\cap$	property name
$A \cup A = A$	$A \cap A = A$	(IDEMPOTENT)
$A \cup B = B \cup A$	$A \cap B = B \cap A$	(COMMUTATIVE)
$A \cup (B \cup C) = (A \cup B) \cup C$	$A \cap (B \cap C) = (A \cap B) \cap C$	(ASSOCIATIVE)
$A \cup (A \cap B) = A$	$A \cap (A \cup B) = A$	(ABSORPTIVE)
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	(DISTRIBUTIVE)

✎PROOF:

1. By Definition 1.1 (page 3), T is a *topology*.
2. By Theorem A.3 (page 272), $(T, \cup, \cap; \subseteq)$ is a *distributive lattice*.
3. The properties listed are all properties of *distributive lattices*.

⇒

Proposition A.5. *Let A be a* SET STRUCTURE (Definition A.2 page 265) *on a set X.*

PRP

$$\left\{ \begin{array}{l} A \text{ is an} \\ \textbf{algebra of sets} \end{array} \right\} \implies \left\{ \begin{array}{llllll} 1. & \varnothing & \in A & & (A \text{ includes the } \cup \text{ identity element}) \\ 2. & X & \in A & & (A \text{ includes the } \cap \text{ identity element}) \\ 3. & A^c & \in A & \forall A \in A & (A \text{ is closed under } c) \\ 4. & A \cup B & \in A & \forall A, B \in A & (A \text{ is closed under } \cup) \\ 5. & A \cap B & \in A & \forall A, B \in A & (A \text{ is closed under } \cap) \\ 6. & A \setminus B & \in A & \forall A, B \in A & (A \text{ is closed under } \setminus) \\ 7. & A \triangle B & \in A & \forall A, B \in A & (A \text{ is closed under } \triangle) \end{array} \right\}$$

✎PROOF:

$$
\begin{aligned}
\varnothing &= A \cap A^c \\
X &= c\varnothing \\
A \cup B &= c(A^c \cap B^c) && \text{by } \textit{de Morgan's Law} \text{ (Theorem A.6 page 278)} \\
A \setminus B &= A \cap B^c \\
A \triangle B &= (A \setminus B^c) \cup (B \setminus A)
\end{aligned}
$$

$(\boldsymbol{A}, \cup, \setminus)$ is a ring of sets because $\cup$ and $\setminus$ are closed in $\boldsymbol{A}$ (as shown above). ✏

Theorem A.5. [16] *Let* $\boldsymbol{A}$ *be a* SET STRUCTURE (Definition A.2 page 265) *on a set* X.

THM

$\boldsymbol{A}$ *is an* ***algebra of sets*** $\implies \forall A, B, C \in \boldsymbol{A}$

property emphasizing $\cup$	dual property emphasizing $\cap$	property name
$A \cup A = A$	$A \cap A = A$	(IDEMPOTENT)
$A \cup B = B \cup A$	$A \cap B = B \cap A$	(COMMUTATIVE)
$A \cup (B \cup C) = (A \cup B) \cup C$	$A \cap (B \cap C) = (A \cap B) \cap C$	(ASSOCIATIVE)
$A \cup (A \cap B) = A$	$A \cap (A \cup B) = A$	(ABSORPTIVE)
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	(DISTRIBUTIVE)
$A \cup \varnothing = A$	$A \cap X = A$	(IDENTITY)
$A \cup X = X$	$A \cap \varnothing = \varnothing$	(BOUNDED)
$A \cup A^c = X$	$A \cap A^c = \varnothing$	(COMPLEMENTED)
$(A^c)^c = A$		(UNIQUELY COMPLEMENTED)
$(A \cup B)^c = A^c \cap B^c$	$(A \cap B)^c = A^c \cup B^c$	(DE MORGAN)

✎PROOF:

1. By Definition A.9 (page 272), $\boldsymbol{S}$ is an *algebra of sets*.
2. By the *Stone Representation Theorem* (Theorem A.3 page 272), $(\boldsymbol{S}, \cup, \cap, \varnothing, X ; \subseteq)$ is a *Boolean algebra*.
3. The properties listed are all properties of *Boolean algebras*.

✏

Theorem A.6. [17] *Let* $\boldsymbol{A}$ *be an* ALGEBRA OF SETS (Definition A.9 page 272) *on a set* X.

THM

$\boldsymbol{A}$ *is an* ***algebra of sets*** $\implies \forall A_1, A_2, \ldots, A_N, B \in \boldsymbol{A}$ *and* $\forall N \in \mathbb{N}$

property emphasizing $\cup$	dual property emphasizing $\cap$	property name
$\left(\bigcup_{n=1}^{N} A_n\right)^c = \bigcap_{n=1}^{N} A^c{}_n$	$\left(\bigcap_{n=1}^{N} A_n\right)^c = \bigcup_{n=1}^{N} A^c{}_n$	(DE MORGAN)
$\left(\bigcup_{n=1}^{N} A_n\right) \cap B = \bigcup_{n=1}^{N} (A_n \cap B)$	$\left(\bigcap_{n=1}^{N} A_n\right) \cup B = \bigcap_{n=1}^{N} (A_n \cup B)$	(DISTRIBUTIVE *with respect to* $\cup$ *and* $\cap$)
$\left(\bigcup_{n=1}^{N} A_n\right) \setminus B = \bigcup_{n=1}^{N} (A_n \setminus B)$	$\left(\bigcap_{n=1}^{N} A_n\right) \setminus B = \bigcap_{n=1}^{N} (A_n \setminus B)$	(DISTRIBUTIVE *with respect to* $\setminus$ *and* $\cap$)

✎PROOF:

1. By Theorem A.3 (page 272), the lattice $(X, \cup, \cap ; \subseteq)$ is *Boolean*.
2. The first four properties are true any Boolean system.

[16] Dieudonné (1969) pages 3–4, Copson (1968) page 9

[17] Michel and Herget (1993) page 12, Aliprantis and Burkinshaw (1998) page 4, Vaidyanathaswamy (1960) pages 3–4

3. Proof for the remaining two:

$$\begin{aligned}
\left(\bigcap_{n=1}^{N} A_n\right)\setminus B &= \left(\bigcap_{n=1}^{N} A_n\right)\cap B^c && \text{by Theorem A.1 page 267}\\
&= \bigcap_{n=1}^{N}(A_n \cap B^c) && \text{by previous result}\\
&= \bigcap_{n=1}^{N}(A_n \setminus B) && \text{by Theorem A.1 page 267}
\end{aligned}$$

$$\begin{aligned}
\left(\bigcup_{n=1}^{N} A_n\right)\setminus B &= \left(\bigcup_{n=1}^{N} A_n\right)\cap B^c && \text{by Theorem A.1 page 267}\\
&= \bigcup_{n=1}^{N}(A_n \cap B^c) && \text{by previous result}\\
&= \bigcup_{n=1}^{N}(A_n \setminus B) && \text{by Theorem A.1 page 267}
\end{aligned}$$

⇨

Proposition A.6. [18] *Let $\mathbf{R}$ be a* SET STRUCTURE (Definition A.2 page 265) *on a set X.*

PRP

$$\left\{\begin{array}{l}\mathbf{R} \text{ is a}\\ \textbf{ring of sets}\\ \text{on } X\end{array}\right\} \implies \left\{\begin{array}{llllll}
1. & \varnothing & \in \mathbf{R} & & (\mathbf{R} \text{ includes the } \cup \text{ identity element}) & \text{and}\\
2. & A\cup B & \in \mathbf{R} & \forall A,B\in\mathbf{R} & (\mathbf{R} \text{ is closed under } \cup) & \text{and}\\
3. & A\cap B & \in \mathbf{R} & \forall A,B\in\mathbf{R} & (\mathbf{R} \text{ is closed under } \cap) & \text{and}\\
4. & A\setminus B & \in \mathbf{R} & \forall A,B\in\mathbf{R} & (\mathbf{R} \text{ is closed under } \setminus) & \text{and}\\
5. & A\triangle B & \in \mathbf{R} & \forall A,B\in\mathbf{R} & (\mathbf{R} \text{ is closed under } \triangle) &
\end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
A\triangle B &= (A\setminus B)\cup(B\setminus A)\\
A\cap B &= (A\cup B)\setminus(A\triangle B)\\
A\setminus A &= \varnothing
\end{aligned}$$

⇨

Theorem A.7. [19] *Let $\mathbf{R}$ be a* SET STRUCTURE (Definition A.2 page 265) *on a set X.*

THM

*If $\mathbf{R}$ is an **ring of sets** on X, then $(\mathbf{R}, \triangle, \cap)$ is an* ALGEBRAIC RING; *in particular,*

$$\begin{array}{lllll|llll}
A\triangle\varnothing & = & A & \forall A\in\mathbf{R} & & A\cap\varnothing & = & \varnothing & \forall A\in\mathbf{R}\\
A\triangle X & = & A^c & \forall A\in\mathbf{R} & & A\cap X & = & A & \forall A\in\mathbf{R}\\
A\triangle\varnothing & = & A & \forall A\in\mathbf{R} & & A\cap A & = & A & \forall A\in\mathbf{R}
\end{array}$$

$$A\cap(B\triangle C) = (A\cap B)\triangle(A\cap C) \qquad \forall A,B,C\in\mathbf{R}$$

properties emphasizing $\triangle$ | *properties emphasizing $\cap$*

✎PROOF:

[18] Berezansky et al. (1996) page 4, Halmos (1950) pages 19–20

[19] Vaidyanathaswamy (1960) pages 17–18, Kelley and Srinivasan (1988) page 22, Wilker (1982), page 211, Vaidyanathaswamy (1960) page 19

1. Proof that $(\boldsymbol{R}, \cup, \setminus)$ is an *algebraic ring*: by Theorem A.7 (page 279)
2. Proof that a ring of sets is equivalent to $(\boldsymbol{R}, \cup, \setminus)$: This is proven simply by noting that $\cup$ and $\setminus$ (the two operations in a ring of sets $(\boldsymbol{R}, \cup, \setminus)$) can be expressed in terms of $\triangle$ and $\cap$ (the two operations in the algebraic ring $(\boldsymbol{R}, \triangle, \cap)$) and vice-versa. And this is demonstrated by Theorem A.1 (page 267).

The definition of an algebraic ring is given in Definition F.2 (page 345).

1. Proof that $(\boldsymbol{S}, \triangle)$ is a group: see Proposition A.1 (page 267).

2. Proof that $A \cap (B \cap C) = (A \cap B) \cap C$:

$$\begin{aligned}
A \cap (B \cap C) &= \{x \in X | (x \in A) \wedge [(x \in B) \wedge (x \in C)]\} && \text{by definition of } \cap \text{ page 266} \\
&= \{x \in X | [(x \in A) \wedge (x \in B)] \wedge (x \in C)\} \\
&= (A \cap B) \cap C && \text{by definition of } \cap \text{ page 266}
\end{aligned}$$

3. Proof that $A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$:

$$\begin{aligned}
A \cap (B \triangle C) &= \{x \in X | (x \in A) \wedge [(x \in B) \oplus (x \in C)]\} && \text{by definition of } \cap, \triangle \text{ page 266} \\
&= \{x \in X | [(x \in A) \wedge (x \in B)] \oplus [(x \in A) \wedge (x \in C)]\} \\
&= (A \cap B) \triangle (A \cap C) && \text{by definition of } \cap, \triangle \text{ page 266}
\end{aligned}$$

4. Proof that $(A \triangle B) \cap C = (A \cap C) \triangle (B \cap C)$:

$$\begin{aligned}
(A \triangle B) \cap C &= \{x \in X | [(x \in A) \oplus (x \in B)] \wedge (x \in C)\} && \text{by definition of } \cap, \triangle \text{ page 266} \\
&= \{x \in X | [(x \in A) \wedge (x \in C)] \oplus [(x \in B) \wedge (x \in C)]\} \\
&= (A \cap C) \triangle (B \cap C) && \text{by definition of } \cap, \triangle \text{ page 266}
\end{aligned}$$

✏

A.5 Lattices of set structures

The *set inclusion* relation $\subseteq$ (Definition A.12 page 280) is an *order relation* (Definition B.2 page 290) on set structures, as demonstrated by Proposition A.7 (next proposition).

Definition A.12. *Let $\boldsymbol{S}$ be a* SET STRUCTURE *(Definition A.2 page 265) on a set X.*

DEF: *The relation $\subseteq \in 2^{\boldsymbol{SS}}$ is defined as*

$$A \subseteq B \quad \text{if} \quad x \in A \implies x \in B \quad \forall x \in X$$

Proposition A.7 (order properties)**.** *Let $\boldsymbol{S}$ be a* SET STRUCTURE *(Definition A.2 page 265) on a set X.*

PRP: *The pair $(\boldsymbol{S}, \subseteq)$ is an* ORDERED SET. *In particular,*

$$\begin{array}{lllllllll}
A \subseteq A & & & & \forall A \in \boldsymbol{S} & (\text{REFLEXIVE}) & \text{and} \\
A \subseteq B & \text{and} & B \subseteq C & \implies & A \subseteq C & \forall A,B,C \in \boldsymbol{S} & (\text{TRANSITIVE}) & \text{and} \\
A \subseteq B & \text{and} & B \subseteq A & \implies & A = B & \forall A,B \in \boldsymbol{S} & (\text{ANTI-SYMMETRIC}). &
\end{array}$$

✎PROOF: By Definition B.2 (page 290), a relation is an *order relation* if it is *reflexive*, *transitive*, and *anti-symmetric*.

1. Proof that $\subseteq$ is *reflexive* on 2^X:

$$\begin{aligned}
x \in A &\implies x \in A \\
&\implies A \subseteq A
\end{aligned}$$

2. Proof that $\subseteq$ is *transitive* on 2^X:

$$\begin{aligned}
x \in A &\implies x \in B && \text{by first left hypothesis}\\
&\implies x \in C && \text{by second left hypothesis}\\
&\implies A \subseteq C
\end{aligned}$$

3. Proof that $\subseteq$ is *anti-symmetric* on 2^X:

$$\begin{aligned}
A \subseteq B &\implies (x \in A \implies x \in B)\\
B \subseteq A &\implies (x \in B \implies x \in A)\\
A \subseteq B \text{ and } B \subseteq A &\implies (x \in A \iff x \in B)\\
&\implies A = B
\end{aligned}$$

✏

In a set structure that is *closed* under the *union* operation $\cup$ and *intersection* operation $\cap$, the *greatest lower bound* of any two elements A and B is simply $A \cap B$ and *least upper bound* is simply $A \cup B$ (Proposition A.8 page 281). However, this may not be true for a set structure that is *not* closed under these operations (Example A.8 page 282).

Proposition A.8. *Let S be a* SET STRUCTURE *(Definition A.2 page 265) on a set X.*

PRP

If S is closed under $\cup$ and $\cap$ ***then***

$A \cup B$	*is the* LEAST UPPER BOUND	*of A and B in $(S, \subseteq)$*	$(\cup = \vee)$	*and*
$A \cap B$	*is the* GREATEST LOWER BOUND	*of A and B in $(S, \subseteq)$*	$(\cap = \wedge)$.	

✎PROOF:

1. Proof that $A \cup B$ is the least upper bound:

$$\begin{aligned}
A &= \{x \in X | x \in A\}\\
&\subseteq \{x \in X | x \in A \text{ or } x \in B\}\\
&= A \cup B && \text{by Definition A.5 page 266}\\
B &= \{x \in X | x \in B\}\\
&\subseteq \{x \in X | x \in A \text{ or } x \in B\}\\
&= A \cup B && \text{by Definition A.5 page 266}
\end{aligned}$$

$$\begin{aligned}
A \subseteq C \text{ and } B \subseteq C &\implies \left\{x \in A \text{ and } y \in B \implies x, y \in C\right\}\\
&\implies \{x \in A \text{ or } x \in B \implies x \in C\}\\
&\implies \{x \in A \cup B \implies x \in C\}\\
&\implies A \cup B \subseteq C
\end{aligned}$$

2. Proof that $A \cap B$ is the greatest lower bound:

$$\begin{aligned}
A \cap B &= \left\{x \in X | x \in A \text{ and } x \in B\right\} && \text{by Definition A.5 page 266}\\
&\subseteq \{x \in X | x \in A\}\\
&= A\\
A \cap B &= \left\{x \in X | x \in A \text{ and } x \in B\right\} && \text{by Definition A.5 page 266}\\
&\subseteq \{x \in X | x \in B\}\\
&= B
\end{aligned}$$

$$\begin{aligned}
C \subseteq A \text{ and } C \subseteq B &\implies \{x \in C \implies x \in A \quad\text{and}\quad x \in C \implies x \in B\} \\
&\implies \{x \in C \quad\implies\quad x \in A \text{ or } x \in B\} \\
&\implies \{x \in C \quad\implies\quad x \in A \cap B\} \\
&\implies C \subseteq A \cap B
\end{aligned}$$

✏

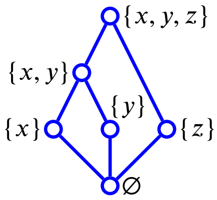
Example A.8. The set structure

$$\boldsymbol{S} \triangleq \{\varnothing,\ \{x\},\ \{y\},\ \{z\},\ \{x,y\},\ \{x,y,z\}\}$$

ordered by the set inclusion relation $\subseteq$ is illustrated by the Hasse diagram to the right. Note that

$$\{x\} \vee \{z\} = \{x,y,z\} \neq \{x,z\} = \{x\} \cup \{z\}.$$

That is, the set union operation $\cup$ is *not* equivalent to the order join operation $\vee$.



A.5.1 Lattices of topologies

See Section 1.1.6 (page 10)

A.5.2 Lattices of algebra of sets

Example A.9. The following table lists some algebras of sets on a finite set X. Lattices of algebras of sets are illustrated in Figure A.7 (page 284) and Figure A.5 (page 283).

algebra of sets $\mathcal{A}(X)$ on a set X		
$\mathcal{A}(\varnothing)$	$=$	$\left\{ \boldsymbol{A}_1 = \{\varnothing\} \right\}$
$\mathcal{A}(\{x\})$	$=$	$\left\{ \boldsymbol{A}_1 = \{\varnothing,\ \{x\}\} \right\}$
$\mathcal{A}(\{x,y\})$	$=$	$\left\{ \begin{array}{lcl} \boldsymbol{A}_1 &=& \{\varnothing,\ X\} \\ \boldsymbol{A}_2 &=& \{\varnothing,\ \{x\},\ \{y\},\ X\} \end{array} \right\}$
$\mathcal{A}(\{x,y,z\})$	$=$	$\left\{ \begin{array}{lcllllllllll} \boldsymbol{A}_1 &=& \{ & \varnothing, & & & & & & & X & \} \\ \boldsymbol{A}_2 &=& \{ & \varnothing, & \{x\}, & & & & & \{y,z\}, & X & \} \\ \boldsymbol{A}_3 &=& \{ & \varnothing, & & \{y\}, & & & \{x,z\}, & & X & \} \\ \boldsymbol{A}_4 &=& \{ & \varnothing, & & & \{z\}, & \{x,y\}, & & & X & \} \\ \boldsymbol{A}_5 &=& \{ & \varnothing, & \{x\}, & \{y\}, & \{z\}, & \{x,y\}, & \{x,z\}, & \{y,z\}, & X & \} \end{array} \right\}$
$\mathcal{A}(\{w,x,y,z\})$	$=$	

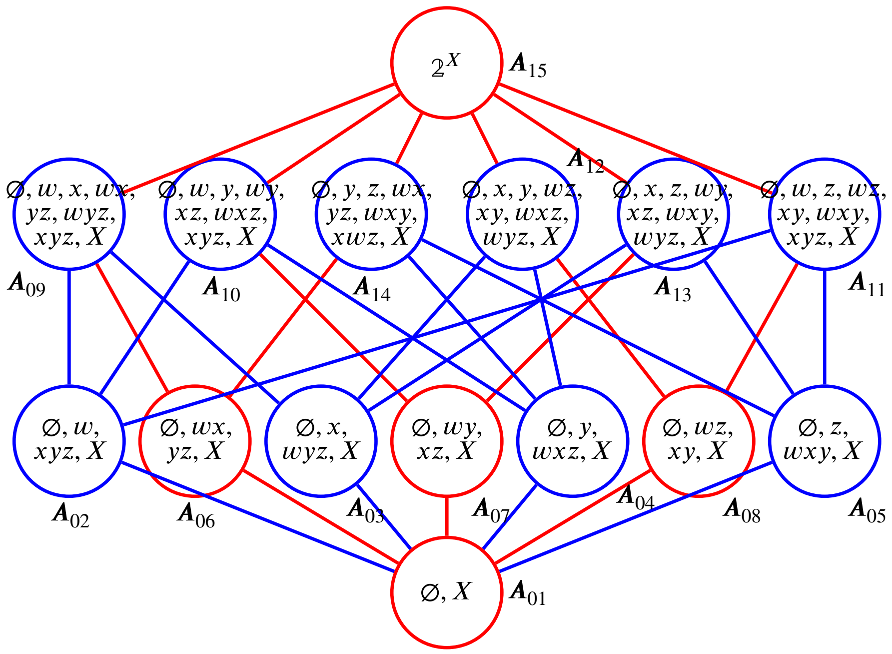


Figure A.5: lattice of *algebras of sets* on $\{w, x, y, z\}$ (Example A.9 page 282)

$$\left\{\begin{array}{lllllllll}
\boldsymbol{A}_1 & = & \{ & \varnothing, & & & & X & \} \\
\boldsymbol{A}_2 & = & \{ & \varnothing, & \{w\}, & & \{x,y,z\}, & X & \} \\
\boldsymbol{A}_3 & = & \{ & \varnothing, & \{x\}, & & \{w,y,z\}, & X & \} \\
\boldsymbol{A}_4 & = & \{ & \varnothing, & \{y\}, & & \{w,x,z\}, & X & \} \\
\boldsymbol{A}_5 & = & \{ & \varnothing, & \{z\}, & & \{w,x,y\}, & X & \} \\
\boldsymbol{A}_6 & = & \{ & \varnothing, & & \{w,x\},\{y,z\}, & & X & \} \\
\boldsymbol{A}_7 & = & \{ & \varnothing, & & \{w,y\},\{x,z\}, & & X & \} \\
\boldsymbol{A}_8 & = & \{ & \varnothing, & & \{w,z\},\{x,y\}, & & X & \} \\
\boldsymbol{A}_9 & = & \{ & \varnothing, & \{w\},\{x\}, & \{w,x\},\{y,z\}, & \{w,y,z\},\{x,y,z\}, & X & \} \\
\boldsymbol{A}_{10} & = & \{ & \varnothing, & \{w\},\{y\}, & \{w,y\},\{x,z\}, & \{w,x,z\},\{x,y,z\}, & X & \} \\
\boldsymbol{A}_{11} & = & \{ & \varnothing, & \{w\},\{z\}, & \{w,z\},\{x,y\}, & \{w,x,y\},\{x,y,z\}, & X & \} \\
\boldsymbol{A}_{12} & = & \{ & \varnothing, & \{x\},\{y\}, & \{w,z\},\{x,y\}, & \{w,x,z\},\{w,y,z\}, & X & \} \\
\boldsymbol{A}_{13} & = & \{ & \varnothing, & \{x\},\{z\}, & \{w,y\},\{x,z\}, & \{w,x,y\},\{w,y,z\}, & X & \} \\
\boldsymbol{A}_{14} & = & \{ & \varnothing, & \{y\},\{z\}, & \{w,x\},\{y,z\}, & \{w,x,y\},\{w,x,z\}, & X & \} \\
\boldsymbol{A}_{15} & = & 2^X & & & & & &
\end{array}\right\}$$

A.5.3 Lattices of rings of sets

Example A.10. There are a total of **15** rings of sets on the set $X \triangleq \{x, y, z\}$. These rings of sets are listed in Example A.3 (page 273) and illustrated in Figure A.6 (page 284). The five rings containing X ($\boldsymbol{R}_{11}$–$\boldsymbol{R}_{15}$) are also *algebras of sets* (Proposition A.10 page 287), and thus also *Boolean algebras* (Theorem A.3 page 272). The five algebras of sets are shaded Figure A.6.

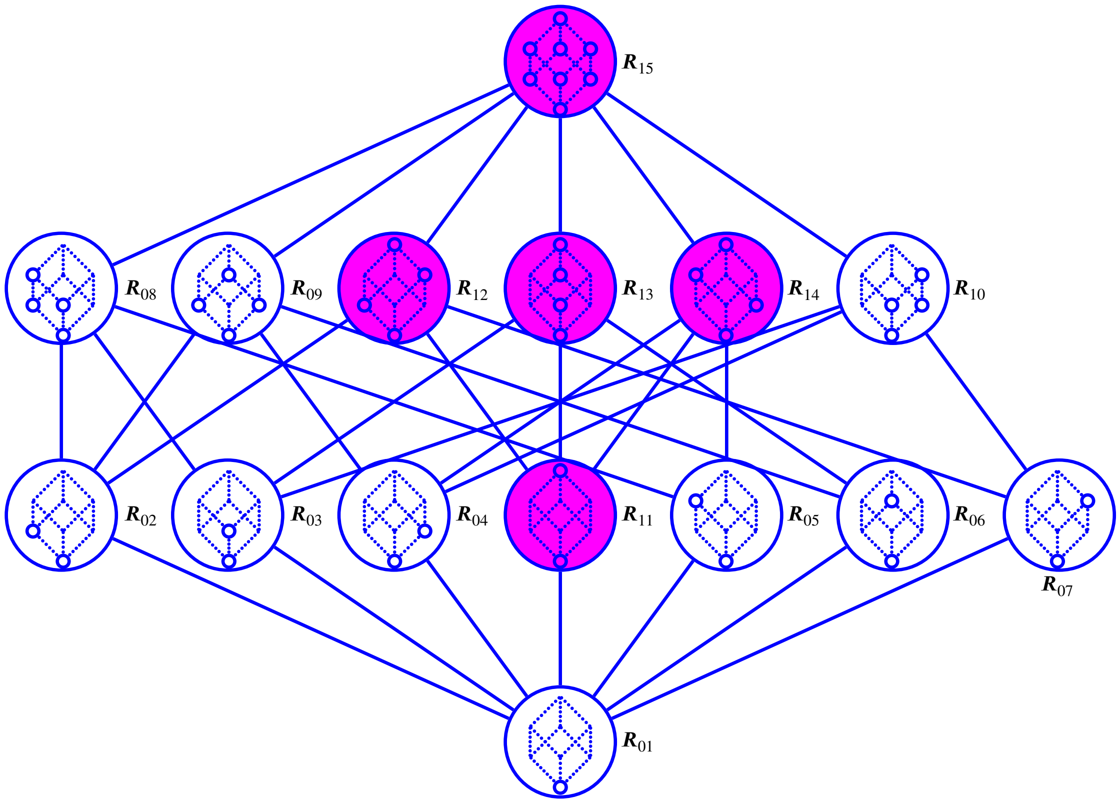


Figure A.6: Lattice of rings of sets on $X \triangleq \{x, y, z\}$ (Example A.10 page 283)

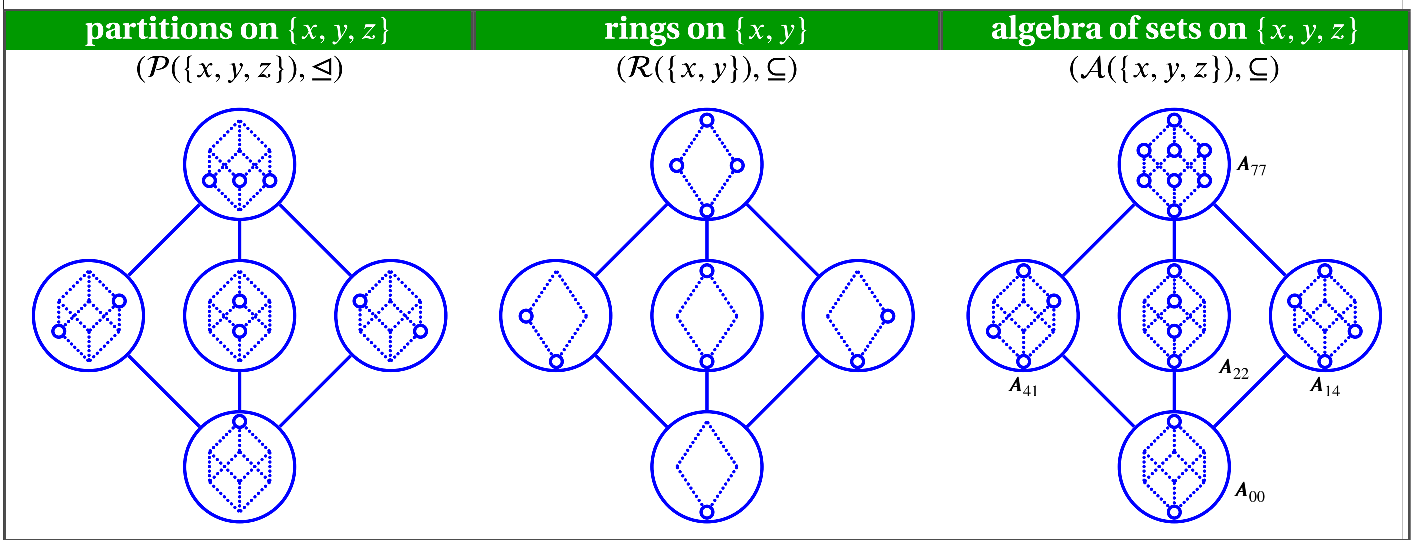


Figure A.7: Lattices of set structures (see Example A.11 (page 285), Example A.3 (page 273), and Example A.9 (page 282))

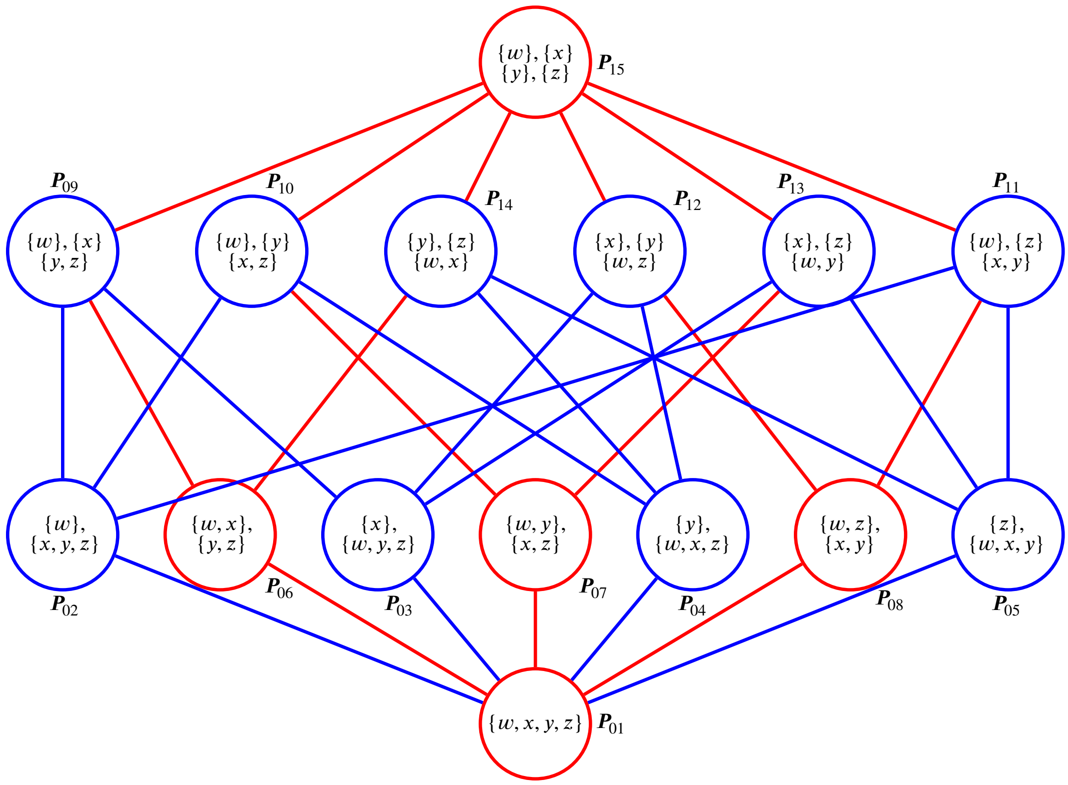


Figure A.8: Lattice of partitions of sets on $X \triangleq \{w, x, y, z\}$ (Example A.12 page 285)

A.5.4 Lattices of partitions of sets

Example A.11. There are a total of **5** partitions of sets on the set $X \triangleq \{x, y, z\}$. These sets are listed in Example A.7 (page 275) and illustrated in Figure A.7 (page 284).

Example A.12. There are a total of **15** partitions of sets on the set $X \triangleq \{w, x, y, z\}$. These sets are listed in Example A.7 (page 275) and illustrated in Figure A.8 (page 285).

In 1946, Philip Whitman proposed an amazing conjecture—that all finite lattices are isomorphic to a lattice of partitions. A proof for this was published some 30 years later by Pavel Pudlák and Jiří Tůma (next theorem).

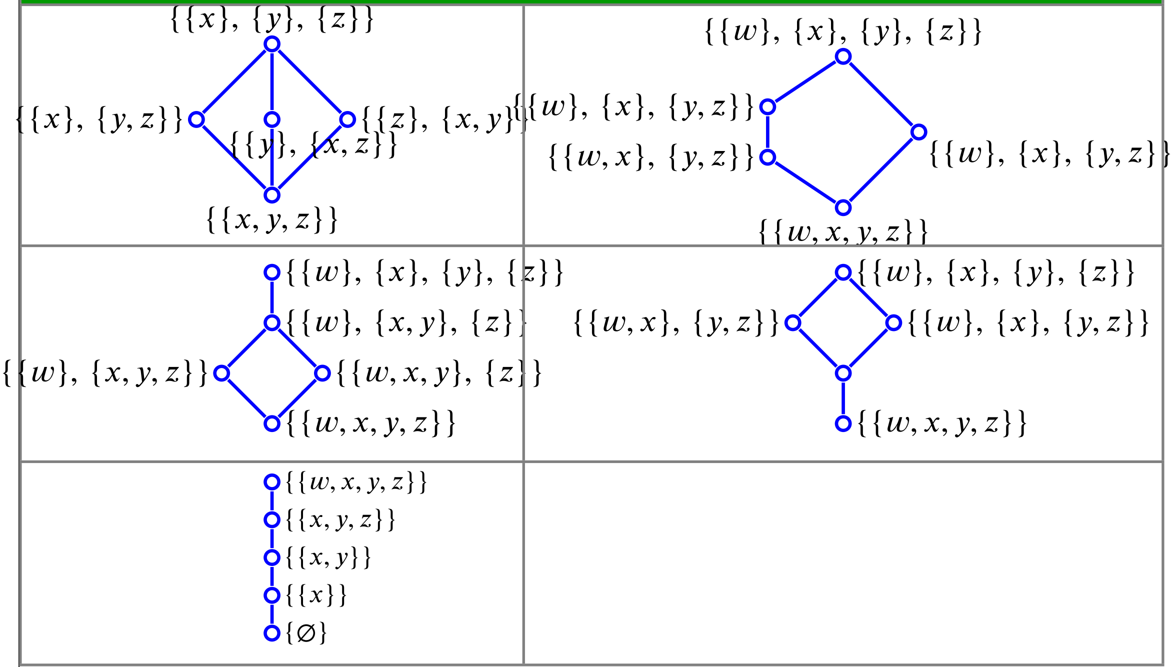
Theorem A.8. [20] *Let* $\mathbf{L}$ *be a lattice.*

THM	
	$\mathbf{L}$ *is* FINITE $\implies$ $\mathbf{L}$ *is isomorphic to a* LATTICE OF PARTITIONS

Example A.13. There are five unlabeled lattices on a five element set as stated in Proposition C.2 (page 311) and illustrated in Example C.11 (page 312). All of these lattices are isomorphic to a lattice of partitions (Theorem A.8 page 285), as illustrated next.

[20] Pudlák and Tůma (1980) ⟨improved proof⟩, Pudlák and Tůma (1977) ⟨proof⟩, Whitman (1946) ⟨conjecture⟩, Saliĭ (1988) page vii ⟨list of lattice theory breakthroughs⟩

lattices on 5 element set as lattices of partitions



A.6 Relationships between set structures

Proposition A.9. [21]

PRP
$$\left\{\begin{array}{l}\boldsymbol{R} \text{ is a } \textbf{ring of sets}\\ \text{on a set } X\end{array}\right\} \implies \left\{\begin{array}{l}\boldsymbol{R}\cup X \text{ is an } \textbf{algebra of sets}\\ \text{on } X\end{array}\right\}$$

Theorem A.9. *Let X be a set.*

THM
$$\left\{\begin{array}{l}\boldsymbol{A} \text{ is an } \textbf{algebra of sets}\\ \text{on } X\end{array}\right\} \iff \left\{\begin{array}{lll}1. & \boldsymbol{A} \text{ is a } \textbf{topology} \text{ on } X & \text{and}\\ 2. & \boldsymbol{A} \text{ is a } \textbf{ring of sets} \text{ on } X & \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\boldsymbol{A} \text{ is an algebra of sets on } X &\implies \boldsymbol{A} \text{ is closed under } \cup, \cap, \mathsf{c}, \setminus, \varnothing, X && \text{by Theorem A.4 page 277}\\
&\implies \left\{\begin{array}{ll}1. & \boldsymbol{A} \text{ is a topology on } X\\ & \text{AND}\\ 2. & \boldsymbol{A} \text{ is a ring of sets on } X\end{array}\right\}
\end{aligned}$$

$$\begin{aligned}
\left\{\begin{array}{ll}1. & \boldsymbol{A} \text{ is a topology on } X\\ & \text{AND}\\ 2. & \boldsymbol{A} \text{ is a ring of sets on } X\end{array}\right\} &\implies \boldsymbol{A} \text{ is closed under } \mathsf{c} \text{ and } \cap && \text{by Theorem A.4 page 277}\\
&\implies \boldsymbol{A} \text{ is a ring of sets}
\end{aligned}$$

⇨

Corollary A.1. *Let X be a set and 2^X the power set of X.*

COR
$$\begin{aligned}&\left\{\boldsymbol{A}\subseteq 2^X \mid \boldsymbol{A} \text{ is an algebra of sets on } X\right\}\\ &\quad = \left\{\boldsymbol{T}\subseteq 2^X \mid \boldsymbol{T} \text{ is a topology on } X\right\} \cap \left\{\boldsymbol{R}\subseteq 2^X \mid \boldsymbol{R} \text{ is a ring of sets on } X\right\}\end{aligned}$$

[21] Berezansky et al. (1996) page 4, Halmos (1950) page 21

✎PROOF:

$$\begin{aligned}
&\{T|T \text{ is a topology}\} \cap \{R|R \text{ is a ring of sets}\} \\
&= \{Y|Y \text{ is a topology AND a ring of sets}\} && \text{by Definition A.5 page 266} \\
&= \{Y|Y \text{ is an algebra of sets}\} && \text{by Theorem A.9 page 286} \\
&= \{A|A \text{ is an algebra of sets}\} && \text{by change of variable}
\end{aligned}$$

✏

Example A.14. Note that the *intersection* of the lattice of topologies on $\{x, y, z\}$ (Figure 1.1 page 11) *and* the lattice of rings of sets on $\{x, y, z\}$ (Figure A.6 page 284) is *equal to* the lattice of algebras of sets on $\{x, y, z\}$ (Figure A.7 page 284).

Proposition A.10. *Let $\mathcal{R}(X)$ be the set of* RINGS OF SETS *(Definition A.10 page 273) and $\mathcal{A}(X)$ the set of* ALGEBRAS OF SETS *(Definition A.9 page 272) on a set X.*

PRP
$$\left\{\begin{array}{ll} 1. & R \text{ is a } \textbf{ring of sets} \quad \text{and} \\ 2. & X \in R \end{array}\right\} \iff \left\{ R \text{ is an } \textbf{algebra of sets} \right\}$$

✎PROOF:

$$\begin{aligned}
A^c &= X \setminus A && \text{by Theorem A.1 page 267} \\
A \cap B &= A \setminus (A \setminus B) && \text{by Theorem A.1 page 267}
\end{aligned}$$

Therefore, $(R \cup X)$ is closed under c and $\cap$, and thus by the definition of algebras of sets (Definition A.9 page 272), $(R \cup X)$ is an algebra of sets.

✏

Definition A.13. [22]

DEF
*The **Borel set** $\mathrm{B}(X, T)$ generated by the topological space (X, T) is the σ-algebra generated by the topology T.*

Example A.15. Suppose we have a dice with the standard six possible outcomes X. Suppose also we construct the following topology T on X, and this in turn generates the following Borel set (σ-algebra) B on X:

EX

$$\begin{aligned}
X &= \{⚀,⚁,⚂,⚃,⚄,⚅\} \\
T &= \left\{ \underbrace{\{\ \}}_{\varnothing}, \underbrace{\{⚀,⚁,⚂,⚃,⚄,⚅\}}_{\Omega}, \underbrace{\{⚀,⚁,⚂,⚃\}}_{\text{first four}}, \underbrace{\{⚃,⚄,⚅\}}_{\text{last three}}, \underbrace{\{⚃\}}_{\{1234\}\cap\{456\}} \right\} \\
B &= \left\{ \underbrace{\{\ \}}_{\varnothing}, \underbrace{\{⚀,⚁,⚂,⚃,⚄,⚅\}}_{\Omega}, \underbrace{\{⚀,⚁,⚂,⚃\}}_{\text{first four}}, \underbrace{\{⚃,⚄,⚅\}}_{\text{last three}}, \underbrace{\{⚃\}}_{\{1234\}\cap\{456\}}, \right. \\
&\qquad \left. \underbrace{\{⚀,⚁,⚂,⚄,⚅\}}_{\{4\}}, \underbrace{\{⚄,⚅\}}_{(\{4\})\cap\{456\}}, \underbrace{\{⚀,⚁,⚂\}}_{\{1234\}\cap\{4\}} \right\}
\end{aligned}$$

Example A.16. There are a total of 29 *topologies* on the set $X \triangleq \{x, y, z\}$ (Theorem 1.2 page 6); and of these, 5 are also *algebras of sets*, 24 are not. Figure A.9 (page 288) illustrates the 24 topologies on the set $\{x, y, z\}$ that are *not* algebras of sets and the 5 algebras of sets that they generate.

[22] Aliprantis and Burkinshaw (1998) page 97

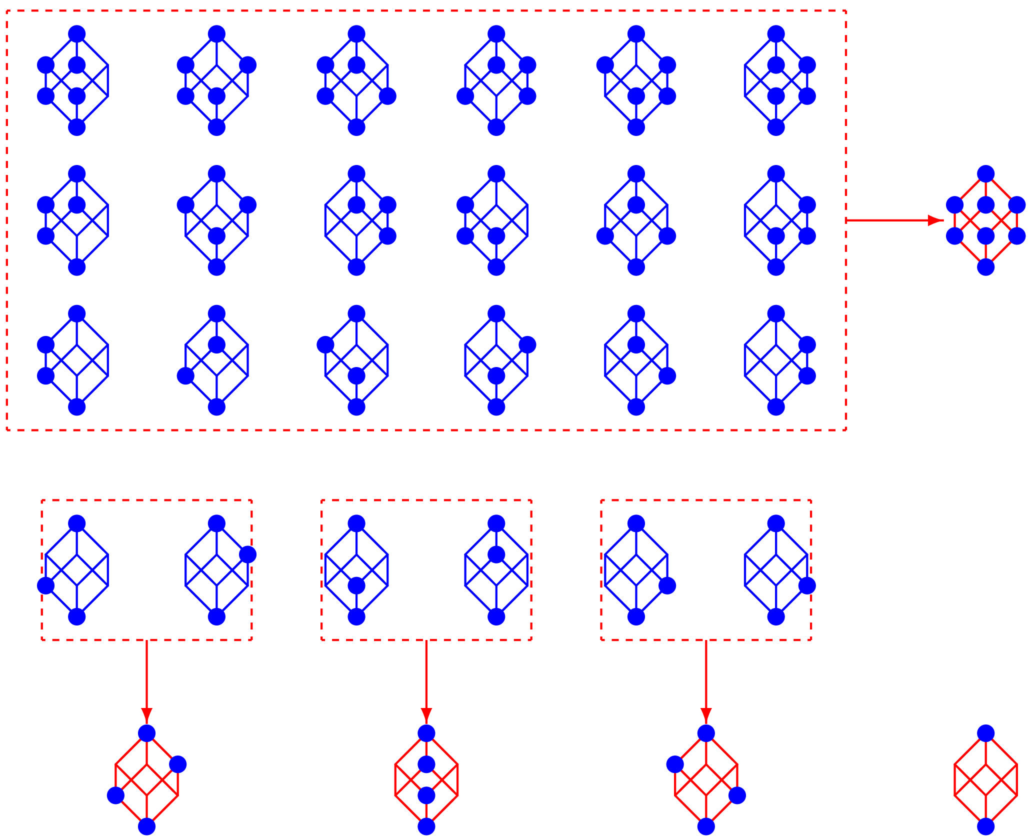


Figure A.9: Algebras of sets generated by topologies on the set $X \triangleq \{x, y, z\}$ (see Example A.16 page 287)

APPENDIX B

ORDER

Equivalence relations require *symmetry* ($x \asymp y \iff y \asymp x$). However another very important type of relation, the *order relation*, actually requires *anti-symmetry*. This chapter presents some useful structures regarding order relations. Ordering relations on a set allow us to *compare* some pairs of elements in a set and determine whether or not one element is *less than* another. In this case, we say that those two elements are *comparable*; otherwise, they are *incomparable*. A set together with an order relation is called an *ordered set*, a *partially ordered set*, or a *poset* (Definition B.2 page 290).

