A Book Concerning Statistical Signal Processing

VERSION 0.15

Daniel J. Greenhoe

***Signal Processing ABCs* series**
volume 2

title:	*A Book Concerning Statistical Signal Processing*
document type:	`book`
series:	*Signal Processing ABCs*
volume:	2
author:	Daniel J. Greenhoe
version:	VERSION 0.15
time stamp:	2019 August 21 (Wednesday) 07:11pm 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/Signal-Processing-ABCs

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 appearing throughout this text 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

COPYRIGHT © 2019 DANIEL J. GREENHOE

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	
$\circledwedge$	relational and	
$X \times Y$	Cartesian product of X and Y	
$(\triangle, \triangledown)$	ordered pair	
$\lvert z \rvert$	absolute value of a complex number z	
$=$	equality relation	
$\triangleq$	equality by definition	
$\rightarrow$	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...

COPYRIGHT © 2019 DANIEL J. GREENHOE

symbol	description	
$\wedge$	logical AND operation	
$\vee$	logical inclusive OR operation	
$\oplus$	logical exclusive OR operation	
$\Longrightarrow$	"implies";	"only if"
$\Longleftarrow$	"implied by";	"if"
$\Longleftrightarrow$	"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	
$\leq$	reflexive ordering relation	"less than or equal to"
$\geq$	reflexive ordering relation	"greater than or equal to"
$<$	irreflexive ordering relation	"less than"
$>$	irreflexive ordering relation	"greater than"
measures on sets:		
$\mid X \mid$	order or counting measure of a set X	
distance spaces:		
d	metric or distance function	
linear spaces:		
$\lVert\cdot\rVert$	vector norm	
$\lvert\lvert\lvert\cdot\rvert\rvert\rvert$	operator norm	
$\langle\triangle \mid \triangledown\rangle$	inner-product	
$\mathsf{span}(\boldsymbol{V})$	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	
$\mathcal{S}_{\mathsf{j}}(X,Y)$	set of *surjective* functions from X to Y	
$\mathcal{I}_{\mathsf{j}}(X,Y)$	set of *injective* functions from X to Y	
$\mathcal{B}_{\mathsf{j}}(X,Y)$	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 (Definition P.2 page 349)	
$\hat{\mathbf{F}}$	*Fourier Series* operator	

...continued on next page...

symbol	description
$\check{\mathbf{F}}$	*Discrete Time Fourier Series* operator (Definition Q.1 page 359)
$\mathbf{Z}$	*Z-Transform* operator (Definition R.4 page 370)
$\tilde{\mathsf{f}}(\omega)$	*Fourier Transform* of a function $\mathsf{f}(x) \in \boldsymbol{L}^2_{\mathbb{R}}$
$\check{\mathsf{x}}(\omega)$	*Discrete Time Fourier Transform* of a sequence $(\!(x_n \in \mathbb{C})\!)_{n\in\mathbb{Z}}$
$\check{\mathsf{x}}(z)$	*Z-Transform* of a sequence $(\!(x_n \in \mathbb{C})\!)_{n\in\mathbb{Z}}$

COPYRIGHT © 2019 DANIEL J. GREENHOE

SYMBOL INDEX

$\bar{\delta}_n$, 283
$(\boldsymbol{A}, \|\cdot\|, *)$, 269
$\|\cdot\|$, 287
$\perp$, 218
$\star$, 369
σ, 383
$\tilde{\mathbf{F}}$, 350
$|\boldsymbol{x}|$, 216
$\boldsymbol{x}^-$, 216
$\boldsymbol{x}^+$, 216
d, 219
$\mathbf{I}_{\mathsf{m}}$, 267
$\mathbf{I}$, 320
$\mathbf{R}_{\mathsf{e}}$, 267
$\mathbf{Z}$, 370
$\mathsf{N}_n(x)$, 383
$\mathsf{epi}(\mathsf{f})$, 299
$\mathsf{hyp}(\mathsf{f})$, 299
span, 221
$*$, 266
$|||\cdot|||$, 324
$\textcircled{\perp}$, 294, 296
$\star$, 352
$\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$, 327
$(\cdot : \cdot)$, 297, 298
$(\cdot : \cdot]$, 297, 298
$[\cdot : \cdot)$, 297, 298
$[\cdot : \cdot]$, 297, 298
$\boldsymbol{Y}^{\boldsymbol{X}}$, 320
ρ, 192, 265
σ_{c}, 192
σ_{p}, 192
σ_{r}, 192
σ, 192, 265
r, 265