B.1 Preordered sets

Definition B.1. [1] *Let X be a set.*

DEF

*A relation $\sqsubseteq$ is a **preorder relation** on X if*

1. $x \sqsubseteq x \qquad \forall x \in X \qquad$ (REFLEXIVE) *and*
2. $x \sqsubseteq y$ *and* $y \sqsubseteq z \implies x \sqsubseteq z \qquad \forall x,y,z \in X \qquad$ (TRANSITIVE)

*A **preordered set** is the pair $(X, \sqsubseteq)$.*

Example B.1. [2]

EX

$\sqsubseteq$ is a *preorder relation* on the set of *positive integers* $\mathbb{N}$ if

$n \sqsubseteq m \iff (p \text{ is a prime factor of } n \implies p \text{ is a prime factor of } m)$

[1] Schröder (2003) page 115, Brown and Watson (1991), page 317

[2] Shen and Vereshchagin (2002) page 43

B.2 Order relations

Definition B.2. [3] *Let X be a set. Let 2^{XX} be the set of all relations on X.*

DEF

*A relation $\le$ is an **order relation** in 2^{XX} if*

1. $x \le x \quad \forall x \in X$ (REFLEXIVE) and
2. $x \le y$ and $y \le z \implies x \le z \quad \forall x,y,z \in X$ (TRANSITIVE) and
3. $x \le y$ and $y \le x \implies x = y \quad \forall x,y \in X$ (ANTI-SYMMETRIC)

(1 and 2: preorder)

*An **ordered set** is the pair $(X, \le)$. The set X is called the **base set** of $(X, \le)$. If $x \le y$ or $y \le x$, then elements x and y are said to be **comparable**, denoted $x \sim y$. Otherwise they are **incomparable**, denoted $x \| y$. The relation $\lneq$ is the relation $\le \setminus =$ ("less than but not equal to"), where $\setminus$ is the SET DIFFERENCE operator, and $=$ is the equality relation. An order relation is also called a **partial order relation**. An ordered set is also called a **partially ordered set** or **poset**.*

The familiar relations $\ge$, $<$, and $>$ (next) can be defined in terms of the order relation $\le$ (Definition B.2—previous).

Definition B.3. [4] *Let $(X, \le)$ be an ordered set.*

DEF

The relations $\ge$, $<$, $> \in 2^{XX}$ are defined as follows:

$$x \ge y \overset{\text{def}}{\iff} y \le x \qquad \forall x,y \in X$$

$$x \lneq y \overset{\text{def}}{\iff} x \le y \text{ and } x \ne y \qquad \forall x,y \in X$$

$$x \gneq y \overset{\text{def}}{\iff} x \ge y \text{ and } x \ne y \qquad \forall x,y \in X$$

*The relation $\ge$ is called the **dual** of $\le$.*

Theorem B.1. [5] *Let X be a set.*

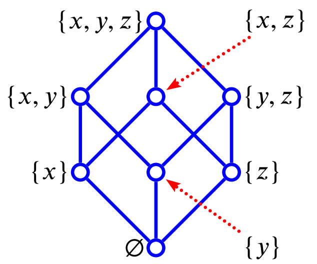
THM

$$(X, \le) \text{ is an ordered set} \iff (X, \ge) \text{ is an ordered set}$$

Example B.2.

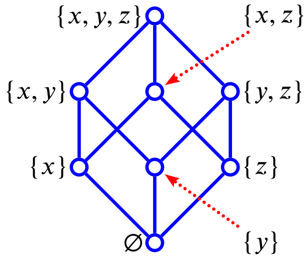
EX

	order relation		dual order relation
$\le$	(integer less than or equal to)	$\ge$	(integer greater than or equal to)
$\subseteq$	(subset)	$\supseteq$	(super set)
$\mid$	(divides)		(divided by)
$\implies$	(implies)	$\impliedby$	(implied by)



Example B.3. The Hasse diagram to the left illustrates the ordered set $\left(2^{\{x,y,z\}}, \subseteq\right)$ and the Hasse diagram to the right illustrates its dual $\left(2^{\{x,y,z\}}, \supseteq\right)$.

[3] MacLane and Birkhoff (1999) page 470, Beran (1985) page 1, Korselt (1894) page 156 ⟨I, II, (1)⟩, Dedekind (1900) page 373 ⟨I–III⟩
[4] Peirce (1880) page 2
[5] Grätzer (1998), page 3



B.3 Linearly ordered sets

In an ordered set we can say that some element is less than or equal to some other element. That is, we can say that these two elements are *comparable*—we can *compare* them to see which one is lesser or equal to the other. But it is very possible that there are two elements that are not comparable, or *incomparable*. That is, we cannot say that one element is less than the other—it is simply not possible to compare them because their ordered pair is not an element of the order relation.

For example, in the ordered set $(2^{\{x,y,z\}}, \subseteq)$ of Example B.9, we can say that $\{x\} \subseteq \{x,z\}$ (we can compare these two sets with respect to the order relation $\subseteq$), but we cannot say $\{y\} \subseteq \{x,z\}$, nor can we say $\{x,z\} \subseteq \{y\}$. Rather, these two elements $\{y\}$ and $\{x,z\}$ are simply *incomparable*.

However, there are some ordered sets in which every element is comparable with every other element; and in this special case we say that this ordered set is a *totally ordered* set or is *linearly ordered* (next definition).

Definition B.4. [6]

DEF

A relation $\le$ is a ***linear order relation*** *on X if*

1. $\le$ *is an* ORDER RELATION (*Definition B.2 page 290*) *and*
2. $x \le y$ *or* $y \le x \quad \forall x,y \in X$ (COMPARABLE).

A ***linearly ordered set*** *is the pair $(X, \le)$.*
A linearly ordered set is also called a ***totally ordered set***, *a* ***fully ordered set***, *and a* ***chain***.

Definition B.5 (poset product). [7]

DEF

The ***product*** *$P \times Q$ of ordered pairs $P \triangleq (X, \precsim)$ and $Q \triangleq (Y, \trianglelefteq)$ is the ordered pair $(X \times Y, \le)$ where*

$$(x_1, y_1) \le (x_2, y_2) \quad \stackrel{\text{def}}{\iff} \quad x_1 \precsim x_2 \text{ and } y_1 \trianglelefteq y_2 \qquad \forall x_1,x_2 \in X;\ y_1,y_2 \in Y$$

B.4 Representation

Definition B.6. [8]

DEF

y ***covers*** *x in the ordered set $(X, \le)$ if*

1. $x \le y$ (*y is greater than x*) *and*
2. $(x \le z \le y) \implies (z = x \text{ or } z = y)$ (*there is no element between x and y*).

The case in which y covers x is denoted

$x \prec y$.

[6] MacLane and Birkhoff (1999) page 470, Ore (1935) page 410
[7] Birkhoff (1948) page 7, MacLane and Birkhoff (1967), page 489
[8] Birkhoff (1933a) page 445

Example B.4. Let $(\{x, y, z\}, \leq)$ be an ordered set with cover relation $\prec$.

EX $\{x < y < z\} \implies \left\{\begin{array}{lll} y & \text{covers} & x \\ z & \text{covers} & y \\ z & \text{does } \textbf{not} \text{ cover} & x \end{array}\right\}$

An ordered set can be represented in four ways:

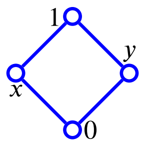
1. Hasse diagram
2. tables
3. set of ordered pairs of order relations
4. set of ordered pairs of cover relations

Definition B.7. *Let $(X, \leq)$ be an ordered pair.*

DEF *A diagram is a* ***Hasse diagram*** *of $(X, \leq)$ if it satisfies the following criteria:*
- *Each element in X is represented by a dot or small circle.*
- *For each $x, y \in X$, if $x \prec y$, then y appears at a higher position than x and a line connects x and y.*

Example B.5. Here are three ways of representing the ordered set $\left(2^{\{x,y\}}, \subseteq\right)$;

1. **Hasse diagrams**: If two elements are comparable, then the lesser of the two is drawn lower on the page than the other with a line connecting them.

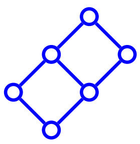


2. Sets of ordered pairs specifying *order relations* (Definition B.2 page 290):

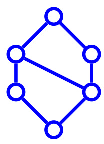
$$\subseteq = \left\{\begin{array}{lllll} (\varnothing, \varnothing), & (\{x\}, \{x\}), & (\{y\}, \{y\}), & (\{x, y\}, \{x, y\}), & \\ (\varnothing, \{x\}), & (\varnothing, \{y\}), & (\varnothing, \{x, y\}), & (\{x\}, \{x, y\}), & (\{y\}, \{x, y\}) \end{array}\right\}$$

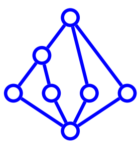
3. Sets of ordered pairs specifying *covering relations*:

$$\prec = \left\{ (\varnothing, \{x\}), \ (\varnothing, \{y\}), \ (\{x\}, \{x, y\}), (\{y\}, \{x, y\}) \right\}$$

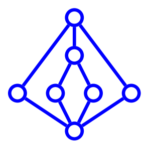


Example B.6. The Hasse diagrams to the left and right represent *equivalent* ordered sets. They are simply drawn differently.

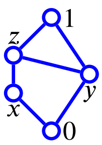




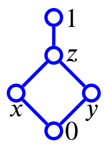
Example B.7. The Hasse diagrams to the left and right represent *equivalent* ordered sets. They are simply drawn differently.



Example B.8. The Hasse diagrams to the left and right represent *equivalent* ordered sets.



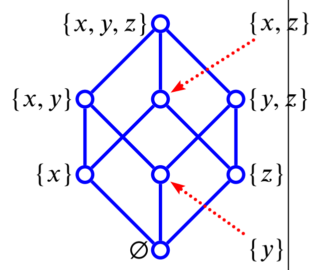
In particular, the line extending from 1 to y in the diagram to the left is redundant because other lines already indicate that $z \leq 1$ and $y \leq z$; and thus by the *transitive* property (Definition B.2 page 290), these two relations imply $1 \leq y$. A more concise explanation is that both have the same convering relation:



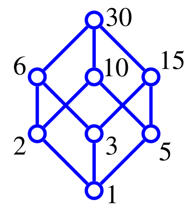
$$\prec = \{(z, 1), (x, z), (0, x), (y, z), (0, y)\}$$

B.5 Examples

Examples of order relations include the following:



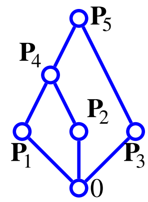
Example B.9 (Set inclusion order relation). [9]Let X be a set, 2^X the power set of X, and $\subseteq$ the set inclusion relation. Then, $\subseteq$ is an *order relation* on the set 2^X and the pair $(2^X, \subseteq)$ is an ordered set. The ordered set $(2^{\{x,y,z\}}, \subseteq)$ is illustrated to the right by its *Hasse diagram*.



Example B.10 (Integer divides order relation). [10]Let $|$ be the "divides" relation on the set $\mathbb{N}$ of positive integers such that $n|m$ represents m divides n. Then $|$ is an *order relation* on $\mathbb{N}$ and the pair $(\mathbb{N}, |)$ is an *ordered set*. The ordered set $(\{n \in \mathbb{N} | n|2 \text{ or } n|3 \text{ or } n|5\}, |)$ is illustrated by a *Hasse diagram* to the right.

Example B.11 (Operator order relation). [11] Let $\boldsymbol{X}$ be an inner-product space. We can define the order relation $\lessapprox$ on the linear operators $\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3 \ldots \in X^X$ as follows:

$$\mathbf{L}_1 \lessapprox \mathbf{L}_2 \quad \stackrel{\text{def}}{\iff} \quad \langle \mathbf{L}_2\boldsymbol{x} - \mathbf{L}_1\boldsymbol{x} \,|\, \boldsymbol{x} \rangle \geq 0 \quad \forall \boldsymbol{x} \in \boldsymbol{X}$$



Example B.12 (Projection operator order relation). [12]Let $(\!(\boldsymbol{V}_n)\!)$ be a sequence of subspaces in a Hilbert space $\boldsymbol{X}$. We can define a projection operator $\mathbf{P}_n$ for every subspace $\boldsymbol{V}_n \subseteq \boldsymbol{X}$ in a subspace lattice such that

$$\boldsymbol{V}_n = \mathbf{P}_n \boldsymbol{X} \qquad \forall n \in \mathbb{Z}.$$

Each projection operator $\mathbf{P}_n$ in the lattice "projects" the range space $\boldsymbol{X}$ onto a subspace $\boldsymbol{V}_n$. We can define an order relation on the projection operators as follows:

$$\mathbf{P}_1 \leq \mathbf{P}_2 \quad \stackrel{\text{def}}{\iff} \quad \mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_2\mathbf{P}_1 = \mathbf{P}_1$$

---

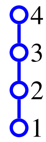
[9] Menini and Oystaeyen (2004) pages 56–57

[10] MacLane and Birkhoff (1999) page 484, Sheffer (1920) page 310 ⟨footnote 1⟩

[11] Michel and Herget (1993) page 429, Pedersen (2000) page 87

[12] Isham (1999) pages 21–22, Dunford and Schwartz (1957), page 481, Svozil (1994) page 72

Example B.13 (Integer order relation). Let $\le$ be the standard order relation on the set of integers $\mathbb{Z}$. Then the ordered pair $(\mathbb{Z}, \le)$ is a totally ordered set. The totally ordered set $(\{1,2,3,4\}, \le)$ is illustrated to the right. Other familiar examples of totally ordered sets include the pair $(\mathbb{Q}, \le)$ (where $\mathbb{Q}$ is the set of rational numbers) and $(\mathbb{R}, \le)$ (where $\mathbb{R}$ is the set of real numbers).



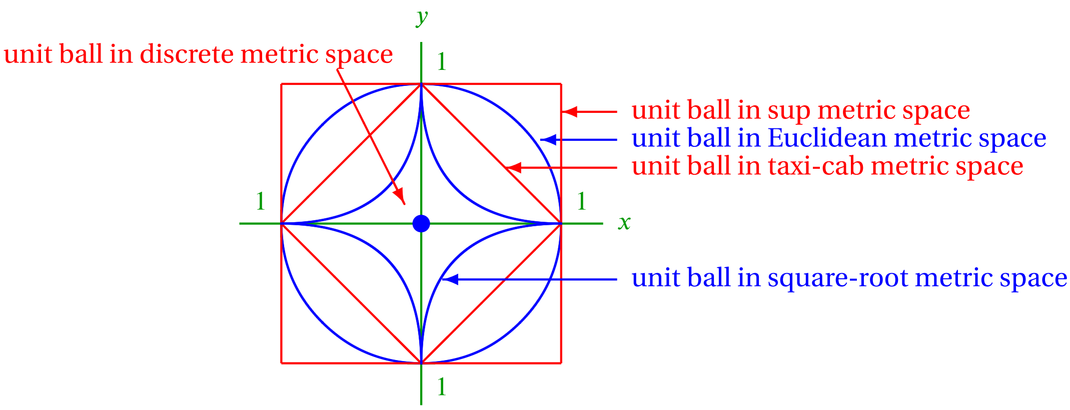
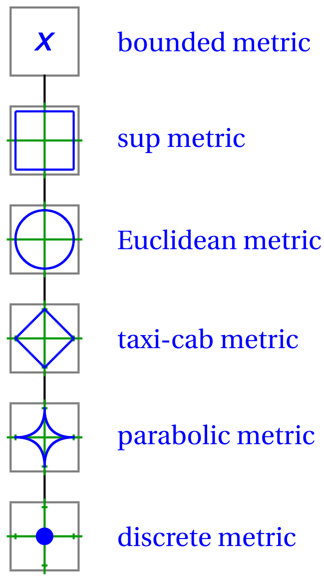


Figure B.1: Balls on the set $\mathbb{R}^2$ using different metrics

Example B.14 (Metric order relation). [13] Let d_n be a metric on the set X and B_n be the unit ball centered at "0" in the metric space (X, d_n). Define an order relation $\le$ on the set of metric spaces $\{(X, \mathsf{d}_n) \mid n = 1, 2, \ldots\}$ such that

$$(X, \mathsf{d}_n) \le (X, \mathsf{d}_m) \quad\iff\quad \mathsf{B}_n \subseteq \mathsf{B}_m.$$

The the tuple $(\{(X, \mathsf{d}_n) \mid n = 1, 2, \ldots\}, \le)$ is an ordered set. The ordered set of several common metric spaces is a *totally ordered* set, as illustrated to the right and with associated unit balls illustrated in Figure B.1 (page 294).



Example B.15 (Coordinatewise order relation). [14] Let $(X, \le)$ be an ordered set.
Let $\boldsymbol{x} \triangleq (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} \triangleq (y_1, y_2, \ldots, y_n)$.

EX The **coordinatewise order relation** $\lesssim$ on the Cartesian product X^n is defined for all $\boldsymbol{x}, \boldsymbol{y} \in X^n$ as

$$\boldsymbol{x} \lesssim \boldsymbol{y} \quad\overset{\text{def}}{\iff}\quad \{x_1 \le y_1 \text{ and } x_2 \le y_2 \text{ and } \ldots \text{ and } x_n \le y_n\}$$

Example B.16 (Lexicographical order relation). [15] Let $(X, \le)$ be an ordered set.
Let $\boldsymbol{x} \triangleq (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} \triangleq (y_1, y_2, \ldots, y_n)$.

[13] Michel and Herget (1993) page 354, Giles (1987) page 29
[14] Shen and Vereshchagin (2002) page 43
[15] Shen and Vereshchagin (2002) page 44, Halmos (1960) page 58, Hausdorff (1937) page 54

EX The **lexicographical order relation** $\precsim$ on the Cartesian product X^n is defined forall $\boldsymbol{x}, \boldsymbol{y} \in X^n$ as

$$\boldsymbol{x} \precsim \boldsymbol{y} \overset{\text{def}}{\iff} \left\{\begin{array}{llll} (\; x_1 < y_1 & & &) \text{ or} \\ (\; x_2 < y_2 & \text{and} & x_1 = y_1 &) \text{ or} \\ (\; x_3 < y_3 & \text{and} & (x_1, x_2) = (y_1, y_2) &) \text{ or} \\ (\; \cdots & \cdots & \cdots &) \text{ or} \\ (\; x_{n-1} < y_{n-1} & \text{and} & (x_1, x_2, \ldots, x_{n-2}) = (y_1, y_2, \ldots, y_{n-2}) &) \text{ or} \\ (\; x_n \le y_n & \text{and} & (x_1, x_2, \ldots, x_{n-1}) = (y_1, y_2, \ldots, y_{n-1}) &) \end{array}\right\}$$

The lexicographical order relation is also called the **dictionary order relation** or **alphabetic order relation**.

Definition B.8.

DEF *An ordered set is* ***labeled*** *if the labels on the elements are significant.*
An ordered set is ***unlabeled*** *if the labels on the elements are not significant.*

Proposition B.1. [16] *Let X_n be a finite set with order $n = |X_n|$. Let P_n be the number of labeled ordered sets on X_n and p_n the number of unlabeled ordered sets.*

PRP

n	0	1	2	3	4	5	6	7	8	9
P_n	1	1	3	19	219	4231	130,023	6,129,859	431,723,379	44,511,042,511
p_n	1	1	2	5	16	63	318	2045	16,999	183,231

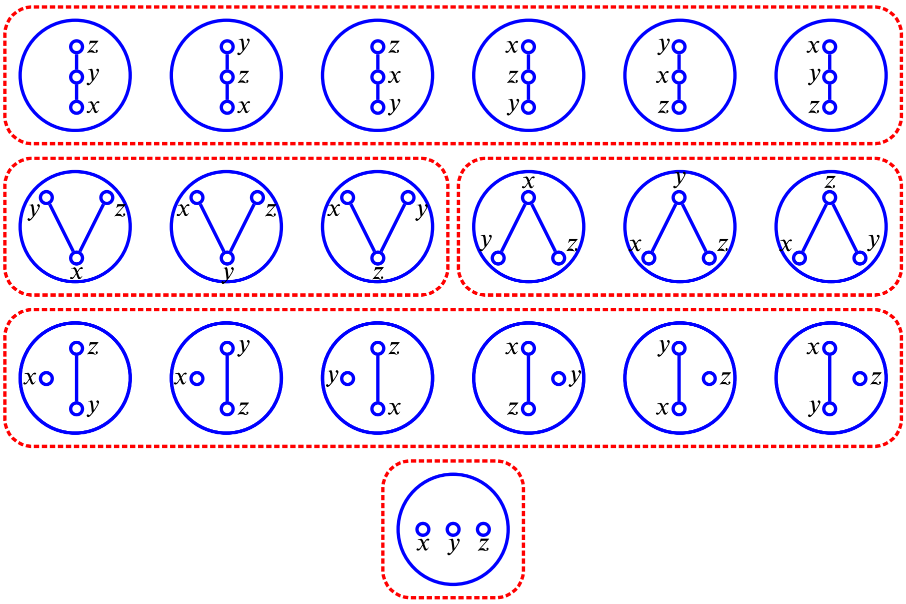


Figure B.2: All possible orderings of the set $\{x, y, z\}$ (Example B.17 page 295).

Example B.17. Proposition B.1 (page 295) indicates that there are exactly 19 labeled order relations on the set $\{x, y, z\}$ and 5 unlabeled order relations.

EX The 19 labeled order relations on $\{x, y, z,\}$ are represented here using three methods:

1. Hasse diagrams: Figure B.2 page 295
2. order relations: Table B.2 page 296
3. covering relations: Table B.3 page 296

In each of these three methods, the 19 *labeled* order relations are arranged into 5 groups, each group representing one of the 5 *unlabeled* order relations.

[16] Sloane (2014) ⟨http://oeis.org/A001035⟩, Sloane (2014) ⟨http://oeis.org/A000112⟩, Comtet (1974) page 60, Brinkmann and McKay (2002)

labeled order relations on $\{x, y, z\}$							
$\le_1$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)$				$\}$
$\le_2$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	(y,z)			$\}$
$\le_3$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	(z,y)			$\}$
$\le_4$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	(x,z)			$\}$
$\le_5$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	(z,x)			$\}$
$\le_6$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	(x,y)			$\}$
$\le_7$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	(y,x)			$\}$
$\le_8$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(x,y)\,,$	(x,z)		$\}$
$\le_9$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(x,y)\,,$	(y,z)		$\}$
$\le_{10}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(z,x)\,,$	(z,y)		$\}$
$\le_{11}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(y,x)\,,$	(z,x)		$\}$
$\le_{12}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(x,y)\,,$	(z,y)		$\}$
$\le_{13}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(x,z)\,,$	(y,z)		$\}$
$\le_{14}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(x,y)\,,$	$(y,z)\,,$	(x,z)	$\}$
$\le_{15}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(x,z)\,,$	$(x,y)\,,$	(z,y)	$\}$
$\le_{16}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(y,x)\,,$	$(y,z)\,,$	(x,z)	$\}$
$\le_{17}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(y,z)\,,$	$(y,x)\,,$	(z,x)	$\}$
$\le_{18}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(z,x)\,,$	$(z,y)\,,$	(x,y)	$\}$
$\le_{19}$	$=$	$\{$	$(x,x)\,,\,(y,y)\,,\,(z,z)\,,$	$(z,y)\,,$	$(z,x)\,,$	(y,x)	$\}$

Table B.2: labeled order relations on $\{x, y, z\}$

labeled cover relations on $\{x, y, z\}$												
$\prec_1$	$=$	$\varnothing$				$\prec_{11}$	$=$	$\{$	$(y,x)\,,$	(z,x)	$\}$
$\prec_2$	$=$	$\{$	(y,z)		$\}$	$\prec_{12}$	$=$	$\{$	$(x,y)\,,$	(z,y)	$\}$
$\prec_3$	$=$	$\{$	(z,y)		$\}$	$\prec_{13}$	$=$	$\{$	$(x,z)\,,$	(y,z)	$\}$
$\prec_4$	$=$	$\{$	(x,z)		$\}$	$\prec_{14}$	$=$	$\{$	$(x,y)\,,$	(y,z)	$\}$
$\prec_5$	$=$	$\{$	(z,x)		$\}$	$\prec_{15}$	$=$	$\{$	$(x,z)\,,$	(x,y)	$\}$
$\prec_6$	$=$	$\{$	(x,y)		$\}$	$\prec_{16}$	$=$	$\{$	$(y,x)\,,$	(y,z)	$\}$
$\prec_7$	$=$	$\{$	(y,x)		$\}$	$\prec_{17}$	$=$	$\{$	$(y,z)\,,$	(y,x)	$\}$
$\prec_8$	$=$	$\{$	$(x,y)\,,$	(x,z)	$\}$	$\prec_{18}$	$=$	$\{$	$(z,x)\,,$	(z,y)	$\}$
$\prec_9$	$=$	$\{$	$(x,y)\,,$	(y,z)	$\}$	$\prec_{19}$	$=$	$\{$	$(z,y)\,,$	(z,x)	$\}$
$\prec_{10}$	$=$	$\{$	$(z,x)\,,$	(z,y)	$\}$						

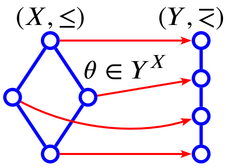
Table B.3: labeled cover relations on $\{x, y, z\}$

B.6 Functions on ordered sets

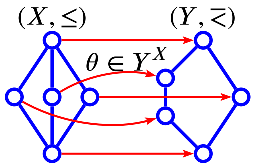
Definition B.9. [17] *Let $(X, \leq)$ and $(Y, \precsim)$ be ordered sets.*

*A function $\theta \in Y^X$ is **order preserving** with respect to $\leq$ and $\precsim$ if*

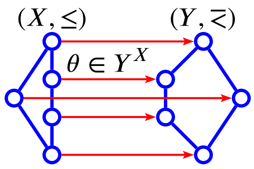
$$x \leq y \implies \theta(x) \precsim \theta(y) \qquad \forall x, y \in X.$$



Example B.18. [18] In the diagram to the right, the function $\theta \in Y^X$ is *order preserving* with respect to $\leq$ and $\precsim$. Note that θ^{-1} is *not* order preserving. This example also illustrates the fact that that order preserving does not imply *isomorphic*.



Example B.19. In the diagram to the right, the function $\theta \in Y^X$ is *order preserving* with respect to $\leq$ and $\precsim$. Note that θ^{-1} is *not* order preserving. Like Example B.18 (page 297), this example also illustrates the fact that that order preserving does not imply *isomorphic*.



Example B.20. In the diagram to the right, the function $\theta \in Y^X$ is *order preserving* with respect to $\leq$ and $\precsim$. Note that θ^{-1} *is also* order preserving. In this case, θ is an *isomorphism* and the ordered sets $(X, \leq)$ and $(Y, \precsim)$ are *isomorphic*.

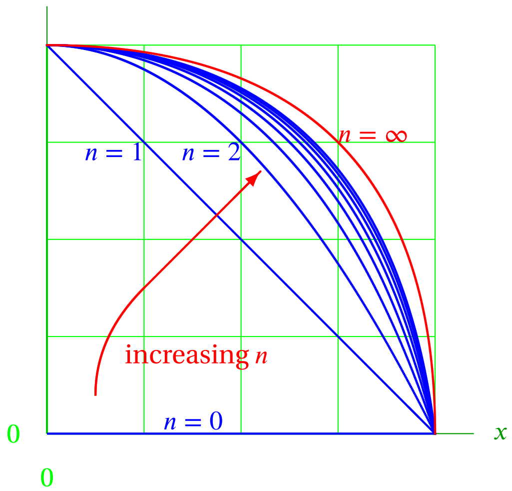
Example B.21. [19]

The function $\mathsf{f}(x) \triangleq \dfrac{x}{1-x^2}$ in $\mathbb{R}^{(-1:1)}$ is *bijective* and *order preserving*.

Theorem B.2 (Pointwise ordering relation). [20] *Let X be a set, $(Y, \leq)$ an ordered set, and $\mathsf{f}, \mathsf{g} \in Y^X$.*

$$\mathsf{f}(x) \leq \mathsf{g}(x) \forall x \in X \implies \left(Y^X, \precsim\right) \text{ is an ordered set.}$$

In this case we say f is "dominated by" g in X, or we say g "dominates" f in X.



Example B.22 (Pointwise ordering relation). [21] Let $\mathsf{f} \precsim \mathsf{g}$ represent that $\mathsf{f}(x) \leq \mathsf{g}(x)$ for all $0 \leq x \leq 1$ (we say f is "dominated by" g in the region $[0,1]$, or we say g "dominates" f in the region $[0,1]$). The pair $\left(\left\{\mathsf{f}_n(x) = 1 - x^n | n \in \mathbb{N}\right\}, \precsim\right)$ is a totally ordered set.

---

[17] Burris and Sankappanavar (2000), page 10
[18] Burris and Sankappanavar (2000), page 10
[19] Munkres (2000) page 25 ⟨Example 1§3.9⟩
[20] Shen and Vereshchagin (2002), page 43, Giles (2000), page 252
[21] Shen and Vereshchagin (2002), page 43, Giles (2000), page 252, Aliprantis and Burkinshaw (2006) page 2

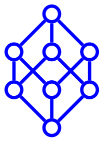
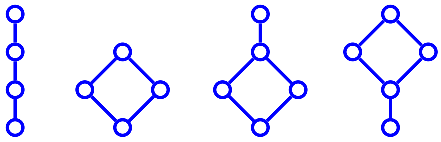
B.7 Decomposition

B.7.1 Subposets

Definition B.10. [22]

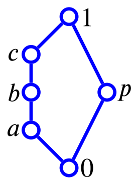
DEF *The tupple $(Y, \lesssim)$ is a **subposet** of the ordered set $(X, \le)$ if*
1. $Y \subseteq X$ *(Y is a subset of X)* *and*
2. $\lesssim = \le \cap Y^2$ *($\lesssim$ is the relation $\le$ restricted to $Y \times Y$)*

Example B.23.

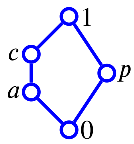
Subposets of  include 

Example B.24. Let

$$(X, \le) \triangleq \Big(\{0, a, b, c, p, 1\}, \quad \big\{ (0,0), (a,a), (b,b), (c,c), (p,p), (1,1), \\ (0,a), (0,b), (0,c), (0,p), (0,1), \\ (a,b), (a,c), (a,1), (p,1), \\ (b,c), (b,1), (c,1), (p,1) \big\} \Big)$$



$$(Y, \lesssim) \triangleq \Big(\{0, a, c, p, 1\}, \quad \big\{ (0,0), (a,a), (c,c), (p,p), (1,1), \\ (0,a), (0,c), (0,p), (0,1), \\ (a,c), (a,1), (p,1), (c,1), (p,1) \big\} \Big).$$

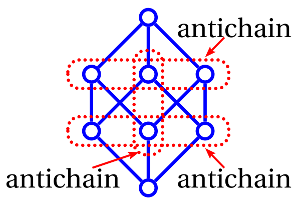


Then $(Y, \lesssim)$ is a subposet of $(X, \le)$ because $Y \subseteq X$ and $\lesssim = (\le \cap Y^2)$.

A *chain* is an ordered set in which every pair of elements is *comparable* (Definition B.4 page 291). An *antichain* is just the opposite—it is an ordered set in which *no* pair of elements is comparable (next definition).

Definition B.11. [23]

DEF *The subposet $(A, \le)$ in the ordered set $(X, \le)$ is an **antichain** if*
$a \| b \qquad \forall a, b \in A$
(all elements in A are INCOMPARABLE*).*



Definition B.12. [24]

DEF
- *The **length** of a chain $(C, \le)$ equals $|C| - 1$.*
- *The **length** of a poset $(X, \le)$ is the length of the longest chain in the ordered set.*
- *The **width** of a poset $(X, \le)$ is number of elements in the largest antichain in the ordered set.*

Theorem B.3 (Dilworth's theorem). [25] *Let $(X, \le)$ be an ordered set with width n.*

---

[22] Grätzer (2003) page 2
[23] Grätzer (2003) page 2
[24] Grätzer (2003) page 2, Birkhoff (1967) page 5
[25] Dilworth (1950a) page 161, Dilworth (1950b), Farley (1997) page 4

THM $\left\{\begin{array}{l}\text{WIDTH } n \text{ of } (X, \leq) \\ \text{is FINITE}\end{array}\right\} \implies \left\{\begin{array}{ll}1. & \text{there exists a PARTITION of } (X, \leq) \text{ into } n \text{ chains} \quad \text{and} \\ 2. & \text{there does not exist any PARTITION} \\ & \text{of } (X, \leq) \text{ into less than } n \text{ chains}\end{array}\right\}$

B.7.2 Operations on posets

Definition B.13. [26] *Let X and Y be disjoint sets. Let $\boldsymbol{P} \triangleq (X, \lessapprox)$ and $\boldsymbol{Q} \triangleq (Y, \trianglelefteq)$ be ordered sets on X and Y.*

DEF *The **direct sum** of $\boldsymbol{P}$ and $\boldsymbol{Q}$ is defined as*

$$\boldsymbol{P} + \boldsymbol{Q} \triangleq (X \cup Y, \leq)$$

where $x \leq y$ if

1. $x, y \in X$ *and* $x \lessapprox y$ *or*
2. $x, y \in Y$ *and* $x \trianglelefteq y$

*The direct sum operation is also called the **disjoint union**. The notation $n\boldsymbol{P}$ is defined as*

$$n\boldsymbol{P} \triangleq \underbrace{\boldsymbol{P} + \boldsymbol{P} + \cdots + \boldsymbol{P}}_{n-1 \text{ “}+\text{” operations}}.$$

Definition B.14. [27] *Let X and Y be disjoint sets. Let $\boldsymbol{P} \triangleq (X, \lessapprox)$ and $\boldsymbol{Q} \triangleq (Y, \trianglelefteq)$ be ordered sets on X and Y.*

DEF *The **direct product** of $\boldsymbol{P}$ and $\boldsymbol{Q}$ is defined as*

$$\boldsymbol{P} \times \boldsymbol{Q} \triangleq (X \times Y, \leq)$$

where $(x_1, y_1) \leq (x_2, y_2)$ if $x_1 \lessapprox x_2$ and $y_1 \lessapprox y_2$.

*The direct product operation is also called the **cartesian product**. The order relation $\leq$ is called a **coordinate wise order relation**. The notation $\boldsymbol{P}^n$ is defined as*

$$\boldsymbol{P}^n \triangleq \underbrace{\boldsymbol{P} \times \boldsymbol{P} \times \cdots \times \boldsymbol{P}}_{n-1 \text{ “}\times\text{” operations}}.$$

Definition B.15. [28] *Let X and Y be disjoint sets. Let $\boldsymbol{P} \triangleq (X, \lessapprox)$ and $\boldsymbol{Q} \triangleq (Y, \trianglelefteq)$ be ordered sets on X and Y.*

DEF *The **ordinal sum** of $\boldsymbol{P}$ and $\boldsymbol{Q}$ is defined as*

$$\boldsymbol{P} \oplus \boldsymbol{Q} \triangleq (X \cup Y, \leq)$$

where $x \leq y$ if

1. $x, y \in X$ *and* $x \lessapprox y$ *or*
2. $x, y \in Y$ *and* $x \trianglelefteq y$ *or*
3. $x \in X$ *and* $y \in Y$.

Definition B.16. [29] *Let X and Y be disjoint sets. Let $\boldsymbol{P} \triangleq (X, \lessapprox)$ and $\boldsymbol{Q} \triangleq (Y, \trianglelefteq)$ be ordered sets on X and Y.*

DEF *The **ordinal product** of $\boldsymbol{P}$ and $\boldsymbol{Q}$ is defined as*

$$\boldsymbol{P} \otimes \boldsymbol{Q} \triangleq (X \times Y, \leq)$$

where $(x_1, y_1) \leq (x_2, y_2)$ if

1. $x_1 \neq x_2$ *and* $x_1 \lessapprox x_2$ *or*
2. $x_1 = x_2$ *and* $y_1 \trianglelefteq y_2$

*The order relation $\leq$ is called a **lexicographical** order relation, **dictionary order relation**, or **alphabetic order relation**.*

---

[26] Stanley (1997) page 100

[27] Stanley (1997) pages 100–101, Shen and Vereshchagin (2002) page 43

[28] Stanley (1997) page 100

[29] Stanley (1997) page 101, Shen and Vereshchagin (2002) page 44, Halmos (1960) page 58, Hausdorff (1937) page 54

Definition B.17. [30] *Let $\mathbf{P} \triangleq (X, \leq)$ be an ordered set. Let $\geq$ be the dual order relation of $\leq$.*

DEF: *The **dual** of $\mathbf{P}$ is defined as*
$$\mathbf{P}^* \triangleq (X, \geq)$$

Definition B.18. [31] *Let X and Y be disjoint sets. Let $\mathbf{P} \triangleq (X, \precsim)$ and $\mathbf{Q} \triangleq (Y, \trianglelefteq)$ be ordered sets on X and Y.*

DEF: *The **ordinal product** of $\mathbf{P}$ and $\mathbf{Q}$ is defined as*
$$\mathbf{Q}^{\mathbf{P}} \triangleq \left(\left\{ \mathsf{f} \in Y^X \middle| \mathsf{f} \text{ is ORDER PRESERVING} \right\}, \leq \right)$$
where $\mathsf{f} \leq \mathsf{g}$ iff $\mathsf{f}(x) \leq \mathsf{g}(x) \quad \forall x \in X$.
*The order relation $\leq$ is called a **pointwise order relation*** (Example B.22 page 297).

Theorem B.4 (cardinal arithmetic). [32] *Let $\mathbf{P} \triangleq (X, \leq)$ be an ordered set.*

THM:

1.	$\mathbf{P}+\mathbf{Q}$	$=$	$\mathbf{Q}+\mathbf{P}$	*commutative*
2.	$\mathbf{P}\times\mathbf{Q}$	$=$	$\mathbf{Q}\times\mathbf{P}$	*commutative*
3.	$(\mathbf{P}+\mathbf{Q})+(\mathbb{R}, \leq)$	$=$	$\mathbf{P}+(\mathbf{Q}+(\mathbb{R}, \leq))$	*associative*
4.	$(\mathbf{P}\times\mathbf{Q})\times(\mathbb{R}, \leq)$	$=$	$\mathbf{P}\times(\mathbf{Q}\times(\mathbb{R}, \leq))$	*associative*
5.	$\mathbf{P}\times(\mathbf{Q}+(\mathbb{R}, \leq))$	$=$	$(\mathbf{P}\times\mathbf{Q})+(\mathbf{P}\times(\mathbb{R}, \leq))$	*distributive*
6.	$(\mathbb{R}, \leq)^{\mathbf{P}+\mathbf{Q}}$	$=$	$(\mathbb{R}, \leq)^{\mathbf{P}}\times(\mathbb{R}, \leq)^{\mathbf{Q}}$	
7.	$\left(\mathbf{P}^{\mathbf{Q}}\right)^{(\mathbb{R}, \leq)}$	$=$	$\mathbf{P}^{\mathbf{Q}\times(\mathbb{R}, \leq)}$	

B.7.3 Primitive subposets

Definition B.19.

DEF: *The ordered set $\mathbf{L}_1$ is defined as $(\{x\}, \leq)$, for some value x.*

The $\mathbf{L}_1$ ordered set is illustrated by the Hasse diagram to the right.

Definition B.20.

DEF: *The ordered set $\mathbb{2}$ is defined as $\mathbb{2} \triangleq \mathbb{1}^2$.*

The $\mathbb{2}$ ordered set is illustrated by the Hasse diagram to the right.

B.7.4 Decomposition examples

Example B.25. Figure B.3 (page 301) illustrates the four ordered set operations $+$, $\times$, $\oplus$, and $\otimes$.

Example B.26. [33]The ordered set $n\mathbb{1}$ is the *anti-chain* with n elements. The ordered set $4\mathbb{1}$ is illustrated to the right.

[30] Stanley (1997) page 101
[31] Stanley (1997) page 101
[32] Stanley (1997) page 102
[33] Stanley (1997) page 100

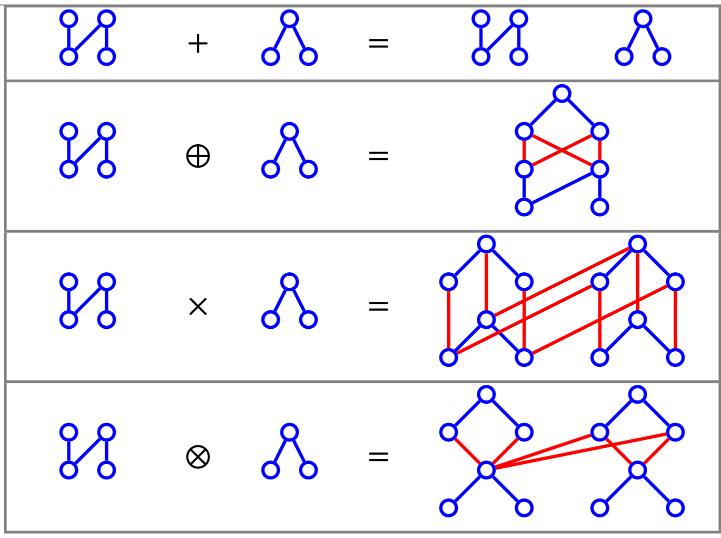
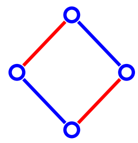
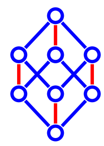


Figure B.3: Operations on ordered sets (Example B.25 page 300)

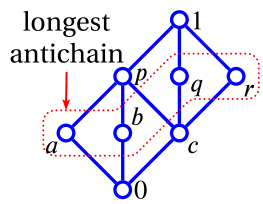
Example B.27. The ordered set $\mathbb{1}^n$ is the *chain* with n elements. The ordered set $\mathbb{1}^4$ is illustrated to the right.



Example B.28. The ordered set $\mathbb{2}^2$ is the 4 element *Boolean algebra* illustrated to the right.



Example B.29. The ordered set $\mathbb{2}^3$ is the 8 element *Boolean algebra* illustrated to the right.



Example B.30. [34]The longest *antichain* (Definition B.11 page 298) in the figure to the right has 4 elements giving this ordered set a *width* (Definition B.12 page 298) of 4. The longest chain also has 4 elements, giving the ordered set a *length* (Definition B.12 page 298) of 3. By *Dilworth's theorem* (Theorem B.3 page 298), the smallest *partition* consists of four *chains* (Definition B.4 page 291). One such *partition* is

$$\{\{0, a, p, 1\}, \{b\}, \{c, q\}, \{r\}\}.$$

B.8 Bounds on ordered sets

In an *ordered set* (Definition B.2 page 290), a pair of elements $\{x, y\}$ may not be *comparable*. Despite this, we may still be able to find elements that *are* comparable to both x and y and are "*greater*" than both of them. Such a greater element is called an *upper bound* of x and y. There may be many elements that are upper bounds of x and y. But if one of these upper bounds is comparable with and is smaller than all the other upper bounds, than this "smallest" of the "greater" elements is called the *least upper bound* (*lub*) of x and y, and is denoted $x \vee y$ (Definition B.21 page 302). Likewise,

[34] Farley (1997) page 4

we may also be able to find elements that are comparable to $\{x, y\}$ and are "*lesser*" than both of them. Such a lesser element is called a *lower bound* of x and y. If one of these lower bounds is comparable with and is larger than all the other lower bounds, than this "largest" of the "lesser" elements is called the *greatest lower bound* (*glb*) of $\{x, y\}$ and is denoted $x \wedge y$ (Definition B.22 page 302). If every pair of elements in an ordered set has both a least upper bound and a greatest lower bound in the ordered set, then that ordered set is a *lattice* (Definition C.3 page 305).