Copyright © 2019 Daniel J. Greenhoe

CONTENTS

Title page . . . v
Typesetting . . . vi
Quotes . . . vii
Symbol list . . . ix
Symbol index . . . xiii
Contents . . . xv

I Statistical Analysis 1

1 Expectation operator 3
1.1 Definitions . . . 3
1.2 Expectation as a linear operator . . . 4
1.3 Expectation inequalities . . . 6
1.4 Joint and conditional probability spaces . . . 7
1.5 Expectation inner product space . . . 8

2 Random Sequences 11
2.1 Definitions . . . 11
2.2 Properties . . . 12
2.3 Wide Sense Stationary processes . . . 12
2.4 Spectral density . . . 14
2.5 Spectral Power . . . 16

3 Continuous Random Processes 17
3.1 Definitions . . . 17

4 Random Process Eigen-Analysis 19
4.1 Definitions . . . 19
4.2 Properties . . . 20

II Statistical Processing 25

5 Operations on Random Variables 27
5.1 Functions of one random variable . . . 27
5.2 Functions of two random variables . . . 33

6 Operators on Discrete Random Sequences 37
6.1 LTI operators on random sequences . . . 37
6.2 LTI operators on WSS random sequences . . . 38
6.3 Parallel operators on WSS random sequences . . . 39
6.4 Whitening discrete random sequences . . . 40

7 Operators on Continuous Random Sequences 43
7.1 LTI Operations on non-stationary random processes . . . 43

7.2 LTI Operations on WSS random processes . . . 45
7.3 Whitening continuous random sequences . . . 50

8 Additive noise on random sequences **53**
8.1 Additive noise and correlation . . . 53
8.2 Additive noise and operators . . . 55
8.3 Additive noise and LTI operators . . . 58
8.4 Additive noise and dual operators . . . 60

III Statistical Estimation **63**

9 Estimation Overview **65**
9.1 Estimation types . . . 65
9.2 Estimation criterion . . . 66
9.3 Measures of estimator quality . . . 67
9.4 Estimation techniques . . . 68
9.5 Sequential decoding . . . 69

10 Norm Minimization **71**
10.1 Minimum mean square estimation . . . 72
10.2 Least squares . . . 74

11 Gradient Search Techniques **77**
11.1 Gradient search techniques . . . 77
11.2 Direct search . . . 79

12 KL-expansion application **81**
12.1 Sufficient statistics . . . 81
12.2 Optimal symbol estimation . . . 86
12.3 Colored noise . . . 94
12.4 Signal matching . . . 94

13 Moment Estimation **97**
13.1 Mean Estimation . . . 97
13.2 Variance Estimation . . . 100
13.3 Estimates in terms of moment estimates . . . 103

14 Correlation Estimation **105**

15 Spectral Estimation **107**

16 Density Estimation **109**

17 System Identification **111**
17.1 Estimation techniques . . . 111
17.2 Additive noise system models . . . 112
17.3 Transfer function estimate definitions and interpretation . . . 113
17.4 Estimator relationships . . . 118
17.5 Alternate forms . . . 123
17.6 Least squares estimates of non-linear systems . . . 125
17.7 Least squares estimates of linear systems . . . 129
17.8 Coherence . . . 133
17.8.1 Application . . . 133
17.8.2 Definitions . . . 134
17.8.3 A warning . . . 134

18 Estimating Noise **137**

IV Statistical Detection **139**

19 Communication channels **141**

Copyright © 2019 Daniel J. Greenhoe

19.1 System model . . . 141
19.1.1 Channel operator . . . 142
19.1.2 Receive operator . . . 142
19.2 Optimization in the case of additional operations . . . 143
19.3 Channel Statistics . . . 144