Definition B.21. *Let $(X, \leq)$ be an ordered set and 2^X the power set of X.*

DEF

*For any set $A \in 2^X$, c is an **upper bound** of A in $(X, \leq)$ if*

1. $x \leq c \quad \forall x \in A.$

*An element b is the **least upper bound**, or **lub**, of A in $(X, \leq)$ if*

2. *b and c are* UPPER BOUNDS *of A* $\implies b \leq c.$

*The least upper bound of the set A is denoted $\bigvee A$. It is also called the **supremum** of A, which is denoted $\sup A$. The **join** $x \vee y$ of x and y is defined as $x \vee y \triangleq \bigvee \{x, y\}$.*

Definition B.22. *Let $(X, \leq)$ be an ordered set and 2^X the power set of X.*

DEF

*For any set $A \in 2^X$, p is a **lower bound** of A in $(X, \leq)$ if*

1. $p \leq x \quad \forall x \in A.$

*An element a is the **greatest lower bound**, or **glb**, of A in $(X, \leq)$ if*

2. *a and p are* LOWER BOUNDS *of A* $\implies p \leq a.$

*The greatest lower bound of the set A is denoted $\bigwedge A$. It is also called the **infimum** of A, which is denoted $\inf A$. The **meet** $x \wedge y$ of x and y is defined as $x \wedge y \triangleq \bigwedge \{x, y\}$.*

Definition B.23 (least upper bound property). [35] *Let X be a set. Let $\sup A$ be the supremum (least upper bound) of a set A.*

DEF

*A set X satisfies the **least upper bound property** if*

$$\left.\begin{array}{lll} 1. & A \subseteq X & \textit{and} \\ 2. & A \neq \varnothing & \textit{and} \\ 3. & \exists b \in X \text{ such that } \forall a \in A,\ a \leq b & (A \text{ is bounded above in } X) \end{array}\right\} \implies \exists \sup A \in X$$

*A set X that satisfies the least upper bound property is also said to be **complete**.*

Proposition B.2. *Let $(X, \vee, \wedge; \leq)$ be an* ORDERED SET *(Definition B.2 page 290).*

PRP

$$x \leq y \iff \left\{\begin{array}{llll} 1. & x \wedge y = x & \textit{and} \\ 2. & x \vee y = y \end{array}\right\} \quad \forall x,y \in X$$

Proposition B.3. *Let 2^X be the* POWER SET *of a set X.*

PRP

$$A \subseteq B \implies \left\{\begin{array}{lll} 1. & \bigvee A \leq \bigvee B & \textit{and} \\ 2. & \bigwedge A \leq \bigwedge B \end{array}\right\} \quad \forall A,B \in 2^X$$

[35] Pugh (2002) page 13, Rudin (1976) page 4

APPENDIX C

LATTICES

C.1 Semi-lattices

Definition B.21 (page 302) defined the least upper bound $\vee$ of pairs of elements in terms of an ordering relation $\le$. However, the converse development is also possible— we can first define a binary operation $\ovee$ with a handful of "least upper bound like properties", and then define an ordering relation $\lesssim$ in terms of $\ovee$ (Definition C.1 page 303). In fact, Theorem C.1 (page 303) shows that under Definition C.1, $(X, \lesssim)$ *is* a partially ordered set and $\ovee$ is a least upper bound on that ordered set.

The same development is performed with regards to a greatest lower bound $\owedge$ with the result that $(X, \lesssim)$ *is* a partially ordered set and $\owedge$ is a greatest lower bound on that ordered set (Theorem C.2 page 304).

Definition C.1. [1] *Let* $\ovee, \lesssim \colon X^2 \to X$ *be binary operators on a set* X.

DEF: *The algebraic structure* $(X, \lesssim, \ovee)$ *is a* ***join semilattice*** *if*

$$\begin{array}{llcllll} 1. & x \ovee x & = & x & \forall x \in X & (\text{IDEMPOTENT}) & \text{and} \\ 2. & x \ovee y & = & y \ovee x & \forall x,y \in X & (\text{COMMUTATIVE}) & \text{and} \\ 3. & (x \ovee y) \ovee z & = & x \ovee (y \ovee z) & \forall x,y,z \in X & (\text{ASSOCIATIVE}). & \end{array}$$

Definition C.2. [2] *Let* $\owedge, \lesssim \colon X^2 \to X$ *be binary operators on a set* X.

DEF: *The algebraic structure* $(X, \lesssim, \owedge)$ *is a* ***meet semilattice*** *if*

$$\begin{array}{llcllll} 1. & x \owedge x & = & x & \forall x \in X & (\text{IDEMPOTENT}) & \text{and} \\ 2. & x \owedge y & = & y \owedge x & \forall x,y \in X & (\text{COMMUTATIVE}) & \text{and} \\ 3. & (x \owedge y) \owedge z & = & x \owedge (y \owedge z) & \forall x,y,z \in X & (\text{ASSOCIATIVE}). & \end{array}$$

Theorem C.1. [3] *Let* $\ovee, \lesssim \colon X^2 \to X$ *be binary operators over a set* X.

THM:

$$\left\{ \begin{array}{l} (X, \lesssim, \ovee) \text{ is a} \\ \text{JOIN SEMILATTICE} \end{array} \right\} \implies \left\{ \begin{array}{lll} 1. & (X, \lesssim) \text{ is a PARTIALLY ORDERED SET} & \text{and} \\ 2. & x \ovee y \text{ is a LEAST UPPER BOUND of } x \text{ and } y & \forall x, y \in X. \end{array} \right\}$$

✎PROOF: In order for $(X, \le)$ to be an ordered set, $\le$ must be, according to Definition B.2 (page 290), *reflexive*, *antisymmetric*, and *transitive*;

[1] MacLane and Birkhoff (1999) page 475, Birkhoff (1967) page 22
[2] MacLane and Birkhoff (1999) page 475
[3] MacLane and Birkhoff (1999) page 475

Proof that $\le$ is reflexive:

$$\begin{aligned}
& x = x \ovee x && \text{by idempotent hypothesis}\\
\iff\ & x \le x && \text{by definition of } \le\\
\implies\ & \le \text{ is reflexive}
\end{aligned}$$

Proof that $\le$ is antisymmetric:

$$\begin{aligned}
x \le y \text{ and } y \le x \iff\ & x \ovee y = y \text{ and } y \ovee x = x && \text{by definition of } \le\\
\implies\ & x \ovee y = y \text{ and } x \ovee y = x && \text{by commutative hypothesis}\\
\implies\ & x = y\\
\implies\ & \le \text{ is antisymmetric}
\end{aligned}$$

Proof that $\le$ is transitive:

$$\begin{aligned}
x \le y \text{ and } y \le z \iff\ & x \ovee y = y \text{ and } y \ovee z = z && \text{by definition of } \le\\
\implies\ & (x \ovee y) \ovee z = z\\
\iff\ & x \ovee (y \ovee z) = z && \text{by associative hypothesis}\\
\implies\ & x \ovee z = z\\
\iff\ & x \le z\\
\iff\ & \le \text{ is transitive}
\end{aligned}$$

Proof that $x \ovee y$ is a lub of x and y:

$$\begin{aligned}
x \ovee y = y \iff\ & x \le y && \text{by definition of } \le\\
\iff\ & x \vee y = y && \text{by definition of } \vee\\
\implies\ & x \ovee y = x \vee y\\
\implies\ & x \ovee y \text{ is the lub of } x \text{ and } y
\end{aligned}$$

Theorem C.2. [4] *Let $\owedge, \lesssim \colon X^2 \to X$ be binary operators over a set X.*

THM
$$\left\{ \begin{array}{l} (X, \lesssim, \owedge) \text{ is a}\\ \text{MEET SEMILATTICE} \end{array} \right\} \implies \left\{ \begin{array}{lll} 1. & (X, \lesssim) \text{ is a PARTIALLY ORDERED SET} & \text{and}\\ 2. & x \owedge y \text{ is a GREATEST LOWER BOUND of } x \text{ and } y & \forall x, y \in X. \end{array} \right\}$$

✎PROOF: In order for $(X, \le)$ to be an ordered set, $\le$ must be, according to Definition B.2 (page 290), *reflexive*, *antisymmetric*, and *transitive*;

Proof that $\le$ is reflexive:

$$\begin{aligned}
& x = x \owedge x && \text{by idempotent hypothesis}\\
\iff\ & x \le x && \text{by definition of } \le\\
\implies\ & \le \text{ is reflexive}
\end{aligned}$$

Proof that $\le$ is antisymmetric:

$$\begin{aligned}
x \le y \text{ and } y \le x \iff\ & x \owedge y = x \text{ and } y \owedge x = y && \text{by definition of } \le\\
\implies\ & x \owedge y = x \text{ and } x \owedge y = y && \text{by commutative hypothesis}\\
\implies\ & x = y\\
\implies\ & \le \text{ is antisymmetric}
\end{aligned}$$

[4] MacLane and Birkhoff (1999) page 475

Proof that $\leq$ is transitive:

$$\begin{aligned}
x \leq y \text{ and } y \leq z &\iff x \oslash y = x \text{ and } y \oslash z = y && \text{by definition of } \leq \\
&\implies x \oslash (y \oslash z) = x \\
&\iff (x \oslash y) \oslash z = x && \text{by associative hypothesis} \\
&\implies x \oslash z = x \\
&\iff x \leq z \\
&\iff \leq \text{ is transitive}
\end{aligned}$$

Proof that $x \oslash y$ is a glb of x and y:

$$\begin{aligned}
x \oslash y = x &\iff x \leq y && \text{by definition of } \leq \\
&\iff x \wedge y = x && \text{by definition of } \wedge \\
&\implies x \oslash y = x \wedge y \\
&\implies x \oslash y \text{ is the glb of } x \text{ and } y
\end{aligned}$$

C.2 Lattices

An *ordered set* is a set together with the additional structure of an ordering relation (Definition B.2 page 290). However, this amount of structure tends to be insufficient to ensure "well-behaved" mathematical systems. This situation is greatly remedied if every pair of elements in an ordered set (partially or linearly ordered) has both a *least upper bound* and a *greatest lower bound* (Definition B.22 page 302) in the ordered set; in this case, that ordered set is a *lattice* (next definition). Gian-Carlo Rota (1932–1999) illustrates the advantage of lattices over simple ordered sets by pointing out that the *ordered set* of partitions of an integer "is fraught with pathological properties", while the *lattice* of partitions of a set "remains to this day rich in pleasant surprises".[5] Further examples of lattices follow in Section C.3 (page 310).

Definition C.3. [6]

DEF *An algebraic structure $\mathbf{L} \triangleq (X, \vee, \wedge; \leq)$ is a **lattice** if*

1. $(X, \leq)$ *is an ordered set* *and*
2. $x, y \in X \implies x \vee y \in X$ *and*
3. $x, y \in X \implies x \wedge y \in X$

*The algebraic structure $\mathbf{L}^* \triangleq (X, \ovee, \owedge; \geq)$ is the **dual** lattice of $\mathbf{L}$, where $\ovee$ and $\owedge$ are determined by $\geq$. The* LATTICE *$\mathbf{L}$ is **linear** if $(X, \leq)$ is a* CHAIN *(Definition B.4 page 291).*

Definition C.3 (previous) characterizes lattices in terms of *order properties*. Under this definition, lattices have an equivalent characterization in terms of *algebraic properties*. In particular, all lattices have four basic algebraic properties: all lattices are *idempotent*, *commutative*, *associative*, and *absorptive*. Conversely, any structure that possesses these four properties *is* a lattice. These results are demonstrated by Theorem C.3 (next). However, note that the four properties are not *independent*, as it is possible to prove that any structure $\mathbf{L} \triangleq (X, \vee, \wedge; \leq)$ that is *commutative*, *associative*, and *absorptive*, is also *idempotent* (Theorem C.8 page 314). Thus, when proving that $\mathbf{L}$ is a lattice, it is only necessary to prove that it is *commutative*, *associative*, and *absorptive*.

[5] Rota (1997) page 1440 ⟨Introduction⟩, Rota (1964) page 498 ⟨partitions of a set⟩

[6] MacLane and Birkhoff (1999) page 473, Birkhoff (1948) page 16, Ore (1935), Birkhoff (1933a) page 442, Maeda and Maeda (1970), page 1

Theorem C.3. [7]

$(X, \vee, \wedge; \leq)$ *is a* LATTICE $\iff$

$$\left\{\begin{array}{rcl|rcl|lll} x \vee x &=& x & x \wedge x &=& x & \forall x \in X & \text{(IDEMPOTENT)} & \text{and} \\ x \vee y &=& y \vee x & x \wedge y &=& y \wedge x & \forall x,y \in X & \text{(COMMUTATIVE)} & \text{and} \\ (x \vee y) \vee z &=& x \vee (y \vee z) & (x \wedge y) \wedge z &=& x \wedge (y \wedge z) & \forall x,y,z \in X & \text{(ASSOCIATIVE)} & \text{and} \\ x \vee (x \wedge y) &=& x & x \wedge (x \vee y) &=& x & \forall x,y \in X & \text{(ABSORPTIVE).} & \end{array}\right\}$$

✎PROOF:

1. Proof that $(X, \vee, \wedge; \leq)$ is a lattice $\implies$ 4 properties:
   These follow directly from the definitions of least upper bound $\vee$ and greatest lower bound $\wedge$. For the absorptive property,

$$\begin{aligned} x \leq y &\implies x \vee (x \wedge y) = x \vee x = x \\ y \leq x &\implies x \vee (x \wedge y) = x \vee y = x \\ x \leq y &\implies x \wedge (x \vee y) = x \wedge y = x \\ y \leq x &\implies x \wedge (x \vee y) = x \wedge x = x \end{aligned}$$

2. Proof that $(X, \vee, \wedge; \leq)$ is a lattice $\impliedby$ 4 properties:

   According to Definition C.3 (page 305), in order for $(X, \vee, \wedge; \leq)$ to be a lattice, $(X, \vee, \wedge; \leq)$ must be an ordered set, $x \vee y$ must be the least upper bound for any $x, y \in X$ and $x \wedge y$ must be the greatest lower bound for any $x, y \in X$.

   (a) By Theorem C.1 (page 303), $(X, \vee, \wedge; \leq)$ is an ordered set.
   (b) By Theorem C.1 (page 303), $x \vee y$ is the least upper bound for any $x, y \in X$.
   (c) Proof that $x \wedge y$ is the greatest lower bound for any $x, y \in X$: To prove this, we must show that $x \leq y \iff x \wedge y = x$.

   Proof that $x \leq y \implies x \wedge y = x$:

$x = x \wedge (x \vee y)$	by absorptive hypothesis
$= x \wedge y$	by $x \leq y$ hypothesis and definition of $\leq$

   Proof that $x \leq y \impliedby x \wedge y = x$:

$y = y \vee (y \wedge x)$	by absorptive hypothesis
$= y \vee (x \wedge y)$	by commutative hypothesis
$= y \vee x$	by $x \wedge y = x$ hypothesis
$= x \vee y$	by commutative hypothesis
$\implies x \leq y$	by definition of $\leq$

✏

[7] MacLane and Birkhoff (1999) pages 473–475 ⟨LEMMA 1, THEOREM 4⟩, Burris and Sankappanavar (1981) pages 4–7, Birkhoff (1938), pages 795–796, Ore (1935) page 409 ⟨(α)⟩, Birkhoff (1933a) page 442, Dedekind (1900) pages 371–372 ⟨(1)–(4)⟩. Peirce (1880) credits Boole and Jevons with the *commutative* property: Peirce (1880), page 33 ⟨"(5)"⟩. Peirce (1880) credits Boole and Jevons with the *associative* property. Peirce (1880) credits Jevons (1864) with the *idempotent* property: Jevons (1864), page 41
$A + A = A$ "Law of Unity"
$AA = A$ "Law of Simplicity"

Lemma C.1. [8] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge; \leq)$ *be a* LATTICE *(Definition C.3 page 305)*.

LEM
$$x \leq y \iff x = x \wedge y \qquad \forall x,y \in \mathbf{L}$$

✎PROOF:

1. Proof for $\Longrightarrow$ case: by left hypothesis and definition of $\wedge$ (Definition B.22 page 302).
2. Proof for $\Longleftarrow$ case: by right hypothesis and definition of $\wedge$ (Definition B.22 page 302).

✏

The identities of Theorem C.3 (page 306) occur in pairs that are *duals* of each other. That is, for each identity, if you swap the join and meet operations, you will have the other identity in the pair. Thus, the characterization of lattices provided by Theorem C.3 (page 306) is called *self-dual*. And because of this, lattices support the *principle of duality* (next theorem).

Theorem C.4 (Principle of duality). [9] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge; \leq)$ *be a lattice.*

THM
$$\left\{\begin{array}{l}\phi \text{ is an identity on } \mathbf{L} \text{ in terms}\\ \text{of the operations} \vee \text{ and } \wedge\end{array}\right\} \implies \mathbf{T}\phi \text{ is also an identity on } \mathbf{L}$$
where the operator $\mathbf{T}$ *performs the following mapping on the operations of* ϕ:
$$\vee \to \wedge, \qquad \wedge \to \vee$$

✎PROOF: For each of the identities in Theorem C.3 (page 306), the operator $\mathbf{T}$ produces another identity that is also in the set of identities:

$$\begin{array}{rclclclclcl}
\mathbf{T}(1a) &=& \mathbf{T}[x \vee y &=& y \vee x] &=& [x \wedge y &=& y \wedge x] &=& (1b)\\
\mathbf{T}(1b) &=& \mathbf{T}[x \wedge y &=& y \wedge x] &=& [x \vee y &=& y \vee x] &=& (1a)\\
\mathbf{T}(2a) &=& \mathbf{T}[x \vee (y \wedge z) &=& (x \vee y) \wedge (x \vee z)] &=& [x \wedge (y \vee z) &=& (x \wedge y) \vee (x \wedge z)] &=& (2b)\\
\mathbf{T}(2b) &=& \mathbf{T}[x \wedge (y \vee z) &=& (x \wedge y) \vee (x \wedge z)] &=& [x \vee (y \wedge z) &=& (x \vee y) \wedge (x \vee z)] &=& (2a)
\end{array}$$

Therefore, if the statement ϕ is consistent with regards to the lattice $\mathbf{L}$, then $\mathbf{T}\phi$ is also consistent with regards to the lattice $\mathbf{L}$. ✏

Proposition C.1 (Monotony laws). [10] *Let* $(X, \vee, \wedge; \leq)$ *be a lattice.*

PRP
$$\left.\begin{array}{rcl} a &\leq& b \quad \text{and}\\ x &\leq& y.\end{array}\right\} \implies \left\{\begin{array}{rcl} a \wedge x &\leq& b \wedge y \quad \text{and}\\ a \vee x &\leq& b \vee y.\end{array}\right.$$

[8] Holland (1970), page ???
[9] Padmanabhan and Rudeanu (2008) pages 7–8, Beran (1985) pages 29–30
[10] Givant and Halmos (2009) page 39, Doner and Tarski (1969) pages 97–99

✎PROOF:

$$
\begin{array}{rll}
1.(a \wedge x) & \leq a & \text{by definition of } \textit{meet} \text{ operation } \wedge \text{ Definition B.22 page 302} \\
& \leq b & \text{by left hypothesis} \\
2.(a \wedge x) & \leq x & \text{by definition of } \textit{meet} \text{ operation } \wedge \text{ Definition B.22 page 302} \\
& \leq y & \text{by left hypothesis} \\
3.(a \wedge x) & = \underbrace{(a \wedge x)}_{\leq b} \wedge \underbrace{(a \wedge x)}_{\leq y} & \text{by } \textit{idempotent} \text{ property Theorem C.3 page 306} \\
& \leq b \wedge y & \text{by 1 and 2} \\
4.(a \vee x) & = \underbrace{(a \vee x)}_{\leq b} \vee \underbrace{(a \vee x)}_{\leq y} & \text{by } \textit{idempotent} \text{ property Theorem C.3 page 306} \\
& \leq b \vee y & \text{by 1 and 2}
\end{array}
$$

✏

Minimax inequality. Suppose we arrange a finite sequence of values into m groups of n elements per group. This could be represented as an $m \times n$ matrix. Suppose now we find the minimum value in each row, and the maximum value in each column. We can call the maximum of all the minimum row values the *maximin*, and the minimum of all the maximum column values the *minimax*. Now, which is greater, the maximin or the minimax? The *minimax inequality* demonstrates that the maximin is always less than or equal to the minimax. The minimax inequality is illustrated below and stated formerly in Theorem C.5 (page 308).

$$
\underbrace{\bigvee_1^m \left\{ \begin{array}{l} \bigwedge_1^n \{ x_{11} \quad x_{12} \quad \cdots \quad x_{1n} \} \\ \hline \bigwedge_1^n \{ x_{21} \quad x_{22} \quad \cdots \quad x_{2n} \} \\ \hline \bigwedge_1^n \{ \vdots \quad \ddots \quad \ddots \quad \vdots \} \\ \hline \bigwedge_1^n \{ x_{m1} \quad x_{m2} \quad \cdots \quad x_{mn} \} \end{array} \right\}}_{\text{maximin}}
\leq
\underbrace{\bigwedge_1^n \left\{ \begin{array}{c|c|c|c} \bigvee_1^m & \bigvee_1^m & \bigvee_1^m & \bigvee_1^m \\ x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{array} \right\}}_{\text{minimax}}
$$

Theorem C.5 (Minimax inequality). [11] *Let* $(X, \vee, \wedge; \leq)$ *be a lattice.*

THM

$$
\underbrace{\bigvee_{i=1}^{m} \bigwedge_{j=1}^{n} x_{ij}}_{\textit{maxmini: largest of the smallest}} \leq \underbrace{\bigwedge_{j=1}^{n} \bigvee_{i=1}^{m} x_{ij}}_{\textit{minimax: smallest of the largest}} \qquad \forall x_{ij} \in X
$$

[11] Birkhoff (1948) pages 19–20

✎PROOF:

$$\underbrace{\left(\bigwedge_{k=1}^{n} x_{ik}\right)}_{\text{smallest for any given } i} \le x_{ij} \le \underbrace{\left(\bigvee_{k=1}^{n} x_{kj}\right)}_{\text{largest for any given } j} \qquad \forall i,j$$

$$\implies \underbrace{\bigvee_{i=1}^{m}\left(\bigwedge_{k=1}^{n} x_{ik}\right)}_{\text{largest amoung all } i\text{s of the smallest values}} \le \underbrace{\bigwedge_{j=1}^{n}\left(\bigvee_{k=1}^{m} x_{kj}\right)}_{\text{smallest amoung all } j\text{s of the largest values}}$$

$$\implies \underbrace{\bigvee_{i=1}^{m}\left(\bigwedge_{j=1}^{n} x_{ij}\right)}_{\text{maxmini}} \le \underbrace{\bigwedge_{j=1}^{n}\left(\bigvee_{i=1}^{m} x_{ij}\right)}_{\text{minimax}} \qquad \text{(change of variables)}$$

✏

Distributive inequalities. Special cases of the minimax inequality include three distributive *inequalities* (next theorem). If for some lattice any *one* of these inequalities is an *equality*, then *all three* are *equalities*; and in this case, the lattice is a called a *distributive* lattice.

Theorem C.6 (distributive inequalities). [12]

THM

$(X, \vee, \wedge; \le)$ *is a lattice* $\implies$ *for all x,y,z*$\in X$

$$\begin{array}{rcll}
x \wedge (y \vee z) & \ge & (x \wedge y) \vee (x \wedge z) & \text{(JOIN SUPER-DISTRIBUTIVE)} \quad \textit{and}\\
x \vee (y \wedge z) & \le & (x \vee y) \wedge (x \vee z) & \text{(MEET SUB-DISTRIBUTIVE)} \quad \textit{and}\\
(x \wedge y) \vee (x \wedge z) \vee (y \wedge z) & \le & (x \vee y) \wedge (x \vee z) \wedge (y \vee z) & \text{(MEDIAN INEQUALITY).}
\end{array}$$

✎PROOF:

1. Proof that $\wedge$ sub-distributes over $\vee$:

$$\begin{aligned}
(x \wedge y) \vee (x \wedge z) &\le (x \vee x) \wedge (y \vee z) && \text{by } \textit{minimax inequality} \text{ (Theorem C.5 page 308)}\\
&= x \wedge (y \vee z) && \text{by } \textit{idempotent} \text{ property of lattices (Theorem C.3 page 306)}
\end{aligned}$$

$$\bigvee \left\{ \begin{array}{c} \bigwedge \left\{ \begin{array}{cc} x & y \end{array} \right\} \\ \hline \bigwedge \left\{ \begin{array}{cc} x & z \end{array} \right\} \end{array} \right\} \le \bigwedge \left\{ \begin{array}{c|c} \bigvee & \bigvee \\ x & y \\ x & z \end{array} \right\}$$

2. Proof that $\vee$ super-distributes over $\wedge$:

$$\begin{aligned}
x \vee (y \wedge z) &= (x \wedge x) \vee (y \wedge z) && \text{by } \textit{idempotent} \text{ property of lattices (Theorem C.3 page 306)}\\
&\le (x \vee y) \wedge (x \vee z) && \text{by } \textit{minimax inequality} \text{ (Theorem C.5 page 308)}
\end{aligned}$$

$$\bigvee \left\{ \begin{array}{c} \bigwedge \left\{ \begin{array}{cc} x & x \end{array} \right\} \\ \hline \bigwedge \left\{ \begin{array}{cc} y & z \end{array} \right\} \end{array} \right\} \le \bigwedge \left\{ \begin{array}{c|c} \bigvee & \bigvee \\ x & x \\ y & z \end{array} \right\}$$

3. Proof that of median inequality: by *minimax inequality* (Theorem C.5 page 308)

✏

[12] Davey and Priestley (2002) page 85, Grätzer (2003) page 38, Birkhoff (1933a) page 444, Korselt (1894) page 157, Müller-Olm (1997) page 13 ⟨terminology⟩

Modular inequalities. Besides the distributive property, another consequence of the minimax inequality is the *modularity inequality* (next theorem). A lattice in which this inequality becomes equality is said to be *modular*.

Theorem C.7 (Modular inequality). [13] *Let* $(X, \vee, \wedge\,;\, \leq)$ *be a* LATTICE *(Definition C.3 page 305).*

THM $\quad x \leq y \quad\implies\quad x \vee (y \wedge z) \leq y \wedge (x \vee z)$

✎PROOF:

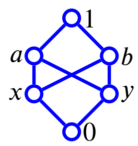
$$\begin{aligned} x \vee (y \wedge z) &= (x \wedge x) \vee (y \wedge z) && \text{by \textit{absorptive} property (Theorem C.3 page 306)} \\ &\leq (x \vee y) \wedge (x \vee z) && \text{by the \textit{minimax inequality} (Theorem C.5 page 308)} \\ &= y \wedge (x \vee z) && \text{by left hypothesis} \end{aligned}$$

$$\bigvee\left\{\begin{array}{c} \bigwedge\left\{\; x \quad x \;\right\} \\ \hline \bigwedge\left\{\; y \quad z \;\right\} \end{array}\right\} \quad\leq\quad \bigwedge\left\{\begin{array}{c|c} \bigvee & \bigvee \\ x & x \\ y & z \end{array}\right\}$$

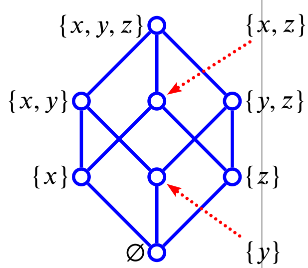
✏

C.3 Examples

Example C.1. [14] the ordered set illustrated to the right is **not** a lattice because, for example, while x and y have *upper bounds* a, b, and 1, x and y have no *least upper bound*. Obviously 1 is not the least upper bound because $a \leq 1$ and $b \leq 1$. And neither a nor b is a least upper bound because $a \nleq b$ and $b \nleq a$; rather, a and b are incomparable ($a \| b$). Note that if we remove either or both of the two lines crossing the center, the ordered set becomes a lattice.



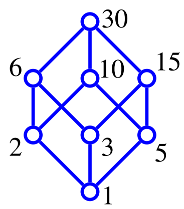
Example C.2 (Discrete lattice). Let 2^A be the power set of a set A, $\subseteq$ the set inclusion relation, $\cup$ the set union operation, and $\cap$ the set intersection operation. Then the tupple $\left(2^{\{x,y,z\}}, \cup, \cap\,;\, \subseteq\right)$ is a lattice.



Examples of least upper bounds	Examples of greatest lower bounds
$\{x\} \cup \{z\} = \{x, z\}$	$\{x\} \cap \{z\} = \varnothing$
$\{x, y\} \cup \{y\} = \{x, y\}$	$\{x, y\} \cap \{y\} = \{y\}$
$\{x, z\} \cup \{y, z\} = \{x, y, z\}$	$\{x, z\} \cap \{y, z\} = \{z\}$

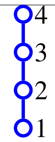
Example C.3 (Integer factor lattice). [15] For any pair of natural numbers $n, m \in \mathbb{N}$, let $n|m$ represent the relation "m divides n", $\operatorname{lcm}(n, m)$ the *least common multiple* of n and m, and $\gcd(n, m)$ the *greatest common divisor* of n and m.

EX $\quad \left(\{1, 2, 3, 5, 6, 10, 15, 30\},\ \gcd,\ \operatorname{lcm}\,;\, |\right)$ is a lattice.

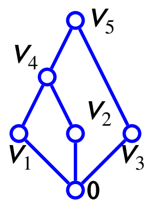


[13] Birkhoff (1948) page 19, Burris and Sankappanavar (1981) page 11, Dedekind (1900) page 374
[14] Oxley (2006) page 54, Farley (1997), page 3, Farley (1996), page 5, Birkhoff (1967) pages 15–16
[15] MacLane and Birkhoff (1999) page 484, Sheffer (1920) page 310 ⟨footnote 1⟩

Example C.4 (Linear lattice). Let $\le$ be the standard counting ordering relation on the set of integers; and for any pair of integers $n, m \in \mathbb{N}$, let $\max(n,m)$ be the maximum of n and m, and $\min(n,m)$ be the minimum of n and m. Then the tupple $(\{1,2,3,4\}, \max, \min; \le)$ is a lattice.



Example C.5 (Subspace lattices). [16]Let $(\!(\boldsymbol{V}_n)\!)$ be a sequence of subspaces, $\subseteq$ be the set inclusion relation, $+$ the subspace addition operator, and $\cap$ the set intersection operator. Then the tuple $(\{\boldsymbol{V}_n\}, +, \cap; \subseteq)$ is a lattice.

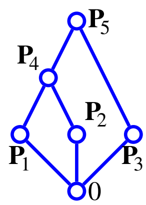


Example C.6 (Projection operator lattices). [17]Let $(\!(\mathbf{P}_n)\!)$ be a sequence of projection operators in a Hilbert space X.

EX

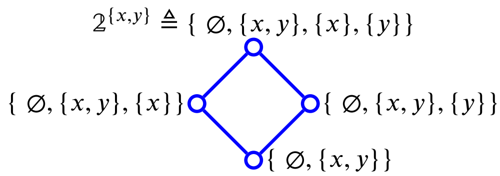
$$\left(\{\mathbf{P}_n\}, \vee, \wedge; \le\right) \text{ is a lattice}$$

$$\begin{array}{rrcl}\text{where} & \mathbf{P}_1 \le \mathbf{P}_2 & \stackrel{\text{def}}{\iff} & \mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_1\\ & \mathbf{P}_1 \vee \mathbf{P}_2 & = & \mathbf{P}_1 + \mathbf{P}_2 - \mathbf{P}_1\mathbf{P}_2\\ & \mathbf{P}_1 \wedge \mathbf{P}_2 & = & \mathbf{P}_1\mathbf{P}_2\end{array}$$



Example C.7 (Lattice of a single topology). [18] Let X be a set, τ a topology on X, $\subseteq$ the set inclusion relation, $\cup$ the set union operator, and $\cap$ the set intersection operator. Then the tuple $(\tau, \cup, \cap; \subseteq)$ is a lattice.

Example C.8 (Lattice of topologies). [19]Let X be a set and $\{\tau_1, \tau_2, \tau_3, \ldots\}$ all the possible topologies on X. Let $\subseteq$ be the set inclusion relation, $\cup$ the set union operator, and $\cap$ the set intersection operator. Then the tuple $\left(\{(X, \tau_n)\}, \cup, \cap; \subseteq\right)$ is a lattice.

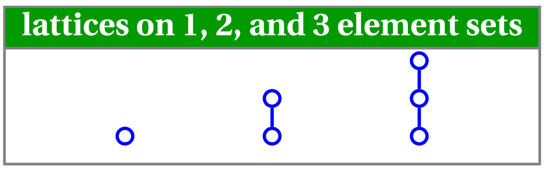


Proposition C.2. [20] *Let X_n be a finite set with order $n = |X_n|$. Let L_n be the number of labeled lattices on X_n, l_n the number of unlabeled lattices, and p_n the number of unlabeled posets.*

PRP

n	0	1	2	3	4	5	6	7	8	9	10
L_n	1	1	2	6	36	380	6390	157962	5396888	243, 179, 064	13, 938, 711, 210
l_n	1	1	1	1	2	5	15	53	222	1078	5994
p_n	1	1	2	5	16	63	318	2045	16, 999	183, 231	2, 567, 284

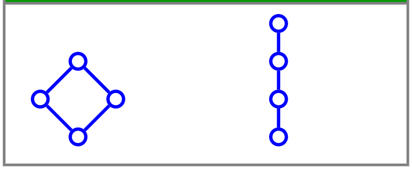
Example C.9 (lattices on 1–3 element sets). [21]There is only one unlabeled lattice for finite sets with 3 or fewer elements (Proposition C.2 page 311). Thus, these lattices are all linearly ordered. These 3 lattices are illustrated to the right.



[16] Isham (1999) pages 21–22
[17] Isham (1999) pages 21–22, Dunford and Schwartz (1957), pages 481–482
[18] Burris and Sankappanavar (1981) page 9, Birkhoff (1936a) page 161
[19] Isham (1999) page 44, Isham (1989), page 1515
[20] Sloane (2014) ⟨http://oeis.org/A055512⟩, Sloane (2014) ⟨http://oeis.org/A006966⟩, Sloane (2014) ⟨http://oeis.org/A000112⟩, Heitzig and Reinhold (2002)
[21] Kyuno (1979), page 412, Stanley (1997), page 102

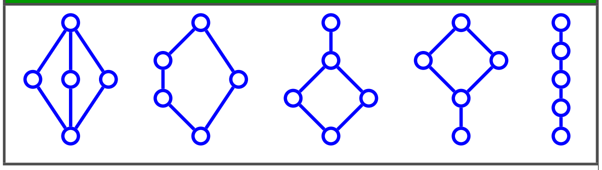
Example C.10 (lattices on a 4 element set). [22]There are 2 unlabeled lattices on a 4 element set (Proposition C.2 page 311). These are illustrated to the right.

lattices on 4 element sets



Example C.11 (lattices on a 5 element set). [23]There are 5 unlabeled lattices on a 5 element set (Proposition C.2 page 311). These are illustrated to the right.

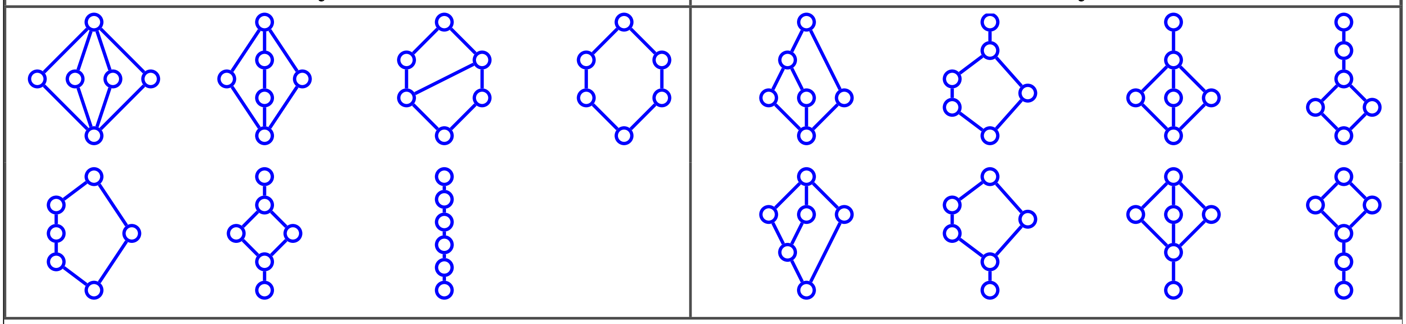
lattices on 5 element sets



Example C.12 (lattices on a 6 element set). [24] There are 15 *unlabeled lattices* on a 6 element set (Proposition C.2 page 311). These are illustrated in the following table. Notice that the lattices in the second row are simply generated from the 5 element lattices (Example C.11 page 312) with a "head" or "tail" added to each one.

lattices on 6 element sets

self-dual | *non-self dual*



Example C.13 (lattices on a 7 element set). [25] There are 53 unlabeled lattices on a 7 element set (Proposition C.2 page 311). These are illustrated in Figure C.1 (page 313).

Example C.14 (lattices on 8 element sets). There are 222 unlabeled lattices on a 8 element set (Proposition C.2 page 311). See Kyuno's paper[26] for Hasse diagrams of all 222 lattices.

C.4 Characterizations

Theorem C.3 (page 306)gave eight equations in three variables and two operators that are true of all lattices. But the converse is also true: that is, if the eight equations of Theorem C.3 are true for all values of the underlying set, then that set together with the two operators are a lattice.

That is, the eight equations in three variables of Theorem C.3 *characterize* lattices, or serve as an *equational basis* for lattices.[27] And this is not the only system of equations that characterize a lattice. There are other systems that use fewer equations in more variables. Here are some examples of lattice characterizations:

[22] Kyuno (1979), page 412, Stanley (1997), page 102
[23] Kyuno (1979), page 413, Stanley (1997), page 102
[24] Kyuno (1979), page 413, Stanley (1997), page 102
[25] Kyuno (1979), pages 413–414
[26] Kyuno (1979), pages 415–421
[27] McKenzie (1970) page 24, Tarski (1966)

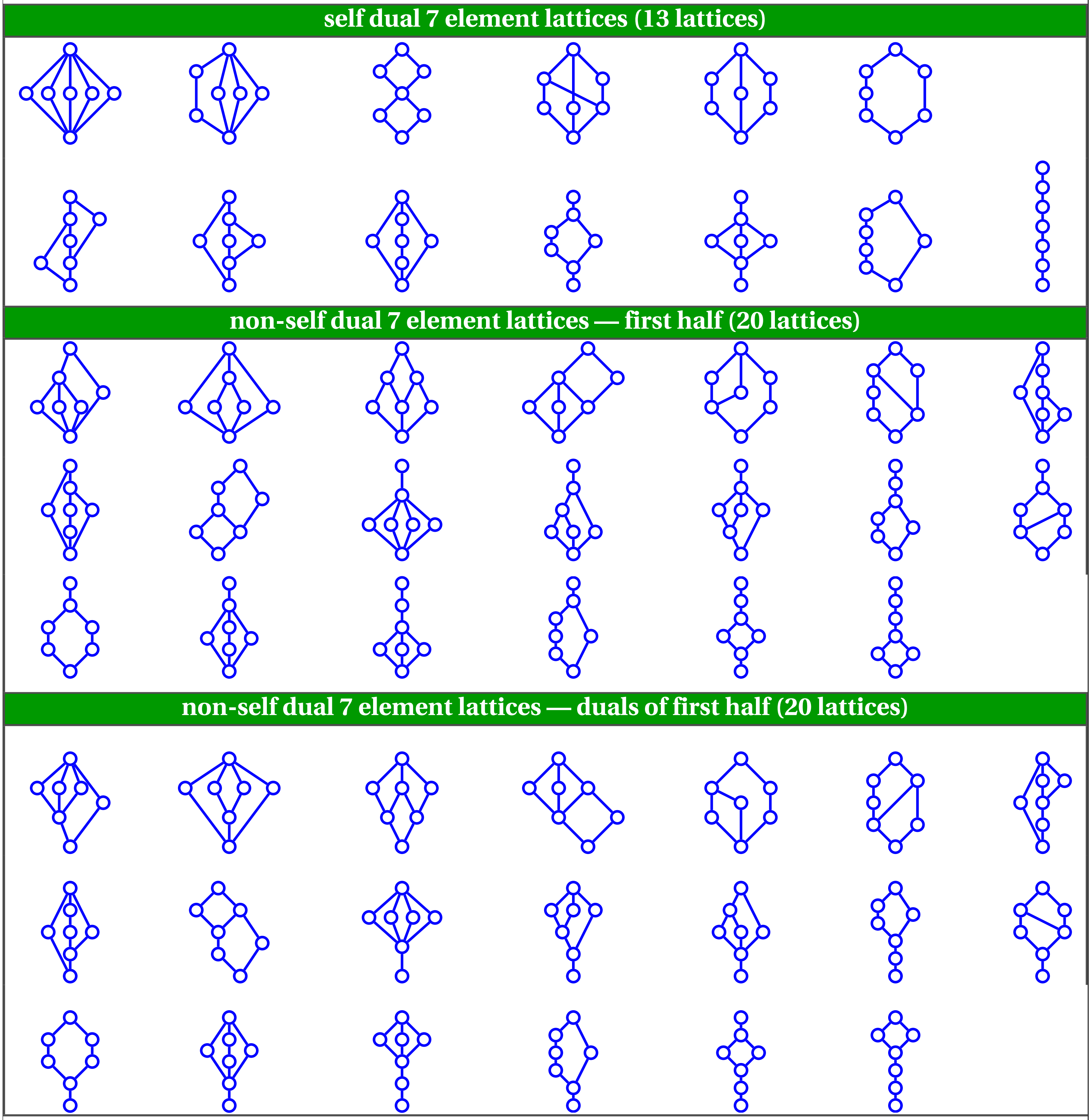


Figure C.1: The 53 unlabeled lattices on a 7 element set (Example C.13 page 312)

Since these characterizations are equivalent to the definition of the lattice, we could in fact change things around and essentially make any of these characterizations into the definition, and make the definition into a theorem.[28]

Theorem C.3 (page 306) gave 4 necessary and sufficient pairs of properties for a structure $(X, \vee, \wedge; \leq)$ to be a *lattice*. However, these 4 pairs are actually *overly* sufficient (they are not *independent*), as demonstrated next.

Theorem C.8. [29]

THM

$(X, \vee, \wedge; \leq)$ *is a lattice* $\iff$

$$\left\{\begin{array}{rcl|rcl|lll}
x \vee y &=& y \vee x & x \wedge y &=& y \wedge x & \forall x,y \in X & (\text{COMMUTATIVE}) & \textit{and}\\
(x \vee y) \vee z &=& x \vee (y \vee z) & (x \wedge y) \wedge z &=& x \wedge (y \wedge z) & \forall x,y,z \in X & (\text{ASSOCIATIVE}) & \textit{and}\\
x \vee (x \wedge y) &=& x & x \wedge (x \vee y) &=& x & \forall x,y \in X & (\text{ABSORPTIVE}) &
\end{array}\right\}$$

✎PROOF: Let $\boldsymbol{L} \triangleq (X, \vee, \wedge; \leq)$.