20 Optimal Symbol Detection **145**
20.1 ML Estimation . . . 145
20.2 Generalized coherent modulation . . . 146
20.3 Frequency Shift Keying (FSK) . . . 147
20.4 Quadrature Amplitude Modulation (QAM) . . . 149
20.4.1 Receiver statistics . . . 149
20.4.2 Detection . . . 150
20.4.3 Probability of error . . . 150
20.5 Phase Shift Keying (PSK) . . . 151
20.5.1 Receiver statistics . . . 151
20.5.2 Detection . . . 153
20.5.3 Probability of error . . . 153
20.6 Pulse Amplitude Modulation (PAM) . . . 154
20.6.1 Receiver statistics . . . 154
20.6.2 Detection . . . 154
20.6.3 Probability of error . . . 155

21 Network Detection **157**
21.1 Detection . . . 157
21.2 Bayesian Estimation . . . 157
21.3 Joint Gaussian Model . . . 158
21.4 2 hypothesis, 2 sensor detection . . . 159

V Appendices **163**

A Probability Space **165**
A.1 Probability functions . . . 166
A.2 Probability Space . . . 167
A.3 Properties . . . 168
A.4 Examples . . . 170
A.5 Probability subspaces . . . 173

B Probability Density Functions **175**
B.1 Random variables . . . 175
B.2 Probability distributions . . . 176
B.3 Properties . . . 176

C Some Probability Density Functions **181**
C.1 Discrete distributions . . . 181
C.2 Continuous distributions . . . 181
C.2.1 Uniform distribution . . . 181
C.2.2 Gaussian distribution . . . 182
C.2.3 Gamma distribution . . . 183
C.2.4 Chi-squared distributions . . . 184
C.2.5 Radial distributions . . . 187
C.3 Joint Gaussian distributions . . . 188

D Spectral Theory **191**
D.1 Operator Spectrum . . . 191
D.2 Fredholm kernels . . . 193

E Matrix Calculus **195**
E.1 First derivative of a vector with respect to a vector . . . 195
E.2 First derivative of a matrix with respect to a scalar . . . 202
E.3 Second derivative of a scalar with respect to a vector . . . 204

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

E.4 Multiple derivatives of a vector with respect to a scalar . . . 205

F Algebraic structures **207**

G Linear spaces **209**
G.1 Definition and basic results . . . 209
G.2 Order on Linear Spaces . . . 214

H Linear Combinations **221**
H.1 Linear combinations in linear spaces . . . 221
H.2 Bases in topological linear spaces . . . 225
H.3 Schauder bases in Banach spaces . . . 226
H.4 Linear combinations in inner product spaces . . . 228
H.5 Orthonormal bases in Hilbert spaces . . . 231
H.6 Riesz bases in Hilbert spaces . . . 239
H.7 Frames in Hilbert spaces . . . 244

I Linear Subspaces **251**
I.1 Subspaces of a linear space . . . 251
I.2 Subspaces of an inner product space . . . 257
I.3 Subspaces of a Hilbert Space . . . 261
I.4 Subspace Metrics . . . 263
I.5 Literature . . . 264

J Normed Algebras **265**
J.1 Algebras . . . 265
J.2 Star-Algebras . . . 266
J.3 Normed Algebras . . . 269
J.4 C* Algebras . . . 269

K Inner Product Spaces **271**
K.1 Definition and basic results . . . 271
K.2 Relationship between norms and inner products . . . 275
K.2.1 Norms induced by inner products . . . 275
K.2.2 Inner products induced by norms . . . 277
K.3 Orthogonality . . . 283

L Normed Linear Spaces **287**
L.1 Definition and basic results . . . 287
L.2 Relationship between metrics and norms . . . 289
L.2.1 Metrics generated by norms . . . 289
L.2.2 Norms generated by metrics . . . 292
L.3 Orthogonality on normed linear spaces . . . 294

M Intervals and Convexity **297**
M.1 Intervals . . . 297
M.2 Convex sets . . . 298
M.3 Convex functions . . . 299
M.4 Literature . . . 302