1. Proof that $\boldsymbol{L}$ is a *lattice* $\implies$ 3 properties: by Theorem C.3 page 306
2. Proof that $\boldsymbol{L}$ is a *lattice* $\impliedby$ 3 properties:
   (a) Proof that 3 properties $\implies$ $\boldsymbol{L}$ is *idempotent*:

$$\begin{aligned}
x \vee x &= x \vee [x \wedge (x \vee y)] && \text{by } \textit{absorptive} \text{ property}\\
&= x \vee [x \wedge z] && \text{where } z \triangleq x \vee y\\
&= x && \text{by } \textit{absorptive} \text{ property}\\
x \wedge x &= x \wedge [x \vee (x \wedge y)] && \text{by } \textit{absorptive} \text{ property}\\
&= x \wedge [x \vee z] && \text{where } z \triangleq x \wedge y\\
&= x && \text{by } \textit{absorptive} \text{ property}
\end{aligned}$$

   (b) By Theorem C.3 page 306 and because $\boldsymbol{L}$ is *commutative*, *associative*, *absorptive*, and *idempotent* with respect to $\vee$ and $\wedge$, $\boldsymbol{L}$ is a *lattice*.

⇨

Theorem C.9 (Lattice characterization in 2 equations and 5 variables). [30] *Let X be a set and $\vee$ and $\wedge$ be two binary operators on X.*

THM

$(X, \leq, \vee, \wedge)$ *is a lattice* ***if and only if***

$$\begin{array}{rcll}
x &=& (x \wedge y) \vee x & \forall x,y \in X \quad \textit{and}\\
[(x \wedge y) \wedge z \vee u] \vee w &=& [(y \wedge z) \wedge x \vee w] \vee (y \vee u) \wedge u & \forall x,y,z,u,w \in X
\end{array}$$

Theorem C.10 (Lattice characterizations in 1 equation). [31] *Let X be a set and $\vee$ and $\wedge$ be two binary*

[28] Burris and Sankappanavar (1981) pages 6–7,
[29] Padmanabhan and Rudeanu (2008) page 8, Beran (1985) page 5, McKenzie (1970) page 24
[30] Tamura (1975) page 137
[31] McCune et al. (2003b) page 2, McCune et al. (2003a), McCune and Padmanabhan (1996) page 144, http://www.cs.unm.edu/%7Everoff/LT/

operators on X.

THM *The following four statements are all equivalent:*

1. $(X, \vee, \wedge; \leq)$ *is a* ***lattice***
2. $(((y \vee x) \wedge x) \vee (((z \wedge (x \vee x)) \vee (u \wedge x)) \wedge v)) \wedge (w \vee ((s \vee x) \wedge (x \vee t))) = x$
   $\forall x,y,z,u,v,w,s,t \in X$ *(1 equation, 8 variables, length 29)*
3. $(((y \vee x) \wedge x) \vee (((z \wedge (x \vee x)) \vee (u \wedge x)) \wedge v)) \wedge (((w \vee x) \wedge (s \vee x)) \vee t) = x$
   $\forall x,y,z,u,v,w,s,t \in X$ *(1 equation, 8 variables, length 29)*
4. $(((x \wedge y) \vee (y \wedge (x \vee y))) \wedge z) \vee (((x \wedge (((x_1 \wedge y) \vee (y \wedge x_2)) \vee y)) \vee (((y \wedge (((x_1 \vee (y \vee x_2)) \wedge (x_3 \vee y)) \wedge y)) \vee (u \wedge (y \vee (((x_1 \vee (y \vee x_2)) \wedge (x_3 \vee y)) \wedge y)))) \wedge (x \vee (((x_1 \wedge y) \vee (y \wedge x_2)) \vee y)))) \wedge (((x \wedge y) \vee (y \wedge (x \vee y))) \vee z)) = y$
   $\forall x,y,z,x_1,x_2,x_3,u \in X$ *(1 equation, 7 variables, length 79)*

C.5 Functions on lattices

C.5.1 Isomorphisms

Lattices and *ordered set* (Definition B.2 page 290) are examples of mathematical *order structures*. Often we are interested in similarities between two lattices $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ with respect to order. Similarities between lattices can be described by defining a function θ that maps from the first lattice to the second. The degree of similarity can be roughly described in terms of the mapping θ as follows:

1. If there exists a mapping that is *bijective* then the number of elements in $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ is the same. However, their order structure may still be very different.

2. Lattices $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are more similar if there exists a mapping that is *bijective* and *order preserving* (Definition B.9 page 297). Despite having this property however, their order structure may still be remarkably different, as illustrated by Example B.18 (page 297) and Example B.19 (page 297).

3. Lattices $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are essentially identical (except possibly for their labeling) if there exists a mapping θ that is not only *bijective* and *order preserving*, but whose *inverse* is *also bijective* (Theorem C.11 page 315). In this case, the lattices $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are *isomorphic* and the mapping θ is an *isomorphism*. An isomorphism between $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ implies that the two lattices have an identical order structure. In particular, the isomorphism θ preserves joins and meets (next definition).

Definition C.4. *Let* $\boldsymbol{L}_1 \triangleq (X, \vee, \wedge; \leq)$ *and* $\boldsymbol{L}_2 \triangleq (Y, \varovee, \varowedge; \precsim)$ *be lattices.*

DEF $\boldsymbol{L}_1$ *and* $\boldsymbol{L}_2$ *are* ***algebraically isomorphic****, or simply* ***isomorphic****, if there exists a function* $\theta \in Y^X$ *such that*

1. $\theta(x \vee y) = \theta(x) \varovee \theta(y) \quad \forall x, y \in X$ *(PRESERVES JOINS)* *and*
2. $\theta(x \wedge y) = \theta(x) \varowedge \theta(y) \quad \forall x, y \in X$ *(PRESERVES MEETS).*

In this case, the function θ *is said to be an* ***isomorphism*** *from* $\boldsymbol{L}_1$ *to* $\boldsymbol{L}_2$*, and the isomorphic relationship between* $\boldsymbol{L}_1$ *and* $\boldsymbol{L}_2$ *is denoted as*

$$\boldsymbol{L}_1 \equiv \boldsymbol{L}_2.$$

Theorem C.11. [32] *Let* $(X, \vee, \wedge; \leq)$ *and* $(Y, \varovee, \varowedge; \precsim)$ *be lattices and* $\theta \in Y^X$ *be a* BIJECTIVE *function with inverse* $\theta^{-1} \in X^Y$. *Let* $(X, \vee, \wedge; \leq) \equiv (Y, \varovee, \varowedge; \precsim)$ *represent the condition that the two lattices*

[32] Burris and Sankappanavar (2000), page 10

are ISOMORPHIC.

THM

$$\underbrace{\left.\begin{array}{llll} x_1 \le x_2 & \implies & \theta(x_1) \precsim \theta(x_2) & \forall x_1,x_2 \in X \\ y_1 \precsim y_2 & \implies & \theta^{-1}(y_1) \precsim \theta^{-1}(y_2) & \forall y_1,y_2 \in Y \end{array}\right\}}_{\theta \text{ and } \theta^{-1} \text{ are ORDER PRESERVING with respect to } \le \text{ and } \precsim^{33}} \iff \underbrace{(X, \vee, \wedge; \le) \equiv (Y, \ovee, \owedge; \precsim)}_{\text{isomorphic}}$$

✎PROOF: Let $\theta \in Y^X$ be the isomorphism between lattices $(X, \vee, \wedge; \le)$ and $(Y, \ovee, \owedge; \precsim)$.

1. Proof that *order preserving* $\implies$ *preserves joins*:

   (a) Proof that $\theta(x_1 \vee x_2) \succsim \theta(x_1) \ovee \theta(x_2)$:

   i. Note that

   $$\begin{aligned} x_1 &\le x_1 \vee x_2 \\ x_2 &\le x_1 \vee x_2. \end{aligned}$$

   ii. Because θ is *order preserving*

   $$\begin{aligned} \theta(x_1) &\precsim \theta(x_1 \vee x_2) \\ \theta(x_2) &\precsim \theta(x_1 \vee x_2). \end{aligned}$$

   iii. We can then finish the proof of item (1a):

   $$\begin{aligned} \theta(x_1) \ovee \theta(x_2) &\precsim \underbrace{\theta(x_1 \vee x_2)}_{x_1 \le x_1 \vee x_2} \ovee \underbrace{\theta(x_1 \vee x_2)}_{x_2 \le x_1 \vee x_2} && \text{by *order preserving* hypothesis} \\ &= \theta(x_1 \vee x_2) && \text{by *idempotent* property page 306} \end{aligned}$$

   (b) Proof that $\theta(x_1 \vee x_2) \precsim \theta(x_1) \ovee \theta(x_2)$:

   i. Just as in item (1a), note that $\theta^{-1}(y_1) \vee \theta^{-1}(y_2) \le \theta^{-1}(y_1 \ovee y_2)$:

   $$\begin{aligned} \theta^{-1}(y_1) \vee \theta^{-1}(y_2) &\le \underbrace{\theta^{-1}(y_1 \ovee y_2)}_{y_1 \precsim y_1 \ovee y_2} \vee \underbrace{\theta^{-1}(y_1 \ovee y_2)}_{y_2 \precsim y_1 \ovee y_2} && \text{by *order preserving* hypothesis} \\ &= \theta^{-1}(y_1 \ovee y_2) && \text{by *idempotent* property page 306} \end{aligned}$$

   ii. Because θ is *order preserving*

   $$\begin{aligned} \theta\left[\theta^{-1}(y_1) \vee \theta^{-1}(y_2)\right] &\precsim \theta\theta^{-1}(y_1 \ovee y_2) && \text{by item (1(b)i) page 316} \\ &= y_1 \ovee y_2 && \text{by definition of inverse function } \theta^{-1} \end{aligned}$$

   iii. Let $u_1 \triangleq \theta(x_1)$ and $u_2 \triangleq \theta(x_2)$.

   iv. We can then finish the proof of item (1b):

   $$\begin{aligned} \theta(x_1 \vee x_2) &= \theta\left[\theta^{-1}\theta(x_1) \vee \theta^{-1}\theta(x_2)\right] && \text{by definition of inverse function } \theta^{-1} \\ &= \theta\left[\theta^{-1}(u_1) \vee \theta^{-1}(u_2)\right] && \text{by definition of } u_1, u_2, \text{ item (1(b)iii)} \\ &\precsim u_1 \ovee u_2 && \text{by item (1(b)ii)} \\ &= \theta(x_1) \ovee \theta(x_2) && \text{by definition of } u_1, u_2, \text{ item (1(b)iii)} \end{aligned}$$

   (c) And so, combining item (1a) and item (1b), we have

   $$\left.\begin{array}{llll} \theta(x_1 \vee x_2) & \succsim & \theta(x_1) \ovee \theta(x_2) & \text{(item (1a) page 316)} \quad \text{and} \\ \theta(x_1 \vee x_2) & \precsim & \theta(x_1) \ovee \theta(x_2) & \text{(item (1b) page 316)} \end{array}\right\} \implies \theta(x_1 \vee x_2) = \theta(x_1) \ovee \theta(x_2)$$

[33] *order preserving*: Definition B.9 page 297

2. Proof that *order preserving* $\Longrightarrow$ *preserves meets*:

   (a) Proof that $\theta(x_1 \wedge x_2) \precsim \theta(x_1) \owedge \theta(x_2)$:

$$\begin{aligned}\theta(x_1) \owedge \theta(x_2) &\circledcirc \underbrace{\theta(x_1 \wedge x_2)}_{x_1 \geq x_1 \wedge x_2} \owedge \underbrace{\theta(x_1 \wedge x_2)}_{x_2 \geq x_1 \wedge x_2} && \text{by } \textit{order preserving} \text{ hypothesis}\\ &= \theta(x_1 \wedge x_2) && \text{by } \textit{idempotent} \text{ property page 306}\end{aligned}$$

   (b) Proof that $\theta(x_1 \wedge x_2) \circledcirc \theta(x_1) \owedge \theta(x_2)$:

      i. Just as in item (2a), note that $\theta^{-1}(y_1) \wedge \theta^{-1}(y_2) \geq \theta^{-1}(y_1 \owedge y_2)$:

$$\begin{aligned}\theta^{-1}(y_1) \wedge \theta^{-1}(y_2) &\geq \underbrace{\theta^{-1}(y_1 \owedge y_2)}_{y_1 \circledcirc y_1 \owedge y_2} \owedge \underbrace{\theta^{-1}(y_1 \owedge y_2)}_{y_2 \circledcirc y_1 \owedge y_2} && \text{by } \textit{order preserving} \text{ hypothesis}\\ &= \theta^{-1}(y_1 \owedge y_2) && \text{by } \textit{idempotent} \text{ property page 306}\end{aligned}$$

      ii. Because θ is *order preserving*

$$\begin{aligned}\theta\left[\theta^{-1}(y_1) \wedge \theta^{-1}(y_2)\right] &\circledcirc \theta\theta^{-1}(y_1 \owedge y_2) && \text{by item (2(b)i)}\\ &= y_1 \owedge y_2\end{aligned}$$

      iii. Let $v_1 \triangleq \theta(x_1)$ and $v_2 \triangleq \theta(x_2)$.

      iv. We can then finish the proof of item (2a):

$$\begin{aligned}\theta(x_1 \wedge x_2) &= \theta\left[\theta^{-1}\theta(x_1) \wedge \theta^{-1}\theta(x_2)\right]\\ &= \theta\left[\theta^{-1}(v_1) \wedge \theta^{-1}(v_2)\right] && \text{by item (2(b)iii)}\\ &\circledcirc v_1 \owedge v_2 && \text{by item (2(b)ii)}\\ &= \theta(x_1) \owedge \theta(x_2) && \text{by item (2(b)iii)}\end{aligned}$$

   (c) And so, combining item (2a) and item (2b), we have

$$\left.\begin{array}{lcll}\theta(x_1 \wedge x_2) & \precsim & \theta(x_1) \owedge \theta(x_2) & \text{(item (2a) page 317)} \quad \text{and}\\ \theta(x_1 \wedge x_2) & \circledcirc & \theta(x_1) \owedge \theta(x_2) & \text{(item (2b) page 317)}\end{array}\right\} \Longrightarrow \theta(x_1 \wedge x_2) = \theta(x_1) \owedge \theta(x_2)$$

3. Proof that *order preserving* $\Longleftarrow$ *isomorphic*:

$$\begin{aligned}x \leq y &\Longrightarrow \theta(y) = \theta(x \vee y) = \theta(x) \ovee \theta(y) && \text{by right hypothesis}\\ &\Longrightarrow \theta(x) \precsim \theta(y)\\ x \leq y &\Longrightarrow \theta(x) = \theta(x \wedge y) = \theta(x) \owedge \theta(y) && \text{by right hypothesis}\\ &\Longrightarrow \theta(x) \precsim \theta(y)\end{aligned}$$

✏

Example C.15. Let $\boldsymbol{L} \equiv \boldsymbol{M}$ represent the condition that a lattice $\boldsymbol{L}$ and a lattice $\boldsymbol{M}$ are *isomorphic*.

EX

$$\left(2^{\{x,y,z\}}, \cup, \cap; \subseteq\right) \equiv \left(\{1,2,3,5,6,10,15,30\}, \text{lcm}, \text{gcd}; \mid\right)$$

with isomorphism

$$\theta(A) = 5^{\mathbb{1}_A(z)} \cdot 3^{\mathbb{1}_A(y)} \cdot 2^{\mathbb{1}_A(x)} \qquad \forall A \in 2^{\{a,b,c\}}$$

Explicit cases are listed below and illustrated in Example B.9 (page 293) and Example B.10 (page 293).

$$\begin{aligned}\theta(\emptyset) &= 5^0 \cdot 3^0 \cdot 2^0 &&= 1 \qquad& \theta(\{z\}) &= 5^1 \cdot 3^0 \cdot 2^0 &&= 5\\ \theta(\{x\}) &= 5^0 \cdot 3^0 \cdot 2^1 &&= 2 & \theta(\{x,z\}) &= 5^1 \cdot 3^0 \cdot 2^1 &&= 10\\ \theta(\{y\}) &= 5^0 \cdot 3^1 \cdot 2^0 &&= 3 & \theta(\{y,z\}) &= 5^1 \cdot 3^1 \cdot 2^0 &&= 15\\ \theta(\{x,y\}) &= 5^0 \cdot 3^1 \cdot 2^1 &&= 6 & \theta(\{x,y,z\}) &= 5^1 \cdot 3^1 \cdot 2^1 &&= 30\end{aligned}$$

✎PROOF:

$$\begin{aligned}
\theta(A \cup B) &= 5^{\mathbb{1}_{A\cup B}(a)} \cdot 3^{\mathbb{1}_{A\cup B}(b)} \cdot 2^{\mathbb{1}_{A\cup B}(c)} \\
&= 5^{\mathbb{1}_A(a)\vee \mathbb{1}_B(a)} \cdot 3^{\mathbb{1}_A(b)\vee \mathbb{1}_B(b)} \cdot 2^{\mathbb{1}_A(c)\vee \mathbb{1}_B(c)} \\
&= \operatorname{lcm}\left(5^{\mathbb{1}_A(a)},\ 5^{\mathbb{1}_B(a)}\right) \cdot \operatorname{lcm}\left(3^{\mathbb{1}_A(b)},\ 3^{\mathbb{1}_B(b)}\right) \cdot \operatorname{lcm}\left(2^{\mathbb{1}_A(c)},\ 2^{\mathbb{1}_B(c)}\right) \\
&= \operatorname{lcm}\left(5^{\mathbb{1}_A(a)} \cdot 3^{\mathbb{1}_A(b)} \cdot 2^{\mathbb{1}_A(c)},\ 5^{\mathbb{1}_B(a)} \cdot 3^{\mathbb{1}_B(b)} \cdot 2^{\mathbb{1}_B(c)}\right) \\
&= \operatorname{lcm}\left(\theta(A),\ \theta(B)\right)
\end{aligned}$$

$$\begin{aligned}
\theta(A \cap B) &= 5^{\mathbb{1}_{A\cap B}(a)} \cdot 3^{\mathbb{1}_{A\cap B}(b)} \cdot 2^{\mathbb{1}_{A\cap B}(c)} \\
&= 5^{\mathbb{1}_A(a)\wedge \mathbb{1}_B(a)} \cdot 3^{\mathbb{1}_A(b)\wedge \mathbb{1}_B(b)} \cdot 2^{\mathbb{1}_A(c)\wedge \mathbb{1}_B(c)} \\
&= \gcd\left(5^{\mathbb{1}_A(a)},\ 5^{\mathbb{1}_B(a)}\right) \cdot \gcd\left(3^{\mathbb{1}_A(b)},\ 3^{\mathbb{1}_B(b)}\right) \cdot \gcd\left(2^{\mathbb{1}_A(c)},\ 2^{\mathbb{1}_B(c)}\right) \\
&= \gcd\left(5^{\mathbb{1}_A(a)} \cdot 3^{\mathbb{1}_A(b)} \cdot 2^{\mathbb{1}_A(c)},\ 5^{\mathbb{1}_B(a)} \cdot 3^{\mathbb{1}_B(b)} \cdot 2^{\mathbb{1}_B(c)}\right) \\
&= \gcd\left(\theta(A),\ \theta(B)\right)
\end{aligned}$$

⇨

C.5.2 Metrics

Definition C.5. [34] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge; \leq)$ *be a lattice.*

DEF

A function $\mathsf{v} \in \mathbb{R}^X$ *is a* ***subvaluation*** *if*

1. $\mathsf{v}(x) \geq 0 \qquad \forall x \in X \qquad$ *and*
2. $\mathsf{v}(x \vee y) + \mathsf{v}(x \wedge y) \leq \mathsf{v}(x) + \mathsf{v}(y) \qquad \forall x,y \in X$

A subvaluation v *is* ***isotone*** *if* $x \leq y \implies \mathsf{v}(x) \leq \mathsf{v}(y)$.

A subvaluation v *is* ***positive*** *if* $x < y \implies \mathsf{v}(x) < \mathsf{v}(y)$.

Definition C.6. [35] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge; \leq)$ *be a lattice.*

DEF

A function $\mathsf{v} \in \mathbb{R}^X$ *is a* ***valuation*** *if*

1. $\mathsf{v}(x) \geq 0 \qquad \forall x \in X \qquad$ *and*
2. $\mathsf{v}(x \vee y) + \mathsf{v}(x \wedge y) = \mathsf{v}(x) + \mathsf{v}(y) \qquad \forall x,y \in X \qquad$ *and*
3. $x \leq y \implies \mathsf{v}(x) \leq \mathsf{v}(y) \qquad \forall x,y \in X \qquad$ (ISOTONE).

Proposition C.3 (lattice subvaluation metric). [36] *Let* $\mathbf{L}$ *be a lattice.*

PRP

$\left\{ \mathsf{v} \text{ is a positive SUBVALUATION on } \mathbf{L} \right\} \implies \left\{ \mathsf{d}(x,y) = 2\mathsf{v}(x \vee y) - \mathsf{v}(x) - \mathsf{v}(y) \text{ is a metric.} \right\}$

Proposition C.4 (lattice valuation metric). [37] *Let* $\mathbf{L}$ *be a lattice.*

PRP

$\left\{ \mathsf{v} \text{ is a positive VALUATION on } \mathbf{L} \right\} \implies \left\{ \mathsf{d}(x,y) = \mathsf{v}(x) + \mathsf{v}(y) - 2\mathsf{v}(x \wedge y) \text{ is a metric.} \right\}$

[34] Deza and Deza (2006) page 143

[35] Deza and Deza (2006) page 143, Istrăţescu (1987) page 127 ⟨differs from Deza⟩, Birkhoff (1948) page 74 ⟨not compatible with Deza⟩

[36] Deza and Deza (2006) page 143

[37] Deza and Deza (2006) page 143

C.5.3 Lattice products

Theorem C.12 (lattice product). [38] *Let* $(X \times Y, \le)$ *be the* POSET PRODUCT[39] *of* $(X, \precsim)$ *and* $(Y, \trianglelefteq)$.

THM

$$\left.\begin{array}{lll} (X, \varovee, \varowedge; \precsim) & \text{is a lattice} & \text{and} \\ (Y, \veebar, \barwedge; \trianglelefteq) & \text{is a lattice} & \end{array}\right\} \implies (X \times Y, \vee, \wedge; \le) \text{ is also a lattice}$$

C.6 Literature

Literature survey:

1. Early lattice theory concepts:
   Dedekind (1900)
   Ore (1935)
2. Garrett Birkhoff's contribution:
   (a) The modern concept of the lattice was introduced by Garrett Birkhoff in 1933:
   Birkhoff (1933a)
   Birkhoff (1933b)
   (b) However, Birkhoff came to realize that the concept of the lattice had actually already been published in 1900 by Richard Dedekind. Birkhoff later remarked in an interview "My ideas about lattice theory developed gradually …It was my father who, when he told Ore at Yale about what I was doing some time in 1933, found out from Ore that my lattices coincided with Dedekind's Dualgruppen …I was lucky to have gone beyond Dedekind before I discovered his work. It would have been quite discouraging if I had discovered all my results anticipated by Dedekind."[40]
   (c) Birkhoff wrote a book in 1940 called *Lattice Theory*. There are basically three editions:
   Birkhoff (1940)
   Birkhoff (1948)
   Birkhoff (1967) With regards to his *Lattice Theory* book and another book entitled *A Survey of Modern Algebra* written with Saunders MacLane, Birkhoff remarked, "Morse had told me that no one under 30 should write a book. So I thought it over and wrote two!"[41]
3. Standard text books of lattice theory:
   Birkhoff (1967)
   Grätzer (1998)
   Crawley and Dilworth (1973)
4. Characterizations / equational bases:
   (a) General discussion:
   Tarski (1966)
   Baker (1969)
   McKenzie (1970)
   McKenzie (1972)
   Pigozzi (1975)
   Taylor (1979)
      Taylor (2008)
   Jipsen and Rose (1992) pages 115–127 ⟨Chapter 5⟩
   Padmanabhan and Rudeanu (2008)
   (b) Characterizations for lattices:
   Kalman (1968)
   Tamura (1975)
   Sobociński (1979)

---

[38] MacLane and Birkhoff (1967), page 489
[39] *poset product*: Definition B.5 page 291
[40] Albers and Alexanderson (1985), page 4
[41] Albers and Alexanderson (1985), page 4

(c) Specific characterizations:
Padmanabhan (1969) ⟨2 equations in 7 variables⟩
McCune and Padmanabhan (1996), page 144 ⟨1 equation, 7 variables, length 79⟩
McCune et al. (2003a) ⟨1 equation, 8 variables, length 29⟩
McCune et al. (2003b) ⟨1 equation, 8 variables, length 29⟩

5. Lattice drawing program:
Ralph Freese, http://www.math.hawaii.edu/~ralph/LatDraw/

APPENDIX D

NEGATION

> ❝ *When we say not-being, we speak, I think, not of something that is the opposite of being, but only of something different. ...Then when we are told that the negative signifies the opposite, we shall not admit it; we shall admit only that the particle "not" indicates something different from the words to which it is prefixed, or rather from the things denoted by the words that follow the negative.* ❞
>
> Plato's the *Sophist* (circa 360 B.C.) [1]

> ❝ *Clearly, then, it is a principle of this kind that is the most certain of all principles.... Let us next state what this principle is. "It is impossible for the same attribute at once to belong and not to belong to the same thing and in the same relation"; ...This is the most certain of all principles,... for it is impossible for anyone to suppose that the same thing is and is not,... it is by nature the starting-point of all the other axioms as well.* ❞
>
> Aristotle (384BC–322BC), Greek philosopher [2]

D.1 Definitions

Definition D.1. [3] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ *be a* BOUNDED LATTICE *(Definition* **??** *page* **??***).*

DEF: *A* FUNCTION $\neg \in X^X$ *is a **subminimal negation** on* $\mathbf{L}$ *if*

$$x \leq y \implies \neg y \leq \neg x \quad \forall x,y \in X \qquad \text{(ANTITONE)}^{4}$$

Remark D.1. [5] In the context of natural language, D. Devidi argues that, *subminimal negation* (Definition D.1 page 321) is "difficult to take seriously as" a negation. He essentially gives this example: Let $x \triangleq$"p is a fish" and $y \triangleq$"p has gills". Suppose "p is a fish" implies "p has gills" ($x \leq y$). Now let $p \triangleq$ "many dogs". Then the *antitone* property and $x \leq y$ tells us ($\implies$) that "Not many dogs have gills" implies that "Not many dogs are fish".

---

[1] Plato (circa 360 B.C.) ⟨257b–257c⟩, Horn (2001), page 5

[2] Aristotle page 4.1005b

[3] Dunn (1996) pages 4–6, Dunn (1999) pages 24–26 ⟨2 THE KITE OF NEGATIONS⟩

[4] The *antitone* property may also be referred to as *antitonic*, *order-reversing*, or *contrapositive*.

[5] Devidi (2010) page 511, Devidi (2006) page 568

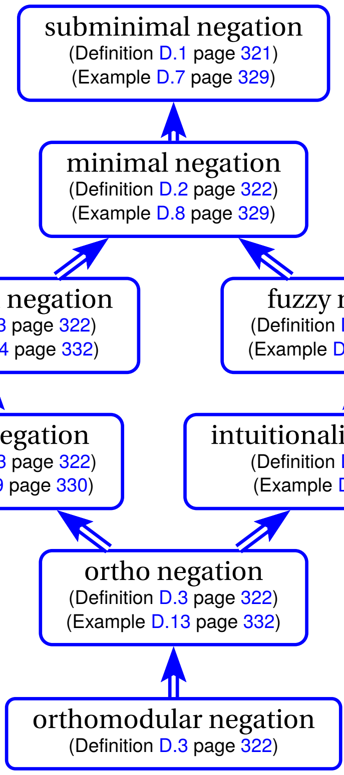


Figure D.1: lattice of negations

Definition D.2. [6] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ *be a* BOUNDED LATTICE *(Definition ?? page ??).*

DEF

A FUNCTION $\neg \in X^X$ *is a **negation**, or **minimal negation**, on* $\mathbf{L}$ *if*

1. $x \leq y \implies \neg y \leq \neg x \quad \forall x,y \in X$ (ANTITONE) *and*
2. $x \leq \neg\neg x \quad \forall x \in X$ (WEAK DOUBLE NEGATION).

A MINIMAL NEGATION $\neg$ *is an **intuitionistic negation** if*

3. $x \wedge \neg x = 0 \quad \forall x,y \in X$ (NON-CONTRADICTION).

A MINIMAL NEGATION $\neg$ *is a **fuzzy negation** if*

4. $\neg 1 = 0$ (BOUNDARY CONDITION).

Definition D.3. [7] *Let* $\mathbf{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ *be a* BOUNDED LATTICE *(Definition ?? page ??).*

DEF

A MINIMAL NEGATION $\neg$ *is a **de Morgan negation** if*

5. $x = \neg\neg x \quad \forall x \in X$ (INVOLUTORY).

A DE MORGAN NEGATION $\neg$ *is a **Kleene negation** if*

6. $x \wedge \neg x \leq y \vee \neg y \quad \forall x,y \in X$ (KLEENE CONDITION).

A DE MORGAN NEGATION $\neg$ *is an **ortho negation** if*

7. $x \wedge \neg x = 0 \quad \forall x,y \in X$ (NON-CONTRADICTION).

A DE MORGAN NEGATION $\neg$ *is an **orthomodular negation** if*

8. $x \wedge \neg x = 0 \quad \forall x,y \in X$ (NON-CONTRADICTION) *and*
9. $x \leq y \implies x \vee (y \wedge \neg x) = y \quad \forall x,y \in X$ (ORTHOMODULAR).

[6] Dunn (1996) pages 4–6, Dunn (1999) pages 24–26 ⟨2 The Kite of Negations⟩, Troelstra and van Dalen (1988) page 4 ⟨1.6 Intuitionism. (b)⟩, de Vries (2007) page 11 ⟨Definition 16⟩, Gottwald (1999) page 21 ⟨Definition 3.3⟩, Novák et al. (1999) page 50 ⟨Definition 2.26⟩, Nguyen and Walker (2006) pages 98–99 ⟨5.4 Negations⟩, Höhle (1978) ⟨???⟩, Bellman and Giertz (1973) pages 155–156 ⟨(N1) $\neg 0 = 1$ and $\neg 1 = 0$, (N3) $\neg\neg x = x$⟩

[7] Dunn (1999) pages 24–26 ⟨2 The Kite of Negations⟩, Jenei (2003) page 283, Kalmbach (1983) page 22, Lidl and Pilz (1998) page 90, Husimi (1937)

Remark D.2. [8] The *Kleene condition* is basically a weakened form of the *non-contradiction* and *excluded middle* properties because $\underbrace{x \wedge \neg x = 0}_{\text{non-contradiction}} \le \underbrace{1 = y \vee \neg y}_{\text{excluded middle}}$.

Definition D.4. [9]

DEF

A MINIMAL NEGATION $\neg \in X^X$ *is* ***strict*** *($\neg$ is a* ***strict negation****) if*

1. $x \lneq y \implies \neg y \lneq \neg x \quad \forall x,y \in X$ (STRICTLY ANTITONE) *and*
2. $\neg$ *is* CONTINUOUS

A STRICT NEGATION $\neg$ *is* ***strong*** *($\neg$ is a* ***strong negation****) if*

3. $\neg\neg x = x \quad \forall x \in X$ (INVOLUTORY).

Definition D.5. *Let* $\mathbf{L} \triangleq (X, \vee, \wedge, \neg, 0, 1; \le)$ *be a* BOUNDED LATTICE *(Definition* **??** *page* **??***) with a function* $\neg$ *in* X^X.

DEF

If $\neg$ *is a* MINIMAL NEGATION, *then* $\mathbf{L}$ *is a* ***lattice with negation****.*

D.2 Properties of negations

Lemma D.1. [10] *Let* $\neg \in X^X$ *be a function on a* BOUNDED LATTICE $\mathbf{L} \triangleq (X, \vee, \wedge, 0, 1; \le)$ *(Definition* **??** *page* **??***).*

LEM

$$\underbrace{x \le y \implies \neg y \le \neg x\}}_{\text{ANTITONE}} \implies \left\{ \begin{array}{lllll} \neg x \vee \neg y & \le & \neg(x \wedge y) & \forall x,y \in X & \text{(CONJUNCTIVE DE MORGAN INEQ.) and} \\ \neg(x \vee y) & \le & \neg x \wedge \neg y & \forall x,y \in X & \text{(DISJUNCTIVE DE MORGAN INEQ.) and} \end{array}\right.$$

✎PROOF:

1. Proof that *antitone* $\implies$ *conjunctive de Morgan*:

$$\begin{aligned} & x \wedge y \le x \text{ and } x \wedge y \le y && \text{by definition of } \wedge \\ \implies & \neg(x \wedge y) \ge \neg x \text{ and } \neg(x \wedge y) \ge \neg y && \text{by } \textit{antitone} \\ \implies & \neg(x \wedge y) \ge \neg x \vee \neg y && \text{by definition of } \vee \end{aligned}$$

2. Proof that *antitone* $\implies$ *disjunctive de Morgan*:

$$\begin{aligned} & x \le x \vee y \text{ and } y \le x \vee y && \text{by definition of } \vee \\ \implies & \neg x \ge \neg(x \vee y) \text{ and } \neg y \ge \neg(x \vee y) && \text{by } \textit{antitone} \\ \implies & \neg x \wedge \neg y \ge \neg(x \vee y) && \text{by definition of } \wedge \\ \implies & \neg(x \vee y) \le \neg x \wedge \neg y && \end{aligned}$$

⇨

[8] Cattaneo and Ciucci (2009) page 78
[9] Fodor and Yager (2000), pages 127–128, Bellman and Giertz (1973)
[10] Beran (1985) page 31 ⟨Theorem 1.2 Proof⟩, Fáy (1967) page 268 ⟨Lemma 1 Proof⟩, de Vries (2007) page 12 ⟨Theorem 18⟩

Lemma D.2. [11] *Let* $\neg \in X^X$ *be a function on a* LATTICE $\boldsymbol{L} \triangleq (X, \vee, \wedge; \leq)$ (Definition C.3 page 305).

LEM

If $x = (\neg\neg x)$ *for all* $x \in X$ (INVOLUTORY)*, then*

$$\underbrace{x \leq y \implies \neg y \leq \neg x\}}_{\text{ANTITONE}} \iff \underbrace{\left\{\begin{array}{lllll} \neg(x \vee y) & = & \neg x \wedge \neg y & \forall x,y \in X & \text{(DISJUNCTIVE DE MORGAN) and} \\ \neg(x \wedge y) & = & \neg x \vee \neg y & \forall x,y \in X & \text{(CONJUNCTIVE DE MORGAN)} \end{array}\right.}_{\text{DE MORGAN}}$$

✎PROOF:

1. Proof that *antitone* $\implies$ *de Morgan* equalities:

   (a) Proof that $\neg(\neg x \wedge \neg y) \geq x \vee y$:

   $$\begin{aligned} \neg(\neg x \wedge \neg y) &\geq \neg\neg x \vee \neg\neg y && \text{by Lemma D.1} \\ &= x \vee y && \text{by } \textit{involutory} \text{ property (Definition D.5 page 323)} \end{aligned}$$

   (b) Proof that $\neg(\neg x \vee \neg y) \leq x \wedge y$:

   $$\begin{aligned} \neg(\neg x \vee \neg y) &\leq \neg\neg x \wedge \neg\neg y && \text{by Lemma D.1} \\ &= x \wedge y && \text{by } \textit{involutory} \text{ property (Definition D.5 page 323)} \end{aligned}$$

   (c) Proof that $\neg(x \wedge y) = \neg x \vee \neg y$:

   $$\begin{aligned} \neg(x \wedge y) &\geq \neg x \vee \neg y && \text{by Lemma D.1} \\ \neg(x \wedge y) &= \neg[\neg\neg x \wedge \neg\neg y] && \text{by } \textit{involutory} \text{ property (Definition D.5 page 323)} \\ &\leq \neg x \vee \neg y && \text{by item (1b)} \end{aligned}$$

   (d) Proof that $\neg(x \vee y) = \neg x \wedge \neg y$:

   $$\begin{aligned} \neg(x \vee y) &\geq \neg x \wedge \neg y && \text{by Lemma D.1} \\ \neg(x \vee y) &= \neg[\neg\neg x \vee \neg\neg y] && \text{by } \textit{involutory} \text{ property (Definition D.5 page 323)} \\ &\leq \neg x \wedge \neg y && \text{by item (1a)} \end{aligned}$$

2. Proof that *antitone* $\impliedby$ *de Morgan*:

   $$\begin{aligned} x \leq y \implies \neg y &= \neg(x \vee y) && \text{because } x \leq y \\ &= \neg x \wedge \neg y && \text{by } \textit{de Morgan} \\ &\leq \neg x && \text{by definition of } \wedge \end{aligned}$$

⇨

Lemma D.3. *Let* $\neg \in X^X$ *be a function on a* LATTICE $\boldsymbol{L} \triangleq (X, \vee, \wedge; \leq)$ (Definition C.3 page 305).

LEM

$$\left\{\begin{array}{llll} 1. & x \leq \neg\neg x \quad \forall x \in X & \text{(WEAK DOUBLE NEGATION)} & \text{and} \\ 2. & \neg 1 = 0 & \text{(BOUNDARY CONDITION)} & \end{array}\right\} \implies \left\{ \neg 0 = 1 \quad \text{(BOUNDARY CONDITION)} \right\}$$

✎PROOF:

$$\begin{aligned} \neg 0 &= \neg\neg 1 && \text{by } \textit{boundary condition} \text{ hypothesis (2)} \\ &\geq 1 && \text{by } \textit{weak double negation} \text{ hypothesis (1)} \\ &\implies \neg 0 = 1 && \text{by } \textit{upper bound} \text{ property (Definition ?? page ??)} \end{aligned}$$

⇨

[11] Beran (1985) pages 30–31 ⟨Theorem 1.2⟩, Fáy (1967) page 268 ⟨Lemma 1⟩, Nakano and Romberger (1971) ⟨cf Beran 1985⟩

Lemma D.4. *Let* $\neg \in X^X$ *be a function on a* LATTICE $\boldsymbol{L} \triangleq (X, \vee, \wedge; \leq)$ (Definition C.3 page 305).

LEM
$$\left\{ (x \wedge \neg x = 0 \quad \forall x \in X \quad \text{(NON-CONTRADICTION)} \right\} \implies \left\{ \neg 1 = 0 \quad \text{(BOUNDARY CONDITION)} \right\}$$

✎PROOF:

$$\begin{aligned} 0 &= 1 \wedge \neg 1 && \text{by *non-contradiction* hypothesis} \\ &= \neg 1 && \text{by definition of g.u.b. } 1 \text{ and } \wedge \end{aligned}$$

⇨

Lemma D.5. [12] *Let* $\neg \in X^X$ *be a function on a* BOUNDED LATTICE $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ (Definition ?? page ??).

LEM
$$\left\{ \begin{array}{lll} \text{(A).} & \neg \text{ is BIJECTIVE} & \text{and} \\ \text{(B).} & x \leq y \implies \neg y \leq \neg x \quad \forall x,y \in X & \text{(ANTITONE)} \end{array} \right\} \implies \underbrace{\left\{ \begin{array}{ll} \text{(1).} & \neg 0 = 1 \quad \text{and} \\ \text{(2).} & \neg 1 = 0 \end{array} \right\}}_{\text{BOUNDARY CONDITIONS}}$$

✎PROOF:

1. Proof that $\neg 0 = 1$:

$$\begin{aligned} & x \leq 1 && \forall x \in X && \text{by definition of l.u.b. } 1 \\ \implies & \neg 1 \leq \neg x && \forall x \in X && \text{by *antitone* hypothesis} \\ \implies & \neg 1 \leq y && \forall y \in X && \text{by *bijective* hypothesis} \\ \implies & \neg 1 = 0 && && \text{by definition of g.l.b. } 0 \end{aligned}$$

2. Proof that $\neg 0 = 1$:

$$\begin{aligned} & 0 \leq x && \forall x \in X && \text{by definition of g.l.b. } 0 \\ \implies & \neg x \leq \neg 0 && \forall x \in X && \text{by *antitone* hypothesis} \\ \implies & \neg x \leq y && \forall y \in X && \text{by *bijective* hypothesis} \\ \implies & \neg 0 = 1 && && \text{by definition of l.u.b. } 1 \end{aligned}$$

⇨

Theorem D.1. *Let* $\neg \in X^X$ *be a function on a* BOUNDED LATTICE $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ (Definition ?? page ??).

THM
$$\left\{ \begin{array}{l} \neg \text{ is an} \\ \text{INTUITIONISTIC NEGATION} \end{array} \right\} \implies \left\{ \neg 1 = 0 \quad \text{(BOUNDARY CONDITION)} \right\}$$

✎PROOF: This follows directly from Definition D.5 (page 323) and Lemma D.4 (page 325). ⇨

Theorem D.2. *Let* $\neg \in X^X$ *be a function on a* BOUNDED LATTICE $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ (Definition ?? page ??).

THM
$$\left\{ \begin{array}{l} \neg \text{ is a} \\ \text{FUZZY NEGATION} \end{array} \right\} \implies \left\{ \neg 0 = 1 \quad \text{(BOUNDARY CONDITION)} \right\}$$

[12] Varadarajan (1985) page 42

✎PROOF: This follows directly from Definition D.2 (page 322) and Lemma D.3 (page 324). ⇨

Theorem D.3. *Let* $\neg \in X^X$ *be a function on a* BOUNDED LATTICE $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1\,;\, \leq)$ *(Definition ?? page ??).*

$$\text{THM}\quad \left\{\begin{array}{l}\neg \textit{ is a}\\ \textit{minimal}\\ \textit{negation}\end{array}\right\} \implies \left\{\begin{array}{lllll}\neg x \vee \neg y & \leq & \neg(x \wedge y) & \forall x,y \in X & \text{(CONJUNCTIVE DE MORGAN INEQUALITY)} \quad \textit{and}\\ \neg(x \vee y) & \leq & \neg x \wedge \neg y & \forall x,y \in X & \text{(DISJUNCTIVE DE MORGAN INEQUALITY)}\end{array}\right\}$$

✎PROOF: This follows directly from Definition D.5 (page 323) and Lemma D.1 (page 323). ⇨

Theorem D.4. *Let* $\neg \in X^X$ *be a function on a* BOUNDED LATTICE $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1\,;\, \leq)$ *(Definition ?? page ??).*

$$\text{THM}\quad \left\{\begin{array}{l}\neg \textit{ is a}\\ \textit{de Morgan negation}\end{array}\right\} \implies \left\{\begin{array}{lllll}\neg(x \vee y) & = & \neg x \wedge \neg y & \forall x,y \in X & \text{(DISJUNCTIVE DE MORGAN)} \quad \textit{and}\\ \neg(x \wedge y) & = & \neg x \vee \neg y & \forall x,y \in X & \text{(CONJUNCTIVE DE MORGAN)}\end{array}\right.$$

✎PROOF: This follows directly from Definition D.5 (page 323) and Lemma D.2 (page 324). ⇨

Theorem D.5. [13] *Let* $\neg \in X^X$ *be a function on a* BOUNDED LATTICE $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1\,;\, \leq)$ *(Definition ?? page ??).*

$$\text{THM}\quad \left\{\begin{array}{l}\neg \textit{ is an}\\ \textit{ortho negation}\end{array}\right\} \implies \left\{\begin{array}{llllll}1. & \neg 0 & = & 1 & & \text{(BOUNDARY CONDITION)} \quad \textit{and}\\ 2. & \neg 1 & = & 0 & & \text{(BOUNDARY CONDITION)} \quad \textit{and}\\ 3. & \neg(x \vee y) & = & \neg x \wedge \neg y & \forall x,y \in X & \text{(DISJUNCTIVE DE MORGAN)} \quad \textit{and}\\ 4. & \neg(x \wedge y) & = & \neg x \vee \neg y & \forall x,y \in X & \text{(CONJUNCTIVE DE MORGAN)} \quad \textit{and}\\ 5. & x \vee \neg x & = & 1 & \forall x \in X & \text{(EXCLUDED MIDDLE)} \quad \textit{and}\\ 6. & x \wedge \neg x & \leq & y \vee \neg y & \forall x,y \in X & \text{(KLEENE CONDITION).}\end{array}\right\}$$

✎PROOF:

1. Proof for $0 = \neg 1$ boundary condition: by Lemma D.4 (page 325)
2. Proof for boundary conditions:

$$\begin{aligned} 1 &= \neg\neg 1 && \text{by } \textit{involutory} \text{ property}\\ &= \neg 0 && \text{by previous result}\end{aligned}$$

3. Proof for *de Morgan* properties:
   (a) By Definition D.5 (page 323), *ortho negation* is *involutory* and *antitone*.
   (b) Therefore by Lemma D.2 (page 324), *de Morgan* properties hold.
4. Proof for *excluded middle* property:

$$\begin{aligned} x \vee \neg x &= (x \vee \neg x)^{\neg\neg} && \text{by } \textit{involutory} \text{ property of } \textit{ortho negation} \text{ (Definition D.5 page 323)}\\ &= \neg(x\neg \wedge x^{\neg\neg}) && \text{by } \textit{disjunctive de Morgan} \text{ property}\\ &= \neg(\neg x \wedge x) && \text{by } \textit{involutory} \text{ property of } \textit{ortho negation} \text{ (Definition D.5 page 323)}\\ &= \neg(x \wedge \neg x) && \text{by } \textit{commutative} \text{ property of } \textit{lattices} \text{ (Definition C.3 page 305)}\\ &= \neg 0 && \text{by } \textit{non-contradiction} \text{ property of } \textit{ortho negation} \text{ (Definition D.5 page 323)}\\ &= 1 && \text{by } \textit{boundary condition} \text{ (item (2) page 326) of } \textit{minimal negation}\end{aligned}$$

[13] Beran (1985) pages 30–31, Birkhoff and Neumann (1936) page 830 ⟨L74⟩, Cohen (1989) page 37 ⟨3B.13. Theorem⟩

5. Proof for *Kleene condition*:

$$\begin{aligned} x \wedge \neg x &= 0 && \text{by } \textit{non-contradiction} \text{ property (Definition D.5 page 323)} \\ &\le 1 && \text{by definition of 0 and 1} \\ &= y \vee \neg y && \text{by } \textit{excluded middle} \text{ property (item (4) page 326)} \end{aligned}$$

D.3 Examples

Example D.1 (discrete negation). [14] Let $\boldsymbol{L} \triangleq (X, \vee, \wedge, \neg, 0, 1; \le)$ be a *bounded lattice* (Definition ?? page ??) with a function $\neg \in X^X$.

EX The function $\neg x$ defined as

$$\neg x \triangleq \begin{cases} 1 & \text{for } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

is an *intuitionistic negation* (Definition D.2 page 322) and a *fuzzy negation* (Definition D.2 page 322).

✎PROOF: To be an *intuitionistic negation*, $\neg x$ must be *antitone*, have *weak double negation*, and have the *non-contradiction* property (Definition D.2 page 322). To be a *fuzzy negation*, $\neg x$ must be *antitone*, have *weak double negation*, and have the *boundary condition* $\neg 1 = 0$.

$$\left\{\begin{array}{lcll} \neg y \le \neg x & \iff & 1 \le 1 & \text{for } 0 = x = y \\ \neg y \le \neg x & \iff & 0 \le 1 & \text{for } 0 = x \lneq y \\ \neg y \le \neg x & \iff & 0 \le 0 & \text{for } 0 \ne x \le y \end{array}\right\} \implies \neg x \text{ is } \textit{antitone}$$

$$\left\{\begin{array}{lllllllll} \neg\neg x & = & \neg 1 & = & 0 & \ge & 0 & = & x & \text{for } x = 0 \\ \neg\neg x & = & \neg 0 & = & 1 & \ge & x & = & x & \text{for } x \ne 0 \end{array}\right\} \implies \neg x \text{ has } \textit{weak double negation}$$

$$\left\{\begin{array}{lllllll} x \wedge \neg x & = & x \wedge 1 & = & 0 \wedge 0 & = & 0 & \text{for } x = 0 \\ x \wedge \neg x & = & x \wedge 0 & = & x \wedge 0 & = & 0 & \text{for } x \ne 0 \end{array}\right\} \implies \neg x \text{ has } \textit{non-contradiction} \text{ property}$$

$$\neg 1 = 0 \implies \neg x \text{ has the } \textit{boundary condition} \text{ property}$$

⇒

Example D.2 (dual discrete negation). [15] Let $\boldsymbol{L} \triangleq (X, \vee, \wedge, \neg, 0, 1; \le)$ be a *bounded lattice* (Definition ?? page ??) with a function $\neg \in X^X$.

EX The function $\neg x$ defined as

$$\neg x \triangleq \begin{cases} 0 & \text{for } x = 1 \\ 1 & \text{otherwise} \end{cases}$$

is a *subminimal negation* (Definition D.1 page 321) but it is *not* a *minimal negation* (Definition D.2 page 322) (and not any other negation defined here).

✎PROOF: To be an *subminimal negation*, $\neg x$ must be *antitone* (Definition D.1 page 321). To be a *minimal negation*, $\neg x$ must be *antitone* and have *weak double negation* (Definition D.2 page 322).

$$\left\{\begin{array}{lcll} \neg y \le \neg x & \iff & 0 \le 0 & \text{for } x = y = 1 \\ \neg y \le \neg x & \iff & 0 \le 1 & \text{for } x \lneq y = 1 \\ \neg y \le \neg x & \iff & 1 \le 1 & \text{for } x \le y \ne 1 \end{array}\right\} \implies \neg x \text{ is } \textit{antitone}$$

$$\left\{\begin{array}{lllllll} \neg\neg x & = & \neg 0 & = & 1 & \ge & x & \text{for } x = 1 \\ \neg\neg x & = & \neg 1 & = & 0 & \le & x & \text{for } x \ne 1 \end{array}\right\} \implies \neg x \text{ does } \textit{not} \text{ have } \textit{weak double negation}$$

⇒

[14] Fodor and Yager (2000) page 128, Yager (1980) pages 256–257, Yager (1979) ⟨cf Fodor⟩
[15] Fodor and Yager (2000) page 128, Ovchinnikov (1983) page 235 ⟨Example 4⟩

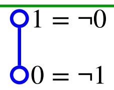
Example D.3. [16] Let $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ be a *bounded lattice*

EX The function $\neg x$ is an *intuitionistic negation* (Definition D.2 page 322) if

$$\neg x \triangleq \begin{cases} 1 & \text{for } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

Example D.4.

EX The function $\neg$ illustrated to the right is an *ortho negation* (Definition D.3 page 322).



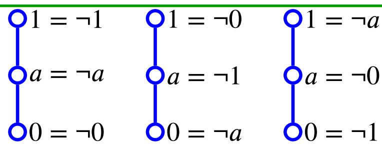
✎PROOF:

1. Proof that $\neg$ is *antitone*: $0 \leq 1 \implies \neg 1 = 0 \leq x = \neg 0 \implies \neg$ is *antitone* over $(0, 1)$
2. Proof that $\neg$ is *involutory*: $1 = \neg 0 = \neg\neg 1$
3. Proof that $\neg$ has the *non-contradiction* property: $\begin{array}{ccccccccc} 1 & \wedge & \neg 1 & = & 1 & \wedge & 0 & = & 0 \\ 0 & \wedge & \neg 0 & = & 0 & \wedge & 1 & = & 0 \end{array}$

✏

Example D.5.

EX The functions $\neg$ illustrated to the right are *not* any negation defined here. In particular, they are *not antitone*.



(a) (b) (c)

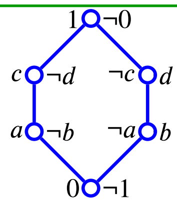
✎PROOF:

1. Proof that (a) is *not antitone*: $a \leq 1 \implies \neg 1 = 1 \nleq a = \neg a$
2. Proof that (b) is *not antitone*: $a \leq 1 \implies \neg 1 = a \nleq 0 = \neg a$
3. Proof that (c) is *not antitone*: $0 \leq a \implies \neg a = 1 \nleq a = \neg 0$

✏

Example D.6.

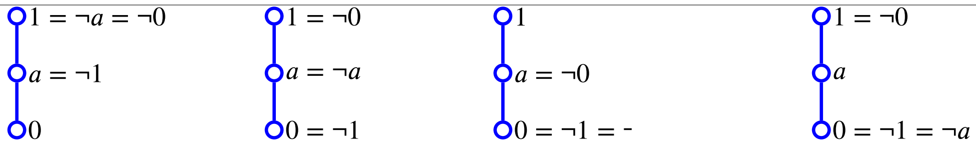
EX The function $\neg$ as illustrated to the right is *not* a *subminimal negation* (it is *not antitone*) and so is *not* any negation defined here. Note however that the problem is *not* the O_6 *lattice*—it is possible to define a negation on an O_6 *lattice* (Example D.16 page 333).



✎PROOF: Proof that $\neg$ is *not antitone*: $a \leq c \implies \neg c = d \nleq b = \neg a$ ✏

Remark D.3. The concept of a *complement* and the concept of a *negation* are fundamentally different. A *complement* is a *relation* on a lattice $\boldsymbol{L}$ and a *negation* is a *function*. In Example D.6 (page 328), b and d are both complements of a, but yet $\neg$ is *not* a negation. In the right side lattice of Example D.16 (page 333), both b and d are complements of a (and so the lattice is *multipy complemented*), but yet only d is equal to the negation of a ($d = \neg a$). It can also be said that complementation is a property *of* a lattice, whereas negation is a function defined *on* a lattice.

[16] Fodor and Yager (2000) page 128

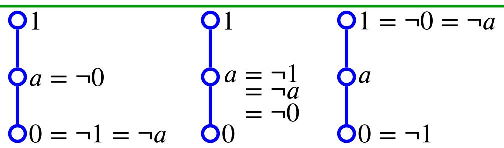


(A) *minimal negation* (Example D.8 page 329) (B) *Kleene negation* (Example D.9 page 330) (C) *intuitionistic negation* (Example D.10 page 330) (D) *fuzzy* and *intuitionistic* negation (Example D.11 page 331)

Figure D.2: negations on $\boldsymbol{L}_3$

Example D.7.

EX Each of the functions $\neg$ illustrated to the right is a *subminimal negation* (Definition D.1 page 321); *none* of them is a *minimal negation* (each fails to have *weak double negation*).



(a) (b) (c)

✎PROOF:

1. Proof that (a) $\neg$ is *antitone*:
$$\begin{array}{ccccccccccccc} a & \le & 1 & \implies & \neg 1 & = & 0 & \le & 0 & = & \neg a & \implies & \neg \text{ is } \textit{antitone} \text{ over } (a,1) \\ 0 & \le & 1 & \implies & \neg 1 & = & 0 & \le & a & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ over } (0,1) \\ 0 & \le & a & \implies & \neg a & = & 0 & \le & a & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ over } (0,a) \end{array}$$

2. Proof that (a) $\neg$ *fails* to have *weak double negation*:
$1 \not\le a = \neg 0 = \neg\neg 1$

3. Proof that (b) $\neg$ is *antitone*:
$$\begin{array}{ccccccccccccc} a & \le & 1 & \implies & \neg 1 & = & a & \le & a & = & \neg a & \implies & \neg \text{ is } \textit{antitone} \text{ over } (a,1) \\ 0 & \le & 1 & \implies & \neg 1 & = & a & \le & a & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ over } (0,1) \\ 0 & \le & a & \implies & \neg a & = & a & \le & a & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ over } (0,a) \end{array}$$

4. Proof that (b) $\neg$ *fails* to have *weak double negation*: $1 \not\le a = \neg a = \neg\neg 1$

5. (c) is a special case of the *dual discrete negation* (Example D.2 page 327).

✏

Example D.8. The function $\neg$ illustrated in Figure D.2 page 329 (A) is a **minimal negation** (Definition D.2 page 322); it is *not* an *intuitionistic negation* (it does not have the *non-contradiction* property), it is *not* a *de Morgan negation* (it is not *involutory*), and it is *not* a *fuzzy negation* ($\neg 1 \ne 0$).

✎PROOF:

1. Proof that $\neg$ is *antitone*:
$$\begin{array}{ccccccccccccc} a & \le & 1 & \implies & \neg 1 & = & a & \le & 1 & = & \neg a & \implies & \neg \text{ is } \textit{antitone} \text{ over } (a,1) \\ 0 & \le & 1 & \implies & \neg 1 & = & a & \le & 1 & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ over } (0,1) \\ 0 & \le & a & \implies & \neg a & = & 1 & \le & 1 & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ over } (0,a) \end{array}$$

2. Proof that $\neg$ is a *weak double negation* (and so is a *minimal negation*, but is *not* a *de Morgan negation*):
$$\begin{array}{ccccccccc} 1 & = & 1 & = & \neg a & = & \neg\neg 1 & \implies & \neg \text{ is } \textit{involutory} \text{ at } 1 \\ a & = & a & = & \neg 1 & = & \neg\neg a & \implies & \neg \text{ is } \textit{involutory} \text{ at } a \\ 0 & \le & a & = & \neg 1 & = & 0^{\neg\neg} & \implies & \neg \text{ is a } \textit{weak double negation} \text{ at } 0 \end{array}$$

3. Proof that $\neg$ does *not* have the *non-contradiction* property (and so is not an *intuitionistic negation*):
$1 \wedge \neg 1 = 1 \wedge a = a \ne 0$

4. Proof that $\neg$ is not a *fuzzy negation*: $\neg 1 = a \ne 0$

Example D.9 (Łukasiewicz 3-valued logic/Kleene 3-valued logic/RM_3 logic). [17] The function $\neg$ illustrated in Figure D.2 page 329 (B) is a **Kleene negation** (Definition D.3 page 322), and is also a *fuzzy negation* (Definition D.2 page 322); but it is *not* an *ortho negation* and is *not* an *intuitionistic negation* (it does not have the *non-contradiction* property).

✎PROOF:

1. Proof that $\neg$ is *antitone*:
$$\begin{array}{lllllllllll} a & \le & 1 & \implies & \neg 1 & = & 0 & \le & a & = & \neg a & \implies & \neg \text{ is } \textit{antitone} \text{ over } (a,1)\\ 0 & \le & 1 & \implies & \neg 1 & = & 0 & \le & 1 & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ over } (0,1)\\ 0 & \le & a & \implies & \neg a & = & a & \le & 1 & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ over } (0,a) \end{array}$$

2. Proof that $\neg$ is *involutory* (and so is a *de Morgan negation*):
$$\begin{array}{llllllll} 1 & = & \neg 0 & = & \neg\neg 1 & \implies & \neg \text{ is } \textit{involutory} \text{ at } 1\\ a & = & \neg a & = & \neg\neg a & \implies & \neg \text{ is } \textit{involutory} \text{ at } a\\ 0 & = & \neg 0 & = & 0^{\neg\neg} & \implies & \neg \text{ is } \textit{involutory} \text{ at } 0 \end{array}$$

3. Proof that $\neg$ does *not* have the *non-contradiction* property (and so is not an *ortho negation*):
$$x \wedge \neg x = x \wedge x = x \neq 0$$

4. Proof that $\neg$ satisfies the *Kleene condition* (and so is a *Kleene negation*):
$$\begin{array}{lllllllllllllllllll} 1 & \wedge & \neg 1 & = & 1 & \wedge & 0 & = & 0 & \le & a & = & a & \vee & a & = & a & \vee & \neg a\\ 1 & \wedge & \neg 1 & = & 1 & \wedge & 0 & = & 0 & \le & 1 & = & 0 & \vee & 1 & = & 0 & \vee & \neg 0\\ a & \wedge & \neg a & = & 1 & \wedge & a & = & a & \le & 1 & = & 1 & \vee & 0 & = & 1 & \vee & \neg 1\\ a & \wedge & \neg a & = & 1 & \wedge & a & = & a & \le & 1 & = & 0 & \vee & 1 & = & 0 & \vee & \neg 0\\ 0 & \wedge & \neg 0 & = & 0 & \wedge & 1 & = & 0 & \le & 1 & = & 1 & \vee & 0 & = & 1 & \vee & \neg 1\\ 0 & \wedge & \neg 0 & = & 0 & \wedge & 1 & = & 0 & \le & a & = & a & \vee & a & = & a & \vee & \neg a \end{array}$$

Example D.10. The function $\neg$ illustrated in Figure D.2 page 329 (C) an **intuitionistic negation** (Definition D.2 page 322); but it is *not* a *fuzzy negation* ($1 \neq \neg 0$), and it is *not* a *de Morgan negation* (it is not *involutory*).

✎PROOF:

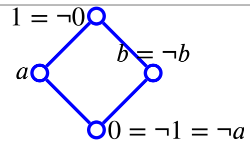
1. Proof that $\neg$ is *antitone*:
$$\begin{array}{lllllllllll} a & \le & 1 & \implies & \neg 1 & = & 0 & \le & 0 & = & \neg a & \implies & \neg \text{ is } \textit{antitone} \text{ at } (a,1)\\ 0 & \le & 1 & \implies & \neg 1 & = & 0 & \le & a & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ at } (0,1)\\ 0 & \le & a & \implies & \neg a & = & 0 & \le & a & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ at } (0,a) \end{array}$$

2. Proof that $\neg$ has *weak double negation* property (and so is a *minimal negation*, but *not* a *de Morgan negation*):
$$\begin{array}{llllllllll} 1 & \le & a & = & \neg 0 & = & \neg\neg 1 & \implies & \neg \text{ has } \textit{weak double negation} \text{ at } 1\\ a & = & \neg 0 & = & \neg\neg a & & & \implies & \neg \text{ has } \textit{weak double negation} \text{ at } a\\ 0 & = & \neg a & = & 0^{\neg\neg} & & & \implies & \neg \text{ is } \textit{involutory} \text{ at } 0 \end{array}$$

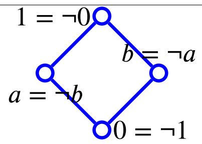
3. Proof that $\neg$ has the *non-contradiction* property (and so is an *intuitionistic negation*):
$$\begin{array}{lllllll} 1 & \wedge & \neg 1 & = & 1 & \wedge & 0 & = & 0\\ a & \wedge & \neg a & = & a & \wedge & 0 & = & 0\\ 0 & \wedge & \neg 0 & = & 0 & \wedge & a & = & 0 \end{array}$$

4. Proof that $\neg$ is *not* a *fuzzy negation*: $\neg 1 \neq 0$

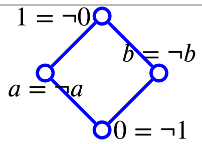
[17] Łukasiewicz (1920), Avron (1991) pages 277–278, Kleene (1938) page 153, Kleene (1952), pages 332–339 (§64. The 3-valued logic), Sobociński (1952)



(A) *fuzzy negation*
(Example D.12 page 331)



(B) *ortho negation*
(Example D.13 page 332)



(C) *de Morgan negation*
(Example D.14 page 332)

Figure D.3: negations on $\boldsymbol{M}_2$

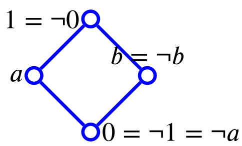
Example D.11 (Heyting 3-valued logic/Jaśkowski's first matrix). [18] The function ¬ illustrated in Figure D.2 page 329 (D) is an **intuitionistic negation** (Definition D.2 page 322), and is also a **fuzzy negation** (Definition D.2 page 322), but it is *not* a *de Morgan negation* (it is not *involutory*).

✎PROOF: This is simply a special case of the *discrete negation* (Example D.1 page 327). ✑

Remark D.4. There is only one linearly ordered (Definition B.4 page 291) 3-element lattice ($\boldsymbol{L}_3$) that is a *fuzzy negation* (Example D.11 page 331). However, this lattice is also an *intuitionistic negation*. There are no $\boldsymbol{L}_3$ lattices that are *fuzzy* but yet not *intuitionistic*. In fact, there are only three linearly ordered 3-element lattices with with $1 = \neg 0$ and $0 = \neg 1$. Of these three, only one is both *fuzzy* and *intuitionistic* (Example D.11 page 331), one is *Kleene* but not *fuzzy* (Example D.9 page 330), and one is *subminimal* but not *fuzzy* (Example D.7 page 329). It can be claimed that the "simplist" *fuzzy negation* that is not *de Morgan* and *not intuitionistic* is the $\boldsymbol{M}_2$ lattice of Example D.12 (next).

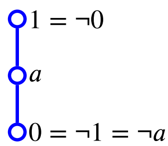
Example D.12. The function ¬ illustrated in Figure D.3 page 331 (A) is a **fuzzy negation** (Definition D.2 page 322). It is not an *intuitionistic negation* (it does not have the *non-contradiction* property) and it is *not* a *de Morgan negation* (it is not *involutory*).

✎PROOF: Note that



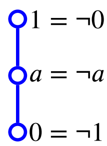
(Example D.12 page 331)

=



fuzzy and *intuitionistic*
(Example D.11 page 331)

+



Kleene negation
(Example D.9 page 330)

1. Proof that ¬ is *antitone*:
$$\begin{array}{ccccccccccccl}
a & \le & 1 & \implies & \neg 1 & = & 0 & \le & 0 & = & \neg a & \implies & \neg \text{ is } \textit{antitone} \text{ at } (a,1)\\
0 & \le & 1 & \implies & \neg 1 & = & 0 & \le & 1 & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ at } (0,1)\\
0 & \le & a & \implies & \neg a & = & 0 & \le & 1 & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ at } (0,a)\\
b & \le & 1 & \implies & \neg 1 & = & 0 & \le & b & = & \neg b & \implies & \neg \text{ is } \textit{antitone} \text{ at } (b,1)\\
0 & \le & b & \implies & \neg b & = & b & \le & 1 & = & \neg 0 & \implies & \neg \text{ is } \textit{antitone} \text{ at } (0,b)
\end{array}$$

2. Proof that ¬ has *weak double negation* property (and so is a *minimal negation*, but *not* a *de Morgan negation*):
$$\begin{array}{cccccccl}
1 & = & \neg 0 & = & \neg\neg 1 & & \implies & \neg \text{ is } \textit{involutory} \text{ at } 1\\
a & \le & 1 & = & \neg 0 & = \neg\neg a & \implies & \neg \text{ has } \textit{weak double negation} \text{ at } a\\
0 & = & \neg 1 & = & 0^{\neg\neg} & & \implies & \neg \text{ is } \textit{involutory} \text{ at } 0\\
b & = & \neg b & = & \neg\neg b & = & \implies & \neg \text{ is } \textit{involutory} \text{ at } b
\end{array}$$

3. Proof that ¬ does *not* have the *non-contradiction* property (and so is *not* an *intuitionistic negation*): $b \wedge \neg b = b \wedge b = b \neq 0$

4. Proof that ¬ is has *boundary conditions* (and so is a *fuzzy negation*): $\neg 1 = 0, \quad \neg 0 = 1$

[18] Karpenko (2006) page 45, Johnstone (1982) page 9 ⟨§1.12⟩, Heyting (1930a), Heyting (1930b), Heyting (1930c), Heyting (1930d), Jaskowski (1936), Mancosu (1998)

Example D.13. [19] The function $\neg$ illustrated in Figure D.3 page 331 (B) is an *ortho negation* (Definition D.3 page 322).

✎PROOF:

1. Proof that $\neg$ is *antitone*:
$$\begin{array}{ccccccccccc} a & \le & 1 & \Longrightarrow & \neg 1 & = & 0 & \le & b & = & \neg a \\ 0 & \le & 1 & \Longrightarrow & \neg 1 & = & 0 & \le & 1 & = & \neg 0 \\ 0 & \le & a & \Longrightarrow & \neg a & = & b & \le & 1 & = & \neg 0 \\ b & \le & 1 & \Longrightarrow & \neg 1 & = & 0 & \le & a & = & \neg b \\ 0 & \le & b & \Longrightarrow & \neg b & = & a & \le & 1 & = & \neg 0 \end{array}$$

2. Proof that $\neg$ is *involutory* (and so is a *de Morgan negation*):
$$\begin{array}{ccccc} 1 & = & \neg 0 & = & \neg\neg 1 \\ a & = & \neg a & = & \neg\neg a \\ b & = & \neg b & = & \neg\neg b \\ 0 & = & \neg 0 & = & 0^{\neg\neg} \end{array}$$

3. Proof that $\neg$ is has the *non-contradiction* property (and so is an *ortho negation*):
$$\begin{array}{ccccccccc} 1 & \wedge & \neg 1 & = & 1 & \wedge & 0 & = & 0 \\ a & \wedge & \neg a & = & a & \wedge & b & = & 0 \\ b & \wedge & \neg b & = & b & \wedge & a & = & 0 \\ 0 & \wedge & \neg 0 & = & 0 & \wedge & 1 & = & 0 \end{array}$$

⇨

Example D.14 (BN_4). [20] The function $\neg$ illustrated in Figure D.3 page 331 (C) is a **de Morgan negation** (Definition D.3 page 322), but it is *not* a *Kleene negation* and not an *ortho negation* (it does *not* satisfy the *Kleene condition*).

✎PROOF:

1. Proof that $\neg$ is *antitone*:
$$\begin{array}{ccccccccccc} a & \le & 1 & \Longrightarrow & \neg 1 & = & 0 & \le & b & = & \neg a \\ 0 & \le & 1 & \Longrightarrow & \neg 1 & = & 0 & \le & 1 & = & \neg 0 \\ 0 & \le & a & \Longrightarrow & \neg a & = & a & \le & 1 & = & \neg 0 \\ b & \le & 1 & \Longrightarrow & \neg 1 & = & 0 & \le & b & = & \neg b \\ 0 & \le & b & \Longrightarrow & \neg b & = & b & \le & 1 & = & \neg 0 \end{array}$$

2. Proof that $\neg$ is *involutory* (and so is a *de Morgan negation*):
$$\begin{array}{ccccc} 1 & = & \neg 0 & = & \neg\neg 1 \\ a & = & \neg a & = & \neg\neg a \\ b & = & \neg b & = & \neg\neg b \\ 0 & = & \neg 0 & = & 0^{\neg\neg} \end{array}$$

3. Proof that $\neg$ does *not* have the *non-contradiction* property (and so is *not* an *ortho negation*):
$$\begin{array}{ccccccccccc} a & \wedge & \neg a & = & a & \wedge & a & = & a & \ne & 0 \\ b & \wedge & \neg b & = & b & \wedge & b & = & b & \ne & 0 \end{array}$$

4. Proof that $\neg$ does *not* satisfy the *Kleene condition* (and so is a *de Morgan negation*):
$$a \wedge \neg a = a \wedge a = a \not\le b \vee \neg b = b$$

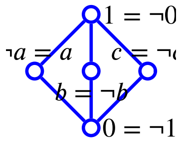
⇨

[19] Belnap (1977) page 13 Restall (2000) page 177 ⟨Example 8.44⟩, Pavičić and Megill (2008) page 28 ⟨Definition 2, *classical implication*⟩

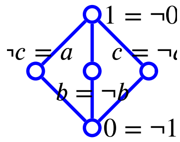
[20] Cignoli (1975) page 270, Restall (2000) page 171 ⟨Example 8.39⟩, de Vries (2007) pages 15–16 ⟨Example 26⟩, Dunn (1976), Belnap (1977)

Example D.15.

EX

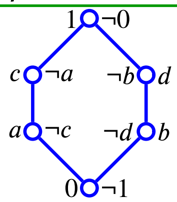


The function $\neg$ illustrated to the left is a *de Morgan negation* (Definition D.3 page 322), but it is *not* a *Kleene negation* and not an *ortho negation* (it does *not* satisfy the *Kleene condition*). The *negation* illustrated to the right is a *Kleene negation* (Definition D.3 page 322), but it is *not* an *ortho negation* (it does *not* have the *non-contradiction* property).

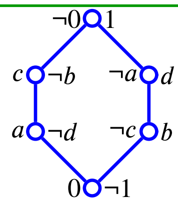


Example D.16.

EX

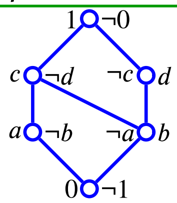


The function $\neg$ illustrated to the left is a *de Morgan negation* (Definition D.3 page 322); it is *not* a *Kleene negation* (it does not satisfy the Kleene condition). The *negation* illustrated to the right is an *ortho negation* (Definition D.3 page 322).

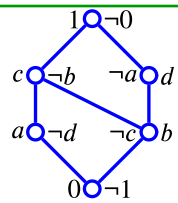


Example D.17.

EX



The function $\neg$ illustrated to the left is *not antitone* and therefore is not a *negation* (Definition D.2 page 322). The function $\neg$ illustrated to the right is a *Kleene negation* (Definition D.3 page 322); it is *not* an *ortho negation* (it does not have the *non-contradiction* property).



✎PROOF:

1. Proof that left $\neg$ is *not antitone*: $a \le c$ but $\neg c \not\le \neg a$.

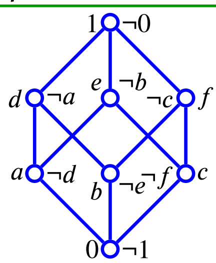
2. Proof that right $\neg$ satisfies the *Kleene condition*:
$$x \wedge \neg x = \begin{cases} b & \text{for } x = b \\ 0 & \text{otherwise} \end{cases} \quad \forall x \in X \qquad \text{and} \qquad y \wedge \neg y = \begin{cases} c & \text{for } y = c \\ 0 & \text{otherwise} \end{cases} \quad \forall y \in X$$
$$\implies \quad x \wedge \neg x \le y \vee \neg y \qquad \forall x, y \in X$$

3. Proof that right $\neg$ does not have the *non-contradiction* property: $b \wedge \neg b = b \wedge c = b \neq 0$

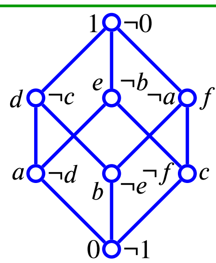
⇒

Example D.18.

EX



The lattices illustrated to the left and right are *Boolean*. The function $\neg$ illustrated to the left is a *Kleene negation* (Definition D.3 page 322), but it is *not* an *ortho negation* (it does *not* have the *non-contradiction* property). The *negation* illustrated to the right is an *ortho negation* (Definition D.3 page 322).



✎PROOF:

1. Proof that left side negation does *not* have *non-contradiction* property (and so is *not* an *ortho negation*):
$a \wedge \neg a = a \wedge d = a \neq 0$

2. Proof that left side negation does *not* satisfy *Kleene condition* (and so is *not* a *Kleene negation*):
$a \wedge \neg a = a \wedge d = a \not\leq f = c \vee f = c \vee \neg c$

✏

APPENDIX E

RELATIONS ON LATTICES WITH NEGATION

The relations in this chapter are typically defined on an *orthocomplemented lattice* (Definition 15.1 page 244). Here, some relations are generalized to a *lattice with negation* (Definition D.5 page 323). A *lattice* (Definition C.3 page 305) with an *ortho negation* negation successfully defined on it is an *orthocomplemented lattice* (Definition 15.1 page 244). In many cases, these relations only work well on an *orthocomplemented lattice*, and thus many results are restricted to orthocomplemented lattices.

E.1 Orthogonality

Proposition E.1. *Let* $(X, \vee, \wedge, 0, 1; \leq)$ *be an* ORTHOCOMPLEMENTED LATTICE *(Definition 15.1 page 244).*

PRP

$$x \leq y \implies \left\{\begin{array}{ccccc} x^\perp & \vee & y & = & 1 \quad \text{and} \\ x & \wedge & y^\perp & = & 0 \end{array}\right\} \qquad \forall x,y \in X$$

✎PROOF:

$$\begin{aligned}
x \leq y &\implies x \vee x^\perp \leq y \vee x^\perp && \text{by } \textit{monotone} \text{ property of } \textit{lattices} \text{ (Proposition C.1 page 307)} \\
&\implies 1 \leq y \vee x^\perp && \text{by } \textit{excluded middle} \text{ property of } \textit{ortho lattices} \text{ (Definition 15.1 page 244)} \\
&\implies x^\perp \vee y = 1 && \text{by } \textit{upper bounded} \text{ property of } \textit{bounded lattices} \text{ (Definition ?? page ??)} \\
x \leq y &\implies x \wedge y^\perp \leq y \wedge y^\perp && \text{by } \textit{monotone} \text{ property of } \textit{lattices} \text{ (Proposition C.1 page 307)} \\
&\implies x \wedge y^\perp \leq 0 && \text{by } \textit{non-contradiction} \text{ property of } \textit{ortho lattices} \text{ (Definition 15.1 page 244)} \\
&\implies x \wedge y^\perp = 0 && \text{by } \textit{lower bounded} \text{ property of } \textit{bounded lattices} \text{ (Definition ?? page ??)}
\end{aligned}$$

✏

Definition E.1. [1] *Let* $(X, \vee, \wedge, \neg, 0, 1; \leq)$ *be a* LATTICE WITH NEGATION *(Definition D.5 page 323).*

DEF

*The **orthogonality** relation* $\perp \in 2^{XX}$ *is defined as*

$$x \perp y \quad \overset{\text{def}}{\iff} \quad x \leq \neg y$$

If $x \perp y$*, we say that* x *is **orthogonal** to* y*.*

[1] Stern (1999) page 12, Loomis (1955) page 3

Lemma E.1. *Let* $(X, \vee, \wedge, \neg, 0, 1; \leq)$ *be a* LATTICE WITH NEGATION *(Definition D.5 page 323).*

LEM

$$\left\{ x \perp y \quad (\text{ORTHOGONAL Definition E.1 page 335}) \right\} \implies \left\{ y \perp x \quad (\text{SYMMETRIC}) \right\}$$

✎PROOF:

$$\begin{aligned}
x \perp y &\implies x \leq \neg y && \text{by definition of } \perp \text{ (Definition E.1 page 335)}\\
&\implies (\neg\neg y) \leq \neg x && \text{by } \textit{antitone} \text{ property (Definition 15.1 page 244)}\\
&\implies y \leq \neg x && \text{by } \textit{weak double negation} \text{ property of } \textit{negation} \text{ (Definition D.2 page 322)}\\
&\implies y \perp x && \text{by definition of } \perp \text{ (Definition E.1 page 335)}
\end{aligned}$$

⇨

Lemma E.2. [2] *Let* $(X, \vee, \wedge, 0, 1; \leq)$ *be an* ORTHOCOMPLEMENTED LATTICE *(Definition 15.1 page 244).*

LEM

$$\underbrace{x \perp y}_{\text{ORTHOGONAL (Definition E.1 page 335)}} \implies \left\{ \begin{array}{lrcll} 1. & x \wedge y & = & 0 & \textit{and} \\ 2. & x^{\perp} \vee y^{\perp} & = & 1 & \end{array} \right\}$$

✎PROOF:

$$\begin{aligned}
x \perp y &\implies x \leq y^{\perp} && \text{by definition of } \perp \text{ (Definition E.1 page 335)}\\
&\implies x \wedge y \leq y^{\perp} \wedge y && \text{by } \textit{monotone} \text{ property of } \textit{lattices} \text{ (Proposition C.1 page 307)}\\
&\implies x \wedge y \leq y \wedge y^{\perp} && \text{by } \textit{commutative} \text{ property of lattices (Theorem C.3 page 306)}\\
&\implies x \wedge y \leq 0 && \text{by } \textit{non-contradiction} \text{ property of } \textit{ortho negation} \text{ (Definition D.3 page 322)}\\
&\implies x \wedge y = 0 && \text{by } \textit{lower bound} \text{ property of } \textit{bounded lattices} \text{ (Definition ?? page ??)}
\end{aligned}$$

$$\begin{aligned}
x \perp y &\implies x \leq y^{\perp} && \text{by definition of } \perp \text{ (Definition E.1 page 335)}\\
&\implies x^{\perp} \vee x \leq x^{\perp} \vee y^{\perp} && \text{by } \textit{monotone} \text{ property of } \textit{lattices} \text{ (Proposition C.1 page 307)}\\
&\implies x \vee x^{\perp} \leq x^{\perp} \vee y^{\perp} && \text{by } \textit{commutative} \text{ property of } \textit{lattices} \text{ (Theorem C.3 page 306)}\\
&\implies 1 \leq x^{\perp} \vee y^{\perp} && \text{by } \textit{excluded middle} \text{ property of } \textit{ortho lattices} \text{ (Theorem D.5 page 326)}\\
&\implies x^{\perp} \vee y^{\perp} && \text{by } \textit{upper bound} \text{ property of } \textit{bounded lattices} \text{ (Definition ?? page ??)}
\end{aligned}$$

⇨

Remark E.1. In an *orthocomplemented lattice* $\boldsymbol{L}$, the *orthogonality* relation $\perp$ is in general *non-associative*. That is

$$\left\{ \begin{array}{ccll} x & \perp & y & \text{and} \\ y & \perp & z & \end{array} \right\} \;\not\Longrightarrow\; x \perp z$$

✎PROOF: Consider the $\boldsymbol{L}_2^4$ *Boolean lattice* in Example 15.2 (page 244).
- $a^{\perp} \perp p$ because $a^{\perp} \leq p^{\perp}$.
- $p \perp r$ because $p \leq r^{\perp}$.
- But yet $a^{\perp}$ is *not* orthogonal to r because $a^{\perp} \not\leq r^{\perp}$.

⇨

Example E.1.

EX

In the O_6 *lattice* (Definition 15.2 page 244), there are a total of $\binom{6}{2} = \frac{6!}{(6-2)!2!} = \frac{6{\times}5}{2} = 15$ distinct unordered (the $\perp$ relation is *symmetric* by Lemma E.1 page 336 so the order doesn't matter) pairs of elements.

Of these 15 pairs, 8 are orthogonal to each other, and 0 is orthogonal to itself, making a total of 9 orthogonal pairs:

$x \perp y$	$x \perp 0$	$y^{\perp} \perp 0$
$x \perp x^{\perp}$	$y \perp 0$	$1 \perp 0$
$y \perp y^{\perp}$	$x^{\perp} \perp 0$	$0 \perp 0$

[2] Holland (1963), page 67

Example E.2.

EX

In lattice 5 of Example 15.2 (page 244), there are a total of $\binom{10}{2} = \frac{10!}{(10-2)!2!} = \frac{10\times 9}{2} = 45$ distinct unordered pairs of elements.

Of these 45 pairs, 18 are orthogonal to each other, and 0 is orthogonal to itself, making a total of 19 orthogonal pairs:

$p \perp p^\perp$	$x \perp x^\perp$	$y \perp z$	$x^\perp \perp 0$
$p \perp x^\perp$	$x \perp y$	$y \perp 0$	$y^\perp \perp 0$
$p \perp y$	$x \perp z$	$z \perp z^\perp$	$z^\perp \perp 0$
$p \perp z$	$x \perp 0$	$z \perp 0$	$0 \perp 0$
$p \perp 0$	$y \perp y^\perp$	$p^\perp \perp 0$	

Example E.3.

EX

In the $\mathbb{R}^3$ **Euclidean space** illustrated in Example 15.3 (page 245),

$$\begin{array}{llll}
\boldsymbol{X} \subseteq \boldsymbol{Y}^\perp \implies \boldsymbol{X} \perp \boldsymbol{Y} & & \boldsymbol{Y} \subseteq \boldsymbol{X}^\perp \implies \boldsymbol{Y} \perp \boldsymbol{X} \\
\boldsymbol{X} \subseteq \boldsymbol{Z}^\perp \implies \boldsymbol{X} \perp \boldsymbol{Z} & & \boldsymbol{Y} \subseteq \boldsymbol{Z}^\perp \implies \boldsymbol{Y} \perp \boldsymbol{Z} \\
\boldsymbol{X} \wedge \boldsymbol{Y} = \boldsymbol{X} \wedge \boldsymbol{Z} = \boldsymbol{Y} \wedge \boldsymbol{Z} = \boldsymbol{0}
\end{array}$$

E.2 Commutativity

The *commutes* relation is defined next. Motivation for the name "commutes" is provided by Proposition E.4 (page 340) which shows that if x commutes with y in a lattice $\boldsymbol{L}$, then x and y commute in the *Sasaki projection* $\phi_x(y)$ on $\boldsymbol{L}$.

Definition E.2. [3] *Let* $\boldsymbol{L} \triangleq (X, \vee, \wedge, \neg, 0, 1; \leq)$ *be a* LATTICE WITH NEGATION (Definition D.5 page 323).

DEF

The ***commutes*** *relation* © *is defined as*

$$x © y \overset{\text{def}}{\iff} x = (x \wedge y) \vee (x \wedge \neg y) \qquad \forall x,y \in X,$$

in which case we say, "x ***commutes*** *with y in $\boldsymbol{L}$".*
That is, © is a relation in 2^{XX} such that

$$© \triangleq \left\{(x,y) \in X^2 | x = (x \wedge y) \vee (x \wedge \neg y)\right\}$$

Proposition E.2. [4] *Let* $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \leq)$ *be an* ORTHOCOMPLEMENTED LATTICE.

PRP

$x©0$ *and* $0©x$	$\forall x \in X$	$x©y \iff x©y^\perp$	$\forall x,y \in X$
$x©1$ *and* $1©x$	$\forall x \in X$	$x \leq y \implies x©y$	$\forall x,y \in X$
$x©x$	$\forall x \in X$	$x \perp y \implies x©y$	$\forall x,y \in X$

✎PROOF:

$$\begin{aligned}
(x \wedge 0) \vee (x \wedge 0^\perp) &= 0 \vee (x \wedge 0^\perp) && \text{by } \textit{lower bound} \text{ property of } \textit{bounded lattices} \text{ (Definition ?? page ??)}\\
&= 0 \vee (x \wedge 1) && \text{by } \textit{boundary condition} \text{ of } \textit{ortho negation} \text{ (Theorem D.5 page 326)}\\
&= 0 \vee (x) && \text{by } \textit{upper bound} \text{ property of } \textit{bounded lattices} \text{ (Definition ?? page ??)}\\
&= x && \text{by } \textit{lower bound} \text{ property of } \textit{bounded lattices} \text{ (Definition ?? page ??)}\\
&\implies x©0 && \text{by definition of © relation (Definition E.2 page 337)}\\
(0 \wedge x) \vee (0 \wedge x^\perp) &= 0 \vee (0) && \text{by } \textit{lower bound} \text{ property of } \textit{bounded lattices} \text{ (Definition ?? page ??)}\\
&= 0 && \text{by } \textit{lower bound} \text{ property of } \textit{bounded lattices} \text{ (Definition ?? page ??)}\\
&\implies 0©x && \text{by definition of © relation (Definition E.2 page 337)}
\end{aligned}$$

[3] Kalmbach (1983) page 20, Holland (1970), page 79 ⟨A. Commutativity⟩, Maeda (1958), page 227 ⟨Hilfssatz (Lemma) XII.1.2⟩, Sasaki (1954) page 301 ⟨Def.5.2, cf Foulis 1962⟩, Birkhoff (1936b) page 833 ⟨"$a = (a \cap x) \cup (a \cap x')$"⟩

[4] Holland (1963), page 67

$(x \wedge 1) \vee (x \wedge 1^{\perp}) = x \vee (x \wedge 1^{\perp})$	by *lower bound* property of *bounded lattices* (Definition **??** page **??**)
$= x \vee (x \wedge 0)$	by *boundary condition* of *ortho negation* (Theorem D.5 page 326)
$= (x) \vee (0)$	by *lower bound* property of *bounded lattices* (Definition **??** page **??**)
$= x$	by *lower bound* property of *bounded lattices* (Definition **??** page **??**)
$\implies x \copyright 1$	by definition of © relation (Definition E.2 page 337)
$(1 \wedge x) \vee (1 \wedge x^{\perp}) = (x) \vee (x^{\perp})$	by *non-contradiction* prop. of *ortho negation* (Definition D.3 page 322)
$= 1$	by *excluded middle* property of *ortho negation* (Theorem D.5 page 326)
$\implies 1 \copyright x$	by definition of © relation (Definition E.2 page 337)
$(x \wedge x) \vee (x \wedge x^{\perp}) = x \vee (x \wedge x^{\perp})$	by *idempotent* property of *lattices* (Theorem C.3 page 306)
$= x \vee (0)$	by *non-contradiction* prop. of *ortho negation* (Definition D.3 page 322)
$= x$	by *lower bound* property of *bounded lattices* (Definition **??** page **??**)
$\implies x \copyright x$	by definition of © relation (Definition E.2 page 337)
$x \copyright y \implies (x \wedge y^{\perp}) \vee (x \wedge y^{\perp\perp})$	by definition of © (Definition E.2 page 337)
$= (x \wedge y^{\perp}) \vee (x \wedge y)$	by *involution* property of $\perp$ (Definition 15.1 page 244)
$= (x \wedge y) \vee (x \wedge y^{\perp})$	by *commutative* property of *lattices* (Definition C.3 page 305)
$= x$	by $x \copyright y$ hypothesis and Definition E.2 page 337
$\implies x \copyright y^{\perp}$	by definition of © relation (Definition E.2 page 337)
$x \copyright y^{\perp} \implies (x \wedge y) \vee (x \wedge y^{\perp})$	by definition of © (Definition E.2 page 337)
$= (x \wedge y^{\perp\perp}) \vee (x \wedge y^{\perp})$	by *involution* property of $\perp$ (Definition 15.1 page 244)
$= (x \wedge y^{\perp}) \vee (x \wedge y^{\perp\perp})$	by *commutative* property of *lattices* (Definition C.3 page 305)
$= x$	by $x \copyright y^{\perp}$ hypothesis and Definition E.2 page 337
$\implies x \copyright y$	by definition of © relation (Definition E.2 page 337)
$x \leq y \implies (x \wedge y) \vee (x \wedge y^{\perp})$	by definition of © (Definition E.2 page 337)
$= x \vee (x \wedge y^{\perp})$	by $x \leq y$ hypothesis
$= x$	by *absorptive* property (Theorem C.3 page 306)
$\implies x \copyright y$	by definition of © (Definition E.2 page 337)
$x \perp y \implies (x \wedge y) \vee (x \wedge y^{\perp})$	by definition of © (Definition E.2 page 337)
$= 0 \vee (x \wedge y^{\perp})$	by Lemma E.2 page 336
$= 0 \vee x$	by $x \perp y$ hypothesis ($x \perp y \implies x \leq y^{\perp}$)
$= x \vee 0$	by *commutative* property (Theorem C.3 page 306)
$= x$	by *identity* property of *bounded lattices*
$\implies x \copyright y$	by definition of © (Definition E.2 page 337)

⇨

Definition E.3. *Let © be the* COMMUTES *relation (Definition E.2 page 337) on a* LATTICE WITH NEGATION $\boldsymbol{L} \triangleq (X, \vee, \wedge, \neg, 0, 1 ; \leq)$ *(Definition D.5 page 323).*

DEF: ***L** is symmetric if*

$$x \copyright y \implies y \copyright x \qquad \forall x,y \in X$$

In general, the commutes relation is not *symmetric*. But Proposition E.3 (next) describes some conditions under which it *is* symmetric.