N Finite Sums **305**
N.1 Summation . . . 305
N.2 Means . . . 306
N.2.1 Weighted ϕ-means . . . 306
N.2.2 Power means . . . 308
N.3 Inequalities on power means . . . 311
N.4 Power Sums . . . 316

O Operators on Linear Spaces **319**
O.1 Operators on linear spaces . . . 319
O.1.1 Operator Algebra . . . 319
O.1.2 Linear operators . . . 320
O.2 Operators on Normed linear spaces . . . 324

Copyright © 2019 Daniel J. Greenhoe

O.2.1 Operator norm . . . 324
O.2.2 Bounded linear operators . . . 327
O.2.3 Adjoints on normed linear spaces . . . 329
O.2.4 More properties . . . 330
O.3 Operators on Inner product spaces . . . 331
O.3.1 General Results . . . 331
O.3.2 Operator adjoint . . . 332
O.4 Special Classes of Operators . . . 334
O.4.1 Projection operators . . . 334
O.4.2 Self Adjoint Operators . . . 336
O.4.3 Normal Operators . . . 337
O.4.4 Isometric operators . . . 339
O.4.5 Unitary operators . . . 342
O.5 Operator order . . . 346

P Fourier Transform **349**
P.1 Definitions . . . 349
P.2 Operator properties . . . 350
P.3 Convolution . . . 352
P.4 Real valued functions . . . 353
P.5 Moment properties . . . 354
P.6 Examples . . . 356

Q Discrete Time Fourier Transform **359**
Q.1 Definition . . . 359
Q.2 Properties . . . 359
Q.3 Derivatives . . . 367

R Operations on Sequences **369**
R.1 Convolution operator . . . 369
R.2 Z-transform . . . 370
R.3 From z-domain back to time-domain . . . 372
R.4 Zero locations . . . 373
R.5 Pole locations . . . 374
R.6 Mirroring for real coefficients . . . 375
R.7 Rational polynomial operators . . . 376
R.8 Filter Banks . . . 377
R.9 Inverting non-minimum phase filters . . . 381

S B-Splines **383**
S.1 Definitions . . . 383
S.2 Algebraic properties . . . 387
S.3 Projection properties . . . 396
S.4 Fourier analysis . . . 397
S.5 Basis properties . . . 399
S.5.1 Uniqueness properties . . . 399
S.5.2 Partition of unity properties . . . 400
S.5.3 Riesz basis properties . . . 401
S.6 Mutiresolution properties . . . 407
S.6.1 Introduction . . . 407
S.6.2 B-spline dyadic decomposition . . . 407
S.6.3 B-spline MRA scaling functions . . . 409
S.6.4 B-spline MRA coefficient sequences . . . 410

T Source Code **415**

Back Matter **419**
References . . . 420
Reference Index . . . 449
Subject Index . . . 453
License . . . 470
End of document . . . 471

Copyright © 2019 Daniel J. Greenhoe

Part I

Statistical Analysis

CHAPTER 1

EXPECTATION OPERATOR

1.1 Definitions

In a *probability space* $(\Omega, \mathbb{E}, \mathsf{P})$ (Definition A.2 page 167), all probability information is contained in the *measure* P. Often times this information is overwhelming and a simpler statistic, which does not offer so much information, is sufficient. Some of the most common statistics can be conveniently expressed in terms of the *expectation operator* E.

Definition 1.1. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE *(Definition A.2 page 167) and* X *a* RANDOM VARIABLE *(Definition B.1 page 176) on* $(\Omega, \mathbb{E}, \mathsf{P})$ *with* PROBABILITY DENSITY FUNCTION p_{X}.

DEF *The* ***expectation operator*** E_{X} *on* X *is defined as*

$$\mathsf{E}_{\mathsf{X}}\mathsf{X} \triangleq \int_{x \in \mathbb{F}} x \mathsf{p}_{\mathsf{X}}(x)\, \mathrm{d}x.$$

We already said that a *random variable* X is neither random nor a variable, but is rather a function of an underlying process that does appear to be random. However, because it is a function of a process that does appear random, the *random variable* X also appears to be random. That is, if we don't know the outcome of of the underlying experimental process, then we also don't know for sure what X is, and so X does indeed appear to be random. However, eventhough X appears to be random, the expected value $\mathsf{E}_{\mathsf{X}}\mathsf{X}$ of X is **not random**. Rather it is a fixed value (like 0 or 7.9 or -2.6).