Proposition E.3. [5] *Let* $(X, \vee, \wedge, 0, 1 ; \leq)$ *be an* ORTHOCOMPLEMENTED LATTICE *(Definition 15.1 page 244).*

[5] Holland (1963) page 68, Nakamura (1957) page 158

PRP

$$\underbrace{\{x \circledcirc y \implies y \circledcirc x\}}_{\circledcirc \text{ is SYMMETRIC at } (x,y)\ (1)} \iff \left\{x \leq y \implies y = x \vee \left(x^{\perp} \wedge y\right)\right\} \quad \text{(ORTHOMODULAR IDENTITY)} \quad (2)$$
$$\iff \left\{x \leq y \implies x = y \wedge \left(x \vee y^{\perp}\right)\right\} \quad (x = \phi_y(x)\ \text{(SASAKI PROJECTION)}) \quad (3)$$
$$\iff \left\{y = (x \wedge y) \vee \left[y \wedge (x \wedge y)^{\perp}\right]\right\} \quad (4)$$
$$\iff \left\{x = (x \vee y) \wedge \left[x \vee (x \vee y)^{\perp}\right]\right\} \quad (5)$$

✎PROOF:

1. Proof that (2) $\iff$ (3):

$$\begin{aligned}
x \leq y \implies y^{\perp} &\leq x^{\perp} && \text{by antitone property (Definition 15.1 page 244)}\\
\implies x^{\perp} &= y^{\perp} \vee \left(y^{\perp\perp} \wedge x^{\perp}\right) && \text{by left hypothesis}\\
\implies \left(x^{\perp}\right)^{\perp} &= \left[y^{\perp} \vee \left(y^{\perp\perp} \wedge x^{\perp}\right)\right]^{\perp} &&\\
\implies x &= \left[y^{\perp} \vee \left(y^{\perp\perp} \wedge x^{\perp}\right)\right]^{\perp} && \text{by involutory property (Definition 15.1 page 244)}\\
&= y^{\perp\perp} \wedge \left(y^{\perp\perp} \wedge x^{\perp}\right)^{\perp} && \text{by de Morgan property (Theorem 15.1 page 246)}\\
&= y \wedge \left(y \wedge x^{\perp}\right)^{\perp} && \text{by involutory property (Definition 15.1 page 244)}\\
&= y \wedge \left(y^{\perp} \vee x^{\perp\perp}\right) && \text{by de Morgan property (Theorem 15.1 page 246)}\\
&= y \wedge \left(y^{\perp} \vee x\right) && \text{by involutory property (Definition 15.1 page 244)}\\
&= y \wedge \left(x \vee y^{\perp}\right) && \text{by commutative property (Theorem C.3 page 306)}
\end{aligned}$$

$$\begin{aligned}
x \leq y \implies y^{\perp} &\leq x^{\perp} && \text{by antitone property (Definition 15.1 page 244)}\\
\implies y^{\perp} &= x^{\perp} \wedge \left(y^{\perp} \vee x^{\perp\perp}\right) && \text{by right hypothesis}\\
\implies \left(y^{\perp}\right)^{\perp} &= \left[x^{\perp} \wedge \left(y^{\perp} \vee x^{\perp\perp}\right)\right]^{\perp} &&\\
\implies y &= \left[x^{\perp} \wedge \left(y^{\perp} \vee x^{\perp\perp}\right)\right]^{\perp} && \text{by involutory property (Definition 15.1 page 244)}\\
&= x^{\perp\perp} \vee \left(y^{\perp} \vee x^{\perp\perp}\right)^{\perp} && \text{by de Morgan property (Theorem 15.1 page 246)}\\
&= x \vee \left(y^{\perp} \vee x\right)^{\perp} && \text{by involutory property (Definition 15.1 page 244)}\\
&= x \vee \left(y^{\perp\perp} \wedge x^{\perp}\right) && \text{by de Morgan property (Theorem 15.1 page 246)}\\
&= x \vee \left(y \wedge x^{\perp}\right) && \text{by involutory property (Definition 15.1 page 244)}\\
&= x \vee \left(x^{\perp} \wedge y\right) && \text{by commutative property (Theorem C.3 page 306)}
\end{aligned}$$

2. Proof that (2) $\iff$ (4):

$$\begin{aligned}
(xy) \vee \left[y(xy)^{\perp}\right] &= u \vee \left[yu^{\perp}\right] && \text{where } u \triangleq xy \leq y\\
&= u \vee \left[u^{\perp}y\right] && \text{by commutative property of lattices (Theorem C.3 page 306)}\\
&= y && \text{by left hypothesis}
\end{aligned}$$

$$\begin{aligned}
x \leq y \implies x \vee \left(x^{\perp}y\right) &= xy \vee \left[(xy)^{\perp}y\right] && \text{by } x \leq y \text{ hypothesis}\\
&= xy \vee \left[y(xy)^{\perp}\right] && \text{by commutative property of lattices (Theorem C.3 page 306)}\\
&= y && \text{by right hypothesis}
\end{aligned}$$

3. Proof that (3) $\iff$ (5):

$$\begin{aligned}
(x \vee y)\left[x \vee (x \vee y)^{\perp}\right] &= u\left[x \vee u^{\perp}\right] && \text{where } x \leq u \triangleq x \vee y\\
&= x && \text{by left hypothesis}
\end{aligned}$$

$$\begin{aligned}
x \leq y \implies y\left(x \vee y^{\perp}\right) &= (x \vee y)\left[x \vee (x \vee y)^{\perp}\right] && \text{by } x \leq y \text{ hypothesis}\\
&= x && \text{by right hypothesis}
\end{aligned}$$

4. Proof that (1) $\implies$ (2):

$$
\begin{aligned}
x \le y &\implies x \copyright y && \text{by Proposition E.2 page 337}\\
&\implies y \copyright x && \text{by } \textit{symmetry} \text{ hypothesis (left hypothesis)}\\
&\implies y = (y \wedge x) \vee (y \wedge x^\perp) && \text{by definition of } \copyright \text{ (Definition E.2 page 337)}\\
&\implies y = x \vee (y \wedge x^\perp) && \text{by } x \le y \text{ hypothesis}\\
&\implies y = x \vee (x^\perp \wedge y) && \text{by } \textit{commutative} \text{ property of lattices (Theorem C.3 page 306)}
\end{aligned}
$$

5. Proof that (2) $\implies$ (4):

   (a) lemma: proof that $x \copyright y \implies x^\perp y = (xy)^\perp y$:

$$
\begin{aligned}
x \copyright y \implies x^\perp y &= (xy \vee xy^\perp)^\perp y && \text{by definition of } \copyright \text{ (Definition E.2 page 337)}\\
&= (xy)^\perp (xy^\perp)^\perp y && \text{by } \textit{de Morgan}\text{'s law (Theorem D.4 page 326)}\\
&= (xy)^\perp \left[(x^\perp \vee y^{\perp\perp}) y \right] && \text{by } \textit{de Morgan}\text{'s law (Theorem D.4 page 326)}\\
&= (xy)^\perp \left[(x^\perp \vee y) y \right] && \text{by } \textit{involutory}\text{'s property (Definition 15.1 page 244)}\\
&= (xy)^\perp y && \text{by } \textit{absorptive} \text{ property of lattices (Theorem C.3 page 306)}
\end{aligned}
$$

   (b) Completion of proof for (2) $\implies$ (4):

$$
\begin{aligned}
x \copyright y \implies xy \vee y(xy)^\perp &= xy \vee (xy)^\perp y && \text{by } \textit{commutative} \text{ property (Theorem C.3 page 306)}\\
&= xy \vee x^\perp y && \text{by } x \copyright y \text{ hypothesis and item (5a)}\\
&= (yx) \vee [yx^\perp] && \text{by } \textit{commutative} \text{ property (Theorem C.3 page 306)}\\
&\implies y \copyright x && \text{by definition of } \copyright \text{ (Definition E.2 page 337)}
\end{aligned}
$$

✏

Theorem E.1. [6] *Let* $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \le)$ *be an* ORTHOCOMPLEMENTED LATTICE *(Definition 15.1 page 244).*

THM

$$\{x \copyright c \quad \forall x \in X\} \iff \{\boldsymbol{L} \text{ is ISOMORPHIC to } [0:c] \times [0:c^\perp]\}$$
with isomorphism $\theta(x) \triangleq ([0:c], [0:c^\perp])$.

Proposition E.4. [7] *Let* $(X, \vee, \wedge, 0, 1; \le)$ *be an* ORTHOMODULAR *lattice.*

PRP

$$x \copyright y \iff \phi_x(y) = \phi_y(x) = x \wedge y \qquad \forall x,y \in X$$

E.3 Center

An element in an *orthocomplemented lattice* (Definition 15.1 page 244) is in the *center* of the lattice if that element *commutes* (Definition E.2 page 337) with every other element in the lattice (next definition). *All* the elements of an *orthocomplemented lattice* are in the *center* if and only if that lattice is *Boolean* (Proposition 15.2 page 251).

Definition E.4. [8] *Let* $\copyright$ *be the* COMMUTES *relation (Definition E.2 page 337) on a* LATTICE WITH NEGATION $\boldsymbol{L} \triangleq (X, \vee, \wedge, \neg, 0, 1; \le)$ *(Definition D.5 page 323).*

DEF

The center of $\boldsymbol{L}$ *is defined as*
$$\{x \in X | x \copyright y \quad \forall y \in X\}$$

[6] Kalmbach (1983) page 20, MacLaren (1964)
[7] Foulis (1962) page 66, Sasaki (1954) ⟨cf Foulis 1962⟩
[8] Holland (1970), page 80

Proposition E.5. *Let* $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \le)$ *be an* ORTHOCOMPLEMENTED LATTICE *(Definition 15.1 page 244).*

PRP | 0 *and* 1 *are in the* ***center*** *of* $\boldsymbol{L}$.

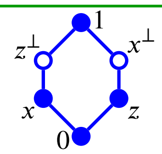
✎PROOF: This follows directly from Definition E.2 (page 337) and Proposition E.2 (page 337). ⇒

Theorem E.2. [9] *Let* $\boldsymbol{L} \triangleq (X, \vee, \wedge, 0, 1; \le)$ *be an* ORTHOCOMPLEMENTED LATTICE *(Definition 15.1 page 244).*

THM | *The* CENTER *of* $\boldsymbol{L}$ *is* BOOLEAN.

Example E.4.

EX | The **center** of the O_6 **lattice** (Definition 15.2 page 244) is the set $\{0, x, z, 1\}$. The elements $x^\perp$ and $z^\perp$ are **not** in the center of $\boldsymbol{L}$. The O_6 lattice is illustrated to the right, with the center elements as solid dots. Note that the center is the *Boolean* lattice $\boldsymbol{L}_2^2$ (Proposition 15.2 page 251).



✎PROOF:

1. Proof that 0 and 1 are in the *center* of $\boldsymbol{L}$: by Proposition E.5 (page 341).
2. Proof that x is in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(x \wedge x) \vee (x \wedge x^\perp) = x \vee 0 &= x &\implies x \copyright x\\
(x \wedge z) \vee (x \wedge z^\perp) = 0 \vee x &= x &\implies x \copyright z
\end{aligned}$$

$x \copyright x$, $x \copyright x^\perp$, $x \copyright z^\perp$, $x \copyright 0$, and $x \copyright 1$ by Proposition E.2 (page 337).

3. Proof that z is in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(z \wedge z) \vee (z \wedge z^\perp) = z \vee 0 &= z &\implies z \copyright z\\
(z \wedge x) \vee (z \wedge x^\perp) = 0 \vee z &= z &\implies z \copyright x
\end{aligned}$$

$z \copyright z$, $z \copyright x^\perp$, $z \copyright z^\perp$, $z \copyright 0$, and $z \copyright 1$ by Proposition E.2 (page 337).

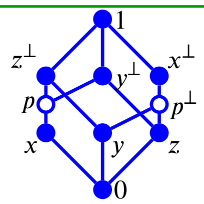
4. Proof that $x^\perp$ and $z^\perp$ are *not* in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(x^\perp \wedge y) \vee (x^\perp \wedge y^\perp) = y \vee 0 &= y &\implies x^\perp \not\copyright y\\
(z^\perp \wedge x) \vee (z^\perp \wedge x^\perp) = x \vee 0 &= x &\implies z^\perp \not\copyright x
\end{aligned}$$

⇒

Example E.5.

EX | The **center** the lattice illustrated to the right (Example 15.2 page 244), with center elements as solid dots, is the set $\{0, 1, p, y, z, x^\perp, y^\perp, z^\perp,\}$. The elements x and $p^\perp$ are *not* in the *center* of $\boldsymbol{L}$. Note that the center is the *Boolean* lattice $\boldsymbol{L}_2^3$ (Proposition 15.2 page 251).



[9] Jeffcott (1972) page 645 ⟨§5. Main theorem⟩

✎PROOF:

1. Proof that 0 and 1 are in the *center* of $\boldsymbol{L}$: by Proposition E.5 (page 341).
2. Proof that x is in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(x \wedge p) \vee (x \wedge p^\perp) &= x \vee 0 && = x && \implies x \copyright p \\
(x \wedge y) \vee (x \wedge y^\perp) &= 0 \vee x && = x && \implies x \copyright y \\
(x \wedge z) \vee (x \wedge z^\perp) &= 0 \vee x && = x && \implies x \copyright z
\end{aligned}$$

$x \copyright x$, $x \copyright x^\perp$, $x \copyright p^\perp$, $x \copyright y^\perp$, $x \copyright z^\perp$, $x \copyright 0$, and $x \copyright 1$ by Proposition E.2 (page 337).

3. Proof that y is in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(y \wedge x) \vee (y \wedge x^\perp) &= 0 \vee y && = y && \implies y \copyright x \\
(y \wedge p) \vee (y \wedge p^\perp) &= 0 \vee y && = y && \implies y \copyright p \\
(y \wedge z) \vee (y \wedge z^\perp) &= 0 \vee y && = y && \implies y \copyright z
\end{aligned}$$

$y \copyright y$, $y \copyright x^\perp$, $y \copyright p^\perp$, $y \copyright y^\perp$, $y \copyright z^\perp$, $y \copyright 0$, and $y \copyright 1$ by Proposition E.2 (page 337).

4. Proof that z is in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(z \wedge x) \vee (z \wedge x^\perp) &= 0 \vee z && = z && \implies z \copyright x \\
(z \wedge p) \vee (z \wedge p^\perp) &= 0 \vee z && = z && \implies z \copyright p \\
(z \wedge y) \vee (z \wedge y^\perp) &= 0 \vee z && = z && \implies z \copyright y
\end{aligned}$$

$z \copyright z$, $z \copyright x^\perp$, $z \copyright p^\perp$, $z \copyright y^\perp$, $z \copyright z^\perp$, $z \copyright 0$, and $z \copyright 1$ by Proposition E.2 (page 337).

5. Proof that $x^\perp$ is in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(p^\perp \wedge x) \vee (p^\perp \wedge x^\perp) &= 0 \vee p^\perp && = p^\perp && \implies p^\perp \copyright x \\
(p^\perp \wedge y) \vee (p^\perp \wedge y^\perp) &= y \vee z && = p^\perp && \implies p^\perp \copyright y \\
(p^\perp \wedge z) \vee (p^\perp \wedge z^\perp) &= z \vee y && = p^\perp && \implies p^\perp \copyright z
\end{aligned}$$

$p^\perp \copyright x^\perp$, $p^\perp \copyright p^\perp$, $p^\perp \copyright y^\perp$, $p^\perp \copyright z^\perp$, $p^\perp \copyright 0$, and $p^\perp \copyright 1$ by Proposition E.2 (page 337).

6. Proof that $y^\perp$ is in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(y^\perp \wedge x) \vee (y^\perp \wedge x^\perp) &= x \vee z && = y^\perp && \implies y^\perp \copyright x \\
(y^\perp \wedge p) \vee (y^\perp \wedge p^\perp) &= p \vee z && = y^\perp && \implies y^\perp \not\copyright p \\
(y^\perp \wedge z) \vee (y^\perp \wedge z^\perp) &= z \vee p && = y^\perp && \implies y^\perp \copyright z
\end{aligned}$$

$p^\perp \copyright x^\perp$, $p^\perp \copyright p^\perp$, $p^\perp \copyright y^\perp$, $p^\perp \copyright z^\perp$, $p^\perp \copyright 0$, and $p^\perp \copyright 1$ by Proposition E.2 (page 337).

7. Proof that $z^\perp$ is in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(z^\perp \wedge x) \vee (z^\perp \wedge x^\perp) &= x \vee y && = z^\perp && \implies z^\perp \copyright x \\
(z^\perp \wedge p) \vee (z^\perp \wedge p^\perp) &= p \vee y && = z^\perp && \implies y^\perp \not\copyright p \\
(z^\perp \wedge y) \vee (z^\perp \wedge y^\perp) &= z \vee p && = z^\perp && \implies y^\perp \copyright z
\end{aligned}$$

$z^\perp \copyright x^\perp$, $z^\perp \copyright p^\perp$, $z^\perp \copyright y^\perp$, $z^\perp \copyright z^\perp$, $z^\perp \copyright 0$, and $z^\perp \copyright 1$ by Proposition E.2 (page 337).

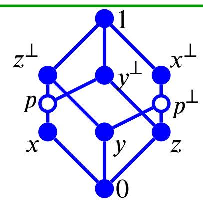
8. Proof that p and $x^\perp$ are *not* in the *center* of $\boldsymbol{L}$:

$$\begin{aligned}
(p \wedge x) \vee (p \wedge x^\perp) = x \vee 0 &\quad = x &&\implies p \not\copyright x\\
(x^\perp \wedge p) \vee (x^\perp \wedge p^\perp) = 0 \vee p^\perp &\quad = p^\perp &&\implies x^\perp \not\copyright p
\end{aligned}$$

Example E.6.

EX

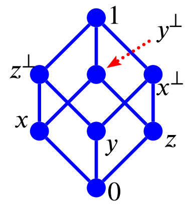
The **center** of the lattice illustrated to the right is illustrated with solid dots. Note that the center is the *Boolean* lattice $\boldsymbol{L}_2^2$ (Proposition 15.2 page 251).



Example E.7.

EX

In a *Boolean* lattice, such as the one illustrated to the right, every element is in the center (Proposition 15.2 page 251).



APPENDIX F

ALGEBRAIC STRUCTURES

❝ *In this book, learned reader, you have the rules of algebra... It unties the knot not only where one term is equal to another or two to one but also where two are equal to two or three to one.... this most abstruse and unsurpassed treasury of the entire arithmetic being brought to light and, as in a theater, exposed to the sight of all...* ❞

Gerolamo Cardano (1501–1576), Italian mathematician, physician, and astrologer [1]

A set together with one or more operations forms several standard mathematical structures:

group $\supseteq$ *ring* $\supseteq$ *commutative ring* $\supseteq$ *integral domain* $\supseteq$ *field*

Definition F.1. [2] *Let X be a set and $\diamond : X \times X \to X$ be an operation on X.*

DEF

*The pair $(X, \diamond)$ is a **group** if*

1.	$\exists e \in X$	such that	$e \diamond x = x \diamond e = x$	$\forall x \in X$	(IDENTITY *element*)	*and*
2.	$\exists (-x) \in X$	such that	$(-x) \diamond x = x \diamond (-x) = e$	$\forall x \in X$	(INVERSE *element*)	*and*
3.			$x \diamond (y \diamond z) = (x \diamond y) \diamond z$	$\forall x,y,z \in X$	(ASSOCIATIVE)	

Definition F.2. [3] *Let $+ : X \times X \to X$ and $* : X \times X \to X$ be operations on a set X. Furthermore, let the operation $*$ also be represented by juxtapostion as in $a * b \equiv ab$.*

DEF

*The triple $(X, +, *)$ is a **ring** if*

1.	$(X, +)$ is a group.			*(additive group)*	*and*
2.	$x(yz) = (xy)z$		$\forall x,y,z \in X$	*(associative with respect to $*$)*	*and*
3.	$x(y+z) = (xy) + (xz)$		$\forall x,y,z \in X$	*($*$ is left distributive over $+$)*	*and*
4.	$(x+y)z = (xz) + (yz)$		$\forall x,y,z \in X$	*($*$ is right distributive over $+$).*	

Definition F.3. [4]

[1] quote: Cardano (1545), page 1
image: http://en.wikipedia.org/wiki/Image:Cardano.jpg

[2] Durbin (2000), page 29

[3] Durbin (2000), pages 114–115

[4] Durbin (2000), page 118

DEF

A triple $(X, +, *)$ *is a* ***commutative ring*** *if*

1.	$(X, +, *)$ is a ring		(ring)	and
2.	$xy = yx$	$\forall x,y \in X$	(commutative).	

Definition F.4. [5] *Let R be a* COMMUTATIVE RING (Definition F.3 page 345).

DEF

A function $|\cdot|$ *in* $\mathbb{R}^{\mathbb{R}}$ *is an* ***absolute value*** (or ***modulus***) *if*

1.	$\|x\| \geq 0$	$x \in \mathbb{R}$	(NON-NEGATIVE)	and
2.	$\|x\| = 0 \iff x = 0$	$x \in \mathbb{R}$	(NONDEGENERATE)	and
3.	$\|xy\| = \|x\| \cdot \|y\|$	$x,y \in \mathbb{R}$	(HOMOGENEOUS / SUBMULTIPLICATIVE)	and
4.	$\|x + y\| \leq \|x\| + \|y\|$	$x,y \in \mathbb{R}$	(SUBADDITIVE / TRIANGLE INEQUALITY)	

Definition F.5. [6]

DEF

The structure $F \triangleq (X, +, \cdot, 0, 1)$ *is a* ***field*** *if*

1.	$(X, +, *)$ is a ring		(ring)	and
2.	$xy = yx$	$\forall x,y \in X$	(commutative with respect to $*$)	and
3.	$(X \setminus \{0\}, *)$ is a group		(group with respect to $*$).	

Definition F.6. [7] *Let* $\boldsymbol{V} = (F, +, \cdot)$ *be a vector space and* $\otimes : \boldsymbol{V} \times \boldsymbol{V} \to \boldsymbol{V}$ *be a vector-vector multiplication operator.*
An ***algebra*** *is any pair* $(\boldsymbol{V}, \otimes)$ *that satisfies* ($\otimes$ *is represented by juxtaposition*)

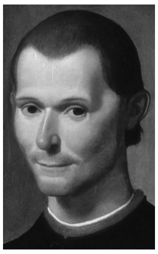
DEF

1.	$(\boldsymbol{ux})\boldsymbol{y} = \boldsymbol{u}(\boldsymbol{xy})$	$\forall \boldsymbol{u},\boldsymbol{x},\boldsymbol{y} \in \boldsymbol{V}$	(ASSOCIATIVE)	and
2.	$\boldsymbol{u}(\boldsymbol{x} + \boldsymbol{y}) = (\boldsymbol{ux}) + (\boldsymbol{uy})$	$\forall \boldsymbol{u},\boldsymbol{x},\boldsymbol{y} \in \boldsymbol{V}$	(LEFT DISTRIBUTIVE)	and
3.	$(\boldsymbol{u} + \boldsymbol{x})\boldsymbol{y} = (\boldsymbol{uy}) + (\boldsymbol{xy})$	$\forall \boldsymbol{u},\boldsymbol{x},\boldsymbol{y} \in \boldsymbol{V}$	(RIGHT DISTRIBUTIVE)	and
4.	$\alpha(\boldsymbol{xy}) = (\alpha\boldsymbol{x})\boldsymbol{y} = \boldsymbol{x}(\alpha\boldsymbol{y})$	$\forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{V}$ and $\alpha \in F$	(SCALAR COMMUTATIVE)	.

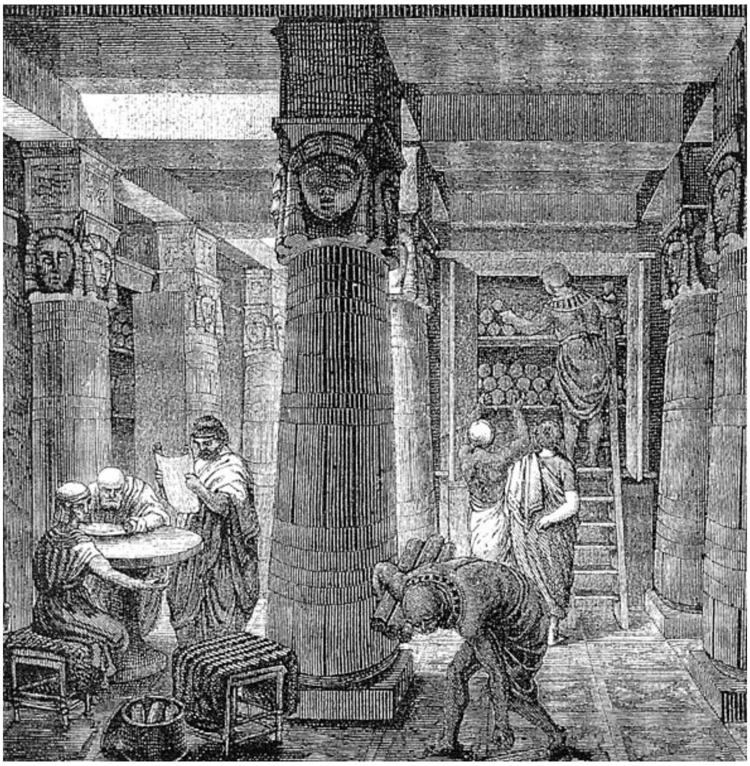
[5] Cohn (2002) page 312
[6] Durbin (2000), page 123, Weber (1893)
[7] Abramovich and Aliprantis (2002), page 3, Michel and Herget (1993), page 56

Back Matter

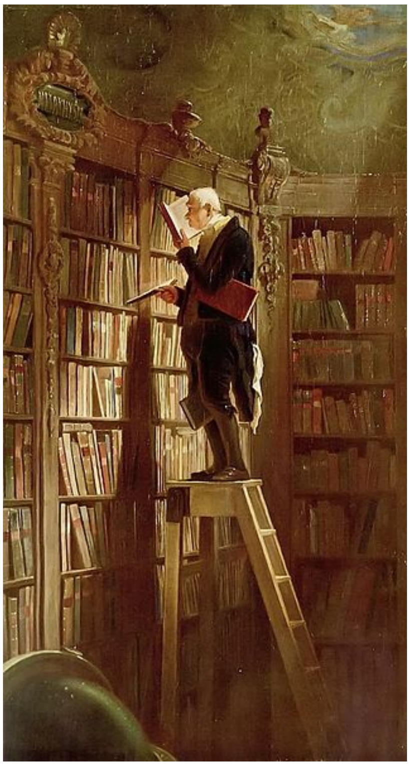
❝It appears to me that if one wants to make progress in mathematics, one should study the masters and not the pupils.❞
Niels Henrik Abel (1802–1829), Norwegian mathematician [8]



❝When evening comes, I return home and go to my study. On the threshold I strip naked, taking off my muddy, sweaty workaday clothes, and put on the robes of court and palace, and in this graver dress I enter the courts of the ancients and am welcomed by them, and there I taste the food that alone is mine, and for which I was born. And there I make bold to speak to them and ask the motives of their actions, and they, in their humanity reply to me. And for the space of four hours I forget the world, remember no vexation, fear poverty no more, tremble no more at death; I pass indeed into their world.❞
Niccolò Machiavelli (1469–1527), Italian political philosopher, in a 1513 letter to friend Francesco Vettori. [9]



ancient library of Alexandria



The Book Worm by *Carl Spitzweg*, circa 1850 [10]

❝To sit alone in the lamplight with a book spread out before you, and hold intimate converse with men of unseen generations—such is a pleasure beyond compare.❞
Yoshida Kenko (Urabe Kaneyoshi) (1283? – 1350?), Japanese author and Buddhist monk [11]

---

[8] quote: Simmons (2007), page 187.
image: http://en.wikipedia.org/wiki/Image:Niels_Henrik_Abel.jpg, public domain

[9] quote: Machiavelli (1961), page 139?.
image: http://commons.wikimedia.org/wiki/File:Santi_di_Tito_-_Niccolo_Machiavelli%27s_portrait_headcrop.jpg, public domain

[10] http://en.wikipedia.org/wiki/File:Ancientlibraryalex.jpg, public domain http://en.wikipedia.org/wiki/File:Carl_Spitzweg_021.jpg,

[11] quote: Kenko (circa 1330)
image: http://en.wikipedia.org/wiki/Yoshida_Kenko

BIBLIOGRAPHY

Jan Łukasiewicz. On three-valued logic. In Storrs McCall, editor, *Polish Logic, 1920–1939*, pages 15–18. Oxford University Press, 1920. ISBN 9780198243045. URL http://books.google.com/books?vid=ISBN0198243049&pg=PA15. collection published in 1967.

Niels Henrik Abel. Untersuchungenüber die reihe $1+\frac{m}{1}x+\frac{m\cdot(m-1)}{2}x^2+\frac{m(m-1)(m-2)}{2\cdot 3}x^3\cdots$ u.s.w. *Journal für die rein und angewandte Mathematik (Crelle's Journal)*, 1.4:311–339, 1826.

Yuri A. Abramovich and Charalambos D. Aliprantis. *An Invitation to Operator Theory*. American Mathematical Society, Providence, Rhode Island, 2002. ISBN 0-8218-2146-6. URL http://books.google.com/books?vid=ISBN0821821466.

Colin Conrad Adams and Robert David Franzosa. *Introduction to Topology: Pure and Applied*. Featured Titles for Topology Series. Pearson Prentice Hall, 2008. ISBN 9780131848696. URL http://books.google.com/books?vid=ISBN0131848690.

N. I. Akhiezer and I. M. Glazman. *Theory of Linear Operators in Hilbert Spaces*, volume 1. Dover, New York, 1993. URL http://books.google.com/books?vid=ISBN0486677486. Translated from the original Russian text *Teoriia lineinykh operatorov v Gil'bertovom prostranstve*.

Donald J. Albers and Gerald L. Alexanderson. *Mathematical People: Profiles and Interviews*. Birkhäuser, Boston, 1985. ISBN 0817631917. URL http://books.google.com/books?vid=ISBN0817631917.

Charalambos D. Aliprantis and Owen Burkinshaw. *Principles of Real Analysis*. Acedemic Press, London, 3 edition, 1998. ISBN 9780120502578. URL http://www.amazon.com/dp/0120502577.

Charalambos D. Aliprantis and Owen Burkinshaw. *Positive Operators*. Springer, Dordrecht, 2006. ISBN 9781402050077. URL http://books.google.com/books?vid=ISBN1402050070. reprint of Academic Press 1985 edition.

Herbert Amann and Joachim Escher. *Analysis II*. Birkhäauser Verlag AG, Basel–Boston–Berlin, 2008. ISBN 978-3-7643-7472-3. URL http://books.google.com/books?vid=ISBN3764374721.

Ichiro Amemiya and Huzihiro Araki. A remark on piron's paper. *Publications of the Research Institute for Mathematical Sciences, Kyoto University*, 2(3):423–427, 1966. ISSN 0034-5318. doi: 10.2977/prims/1195195769. URL http://projecteuclid.org/euclid.prims/1195195769.

Dan Amir. *Characterizations of Inner Product Spaces*, volume 20 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1986. ISBN 3-7643-1774-4. URL http://www.worldcat.org/isbn/3764317744.

Tom M. Apostol. *Mathematical Analysis*. Addison-Wesley series in mathematics. Addison-Wesley, Reading, 2 edition, 1975. ISBN 986-154-103-9. URL http://books.google.com/books?vid=ISBN0201002884.

Aristotle. Metaphysics book iv. In *Aristotle: Metaphysics, Books I–IX*, number 271 in Loeb Classical Library, pages 146–207. Harvard University Press (1933), Cambridge MA. ISBN 0674992997. URL http://www.perseus.tufts.edu/cgi-bin/ptext?lookup=Aristot.+Met.+4.1003a.

Jean-Pierre Aubin. *Applied Functional Analysis*, volume 47 of *Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts*. John Wiley & Sons, 2 edition, September 30 2011. ISBN 9781118030974. URL http://books.google.com/books?vid=ISBN1118030974.

Jean-Pierre Aubin and Hélène Frankowska. *Set-Valued Analysis*. Modern Birkhäuser Classics. Springer, March 2 2009. ISBN 9780817648480. URL http://books.google.com/books?vid=ISBN0817648488.

Léon Autonne. Sur l'hermitien (on the hermitian). In *Comptes Rendus Des SéAnces De L'AcadéMie Des Sciences*, volume 133, pages 209–268. De L'Académie des sciences (Academy of Sciences), Paris, 1901. URL http://visualiseur.bnf.fr/Visualiseur?O=NUMM-3089. Comptes Rendus Des SéAnces De L'AcadéMie Des Sciences (Reports Of the Meetings Of the Academy of Science).

Léon Autonne. Sur l'hermitien (on the hermitian). *Rendiconti del Circolo Matematico di Palermo*, 16:104–128, 1902. Rendiconti del Circolo Matematico di Palermo (Statements of the Mathematical Circle of Palermo).

Arnon Avron. Natural 3-valued logics–characterization and proof theory. *The Journal of Symbolic Logic*, 56(1):276–294, March 1991. URL http://www.jstor.org/stable/2274919.

George Bachman and Lawrence Narici. *Functional Analysis*. Academic Press textbooks in mathematics; Pure and Applied Mathematics Series. Academic Press, 1 edition, 1966. ISBN 9780486402512. URL http://books.google.com/books?vid=ISBN0486402517. "unabridged republication" available from Dover (isbn 0486402517).

George Bachman, Lawrence Narici, and Edward Beckenstein. *Fourier and Wavelet Analysis*. Universitext Series. Springer, 2000. ISBN 9780387988993. URL http://books.google.com/books?vid=ISBN0387988998.

Kirby A. Baker. Equational classes of modular lattices. *Pacific Journal of Mathematics*, 28(1):9–15, 1969. URL http://projecteuclid.org/euclid.pjm/1102983605.

Stefan Banach. Sur les opérations dans les ensembles abstraits et leur applications aux équations intégrales (on abstract operations and their applications to the integral equations). *Fundamenta Mathematicae*, 3:133–181, 1922. URL http://matwbn.icm.edu.pl/ksiazki/fm/fm3/fm3120.pdf.

Stefan Banach. *Théorie des opérations linéaires*. Monografje Matematyczne, Warsaw, Poland, 1932a. URL http://matwbn.icm.edu.pl/kstresc.php?tom=1&wyd=10. (Theory of linear operations).

Stefan Banach. *Theory of Linear Operations*, volume 38 of *North-Holland mathematical library*. North-Holland, Amsterdam, 1932b. ISBN 0444701842. URL http://www.amazon.com/dp/0444701842/. English translation of 1932 French edition, published in 1987.

Robert G. Bartle. *A Modern Theory of Integration*, volume 32 of *Graduate studies in mathematics*. American Mathematical Society, Providence, R.I., 2001. ISBN 0821808451. URL http://books.google.com/books?vid=ISBN0821808451.

Alexander Barvinok. *A Course in Convexity*, volume 54 of *Graduate studies in mathematics*. American Mathematical Society, 2002. ISBN 9780821872314. URL http://books.google.com/books?vid=ISBN0821872311.

Gerald Beer. *Topologies on Closed and Closed Convex Sets*, volume 268 of *Mathematics and Its Applications*. October 31 1993. ISBN 9780792325314. URL http://books.google.com/books?vid=ISBN0792325311.

Rirchard Bellman and Magnus Giertz. On the analytic formalism of the theory of fuzzy sets. *Information Sciences*, 5:149–156, 1973. doi: 10.1016/0020-0255(73)90009-1. URL http://www.sciencedirect.com/science/article/pii/0020025573900091.

Nuel D. Belnap, Jr. A useful four-valued logic. In John Michael Dunn and George Epstein, editors, *Modern Uses of Multiple-valued Logic: Invited Papers from the 5. International Symposium on Multiple-Valued Logic, Held at Indiana University, Bloomington, Indiana, May 13 - 16, 1975 ; with a Bibliography of Many-valued Logic by Robert G. Wolf*, volume 2 of *Episteme*, pages 8–37. D. Reidel, 1977. ISBN 9789401011617. URL http://www.amazon.com/dp/9401011613.

Ladislav Beran. Three identities for ortholattices. *Notre Dame Journal of Formal Logic*, 17(2):251–252, 1976. doi: 10.1305/ndjfl/1093887530. URL http://projecteuclid.org/euclid.ndjfl/1093887530.

Ladislav Beran. *Orthomodular Lattices: Algebraic Approach*. Mathematics and Its Applications (East European Series). D. Reidel Publishing Company, Dordrecht, 1985. ISBN 90-277-1715-X. URL http://books.google.com/books?vid=ISBN902771715X.

Sterling Khazag Berberian. *Introduction to Hilbert Space*. Oxford University Press, New York, 1961. URL http://books.google.com/books?vid=ISBN0821819127.

Yurij M. Berezansky, Zinovij G. Sheftel, and Georgij F. Us. *Functional Analysis: Volume I (Operator Theory, Advances and Applications, Volume 85)*, volume 85 of *Operator Theory Advances and Applications*. Birkhäuser, Basel, 1996. ISBN 3764353449. URL http://books.google.com/books?vid=ISBN3764353449. translated into English from Russian.

Earl Berkson. Some metrics on the subspaces of a banach space. *Pacific Journal of Mathematics*, 13(1):7–22, 1963. URL http://projecteuclid.org/euclid.pjm/1103035953.

Daniel Bernoulli. Expectation of a new theory on the measurement of risk. *Econometrica*, 22(1):23–36, 1783. URL http://msuweb.montclair.edu/~lebelp/BernoulliDRiskEc19541738.pdf. This article was originally published in Latin and has been translated into English. The English version was published January 1954.

D. Besso. Teoremi elementari sui massimi i minimi. *Annuari Ist. Tech. Roma*, pages 7–24, 1879. see Bullen(2003) pages 453, 203.

M. Jales Bienaymé. Société philomatique de paris—extraits des procès-verbaux. *Scéance*, pages 67–68, June 13 1840. URL http://www.archive.org/details/extraitsdesproc46183941soci. see Bullen(2003) pages 453, 203.

Garrett Birkhoff. On the combination of subalgebras. *Mathematical Proceedings of the Cambridge Philosophical Society*, 29:441–464, October 1933a. doi: 10.1017/S0305004100011464. URL http://adsabs.harvard.edu/abs/1933MPCPS..29..441B.

Garrett Birkhoff. On the combination of subalgebras by garrett birkhoff. In Garrett Birkhoff, Gian-Carlo Rota, and Joseph S. Oliveira, editors, *Selected Papers on Algebra and Topology*, Contemporary mathematicians, pages 9–32. Birkhäuser, Boston, 1933b. ISBN 0817631143. URL http://books.google.com/books?vid=ISBN0817631143. This book published in 1987 by Birkhäuser.

Garrett Birkhoff. Orthogonality in linear metric spaces. *Duke Mathematical Journal*, 1(2):169–172, 1935. ISSN 0012-7094. doi: 10.1215/S0012-7094-35-00115-6. URL http://projecteuclid.org/euclid.dmj/1077488974.

Garrett Birkhoff. On the combination of topologies. *Fundamenta Mathematicae*, 26:156–166, 1936a. ISSN 0016-2736. URL http://matwbn.icm.edu.pl/ksiazki/fm/fm26/fm26116.pdf.

Garrett Birkhoff. The logic of quantum mechanics. *Annals of Mathematics*, 37(4):823–843, October 1936b. URL http://www.jstor.org/stable/1968621.

Garrett Birkhoff. Lattices and their applications. *Bulletin of the American Mathematical Society*, 44:1:793–800, 1938. doi: 10.1090/S0002-9904-1938-06866-8. URL http://www.ams.org/bull/1938-44-12/S0002-9904-1938-06866-8/.

Garrett Birkhoff. *Lattice Theory*. American Mathematical Society, New York, 1 edition, 1940. URL http://www.worldcat.org/oclc/1241388.

Garrett Birkhoff. *Lattice Theory*. American Mathematical Society, New York, 2 edition, 1948. URL http://books.google.com/books?vid=ISBN3540120440.

Garrett Birkhoff. *Lattice Theory*, volume 25 of *Colloquium Publications*. American Mathematical Society, Providence, 3 edition, 1967. ISBN 0-8218-1025-1. URL http://books.google.com/books?vid=ISBN0821810251.

Garrett Birkhoff and John Von Neumann. The logic of quantum mechanics. *The Annals of Mathematics*, 37(4):823–843, October 1936. URL http://www.jstor.org/stable/1968621.

Leonard Mascot Blumenthal. *Theory and Applications of Distance Geometry*. Chelsea Publishing Company, Bronx, New York, USA, 2 edition, 1970. ISBN 0-8284-0242-6. URL http://books.google.com/books?vid=ISBN0828402426.

Béla Bollobás. *Linear Analysis; an introductory course*. Cambridge mathematical textbooks. Cambridge University Press, Cambridge, 2 edition, March 1 1999. ISBN 978-0521655774. URL http://books.google.com/books?vid=ISBN0521655773.

Daniel D. Bonar, Michael J. Khoury Jr., and Michael Khoury. *Real Infinite Series*. Mathematical Association of America Textbooks. Mathematical Association of America, 2006. ISBN 9780883857458. URL http://books.google.com/books?vid=ISBN0883857456.

V. Bouniakowsky. Sur quelques inégalités concernant les intégrales ordinaires et les intégrales aux différences finies. *Mémoires de lÁcad. de St.-Pétersbourg*, 1(9), 1859. URL http://www-stat.wharton.upenn.edu/~steele/Publications/Books/CSMC/CSMC_index.html.

Carl Benjamin Boyer and Uta C. Merzbach. *A History of Mathematics*. Wiley, New York, 3 edition, January 11 2011. ISBN 0470525487. URL http://books.google.com/books?vid=ISBN0470525487.

Ronald Newbold Bracewell. *The Fourier transform and its applications*. McGraw-Hill electrical and electronic engineering series. McGraw-Hill, 2, illustrated, international student edition edition, 1978. ISBN 9780070070134. URL http://books.google.com/books?vid=ISBN007007013X.

Gunnar Brinkmann and Brendan D. McKay. Posets on up to 16 points. *Order*, 19(2):147–179, June 2002. ISSN 0167-8094 (print) 1572-9273 (online). doi: 10.1023/A:1016543307592. URL http://www.springerlink.com/content/d4dbce7pmctuenmg/.

Thomas John I'Anson Bromwich. *An Introduction to the Theory of Infinite Series*. Macmillan and Company, 1 edition, 1908. ISBN 9780821839768. URL http://www.archive.org/details/anintroductiont00bromgoog.

Arlen Brown and Carl M. Pearcy. *An Introduction to Analysis*, volume 154 of *Graduate Texts in Mathematics*. Springer, 1995. ISBN 0387943692. URL http://books.google.com/books?vid=ISBN0387943692.

Jason I. Brown and Stephen Watson. Self complementary topologies and preorders. *Order*, 7(4): 317–328, 1991. ISSN 0167-8094 (print) 1572-9273 (online). doi: 10.1007/BF00383196. URL http://www.springerlink.com/content/t164x9114754w41q/.

Jason I. Brown and Stephen Watson. The number of complements of a topology on n points is at least 2^n (except for some special cases). *Discrete Mathematics*, 154(1–3):27–39, 15 June 1996. doi: 10.1016/0012-365X(95)00004-G. URL http://dx.doi.org/10.1016/0012-365X(95)00004-G.

Andrew M. Bruckner, Judith B. Bruckner, and Brian S. Thomson. *Real Analysis*. Prentice-Hall, Upper Saddle River, N.J., 1997. ISBN 9780134588865. URL http://books.google.com/books?vid=ISBN013458886X.

Victor Bryant. *Metric Spaces: Iteration and Application*. Cambridge University Press, Cambridge, illustrated, reprint edition, 1985. ISBN 9780521318976. URL http://books.google.com/books?vid=ISBN0521318971.

P. S. Bullen. *Handbook of Means and Their Inequalities*, volume 560 of *Mathematics and Its Applications*. Kluwer Academic Publishers, Dordrecht, Boston, 2 edition, 2003. ISBN 9781402015229. URL http://books.google.com/books?vid=ISBN1402015224.

Viktor Yakovlevich Bunyakovsky. Sur quelques inégalités concernant les intégrales ordinaires et les intégrales aux différences finies. *Mémoires de L'Académie Impériale des Sciences de St.-Pétersbourg*, 1(9):1–18, 1859. URL http://www-stat.wharton.upenn.edu/~steele/Publications/Books/CSMC/bunyakovsky.pdf. (Mémoires of the Imperial Academy of Sciences of Saint Petersburg); (On Some Inequalities concerning the Ordinary Integrals and Integrals with the Finite Differences).

Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A Course in Metric Geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, 2001. ISBN 978-0821821299. URL http://books.google.com/books?vid=ISBN0821821296.

Stanley Burris and Hanamantagida Pandappa Sankappanavar. *A Course in Universal Algebra*. Number 78 in Graduate texts in mathematics. Springer-Verlag, New York, 1 edition, 1981. ISBN 0-387-90578-2. URL http://books.google.com/books?vid=ISBN0387905782. 2000 edition available for free online.

Stanley Burris and Hanamantagida Pandappa Sankappanavar. A course in universal algebra. Retypeset and corrected version of the 1981 edition, 2000. URL http://www.math.uwaterloo.ca/~snburris/htdocs/ualg.html.

Herbert Busemann. *The Geometry of Geodesics*. Academic Press, 2 edition, 1955a. ISBN 0486154629. URL http://books.google.com/books?vid=ISBN0486154629. a Dover 2005 edition has been published which “is an unabridged republication of the work originally published in 1955”.

Herbert Busemann. *The geometry of geodesics*. Dover (2005), 2 edition, 1955b. ISBN 0486442373. URL http://books.google.com/books?vid=ISBN0486442373. Dover reprint of 1955 edition.

Charles L. Byrne. *Signal Processing: A Mathematical Approach*. AK Peters Series. A K Peters, 2005. ISBN 1568812426.

Florian Cajori. A history of mathematical notations; notations mainly in higher mathematics. In *A History of Mathematical Notations; Two Volumes Bound as One*, volume 2. Dover, Mineola, New York, USA, 1993. ISBN 0-486-67766-4. URL http://books.google.com/books?vid=ISBN0486677664. reprint of 1929 edition by *The Open Court Publishing Company*.

Gerolamo Cardano. *Ars Magna or the Rules of Algebra*. Dover Publications, Mineola, New York, 1545. ISBN 0486458733. URL http://www.amazon.com/dp/0486458733. English translation of the Latin *Ars Magna* edition, published in 2007.

N. L. Carothers. *A Short Course on Banach Space Theory*. Number 64 in London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 2005. ISBN 0521842832. URL http://www.comm.musashi-tech.ac.jp/~arimoto/Drarimoto/preprints/calothers/Assemble.ps.

N.L. Carothers. *Real Analysis*. Cambridge University Press, Cambridge, 2000. ISBN 978-0521497565. URL http://books.google.com/books?vid=ISBN0521497566.

J.C. Carrega. Exclusion d'algèbres. *Comptes Rendus des Seances de l'Academie des Sciences*, 295: 43–46, 1982. Sirie I: Mathematique.

Gianpiero Cattaneo and Davide Ciucci. Lattices with interior and closure operators and abstract approximation spaces. In James F. Peters and Andrzej Skowron, editors, *Transactions on Rough Sets X*, volume 5656 of *Lecture notes in computer science*, pages 67–116. Springer, 2009. ISBN 9783642032813.

Augustin-Louis Cauchy. *Part 1: Analyse Algebrique*. Cours D'Analyse de L'école Royale Polytechnique. Debure frères, Paris, 1821. ISBN 2-87647-053-5. URL http://www.archive.org/details/coursdanalysede00caucgoog. Systems design course of the Polytechnic Royal School; 1st Part: Algebraic analysis).

Ernest Cesàro. Sur la multiplication des séries. *Bulletin des Sciences Mathématiques*, 14(2):114–120, 1890.

S. D. Chatterji. The number of topologies on n points. Technical Report N67-31144, National Aeronautics and Space Administration, July 1967. URL http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19670021815_1967021815.pdf. techreport.

Gustave Choquet. Theory of capacities. *Annales de l'institut Fourier*, 5:131–295, 1954. doi: 10.5802/aif.53. URL http://aif.cedram.org/item?id=AIF_1954__5__131_0.

Ole Christensen. *An Introduction to Frames and Riesz Bases*. Applied and Numerical Harmonic Analysis. Birkhäuser, Boston/Basel/Berlin, 2003. ISBN 0-8176-4295-1. URL http://books.google.com/books?vid=ISBN0817642951.

Ole Christensen. *Frames and bases: An Introductory Course*. Applied and Numerical Harmonic Analysis. Birkhäuser, Boston/Basel/Berlin, 2008. ISBN 9780817646776. URL http://books.google.com/books?vid=ISBN0817646779.

Roberto Cignoli. Injective de morgan and kleene algebras. *Proceedings of the American Mathematical Society*, 47(2):269–278, February 1975. URL http://www.ams.org/journals/proc/1975-047-02/S0002-9939-1975-0357259-4/S0002-9939-1975-0357259-4.pdf.

Erhan Çinlar and Robert J Vanderbei. *Real and Convex Analysis*. Undergraduate Texts in Mathematics. Springer, January 4 2013. ISBN 1461452570. URL http://books.google.com/books?vid=ISBN1461452570.

James A. Clarkson. Uniformly convex spaces. *Transactions of the American Mathematical Society*, 40(3):396–414, December 1936. URL http://www.jstor.org/stable/1989630.

David W. Cohen. *An Introduction to Hilbert Space and Quantum Logic*. Problem Books in Mathematics. Springer-Verlag, New York, 1989. ISBN 0-387-96870-9. URL http://books.google.com/books?vid=ISBN1461388430.

Paul M. Cohn. *Basic Algebra; Groups, Rings and Fields*. Springer, December 6 2002. ISBN 1852335874. URL http://books.google.com/books?vid=isbn1852335874.

Louis Comtet. Recouvrements, bases de filtre et topologies d'un ensemble fini. *Comptes rendus de l'Acade'mie des sciences*, 262(20):A1091–A1094, 1966. Recoveries, bases and filter topologies of a finite set.

Louis Comtet. *Advanced combinatorics: the art of finite and infinite*. D. Reidel Publishing Company, Dordrecht, 1974. ISBN 978-9027704412. URL http://books.google.com/books?vid=ISBN9027704414. translated and corrected version of the 1970 French edition.

John B. Conway. *A Course in Functional Analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer, 2 edition, 1990. ISBN 0-387-97245-5. URL http://books.google.com/books?vid=ISBN0387972455.

Edward Thomas Copson. *Metric Spaces*. Number 57 in Cambridge tracts in mathematics and mathematical physics. Cambridge University Press, London, 1968. ISBN 978-0521047227. URL http://books.google.com/books?vid=ISBN0521047226.

Paul Corazza. Introduction to metric-preserving functions. *The American Mathematical Monthly*, 104(4):309–323, April 1999. URL http://pcorazza.lisco.com/papers/metric-preserving.pdf.

T.M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., New York, 1991. ISBN 0-471-06259-6. URL http://www.amazon.com/dp/0471062596.

Peter Crawley and Robert Palmer Dilworth. *Algebraic Theory of Lattices*. Prentice-Hall, January 1973. ISBN 0130222690. URL http://books.google.com/books?vid=ISBN0130222690.

Ingrid Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, 1992. ISBN 0-89871-274-2. URL http://www.amazon.com/dp/0898712742.

Ingrid Daubechies, A. Grossman, and Y. Meyer. Painless nonorthogonal expansions. *Journal of Mathematical Physics*, 27(5):1271–1283, May 1986. ISSN 0022-2488. URL link.aip.org/link/?jmp/27/1271.

Brian A. Davey and Hilary A. Priestley. *Introduction to Lattices and Order*. Cambridge mathematical text books. Cambridge University Press, Cambridge, 2 edition, May 6 2002. ISBN 978-0521784511. URL http://books.google.com/books?vid=ISBN0521784514.

Sheldon W. Davis. *Topology*. McGraw Hill, Boston, 2005. ISBN 007-124339-9. URL http://www.worldcat.org/isbn/0071243399.

Mahlon Marsh Day. *Normed Linear Spaces*, volume 21 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, Berlin, Heidelberg, New York, 3 edition, 1973. ISBN 0387061487. URL http://books.google.com/books?id=UZ3vSAAACAAJ.

Pierre Rémond de Montmort, editor. *Essay D'analyse sur Les Jeux de Hazard*, volume 307. American Mathematical Society, Providence, Rhode Island, 2 edition, 1713. ISBN 978-0-8218-3781-8. URL http://books.google.com/books?vid=ISBN9780821837818. "Essays on the analysis of games of chance", 2006 AMS reprint edition, An English translation concerning the *St. Petersburg Paradox* correspondence by Richard J. Pulskamp called "Correspondence of Nicolas Bernoulli concerning the St. Petersburg Game" is available at http://www.cs.xu.edu/math/Sources/Montmort/stpetersburg.pdf.

Andreas de Vries. Algebraic hierarchy of logics unifying fuzzy logic and quantum logic. The registered submission date for this paper is 2007 July 14, but the date appearing on paper proper is 2009 December 6. The latest year in the references is 2006, July 14 2007. URL http://arxiv.org/abs/0707.2161.

Lokenath Debnath and Piotr Mikusiński. *Introduction to Hilbert Spaces with Applications*. Academic Press, 2005. ISBN 9780080455921. URL http://books.google.com/books?vid=ISBN0080455921.

Richard Dedekind. Ueber die von drei moduln erzeugte dualgruppe. *Mathematische Annalen*, 53:371–403, January 8 1900. URL http://resolver.sub.uni-goettingen.de/purl/?GDZPPN002257947. Regarding the Dual Group Generated by Three Modules.

René Descartes. *Regulae ad directionem ingenii*. 1684a. URL http://www.fh-augsburg.de/~harsch/Chronologia/Lspost17/Descartes/des_re00.html.

René Descartes. *Rules for Direction of the Mind*. 1684b. URL http://en.wikisource.org/wiki/Rules_for_the_Direction_of_the_Mind.

D. Devidi. Negation: Philosophical aspects. In Keith Brown, editor, *Encyclopedia of Language & Linguistics*, pages 567–570. Elsevier, 2 edition, April 6 2006. ISBN 9780080442990. URL http://www.sciencedirect.com/science/article/pii/B0080448542012025.

D. Devidi. Negation: Philosophical aspects. In Alex Barber and Robert J Stainton, editors, *Concise Encyclopedia of Philosophy of Language and Linguistics*, pages 510–513. Elsevier, April 6 2010. ISBN 9780080965017. URL http://books.google.com/books?vid=ISBN0080965016&pg=PA510.

Elena Deza and Michel-Marie Deza. *Dictionary of Distances*. Elsevier Science, Amsterdam, 2006. ISBN 0444520872. URL http://books.google.com/books?vid=ISBN0444520872.

Michel-Marie Deza and Elena Deza. *Encyclopedia of Distances*. Springer, 2009. ISBN 3642002331. URL http://www.uco.es/users/ma1fegan/Comunes/asignaturas/vision/Encyclopedia-of-distances-2009.pdf.

Emmanuele DiBenedetto. *Real Analysis*. Birkhäuser Advanced Texts. Birkhäuser, Boston, 2002. ISBN 0817642315. URL http://books.google.com/books?vid=ISBN0817642315.

Jean Alexandre Dieudonné. *Foundations of Modern Analysis*. Academic Press, New York, 1969. ISBN 1406727911. URL http://books.google.com/books?vid=ISBN1406727911.

R.P. Dilworth. On complemented lattices. *Tôhoku Mathematical Journal*, 47:18–23, 1940. ISSN 0040-8735. URL http://projecteuclid.org/tmj.

R.P. Dilworth. A decomposition theorem for partially ordered sets. *Annals of Mathematics*, 51(1): 161–166, January 1950a. doi: 10.2307/1969503. URL http://www.jstor.org/stable/1969503.

R.P. Dilworth. A decomposition theorem for partially ordered sets. In Kenneth P. Bogart, Ralph S. Freese, and Joseph P.S. Kung, editors, *The Dilworth theorems: selected papers of Robert P. Dilworth*, page ? Birkhäuser (1990), Boston, 1950b. ISBN 0817634347. URL http://books.google.com/books?vid=ISBN0817634347.

R.P. Dilworth. On complemented lattices. In Kenneth P. Bogart, Ralph S. Freese, and Joseph P.S. Kung, editors, *The Dilworth theorems: selected papers of Robert P. Dilworth*, pages 73–78? Birkhäuser, Boston, 1990. ISBN 0817634347. URL http://books.google.com/books?vid=ISBN0817634347.

Jozef Doboš. *Metric Preserving Functions*. Štroffek, 1998. ISBN 9788088896302. URL http://prof.jozef.xn--dobo-j6a.eu/files/2012/03/mpf1.pdf.

John Doner and Alfred Tarski. An extended arithmetic of ordinal numbers. *Fundamenta Mathematicae*, 65:95–127, 1969. URL http://matwbn.icm.edu.pl/tresc.php?wyd=1&tom=65.

Hamid Drljević. On the representation of functionals and the stability of mappings in hilbert and banach spaces. In Themistocles M Rassias, editor, *Topics in mathematical analysis: a volume dedicated to the memory of A.L. Cauchy*, volume II of *Series in Pure Mathematics*, pages 231–245. World Scientific Publishing Company, 1989. ISBN 9971506661. URL http://books.google.com/books?vid=ISBN9971506661.

R. J. Duffin and A. C. Schaeffer. A class of nonharmonic fourier series. *Transactions of the American Mathematical Society*, 72(2):341–366, March 1952. ISSN 1088-6850. URL http://www.jstor.org/stable/1990760.

Nelson Dunford and Jacob T. Schwartz. *Linear operators. Part 1, General Theory*, volume 7 of *Pure and applied mathematics*. Interscience Publishers, New York, 1957. ISBN 0471226394. URL http://www.amazon.com/dp/0471608483. with the assistance of William G. Bade and Robert G. Bartle.

J. Michael Dunn. Intuitive semantics for first-degree entailments and `coupled trees'. *Philosophical Studies*, 29(3):149–168, 1976. URL http://link.springer.com/article/10.1007/BF00373152.

J. Michael Dunn. Generalized ortho negation. In Heinrich Wansing, editor, *Negation: A Notion in Focus*, volume 7 of *Perspektiven der Analytischen Philosophie / Perspectives in Analytical Philosophy*, pages 3–26. De Gruyter, January 1 1996. ISBN 9783110876802. URL http://books.google.com/books?vid=ISBN3110876809.

J. Michael Dunn. A comparative study of various model-theoretic treatments of negation: A history of formal negation. In Dov M. Gabbay and Heinrich Wansing, editors, *What is Negation?*, volume 13 of *Applied Logic Series*, pages 23–52. De Gruyter, 1999. ISBN 9780792355694. URL http://books.google.com/books?vid=ISBN0792355695.

John R. Durbin. *Modern Algebra; An Introduction*. John Wiley & Sons, Inc., 4 edition, 2000. ISBN 0-471-32147-8. URL http://www.worldcat.org/isbn/0471321478.

W.D. Duthie. Segments of ordered sets. *Transactions of the American Mathematical Society*, 51(1): 1–14, January 1942. doi: 10.2307/1989978. URL http://www.jstor.org/stable/1989978.

Paul H. Edelman. Abstract convexity and meet-distributive lattices. In Ivan Rival, editor, *Combinatorics and ordered sets: Proceedings of the AMS-IMS-SIAM joint summer research conference, held August 11–17, 1985*, volume 57 of *Contemporary Mathematics*, pages 127–150, Providence RI, 1986. American Mathematical Society. ISBN 0821850512. URL http://books.google.com/books?vid=ISBN0821850512. conference held in Arcata California.

Paul H. Edelman and Robert E. Jamison. The theory of convex geometries. *Geometriae Dedicata*, 19(3):247–270, December 1985. ISSN 0046-5755. doi: 10.1007/BF00149365. URL http://www.springerlink.com/content/n4344856887387gw/.

Robert E. Edwards. *Functional Analysis: Theory and Applications*. Dover books on mathematics. Dover, New York, 1995. ISBN 0-486-68143-2. URL http://books.google.com/books?vid=ISBN0486681432.

Yuli Eidelman, Vitali D. Milman, and Antonis Tsolomitis. *Functional Analysis: An Introduction*, volume 66 of *Graduate Studies in Mathematics*. American Mathematical Society, 2004. ISBN 0821836463. URL http://books.google.com/books?vid=ISBN0821836463.

Per Enflo. A counterexample to the approximation problem in banach spaces. *Acta Mathematica*, 130:309–317, 1973. URL http://link.springer.com/content/pdf/10.1007/BF02392270.

Euclid. *Elements*. circa 300BC. URL http://farside.ph.utexas.edu/euclid.html.

J.W. Evans, Frank Harary, and M.S. Lynn. On the computer enumeration of finite topologies. *Communications of the ACM — Association for Computing Machinery*, 10:295–297, 1967. ISSN 0001-0782. URL http://portal.acm.org/citation.cfm?id=363282.363311.

David Ewen. *The Book of Modern Composers*. Alfred A. Knopf, New York, 1950. URL http://books.google.com/books?id=yHw4AAAAIAAJ.

David Ewen. *The New Book of Modern Composers*. Alfred A. Knopf, New York, 3 edition, 1961. URL http://books.google.com/books?id=bZIaAAAAMAAJ.

Marián Fabian, Petr Habala, Petr Hájek, Vicente Montesinos, Vaclav Zizler, and Václav Zizler. *Banach Space Theory: The Basis for Linear and Nonlinear Analysis*. CMS Books in Mathematics. Springer, 2010. ISBN 1441975144. URL http://books.google.com/books?vid=isbn1441975144.

Jonathan David Farley. Chain decomposition theorems for ordered sets and other musings. *African Americans in Mathematics DIMACS Workshop*, 34:3–14, June 26–28 1996. URL http://books.google.com/books?vid=ISBN0821806785.

Jonathan David Farley. Chain decomposition theorems for ordered sets and other musings. *arXiv.org preprint*, pages 1–12, July 16 1997. URL http://arxiv.org/abs/math/9707220.

Gy. Fáy. Transitivity of implication in orthomodular lattices. *Acta Scientiarum Mathematicarum*, 28(3–4):267–270, 1967. ISSN 0001-6969. URL http://www.acta.hu/acta/.

János Fodor and Ronald R. Yager. Fuzzy set-theoretic operators and quantifiers. In Didier Dubois and Henri Padre, editors, *Fundamentals of Fuzzy Sets*, volume 7 of *The Handbooks of Fuzzy Sets*, pages 125–195. Springer Science & Business Media, 2000. ISBN 9780792377320. URL http://books.google.com/books?vid=ISBN079237732X.

Gerald B. Folland. *A Course in Abstract Harmonic Analysis*. Studies in Advanced Mathematics. CRC Press, Boca Raton, 1995. ISBN 0-8493-8490-7. URL http://books.google.com/books?vid=ISBN0849384907.

David J. Foulis. A note on orthomodular lattices. *Portugaliae Mathematica*, 21(1):65–72, 1962. ISSN 0032-5155. URL http://purl.pt/2387.

Jean-Baptiste-Joseph Fourier. Refroidissement séculaire du globe terrestre". In M. Gaston Darboux, editor, *Œuvres De Fourier*, volume 2, pages 271–288. Ministère de L'instruction Publique, Paris, France, April 1820. URL http://gallica.bnf.fr/ark:/12148/bpt6k33707/f276.image. original paper at pages 58–70.

Maurice René Fréchet. Sur quelques points du calcul fonctionnel (on some points of functional calculation). *Rendiconti del Circolo Matematico di Palermo*, 22:1–74, 1906. Rendiconti del Circolo Matematico di Palermo (Statements of the Mathematical Circle of Palermo).

Maurice René Fréchet. *Les Espaces abstraits et leur théorie considérée comme introduction a l'analyse générale*. Borel series. Gauthier-Villars, Paris, 1928. URL http://books.google.com/books?id=9czoHQAACAAJ. Abstract spaces and their theory regarded as an introduction to general analysis.

Ferdinand Georg Frobenius. Uber lineare substitutionen und bilineare formen. *Journal für die reine und angewandte Mathematik (Crelle's Journal)*, 84:1–63, 1878. ISSN 0075-4102. URL http://www.digizeitschriften.de/home/services/pdfterms/?ID=509796.

Ferdinand Georg Frobenius. Uber lineare substitutionen und bilineare formen. In Jean Pierre Serre, editor, *Gesammelte Abhandlungen (Collected Papers)*, volume I, pages 343–405. Springer, Berlin, 1968. URL http://www.worldcat.org/oclc/253015. reprint of Frobenius' 1878 paper.

Otto Frölich. Das halbordnungssystem der topologischen räume auf einer menge. *Mathematische Annalen*, 156:79–95, 1964. URL http://resolver.sub.uni-goettingen.de/purl?GDZPPN002293501.

Jürgen Fuchs. *Affine Lie Algebras and Quantum Groups: An Introduction, With Applications in Conformal Field Theory*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1995. ISBN 052148412X. URL http://books.google.com/books?vid=ISBN052148412X.

Haim Gaifman. The lattice of all topologies on a denumerable set. *Notices of the American Mathematical Society*, 8(356), 1961. ISSN 0002-9920 (print) 1088-9477 (electronic).

Israel M. Gelfand. Normierte ringe. *Mat. Sbornik*, 9(51):3–24, 1941.

Israel M. Gelfand and Mark A. Naimark. *Normed Rings with an Involution and their Representations*, pages 240–274. Chelsea Publishing Company, Bronx, 1964. ISBN 0821820222. URL http://books.google.com/books?vid=ISBN0821820222.

Israel M. Gelfand and Mark A. Neumark. On the imbedding of normed rings into the ring of operators in hilbert space. *Mat. Sbornik*, 12(54:2):197–217, 1943a.

Israel M. Gelfand and Mark A. Neumark. *On the imbedding of normed rings into the ring of operators in Hilbert Space*, pages 3–19. 1943b. ISBN 0821851756. URL http://books.google.com/books?vid=ISBN0821851756.

Michael C. Gemignani. *Elementary Topology*. Addison-Wesley Series in Mathematics. Addison-Wesley Publishing Company, Reading, Massachusetts, 2 edition, 1972. ISBN 9780486665221. URL http://books.google.com/books?vid=ISBN0486665224. A 1990 Dover "unabridged and corrected" edition has been published.

John Robilliard Giles. *Introduction to the Analysis of Metric Spaces*. Number 3 in Australian Mathematical Society lecture series. Cambridge University Press, Cambridge, 1987. ISBN 978-0521359283. URL http://books.google.com/books?vid=ISBN0521359287.

John Robilliard Giles. *Introduction to the Analysis of Normed Linear Spaces*. Number 13 in Australian Mathematical Society lecture series. Cambridge University Press, Cambridge, 2000. ISBN 0-521-65375-4. URL http://books.google.com/books?vid=ISBN0521653754.

Steven Givant and Paul Halmos. *Introduction to Boolean Algebras*. Undergraduate Texts in Mathematics. Springer, 2009. ISBN 0387402934. URL http://books.google.com/books?vid=ISBN0387402934.

Siegfried Gottwald. Many-valued logic and fuzzy set theory. In Ulrich Höhle and S.E. Rodabaugh, editors, *Mathematics of Fuzzy Sets: Logic, Topology, and Measure Theory*, volume 3 of *The Handbooks of Fuzzy Sets*, pages 5–90. Kluwer Academic Publishers, 1999. ISBN 9780792383888. URL http://books.google.com/books?vid=ISBN0792383885.

Ronald L. Graham, Donald Ervin Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, 2 edition, 1994. ISBN 0201558025. URL http://books.google.com/books?vid=ISBN0201558025.

Ivor Grattan-Guinness. *Convolutions in French mathematics, 1800-1840: from the calculus and mechanics to mathematical analysis and mathematical physics. Volume I: The Settings*, volume 2 of *Science networks. Historical Studies*. Birkhäuser Verlag, Basel, 1 edition, 1990. ISBN 9783764322373. URL http://books.google.com/books?vid=ISBN3764322373.

George A. Grätzer. *General Lattice Theory*. Birkhäuser Verlag, Basel, 2 edition, 1998. ISBN 0-8176-5239-6. URL http://books.google.com/books?vid=ISBN0817652396.

George A. Grätzer. *General Lattice Theory*. Birkhäuser Verlag, Basel, 2 edition, January 17 2003. ISBN 3-7643-6996-5. URL http://books.google.com/books?vid=ISBN3764369965.

George A. Grätzer. Two problems that shaped a century of lattice theory. *Notices of the American Mathematical Society*, 54(6):696–707, June/July 2007. URL http://www.ams.org/notices/200706/.

Larry C. Grove. *Classical Groups and Geometric Algebra*, volume 39 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002. ISBN 9780821820193. URL http://books.google.com/books?vid=ISBN0821820192.

Stanley Gudder. *Stochastic Methods in Quantum Mechanics*. North Holland, 1979. ISBN 0444002995. URL http://books.google.com/books?vid=ISBN0444002995.

Stanley Gudder. *Quantum Probability*. Probability and Mathematical Statistics. Academic Press, August 28 1988. ISBN 0123053404. URL http://books.google.com/books?vid=ISBN0123053404.

Stanley Gudder. *Stochastic Methods in Quantum Mechanics*. Dover, Mineola NY, 2005. ISBN 0486445321. URL http://books.google.com/books?vid=ISBN0486445321.

Amritava Gupta. *Real & Abstract Analysis*. Academic Publishers Calcutta, 1998. ISBN 9788186358443. URL http://books.google.com/books?vid=ISBN8186358447. first published in 1998, reprinted in 2000 and 2006.

Norman B. Haaser and Joseph A. Sullivan. *Real Analysis*. Dover Publications, New York, 1991. ISBN 0-486-66509-7. URL http://books.google.com/books?vid=ISBN0486665097.

Hans Hahn and Arthur Rosenthal. *Set Functions.* University of New Mexico Press, 1948. ISBN 111422295X. URL http://books.google.com/books?vid=ISBN111422295X.

Henry Sinclair Hall and Samuel Ratcliffe Knight. *Higher algebra, a sequel to elementary algebra for schools.* Macmillan, London, 1894. URL http://www.archive.org/details/higheralgebraas00kniggoog.

Paul R. Halmos. *Finite Dimensional Vector Spaces.* Princeton University Press, Princeton, 1 edition, 1948. ISBN 0691090955. URL http://books.google.com/books?vid=isbn0691090955.

Paul R. Halmos. *Measure Theory.* The University series in higher mathematics. D. Van Nostrand Company, New York, 1950. URL http://www.amazon.com/dp/0387900888. 1976 reprint edition available from Springer with ISBN 9780387900889.

Paul R. Halmos. *Finite Dimensional Vector Spaces.* Springer-Verlag, New York, 2 edition, 1958. ISBN 0-387-90093-4. URL http://books.google.com/books?vid=isbn0387900934.

Paul R. Halmos. *Intoduction to Hilbert Space and the Theory of Spectral Multiplicity.* Chelsea Publishing Company, New York, 2 edition, 1998a. ISBN 0821813781. URL http://books.google.com/books?vid=ISBN0821813781.

Paul R. Halmos. *Introduction to Hilbert Space and the Theory of Spectral Multiplicity.* AMS Chelsea, Providence RI, 2 edition, 1998b. ISBN 0821813781. URL http://books.google.com/books?vid=ISBN0821813781.

Paul Richard Halmos. *Naive Set Theory.* The University Series in Undergraduate Mathematics. D. Van Nostrand Company, Inc., Princeton, New Jersey, 1960. ISBN 0387900926. URL http://books.google.com/books?vid=isbn0387900926.

Georg Hamel. Eine basis aller zahlen und die unstetigen lösungen der funktionalgleichung $f(x+y) = f(x) + f(y)$. *Mathematische Annalen*, 60(3):459–462, 1905. URL http://gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002260395&IDDOC=28580.

Deguang Han, Keri Kornelson, David Larson, and Eric Weber. *Frames for Undergraduates*, volume 40 of *Student Mathematical Library*. American Mathematical Society, 2007. ISBN 0821842129. URL http://books.google.com/books?vid=ISBN0821842129. Deguang Han = ???

G.H. Hardy. Prolegomena to a chapter on inequalities. *Journal of the London Mathematical Society*, 1–4:61–78, November 8 1929. URL http://jlms.oxfordjournals.org/content/vols1-4/issue13/index.dtl#PRESIDENTIAL_ADDRESS. "Presidential Address" to the London Mathematical Society.

Godfrey H. Hardy. *A Mathematician's Apology.* Cambridge University Press, Cambridge, 1940. URL http://www.math.ualberta.ca/~mss/misc/A%20Mathematician's%20Apology.pdf.

Godfrey Harold Hardy. *Divergent Series.* Oxford University Press, 1949. URL http://archive.org/details/divergentseries033523mbp.

Godfrey Harold Hardy, John Edensor Littlewood, and George Pólya. *Inequalities.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2 edition, 1952. URL http://books.google.com/books?vid=ISBN0521358809.

Juris Hartmanis. On the lattice of topologies. *Canadian Journal of Mathematics*, 10(4):547–553, 1958. URL http://books.google.com/books?&id=0PDcFxeiBesC.

Felix Hausdorff. *Grundzüge der Mengenlehre*. Von Veit, Leipzig, 1914. URL http://books.google.com/books?id=KTs4AAAAMAAJ. Properties of Set Theory.

Felix Hausdorff. *Set Theory*. Chelsea Publishing Company, New York, 3 edition, 1937. ISBN 0828401195. URL http://books.google.com/books?vid=ISBN0828401195. 1957 translation of the 1937 German *Grundzüge der Mengenlehre*.

Michiel Hazewinkel, editor. *Handbook of Algebras*, volume 2. North-Holland, Amsterdam, 1 edition, 2000. ISBN 044450396X. URL http://books.google.com/books?vid=ISBN044450396X.

Jean Van Heijenoort. *From Frege to Gödel : A Source Book in Mathematical Logic, 1879-1931*. Harvard University Press, Cambridge, Massachusetts, 1967. URL http://www.hup.harvard.edu/catalog/VANFGX.html.

Christopher Heil. *A Basis Theory Primer*. Applied and Numerical Harmonic Analysis. Birkhäuser, Boston, expanded edition edition, 2011. ISBN 9780817646868. URL http://books.google.com/books?vid=ISBN9780817646868.

J. H. Heinbockel. *Introduction to Finite and Infinite Series and Related Topics*. Trafford Publishing, 2010. ISBN 9781426949548. URL http://books.google.com/books?vid=ISBN9781426949545.

Jobst Heitzig and Jürgen Reinhold. Counting finite lattices. *Journal Algebra Universalis*, 48(1):43–53, August 2002. ISSN 0002-5240 (print) 1420-8911 (online). doi: 10.1007/PL00013837. URL http://citeseer.ist.psu.edu/486156.html.

Edwin Hewitt and Karl Robert Stomberg. *Real and Abstract Analysis: A Modern Treatment of the Theory of Functions of a Real Variable*, volume 25 of *Graduate Texts in Mathematics*. Springer, New York, 1965. ISBN 0387901388. URL http://books.google.com/books?vid=ISBN0387901388.

Arend Heyting. Die formalen regeln der intuitionistischen logik i. In *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, pages 42–56, 1930a. English translation of title: "The formal rules of intuitionistic logic I". English translation of text in Mancosu 1998 pages 311–327.

Arend Heyting. Die formalen regeln der intuitionistischen logik ii. In *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, pages 57–71, 1930b. English translation of title: "The formal rules of intuitionistic logic II".

Arend Heyting. Die formalen regeln der intuitionistischen logik iii. In *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, pages 158–169, 1930c. English translation of title: "The formal rules of intuitionistic logic III".

Arend Heyting. Sur la logique intuitionniste. *Bulletin de la Classe des Sciences*, 16:957–963, 1930d. English translation of title: "On intuitionistic logic". English translation of text in Mancosu 1998 pages 306–310.

David Hilbert, Lothar Nordheim, and John von Neumann. über die grundlagen der quantenmechanik (on the bases of quantum mechanics). *Mathematische Annalen*, 98:1–30, 1927. ISSN 0025-5831 (print) 1432-1807 (online). URL http://dz-srv1.sub.uni-goettingen.de/cache/toc/D27776.html.

Ulrich Höhle. Probabilistic uniformization of fuzzy topologies. *Fuzzy Sets and Systems*, 1(4):311–332, October 1978. URL http://dx.doi.org/10.1016/0165-0114(78)90021-0.

Otto Hölder. üeber einen mittelwerthssatz. *Göttingen Nachrichten*, pages 38–47, 1889. URL http://www.digizeitschriften.de/dms/img/?PPN=GDZPPN00252421X.

Samuel S. Holland, Jr. A radon-nikodym theorem in dimension lattices. *Transactions of the American Mathematical Society*, 108(1):66–87, July 1963. URL http://www.jstor.org/stable/1993826.

Samuel S. Holland, Jr. The current interest in orthomodular lattices. In James C. Abbott, editor, *Trends in Lattice Theory*, pages 41–126. Van Nostrand-Reinhold, New York, 1970. URL http://books.google.com/books?id=ZfA-AAAAIAAJ. from Preface: "The present volume contains written versions of four talks on lattice theory delivered to a symposium on Trends in Lattice Theory held at the United States Naval Academy in May of 1966.".

Lars Hörmander. *Notions of Convexity*. Birkhäuser, Boston, 1994. ISBN 0817637990. URL http://books.google.com/books?vid=ISBN0817637990.

Laurence R. Horn. *A Natural History of Negation*. The David Hume Series: Philosophy and Cognitive Science Reissues. CSLI Publications, reissue edition, 2001. URL http://emilkirkegaard.dk/en/wp-content/uploads/A-natural-history-of-negation-Laurence-R.-Horn.pdf.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1990. ISBN 0-521-30586-1. URL http://books.google.com/books?vid=isbn0521305861. Library: QA188H66 1985.

Alfred Edward Housman. *More Poems*. Alfred A. Knopf, 1936. URL http://books.google.com/books?id=rTMiAAAAMAAJ.

John K. Hunter and Bruno Nachtergaele. *Applied Analysis*. World Scientific, 2001. ISBN 9810241917. URL http://books.google.com/books?vid=ISBN9810241917.

K Husimi. Studies on the foundations of quantum mechanics i. *Proceedings of the Physico-Mathematical Society of Japan*, 19:766–789, 1937.

Satoru Igari. *Real Analysis—With an Introduction to Wavelet Theory*, volume 177 of *Translations of mathematical monographs*. American Mathematical Society, 1996. ISBN 9780821821046. URL http://books.google.com/books?vid=ISBN0821821040.

Chris J. Isham. *Modern Differential Geometry for Physicists*. World Scientific Publishing, New Jersey, 2 edition, 1999. ISBN 9810235623. URL http://books.google.com/books?vid=ISBN9810235623.

C.J. Isham. Quantum topology and quantisation on the lattice of topologies. *Classical and Quantum Gravity*, 6:1509–1534, November 1989. doi: 10.1088/0264-9381/6/11/007. URL http://www.iop.org/EJ/abstract/0264-9381/6/11/007.

Vasile I. Istrăţescu. *Inner Product Structures: Theory and Applications*. Mathematics and Its Applications. D. Reidel Publishing Company, 1987. ISBN 9789027721822. URL http://books.google.com/books?vid=ISBN9027721823.

Luisa Iturrioz. Ordered structures in the description of quantum systems: mathematical progress. In *Methods and applications of mathematical logic: proceedings of the VII Latin American Symposium on Mathematical Logic held July 29-August 2, 1985*, volume 69, pages 55–75, Providence Rhode Island, July 29–August 2 1985. Sociedade Brasileira de Lógica, Sociedade Brasileira de Matemática, and the Association for Symbolic Logic, AMS Bookstore (1988). ISBN 0821850768.

Robert C. James. Orthogonality in normed linear spaces. *Duke Mathematical Journal*, 12(2):291–302, 1945. ISSN 0012-7094. doi: 10.1215/S0012-7094-45-01223-3. URL http://projecteuclid.org/euclid.dmj/1077473105.

Robert C. James. Orthogonality and linear functionals in normed linear spaces. *Transactions of the American Mathematical Society*, 61(2):265–292, March 1947. ISSN 1088-6850. doi: 10.2307/1990220. URL http://www.jstor.org/stable/1990220.

S. Jaskowski. Investigations into the system of intuitionistic logic. In Storrs McCall, editor, *Polish Logic, 1920–1939*, pages 259–263. Oxford University Press, 1936. ISBN 9780198243045. URL http://books.google.com/books?vid=ISBN0198243049&pg=PA259. collection published in 1967.

Barbara Jeffcott. The center of an orthologic. *The Journal of Symbolic Logic*, 37(4):641–645, December 1972. doi: 10.2307/2272407. URL http://www.jstor.org/stable/2272407.

S. Jenei. Structure of girard monoids on [0,1]. In Stephen Ernest Rodabaugh and Erich Peter Klement, editors, *Topological and Algebraic Structures in Fuzzy Sets: A Handbook of Recent Developments in the Mathematics of Fuzzy Sets*, volume 20 of *Trends in Logic*, pages 277–308. Springer, 2003. ISBN 9781402015151. URL http://books.google.com/books?vid=ISBN1402015151.

J. L. W. V. Jensen. Sur les fonctions convexes et les ine'galite's entre les valeurs moyennes (on the convex functions and the inequalities between the average values). *Acta Mathematica*, 30(1):175–193, December 1906. ISSN 0001-5962. doi: 10.1007/BF02418571. URL http://www.springerlink.com/content/r55q1411g840j446/.

William Stanley Jevons. *Pure Logic or the Logic of Quality Apart from Quantity; with Remarks on Boole's System and the Relation of Logic and Mathematics*. Edward Stanford, London, 1864. URL http://books.google.com/books?id=WVMOAAAAYAAJ.

Peter Jipsen and Henry Rose. *Varieties of Lattices*. Number 1533 in Lecture notes in mathematics. Springer Verlag, New York, 1992. ISBN 3540563148. URL http://www1.chapman.edu/~jipsen/JipsenRoseVoL.html. available for free online.

Peter Johnstone. *Stone Spaces*. Cambridge University Press, London, 1982. ISBN 0-521-23893-5. URL http://books.google.com/books?vid=ISBN0521337798. Library QA611.

Leonard Benjamin William Jolley. *Summation of Series*. Dover, New York, second revised edition, 1961. URL http://plouffe.fr/simon/math/SummationofSeries.pdf.

P. Jordan and J. von Neumann. On inner products in linear, metric spaces. *Annals of Mathematics*, 36(3):719–723, July 1935. doi: 10.2307/1968653. URL http://www.jstor.org/stable/1968653.

Palle E. T. Jørgensen, Kathy Merrill, Judith Packer, and Lawrence W. Baggett. *Representations, Wavelets and Frames: A Celebration of the Mathematical Work of Lawrence Baggett*. Applied and Numerical Harmonic Analysis. Birkhäuser, 2008. ISBN 9780817646820. URL http://books.google.com/books?vid=ISBN0817646825.

K. D. Joshi. *Foundations of Discrete Mathematics*. New Age International, New Delhi, 1989. ISBN 8122401201. URL http://books.google.com/books?vid=ISBN8122401201.

K. D. Joshi. *Applied Discrete Structures*. New Age International, New Delhi, 1997. ISBN 8122408265. URL http://books.google.com/books?vid=ISBN8122408265.

Kapli D. Joshi. *Introduction To General Topology*. New Age International, 1 edition, 1983. ISBN 9780852264447. URL http://books.google.com/books?vid=ISBN0852264445.

M. Jůza. A note on complete metric spaces. *Matematicko-Fyzikálny Časopis*, 6:143–148, 1956.

Mikhail I. Kadets and Vladimir M. Kadets. *Series in Banach Spaces: Conditional and Unconditional Convergence*, volume 94 of *Operator Theory, Advances and Applications*. Springer/Birkhäuser, 1997. ISBN 9783764354015. URL http://books.google.com/books?vid=ISBN3764354011.

Shizuo Kakutani and George W. Mackey. Ring and lattice characterizations of complex hilbert space. *Bulletin of the American Mathematical Socieity*, 52:727–733, 1946. doi: 10.1090/S0002-9904-1946-08644-9. URL http://www.ams.org/bull/1946-52-08/S0002-9904-1946-08644-9/.

JA Kalman. Two axiom definition for lattices. *Revue Roumaine de Mathematiques Pures et Appliquees*, 13:669–670, 1968. ISSN 0035-3965.

Gudrun Kalmbach. *Orthomodular Lattices*. Academic Press, London, New York, 1983. ISBN 0123945801. URL http://books.google.com/books?vid=ISBN0123945801.

Alexander Karpenko. *Łukasiewicz's Logics and Prime Numbers*. Luniver Press, Beckington, Frome BA11 6TT UK, January 1 2006. ISBN 9780955117039. URL http://books.google.com/books?vid=ISBN0955117038.

Yitzhak Katznelson. *An Introduction to Harmonic Analysis*. Cambridge mathematical library. Cambridge University Press, 3 edition, 2004. ISBN 0521543592. URL http://books.google.com/books?vid=ISBN0521543592.

James P. Keener. *Principles of Applied Mathematics; Transformation and Approximation*. Addison-Wesley Publishing Company, Reading, Massachusets, 1988. ISBN 0-201-15674-1. URL http://www.worldcat.org/isbn/0201156741.

John L. Kelley and T. P. Srinivasan. *Measure and Integral*, volume 116 of *Graduate texts in mathematics*. Springer, New York, 1988. ISBN 0387966331. URL http://books.google.com/books?vid=ISBN0387966331.

John Leroy Kelley. *General Topology*. University Series in Higher Mathematics. Van Nostrand, New York, 1955. ISBN 0387901256. URL http://books.google.com/books?vid=ISBN0387901256. Republished by Springer-Verlag, New York, 1975.

Yoshida Kenko. *The Tsuredzure Gusa of Yoshida No Kaneyoshi. Being the meditations of a recluse in the 14th Century (Essays in Idleness)*. circa 1330. URL http://www.humanistictexts.org/kenko.htm. 1911 translation of circa 1330 text.

Mohamed A. Khamsi and W.A. Kirk. *An Introduction to Metric Spaces and Fixed Point Theory*. John Wiley, New York, 2001. ISBN 978-0471418252. URL http://books.google.com/books?vid=isbn0471418250.

John R. Klauder. *A Modern Approach to Functional Integration*. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, 2010. ISBN 0817647902. URL http://books.google.com/books?vid=isbn0817647902. john.klauder@gmail.com.

Stephen Cole Kleene. On notation for ordinal numbers. *The Journal of Symbolic Logic*, 3(4), December 1938. URL http://www.jstor.org/stable/2267778.

Stephen Cole Kleene. *Introduction to Metamathematics*. North-Holland publishing C°, 1952.

D. Kleitman and B. Rothschild. The number of finite topologies. *Proceedings of the American Mathematical Society*, 25(2):276–282, June 1970. URL http://www.jstor.org/stable/2037205.

Morris Kline. *Mathematical Thought From Ancient To Modern Times*, volume 3. Oxford University Press, New York, 1972. ISBN 9780195061376. URL http://books.google.com/books?vid=ISBN0195061373.

Anthony W Knapp. *Advanced Real Analysis*. Cornerstones. Birkhäuser, Boston, Massachusetts, USA, 1 edition, July 29 2005. ISBN 0817643826. URL http://books.google.com/books?vid=ISBN0817643826.

Andrei Nikolaevich Kolmogorov and Sergei Vasil'evich Fomin. *Introductory Real Analysis*. Dover Publications, New York, 1975. ISBN 0486612260. URL http://books.google.com/books?vid=ISBN0486612260. "unabridged, slightly corrected republication of the work originally published by Prentice-Hall, Inc., Englewood, N.J., in 1970".

A. Korselt. Bemerkung zur algebra der logik. *Mathematische Annalen*, 44(1):156–157, March 1894. ISSN 0025-5831. doi: 10.1007/BF01446978. URL http://www.springerlink.com/content/v681m56871273j73/. referenced by Birkhoff(1948)p.133.

M. G. Krein and M. A. Krasnoselski. Fundamental theorems concerning the extension of hermitian operators and some of their applications to the theory of orthogonal polynomials and the moment problem. *Uspekhi Matematicheskikh Nauk*, 2(3), 1947. URL http://www.turpion.org/php/homes/pa.phtml?jrnid=rm. (Russian Mathematical Surveys).

M. G. Krein, M. A. Krasnoselski, and D. P. Milman. Concerning the deficiency numbers of linear operators in banach space and some geometric questions. *Sbornik Trudov Inst. Matem. AN Ukrainian SSR*, 11:97–112, 1948.

V. Krishnamurthy. On the number of topologies on a finite set. *The American Mathematical Monthly*, 73(2):154–157, February 1966. URL http://www.jstor.org/stable/2313548.

Carlos S. Kubrusly. *The Elements of Operator Theory*. Springer, 1 edition, 2001. ISBN 9780817641740. URL http://books.google.com/books?vid=ISBN0817641742.

Carlos S. Kubrusly. *The Elements of Operator Theory*. Springer, 2 edition, 2011. ISBN 9780817649975. URL http://books.google.com/books?vid=ISBN0817649972.

Kazimierz Kuratowski. Sur l'opération a de l'analysis situs. *Fundamenta Mathematicae*, 3:182–199, 1922. URL http://matwbn.icm.edu.pl/ksiazki/fm/fm3/fm3121.pdf.

Andrew Kurdila and Michael Zabarankin. *Convex Functional Analysis*. Systems & Control: Foundations & Applications. Birkhäuser, Boston, 2005. ISBN 9783764321987. URL http://books.google.com/books?vid=ISBN3764321989.

Shoji Kyuno. An inductive algorithm to construct finite lattices. *Mathematics of Computation*, 33(145):409–421, January 1979. URL https://doi.org/10.1090/S0025-5718-1979-0514837-9.

R.E. Larson and S.J. Andima. The lattice of topologies: a survey. *Rocky Mountain Journal of Mathematics*, 5:177–198, 1975. URL http://rmmc.asu.edu/rmj/rmj.html.

Peter D. Lax. *Functional Analysis*. John Wiley & Sons Inc., USA, 2002. ISBN 0-471-55604-1. URL http://www.worldcat.org/isbn/0471556041. QA320.L345 2002.

Steven R. Lay. *Convex Sets and their Applications*. John Wiley & Sons, New York, 1982. ISBN 0-471-09584-2. URL http://books.google.com/books?vid=isbn0471095842.

J. G. Leathem. *Volume and surface integrals used in physics*, volume 1 of *Cambridge Tracts in Mathematics and Mathematical Physics*. Cambridge University Press, 1 edition, 1905. URL http://archive.org/details/volumesurfaceint01leatuoft.

Gottfried Wilhelm Leibniz. Letter to christian huygens, 1679. In Leroy E. Loemker, editor, *Philosophical Papers and Letters*, volume 2 of *The New Synthese Historical Library*, chapter 27, pages 248–249. Kluwer Academic Press, Dordrecht, 2 edition, September 8 1679. ISBN 902770693X. URL http://books.google.com/books?vid=ISBN902770693X.

Azriel Levy. *Basic Set Theory*. Dover, New York, 2002. ISBN 0486420795. URL http://books.google.com/books?vid=ISBN0486420795.

Rudolf Lidl and Günter Pilz. *Applied Abstract Algebra*. Undergraduate texts in mathematics. Springer, New York, 1998. ISBN 0387982906. URL http://books.google.com/books?vid=ISBN0387982906.

J. Lindenstrauss and L. Tzafriri. On the complemented subspaces problem. *Israel Journal of Mathematics*, 9(2):263–269, 1971. ISSN 0021-2172. doi: 10.1007/BF02771592. URL http://www.springerlink.com/content/wk3121768n46462x/.

Joram Lindenstrauss and Lior Tzafriri. *Classical Banach Spaces I: Sequence Spaces*, volume 92 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, 1977. ISBN 978-3642665592. URL http://www.amazon.com/dp/3642665594.

Lynn H. Loomis. *An Introduction to Abstract Harmonic Analysis*. The University Series in Higher Mathematics. D. Van Nostrand Company, Toronto, 1953. URL http://books.google.com/books?id=aNg-AAAAIAAJ.

Lynn H. Loomis. *The Lattice Theoretic Background of the Dimension Theory of Operator Algebras*, volume 18 of *Memoirs of the American Mathematical Society*. American Mathematical Society, Providence RI, 1955. ISBN 0821812181. URL http://books.google.com/books?id=P3Vl_lXCFRkC.

Proclus Lycaeus. *Proclus: A Commentary on the First Book of Euclid's Elements*. Princeton University Press, circa 450. ISBN 0691020906. URL http://books.google.com/books?vid=OCLC03902909. translation published in 1992. reprint edition, 1992.

Niccolò Machiavelli. *The Literary Works of Machiavelli: Mandragola, Clizia, A Dialogue on Language, and Belfagor, with Selections from the Private Correspondence*. Oxford University Press, 1961. ISBN 0313212481. URL http://www.worldcat.org/isbn/0313212481.

Saunders MacLane and Garrett Birkhoff. *Algebra*. Macmillan, New York, 1 edition, 1967. URL http://www.worldcat.org/oclc/350724.

Saunders MacLane and Garrett Birkhoff. *Algebra*. AMS Chelsea Publishing, Providence, 3 edition, 1999. ISBN 0821816462. URL http://books.google.com/books?vid=isbn0821816462.

M. Donald MacLaren. Atomic orthocomplemented lattices. *Pacific Journal of Mathematics*, 14(2):597–612, 1964. URL http://projecteuclid.org/euclid.pjm/1103034188.

I. J. Maddox. *Elements of Functional Analysis*. Cambridge University Press, Cambridge, 2, revised edition, 1989. ISBN 9780521358682. URL books.google.com/books?vid=ISBN052135868X.

Fumitomo Maeda. *Kontinuierliche Geometrien*, volume 95 of *Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen*. Springer-Verlag, Berlin, 1958.

Fumitomo Maeda and Shûichirô Maeda. *Theory of Symmetric lattices*, volume 173 of *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen*. Springer-Verlag, Berlin/New York, 1970. URL http://books.google.com/books?id=4oeBAAAAIAAJ.

Shûichirô Maeda. On conditions for the orthomodularity. *Proceedings of the Japan Academy*, 42(3):247–251, 1966. ISSN 0021-4280. URL http://joi.jlc.jst.go.jp/JST.Journalarchive/pjab1945/42.247.

Lech Maligranda. A simple proof of the hölder and the minkowski inequality. *The American Mathematical Monthly*, 102(3):256–259, March 1995. URL http://www.jstor.org/stable/2975013.

Stéphane G. Mallat. Multiresolution approximations and wavelet orthonormal bases of $l^2(\mathbb{R})$. *Transactions of the American Mathematical Society*, 315(1):69–87, September 1989. URL http://blanche.polytechnique.fr/~mallat/papiers/math_multiresolution.pdf.

Stéphane G. Mallat. *A Wavelet Tour of Signal Processing*. Elsevier, 2 edition, September 15 1999. ISBN 0-12-466606-X. URL http://books.google.com/books?vid=ISBN012466606X.

Paolo Mancosu, editor. *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*. Oxford University Press, 1998. ISBN 9780195096323. URL http://www.amazon.com/dp/0195096320.

J. L. Massera and J. J. Schäffer. Linear differential equations and functional analysis, i. *The Annals of Mathematics, 2nd Series*, 67(3):517–573, May 1958. URL http://www.jstor.org/stable/1969871.

Stefan Mazur. Sur les anneaux linéaires. *Comptes rendus de l'Académie des sciences*, 207:1025–1027, 1938.

Stefan Mazur and Stanislaus M. Ulam. Sur les transformations isométriques d'espaces vectoriels normées. *Comptes rendus de l'Académie des sciences*, 194:946–948, 1932.

George McCarty. *Topology: An Introduction With Application to Topological Groups*. International Series in Pure and Applied Mathematics. McGraw-Hill Book Company, New York, 1967. URL http://www.amazon.com/dp/0486656330. 1988 Dover edition available.

W. McCune and R. Padmanabhan. *Automated deduction in equational logic and cubic curves*. Number 1095 in Leture Notes in Artificial Intelligence. Springer, Berlin, 1996. ISBN 3540613986. URL http://books.google.com/books?vid=ISBN3540613986.

William McCune, Ranganathan Padmanabhan, and Robert Veroff. Yet another single law for lattices. *Algebra Universalis*, 50(2):165–169, December 2003a. ISSN 0002-5240 (print) 1420-8911 (online). doi: 10.1007/s00012-003-1832-2.

William McCune, Ranganathan Padmanabhan, and Robert Veroff. Yet another single law for lattices. pages 1–5, July 21 2003b. URL http://arxiv.org/abs/math/0307284.

Ralph N. McKenzie. Equational bases for lattice theories. *Mathematica Scandinavica*, 27:24–38, December 1970. ISSN 0025-5521. URL http://www.mscand.dk/article.php?id=1973.

Ralph N. McKenzie. Equational bases and nonmodular lattice varieties. *Transactions of the American Mathematical Society*, 174:1–43, December 1972. URL http://www.jstor.org/stable/1996095.

Karl Menger. Untersuchungen über allgemeine metrik. *Mathematische Annalen*, 100:75–163, 1928. ISSN 0025-5831. URL http://link.springer.com/article/10.1007/BF01455705. (Investigations on general metric).

Claudia Menini and Freddy Van Oystaeyen. *Abstract Algebra; A Comprehensive Treatment*. Marcel Dekker Inc, New York, April 2004. ISBN 0-8247-0985-3. URL http://books.google.com/books?vid=isbn0824709853.

merriamwebsterdictionary. *Merriam Webster Dictionary*. Encyclopedia Britannica Co. URL http://www.merriam-webster.com/.

Franz Mertens. Ueber die multiplicationsregel für zwei unendliche reihen. *Journal f ur die rein und angewandte Mathematik (Crelle's Journal)*, 79:182–184, 1875.

Anthony N. Michel and Charles J. Herget. *Applied Algebra and Functional Analysis*. Dover Publications, Inc., 1993. ISBN 0-486-67598-X. URL http://books.google.com/books?vid=ISBN048667598X. original version published by Prentice-Hall in 1981.

Gradimir V. Milovanović and Igorž. Milovanović. On a generalization of certain results of a. ostrowski and a. lupaş. *Publikacije Elektrotehničkog Fakulteta (Publications Electrical Engineering)*, (643):62–69, 1979. URL http://www.mi.sanu.ac.rs/~gvm/radovi/643.pdf.

Hermann Minkowski. *Geometrie der Zahlen*. Druck und Verlag von B.G. Teubner, Leipzig, 1910. URL http://www.archive.org/details/geometriederzahl00minkrich. Geometry of Numbers.

Dragoslav S. Mitrinović, J. E. Pečarić, and Arlington M. Fink. *Classical and New Inequalities in Analysis*, volume 61 of *Mathematics and its Applications (East European Series)*. Kluwer Academic Publishers, Dordrecht, Boston, London, 2010. ISBN 978-90-481-4225-5. URL http://www.amazon.com/dp/0792320646.

Ilya S. Molchanov. *Theory of Random Sets*. Probability and Its Applications. Springer, 2005. ISBN 185233892X. URL http://books.google.com/books?vid=ISBN185233892X.

M.N. Mukherjee. *Elements of Metric Spaces*. Academic Publishers, 2005. ISBN 9788187504849. URL http://books.google.com/books?vid=isbn8187504846.

H. P. Mulholland. On generalizations of minkowski's inequality in the form of a triangle inequality. *Proceedings of the London Mathematical Society*, s2-51:294–307, 1950. URL http://plms.oxfordjournals.org/content/s2-51/1/294.extract. received 1946 October 10, read 1947 June 19.

Markus Müller-Olm. 2. complete boolean lattices. In *Modular Compiler Verification: A Refinement-Algebraic Approach Advocating Stepwise Abstraction*, volume 1283 of *Lecture Notes in Computer Science*, chapter 2, pages 9–14. Springer, September 12 1997. ISBN 978-3-540-69539-4. URL http://link.springer.com/chapter/10.1007/BFb0027455. Chapter 2.

James R. Munkres. *Topology*. Prentice Hall, Upper Saddle River, NJ, 2 edition, 2000. ISBN 0131816292. URL http://www.amazon.com/dp/0131816292.

Mangesh G Murdeshwar. *General Topology*. New Age International, 2 edition, 1990. ISBN 9788122402469. URL http://books.google.com/books?vid=isbn8122402461.

Masahiro Nakamura. The permutability in a certain orthocomplemented lattice. *Kodai Math. Sem. Rep.*, 9(4):158–160, 1957. doi: 10.2996/kmj/1138843933. URL http://projecteuclid.org/euclid.kmj/1138843933.

H. Nakano and S. Romberger. Cluster lattices. *Bulletin De l'Académie Polonaise Des Sciences*, 19: 5–7, 1971. URL books.google.com/books?id=gkUSAQAAMAAJ.

Hung T. Nguyen and Elbert A. Walker. *A First Course in Fuzzy Logic*. Chapman & Hall/CRC, 3 edition, 2006. ISBN 1584885262. URL http://books.google.com/books?vid=ISBN1584885262.

Nikolaĭ Kapitonovich Nikol'skiĭ, editor. *Functional Analysis I: linear functional analysis*, volume 19 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, Heidelberg, New York, 1992. ISBN 3-540-50584-9. URL http://books.google.com/books?vid=ISBN3540505849.

Ben Noble and James W. Daniel. *Applied Linear Algebra*. Prentice-Hall, Englewood Cliffs, NJ, USA, 3 edition, 1988. ISBN 0-13-041260-0. URL http://www.worldcat.org/isbn/0130412600. Library QA184.N6 1988 512.5 87-11511.

Timothy S. Norfolk. When does a metric generate convex balls? not sure about the year, 1991. URL http://www.math.uakron.edu/~norfolk/.

Vilém Novák, Irina Perfilieva, and Jiří Močkoř. *Mathematical Principles of Fuzzy Logic*. The Springer International Series in Engineering and Computer Science. Kluwer Academic Publishers, Boston, 1999. ISBN 9780792385950. URL http://books.google.com/books?vid=ISBN0792385950.

Timur Oikhberg and Haskell Rosenthal. A metric characterization of normed linear spaces. *Rocky Mountain Journal Of Mathematics*, 37(2):597–608, 2007. URL http://www.ma.utexas.edu/users/rosenthl/pdf-papers/95-oikh.pdf.

Oystein Ore. On the foundation of abstract algebra. i. *The Annals of Mathematics*, 36(2):406–437, April 1935. URL http://www.jstor.org/stable/1968580.

S. V. Ovchinnikov. General negations in fuzzy set theory. *Journal of Mathematical Analysis and Applications*, 92(1):234–239, March 1983. doi: 10.1016/0022-247X(83)90282-2. URL http://www.sciencedirect.com/science/article/pii/0022247X83902822.

James G. Oxley. *Matroid Theory*, volume 3 of *Oxford graduate texts in mathematics*. Oxford University Press, Oxford, 2006. ISBN 0199202508. URL http://books.google.com/books?vid=ISBN0199202508.

R. Padmanabhan. Two identities for lattices. *Proceedings of the American Mathematical Society*, 20(2):409–412, February 1969. doi: 10.2307/2035665. URL http://www.jstor.org/stable/2035665.

R. Padmanabhan and S. Rudeanu. *Axioms for Lattices and Boolean Algebras*. World Scientific, Hackensack, NJ, 2008. ISBN 9812834540. URL http://www.worldscibooks.com/mathematics/7007.html.

Lincoln P. Paine. *Warships of the World to 1900*. Ships of the World Series. Houghton Mifflin Harcourt, 2000. ISBN 9780395984147. URL http://books.google.com/books?vid=ISBN9780395984149.

Endre Pap. *Null-Additive Set Functions*, volume 337 of *Mathematics and Its Applications*. Kluwer Academic Publishers, 1995. ISBN 0792336585. URL http://www.amazon.com/dp/0792336585.

Mladen Pavičić and Norman D. Megill. Is quantum logic a logic? pages 1–24, December 15 2008. URL https://arxiv.org/abs/0812.2698v1. Note: this paper also appears in the collection "*Handbook of Quantum Logic and Quantum Structures: Quantum Logic*" (2009).

Giuseppe Peano. *Calcolo geometrico secondo l'Ausdehnungslehre di H. Grassmann preceduto dalle operazioni della logica deduttiva*. Fratelli Bocca Editori, Torino, 1888a. Geometric Calculus: According to the *Ausdehnungslehre* of H. Grassmann.

Giuseppe Peano. *Geometric Calculus: According to the Ausdehnungslehre of H. Grassmann*. Springer (2000), 1888b. ISBN 0817641262. URL http://books.google.com/books?vid=isbn0817641262. originally published in 1888 in Italian.

Michael Pedersen. *Functional Analysis in Applied Mathematics and Engineering*. Chapman & Hall/CRC, New York, 2000. ISBN 9780849371691. URL http://books.google.com/books?vid=ISBN0849371694. Library QA320.P394 1999.

C.S. Peirce. On the algebra of logic. *American Journal of Mathematics*, 3(1):15–57, March 1880. URL http://www.jstor.org/stable/2369442.

J. E. Pečarić, Frank Proschan, and Yung Liang Tong. *Convex Functions, Partial Orderings, and Statistical Applications*, volume 187 of *Mathematics in Science and Engineering*. Academic Press, San Diego, California, 1992. ISBN 978-0125492508. URL http://books.google.com/books?vid=ISBN0125492502.

Don Pigozzi. Equational logic and equational theories of algebras. Technical Report 135, Purdue University, Indiana, March 1975. URL http://www.cs.purdue.edu/research/technical_reports/#1975. 187 pages.

C Piron. Axiomatique quantique. *Helvetica Physica Acta*, 37:439–468, 1964a. ISSN 0018-0238. English translation completed by M. Cole.

C Piron. Qunatum axiomatics. *Helvetica Physica Acta*, 1964b. English translation of *Axiomatique quantique*, RB4 Technical memo 107/106/104, GPO Engineering Department (London).

Plato. Sophist. In *Plato in Twelve Volumes*, volume 12. Harvard University Press, Cambridge, MA, USA, circa 360 B.C. URL http://data.perseus.org/texts/urn:cts:greekLit:tlg0059.tlg007.perseus-eng1.

Pavel Pudlák and Jiří Tůma. Every finite lattice can be embedded in a finite partition lattice (preliminary communication). *Commentationes Mathematicae Universitatis Carolinae*, 18(2):409–414, 1977. URL http://www.dml.cz/dmlcz/105785.

Pavel Pudlák and Jiří Tůma. Every finite lattice can be embedded in a finite partition lattice. *Algebra Universalis*, 10(1):74–95, December 1980. ISSN 0002-5240 (print) 1420-8911 (online). doi: 10.1007/BF02482893. URL http://www.springerlink.com/content/4r820875g8314806/.

Charles Chapman Pugh. *Real Mathematical Analysis*. Undergraduate texts in mathematics. Springer, New York, 2002. ISBN 0-387-95297-7. URL http://books.google.com/books?vid=ISBN0387952977.

Inder K. Rana. *An Introduction to Measure and Integration*, volume 45 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, R.I., 2 edition, 2002. ISBN 978-0821829745. URL http://books.google.com/books?vid=ISBN0821829742.

E. Renedo, E. Trillas, and C. Alsina. On the law $(a \cdot b')' = b + a' \cdot b'$ in de morgan algebras and orthomodular lattices. *Soft Computing*, 8(1):71–73, October 2003. ISSN 1432-7643. doi: 10.1007/s00500-003-0264-4. URL http://www.springerlink.com/content/7gdjaawe55l11260/.

Greg Restall. *An Introduction to Substructural Logics*. Routledge, 2000. ISBN 9780415215343. URL http://books.google.com/books?vid=ISBN041521534X.

Charles Earl Rickart. *General Theory of Banach Algebras*. University series in higher mathematics. D. Van Nostrand Company, Yale University, 1960. URL http://books.google.com/books?id=PVrvAAAAMAAJ.

Frigyes Riesz. Die genesis des raumbegriffs. *Mathematische und naturwissenschaftliche Berichte aus Ungarn*, 24:309–353, 1906.

Frigyes Riesz. Stetigkeitsbegriff und abstrakte mengenlehre. In Guido Castelnuovo, editor, *Atti del IV Congresso Internazionale dei Matematici*, volume II, pages 18–24, Rome, 1909. Tipografia della R. Accademia dei Lincei. URL http://www.mathunion.org/ICM/ICM1908.2/Main/icm1908.2.0018.0024.ocr.pdf. 1908 April 6–11.

Frigyes Riesz. Zur theorie des hilbertschen raumes. *Acta Scientiarum Mathematicarum*, 7:34–38, 1934. URL http://www.math.u-szeged.hu/acta/. The Theory of Hilbert Space.

A. P. Robertson and Wendy Robertson. *Topological Vector Spaces*, volume 53 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2 edition, 1980. ISBN 9780521298827. URL http://books.google.com/books?vid=ISBN0521298822.

R. Tyrrell Rockafellar. *Convex Analysis*. Princeton landmarks in mathematics. Princeton University Press, 1970. ISBN 9780691015866. URL http://books.google.com/books?vid=ISBN0691015864.

Stefan Rolewicz. *Metric Linear Spaces*. Mathematics and Its Applications. D. Reidel Publishing Company, 1 edition, 1985. ISBN 9789027714800. URL http://books.google.com/books?vid=ISBN9027714800.

Steven Roman. *Lattices and Ordered Sets*. Springer, New York, 1 edition, 2008. ISBN 0387789006. URL http://books.google.com/books?vid=ISBN0387789006.

Maxwell Rosenlicht. *Introduction to Analysis*. Dover Publications, New York, 1968. ISBN 0-486-65038-3. URL http://books.google.com/books?vid=ISBN0486650383.

Gian-Carlo Rota. The number of partitions of a set. *The American Mathematical Monthly*, 71(5): 498–504, May 1964. URL http://www.jstor.org/stable/2312585.

Gian-Carlo Rota. The many lives of lattice theory. *Notices of the American Mathematical Society*, 44 (11):1440–1445, December 1997. URL http://www.ams.org/notices/199711/comm-rota.pdf.

Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, 3 edition, 1976. ISBN 007054235X. URL http://books.google.com/books?vid=ISBN007054235X. Library QA300.R8 1976.

Walter Rudin. *Functional Analysis*. McGraw-Hill, New York, 2 edition, 1991. ISBN 0-07-118845-2. URL http://www.worldcat.org/isbn/0070542252. Library QA320.R83 1991.

Volker Runda. *A Taste of Topology*. Springer, 2005. ISBN 978038725790. URL http://books.google.com/books?vid=ISBN038725790X.

Volker Runde. *A Taste of Topology*. Universitext. Springer, New York, 2005. ISBN 038725790X. URL http://books.google.com/books?vid=ISBN038725790X.

Bryan P. Rynne and Martin A. Youngson. *Linear Functional Analysis*. Springer undergraduate mathematics series. Springer, 2 edition, January 1 2008. ISBN 9781848000056. URL http://books.google.com/books?vid=ISBN1848000057.

Shôichirô Sakai. *C*-Algebras and W*-Algebras*. Springer-Verlag, Berlin, 1 edition, 1998. ISBN 9783540636335. URL http://books.google.com/books?vid=ISBN3540636331. reprint of 1971 edition.

Víačheslav Nikolaevich Saliĭ. *Lattices with Unique Complements*, volume 69 of *Translations of mathematical monographs*. American Mathematical Society, Providence, 1988. ISBN 0821845225. URL http://books.google.com/books?vid=ISBN0821845225. translation of *Reshetki s edinstvennymi dopolneniíam̆i*.

H. Salzmann, T. Grundhöfer, H. Hähl, and R. Löwen. *The Classical Fields: Structural Features of the Real and Rational Numbers*, volume 112 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2007. ISBN 9780521865166. URL http://books.google.com/books?vid=ISBN0521865166.

Usa Sasaki. Orthocomplemented lattices satisfying the exchange axiom. *Journal of Science of the Hiroshima University*, 17:293–302, 1954. ISSN 0386-3034. URL http://journalseek.net/cgi-bin/journalseek/journalsearch.cgi?field=issn&query=0386-3034.

Helmut H. Schaefer and Manfred P. H. Wolff. *Topological Vector Spaces*, volume 3 of *Graduate Texts in Mathematics*. Springer, 2, revised edition, 1999. ISBN 9780387987262. URL http://books.google.com/books?vid=ISBN0387987266.

H. A. Scharz. über ein die flächen kleinsten flächeninhalts betreffendes problem der variationsrechnung. *Acta Soc. Scient. Fen.*, 15:315–362, 1885. URL http://www-stat.wharton.upenn.edu/~steele/Publications/Books/CSMC/CSMC_index.html.

Juljusz Schauder. Zur theorie stetiger abbildungen in funktionalräumen. *Mathematische Zeitschrift*, 26:47–65, 1927. URL http://eudml.org/doc/167911;jsessionid=156A34EBAB6C0E2DDAAC8C1232D23E8F.

Juljusz Schauder. Eine eigenschaft des haarschen orthogonalsystems. *Mathematische Zeitschrift*, 28:317–320, 1928.

Martin Schechter. *Principles of Functional Analysis*, volume 36 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002. ISBN 9780821828953. URL http://books.google.com/books?vid=ISBN0821828959.

Paul S. Schnare. Multiple complementation in the lattice of topologies. *Fundamenta Mathematicae*, 62, 1968. URL http://matwbn.icm.edu.pl/tresc.php?wyd=1&tom=62.

Bernd Siegfried Walter Schröder. *Ordered Sets: An Introduction*. Birkhäuser, Boston, 2003. ISBN 0817641289. URL http://books.google.com/books?vid=ISBN0817641289.

Isaac Schur. Uber die charakterischen wurzeln einer linearen substitution mit enier anwendung auf die theorie der integralgleichungen (over the characteristic roots of one linear substitution with an application to the theory of the integral). *Mathematische Annalen*, 66:488–510, 1909. URL http://dz-srv1.sub.uni-goettingen.de/cache/toc/D38231.html.

Karl Hermann Amandus Schwarz. *Ueber Ein Die Flachen Kleinsten Flächeninhalts; Betreffendes Problem Der Variationsrechnung (Over flattening smallest flat contents;*. October 31 1885. URL http://www-stat.wharton.upenn.edu/~steele/Publications/Books/CSMC/Schwarz.pdf. Festschrift Zum Jubelgeburtstage des Herrn Karl Weierstrass (anniversary publication to comemerate the birthdays of Mr. Karl Weierstrass).

Mícheál Ó Searcóid. *Elements of Abstract Analysis*. Springer Undergraduate Mathematics Series. Springer, 2002. ISBN 9781852334246. URL http://books.google.com/books?vid=ISBN185233424X.

Henry Maurice Sheffer. Review of "a survey of symbolic logic" by c. i. lewis. *The American Mathematical Monthly*, 27(7/9):309–311, July–September 1920. URL http://www.jstor.org/stable/2972257.

Alexander Shen and Nikolai Konstantinovich Vereshchagin. *Basic Set Theory*, volume 17 of *Student mathematical library*. American Mathematical Society, Providence, July 9 2002. ISBN 0821827316. URL http://books.google.com/books?vid=ISBN0821827316. translated from Russian.

George Finlay Simmons. *Calculus Gems: Brief Lives and Memorable Mathematicians*. Mathematical Association of America, Washington DC, 2007. ISBN 0883855615. URL http://books.google.com/books?vid=ISBN0883855615.

Barry Simon. *Convexity: An Analytic Viewpoint*, volume 187 of *Cambridge Tracts in Mathematics*. Cambridge University Press, May 19 2011. ISBN 9781107007314. URL http://books.google.com/books?vid=ISBN1107007313.

Iván Singer. *Bases in Banach Spaces I*, volume 154 of *Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete*. Springer-Verlag New York, New York, 1970. ISBN 9780387048338.

Neil J. A. Sloane. On-line encyclopedia of integer sequences. World Wide Web, 2014. URL http://oeis.org/.

Boleslaw Sobociński. Axiomatization of a partial system of three-value calculus of propositions. *Journal of Computing Systems*, 1:23–55, 1952.

Boleslaw Sobociński. Equational two axiom bases for boolean algebras and some other lattice theories. *Notre Dame Journal of Formal Logic*, 20(4):865–875, October 1979. URL http://projecteuclid.org/euclid.ndjfl/1093882808.

Houshang H. Sohrab. *Basic Real Analysis*. Birkhäuser, Boston, 1 edition, 2003. ISBN 978-0817642112. URL http://books.google.com/books?vid=ISBN0817642110.

Richard P. Stanley. *Enumerative Combinatorics*, volume 49 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 1 edition, 1997. ISBN 0-521-55309-1. URL http://books.google.com/books?vid=ISBN0521663512.

J. Michael Steele. *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. Cambridge University Press, Cambridge, April 26 2004. ISBN 052154677X. URL http://www-stat.wharton.upenn.edu/~steele/Publications/Books/CSMC/CSMC_index.htm.

Lynn Arthur Steen. Highlights in the history of spectral theory. *The American Mathematical Monthly*, 80(4):359–381, April 1973. ISSN 00029890. URL http://www.jstor.org/stable/2319079.

Lynn Arthur Steen and J. Arthur Seebach. *Counterexamples in Topology*. Springer-Verlag, 2, revised edition, 1978. URL http://books.google.com/books?vid=ISBN0486319296. A 1995 "unabridged and unaltered republication" Dover edition is available.

Anne K. Steiner. The lattice of topologies: Structure and complementation. *Transactions of the American Mathematical Society*, 122(2):379–398, April 1966. URL http://www.jstor.org/stable/1994555.

Manfred Stern. *Semimodular Lattices: Theory and Applications*, volume 73 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, May 13 1999. ISBN 0521461057. URL http://books.google.com/books?vid=ISBN0521461057.

M. H. Stone. The theory of representation for boolean algebras. *Transactions of the American Mathematical Society*, 40(1):37–111, July 1936a. URL http://www.jstor.org/stable/1989664.

Marshall Harvey Stone. *Linear transformations in Hilbert space and their applications to analysis*, volume 15 of *American Mathematical Society. Colloquium publications*. American Mathematical Society, New York, 1932. URL http://books.google.com/books?vid=ISBN0821810154. 1990 reprint of the original 1932 edition.

Marshall Harvey Stone. Application of boolean algebras to topology. *Mathematics Sbornik*, 1:765–771, 1936b.

Daniel W. Stroock. *A Concise Introduction to the Theory of Integration*. Birkhäuser, Boston, 3 edition, 1999. ISBN 0817640738. URL http://books.google.com/books?vid=ISBN0817640738.

Wilson Alexander Sutherland. *Introduction to Metric and Topological Spaces*. Oxford University Press, illustrated, reprint edition, 1975. ISBN 9780198531616. URL http://books.google.com/books?vid=ISBN0198531613.

Karl Svozil. *Randomness & Undecidability in Physics*. World Scientific, 1994. ISBN 981020809X. URL http://books.google.com/books?vid=ISBN981020809X.

Otto Szász and Joshua Barlaz. *Introduction to the Theory of Divergent Series*. Department of Mathematics Graduate School of Arts and Sciences University of Cincinnati, revised edition, 1952. URL http://books.google.com/books?id=ELsrAAAAYAAJ.

Gábor J. Székely. *Paradoxes in Probability Theory and Mathematical Statistics*. Mathematics and Its Applications. Springer, D. Reidel Publishing Company, Kluwer Academic Publishers Group, Dordrecht/Boston/Lancaster/Tokyo, January 31 1986. ISBN 978-9027718990.

Saburo Tamura. Two identities for lattices, distributive lattices and modular lattices with a constant. *Notre Dame Journal of Formal Logic*, 16(1):137–140, 1975. URL http://projecteuclid.org/euclid.ndjfl/1093891622.

Terence Tao. *Epsilon of Room, I: Real Analysis: pages from year three of a mathematical blog*, volume 117 of *Graduate Studies in Mathematics*. American Mathematical Society, 2010. ISBN 9780821852781. URL http://books.google.com/books?vid=ISBN0821852787.

Terence Tao. *An Introduction to Measure Theory*, volume 126 of *Graduate Studies in Mathematics*. American Mathematical Society, 2011. ISBN 9780821869192. URL http://books.google.com/books?vid=ISBN0821869191.

Alfred Tarski. Equational logic and equational theories of algebras. In Helmut J. Thiele H. Arnold Schmidt, Kurt Schütte, editor, *Contributions to Mathematical Logic: Proceedings of the Logic Colloquium*, Studies in logic and the foundations of mathematics, pages 275–288, Hannover, August 1966. International Union of the History and Philosophy of Science, Division of Logic, Methodology and Philosophy of Science, North-Holland Publishing Company (1968). URL http://books.google.com/books?id=W7tLAAAAMAAJ.

Walter Taylor. Equational logic. *Houston Journal of Mathematics*, pages i–iii,1–83, 1979.

Walter Taylor. *Equational Logic*, chapter Appendix 4, pages 378–400. Springer, New York, 2008. ISBN 978-0-387-77486-2. doi: 10.1007/978-0-387-77487-9. URL http://www.springerlink.com/content/rp1374214u122546/. an "abridged" version of Taylor 1979.

Brian S. Thomson, Andrew M. Bruckner, and Judith B. Bruckner. *Elementary Real Analysis*. www.classicalrealanalysis.com, 2 edition, 2008. ISBN 9781434843678. URL http://classicalrealanalysis.info/com/Elementary-Real-Analysis.php.

Wolfgang J. Thron. *Topological structures*. Holt, Rinehart and Winston, New York, 1966. URL http://books.google.com/books?id=JRM_AAAAIAAJ.

Heinrich Franz Friedrich Tietze. Beiträge zur allgemeinen topologie i. *Mathematische Annalen*, 88 (3–4):290–312, 1923. URL http://link.springer.com/article/10.1007/BF01579182.

Elmer Tolsted. An elementary derivation of cauchy, hölder, and minkowski inequalities from young's inequality. *Mathematics Magazine*, 37:2–12, 1964. URL http://mathdl.maa.org/mathDL/22/?pa=content&sa=viewDocument&nodeId=3036.

A.S. Troelstra and D. van Dalen. *Constructivism in Mathematics: An Introduction*, volume 121 of *Studies in Logic and the Foundations of Mathematics*. North Holland/Elsevier, Amsterdam/New York/Oxford/Tokyo, 1988. ISBN 0080570887. URL http://books.google.com/books?vid=ISBN0080570887.

Constantin Udriste. *Convex Functions and Optimization Methods on Riemannian Manifolds*, volume 297 of *Mathematics and Its Applications*. Springer, July 31 1994. ISBN 0792330021. URL http://books.google.com/books?vid=ISBN0792330021.

Stanislaw Marcin Ulam. *Adventures of a Mathematician*. University of California Press, Berkeley, 1991. ISBN 0520071549. URL http://books.google.com/books?vid=ISBN0520071549.

R. Vaidyanathaswamy. *Treatise on set topology, Part I*. Indian Mathematical Society, Madras, 1947. MR 9, 367.

R. Vaidyanathaswamy. *Set Topology*. Chelsea Publishing, 2 edition, 1960. ISBN 0486404560. URL http://www.amazon.com/dp/0486404560. note: 978-0486404561 is a Dover edition: "This Dover edition, first published in 1999, is an unabridged republication of the work originally published in 1960 by Chelsea Publishing Company.".

Jussi Väisälä. A proof of the mazur-ulam theorem. *The American Mathematical Monthly*, 110(7): 633–635, August–September 2003. URL http://www.helsinki.fi/~jvaisala/mazurulam.pdf.

Robert W. Vallin. Continuity and differentiability aspects of metric preserving functions. *Real Analysis Exchange*, 25(2):849–868, 1999. URL projecteuclid.org/euclid.rae/1230995419.

M.L.J. van de Vel. *Theory of Convex Structures*, volume 50 of *North-Holland Mathematical Library*. North-Holland, Amsterdam, 1993. ISBN 978-0444815057. URL http://books.google.com/books?vid=ISBN0444815058.

A.C.M. van Rooij. The lattice of all topologies is complemented. *Canadian Journal of Mathematics*, 20(805–807), 1968. URL http://books.google.com/books?id=24hsmjEDbNUC.

V. S. Varadarajan. *Geometry of Quantum Theory*. Springer, 2 edition, 1985. ISBN 9780387493862. URL http://books.google.com/books?vid=ISBN0387493867.

Brani Vidakovic. *Statistical Modeling by Wavelets*. John Wiley & Sons, Inc, New York, 1999. ISBN 0-471-29365-2. URL http://www.amazon.com/dp/0471293652.

John von Neumann. Allgemeine eigenwerttheorie hermitescher funktionaloperatoren. *Mathematische Annalen*, 102(1):49–131, 1929. ISSN 0025-5831 (print) 1432-1807 (online). URL http://resolver.sub.uni-goettingen.de/purl?GDZPPN002273535. General eigenvalue theory of Hermitian functional operators.

John von Neumann. *Continuous Geometry*. Princeton mathematical series. Princeton University Press, Princeton, 1960. URL http://books.google.com/books?id=3bjqOgAACAAJ.

H. Wallman. Lattices and topological spaces. *Annals of Mathematics*, 39(1):112–116, 1938. URL http://www.jstor.org/stable/1968717.

Gilbert G. Walter and XiaoPing Shen. *Wavelets and Other Orthogonal Systems*. Chapman and Hall/CRC, New York, 2 edition, 2001. ISBN 1-58488-227-1. URL http://books.google.com/books?vid=ISBN1584882271. Library QA403.3.W34 2000.

Stephen Watson. The number of complements in the lattice of topologies on a fixed set. *Topology and its Applications*, 55(2):101–125, 26 January 1994. URL http://www.sciencedirect.com/science/journal/01668641.

Heinrich Martin Weber. Die allgemeinen grundlagen der galois'schen gleichungstheorie. *Mathematische Annalen*, 43(4):521–549, December 1893. URL http://resolver.sub.uni-goettingen.de/purl?GDZPPN002254670. The general foundation of Galois' equation theory.

Joseph H. Maclagan Wedderburn. On hypercomplex numbers. *Proceedings of the London Mathematical Society*, 6:77–118, 1907. URL http://plms.oxfordjournals.org/cgi/reprint/s2-6/1/77.

Dirk Werner. *Funktionalanalysis*. Springer, Berlin, Germany, 5 edition, 2004. ISBN 3540213813. URL http://books.google.com/books?vid=ISBN3540213813.

Hermann Weyl. Emmy noether. *Scripta Mathematica...*, 3:201–220, April 26 1935a. URL http://link.springer.com/chapter/10.1007/978-1-4684-0535-4_6. memorial address at Bryn Mawr College.

Hermann Weyl. *Emmy Noether*, pages 112–152. Birkhäuser, Boston, April 26 1935b. ISBN 978-1-4684-0537-8. URL http://link.springer.com/chapter/10.1007/978-1-4684-0535-4_6.

Hermann Weyl. Emmy noether. In Peter Pesic, editor, *Levels of Infinity: Selected Writings on Mathematics and Philosophy*. Courier Dover Publications, April 26 1935c. ISBN 9780486266930. URL http://books.google.com/books?vid=ISBN0486266931.

Hermann Weyl. The method of orthogonal projection in potential theory. *Duke Mathematical Journal*, 7(1):411–444, 1940. URL http://projecteuclid.org/euclid.dmj/1077492266.

Eldon Whitesitt. *Boolean Algebra and Its Applications*. Dover, New York, 1995. ISBN 0486684830. URL http://books.google.com/books?vid=ISBN0486684830.

Philip M. Whitman. Lattices, equivalence relations, and subgroups. *Bulletin of the American Mathematical Society*, 52:507–522, 1946. ISSN 0002-9904. doi: 10.1090/S0002-9904-1946-08602-4. URL http://www.ams.org/bull/1946-52-06/S0002-9904-1946-08602-4/.

Edmund Taylor Whittaker. On the functions which are represented by the expansions of the interpolation theory. *Proceedings of the Royal Society*, 35:181–194, 1915. URL http://dx.doi.org/10.1017/S0370164600017806.

Edmund Taylor Whittaker and G. N. Watson. *A Course of Modern Analysis an Introduction to the General Theory of Infinite Processes and of Analytic Functions*. Cambridge at the University Press, London, 3 edition, 1920. URL http://archive.org/details/courseofmodernan00whit.

Norbert Wiener. *I Am a Mathematician*. The M.I.T. Press, Cambridge, Massachusetts, USA, 1956. ISBN 0-262-73007-3. URL http://www.worldcat.org/oclc/60896484.

W. John Wilbur. Quantum logic and the locally convex spaces. *Transactions of the American Mathematical Society*, 207:343–360, June 1975. URL http://www.jstor.org/stable/1997181.

J. B. Wilker. Rings of sets are really rings. *The American Mathematical Monthly*, 89(3):211–211, March 1982. URL http://www.jstor.org/stable/2320207.

P. Wojtaszczyk. *A Mathematical Introduction to Wavelets*, volume 37 of *London Mathematical Society student texts*. Cambridge University Press, February 13 1997. ISBN 9780521578943. URL http://books.google.com/books?vid=ISBN0521578949.

Ronald R. Yager. On the measure of fuzziness and negation part i: Membership in the unit interval. *International Journal of General Systems*, 5(4):221–229, 1979. doi: 10.1080/03081077908547452. URL http://www.tandfonline.com/doi/abs/10.1080/03081077908547452.

Ronald R. Yager. On the measure of fuzziness and negation ii: Lattices. *Information and Control*, 44(3):236–260, March 1980. doi: 10.1016/S0019-9958(80)90156-4. URL http://www.sciencedirect.com/science/article/pii/S0019995880901564.

Kôsaku Yosida. *Functional Analysis*, volume 123 of *Classics in Mathematics*. Springer, 6 reprint revised edition, 1980. ISBN 9783540586548. URL http://books.google.com/books?vid=ISBN3540586547.

Robert M. Young. *An introduction to nonharmonic Fourier series*, volume 93 of *Pure and applied mathematics*. Academic Press, 1 edition, 1980. ISBN 0127728503. URL http://books.google.com/books?vid=ISBN0127728503.

Robert M. Young. *An introduction to nonharmonic Fourier series*, volume 93 of *Pure and applied mathematics*. Academic Press, revised first edition, May 16 2001. ISBN 0127729550. URL http://books.google.com/books?vid=ISBN0127729550.

William Henry Young. On classes of summable functions and their fourier series. *Proceedings of the Royal Society of London*, 87(594):225–229, August 1912. URL http://www.archive.org/details/philtrans02496252.

Antoni Zygmund. Trigonometric series volume i. In *Trigonometric Series*, volume 1, page 383. Cambridge University Press, London/New York/Melbourne, 3 edition, 2002. ISBN 9780521890533. URL http://books.google.com/books?vid=ISBN0521890535.

REFERENCE INDEX

SUBJECT INDEX

License

…last page …please stop reading …