On the other hand, eventhough $\mathsf{E}\mathsf{X}$ is **not random**, note that $\mathsf{E}(\mathsf{X}|\mathsf{Y})$ **is random**. This is because $\mathsf{E}(\mathsf{X}|\mathsf{Y})$ is a function of Y. That is, once we know that Y equals some fixed value y (like 0 or 2.7 or -5.1) then $\mathsf{E}(\mathsf{X}|\mathsf{Y} = y)$ is also fixed. However, if we don't know the value of Y, then Y is still a *random variable* and the expression $\mathsf{E}(\mathsf{X}|\mathsf{Y})$ is also random (a function of *random variable* Y).

Two common statistics that are conveniently expressed in terms of the expectation operator are the *mean* and *variance*. The mean is an indicator of the "middle" of a probability distribution and the variance is an indicator of the "spread".

Definition 1.2. *Let* X *be a* RANDOM VARIABLE *on the* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$.

DEF
(1). *The* ***mean*** μ_{X} *of* X *is* $\quad \mu_{\mathsf{X}} \triangleq \mathsf{E}_{\mathsf{X}}\mathsf{X}$

(2). *The* ***variance*** $\mathsf{Var}(\mathsf{X})$ *or* σ^2_{X} *of* X *is* $\quad \mathsf{Var}(\mathsf{X}) \triangleq \mathsf{E}_{\mathsf{X}}\left[(\mathsf{X} - \mathsf{E}_{\mathsf{X}}\mathsf{X})^2\right]$

1.2 Expectation as a linear operator

The next theorem demonstrates that the operator E is a *linear operator* (Definition O.3 page 320)—which in turn makes E part of a distinguished club of operators along with fellow member operators differentiation $\frac{\mathsf{d}}{\mathsf{d}x}$, integration $\int \mathsf{d}x$, Laplace $\mathbf{L}$, Fourier $\tilde{\mathbf{F}}$, z-transform $\mathbf{Z}$, etc. Because E is a linear operator, it immediately inherits all the properties that it's linear operator birthright grants it (Corollary 1.1 page 4).

Theorem 1.1 (Linearity of E). [1] *Let* X *be a* RANDOM VARIABLE *on a* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$.

THM

$$\mathsf{E}_{\mathsf{x}}(a\mathsf{X} + b\mathsf{Y} + c) = \left(a\mathsf{E}_{\mathsf{x}}\mathsf{X}\right) + \left(b\mathsf{E}_{\mathsf{x}}\mathsf{Y}\right) + c \quad \forall a,b,c \in \mathbb{R} \quad \text{(LINEAR)}$$

✎PROOF:

$$\begin{aligned}
\boxed{\mathsf{E}_{\mathsf{xy}}(a\mathsf{X} + b\mathsf{Y} + c)} &\triangleq \int_{x\in\mathbb{R}} \int_{y\in\mathbb{R}} [ax + by + c]\mathsf{p}_{\mathsf{xy}}(x,y)\,\mathsf{d}y\,\mathsf{d}x \qquad \text{by definition of } \mathsf{E} \text{ (Definition 1.1 page 3)} \\
&= \int_{x\in\mathbb{R}} \int_{y\in\mathbb{R}} ax\mathsf{p}_{\mathsf{xy}}(x,y)\,\mathsf{d}y\,\mathsf{d}x + \int_{x\in\mathbb{R}} \int_{y\in\mathbb{R}} by\mathsf{p}_{\mathsf{xy}}(x,y)\,\mathsf{d}y\,\mathsf{d}x + \int_{x\in\mathbb{R}} \int_{y\in\mathbb{R}} c\mathsf{p}_{\mathsf{xy}}(x,y)\,\mathsf{d}y\,\mathsf{d}x \\
&= \int_{x\in\mathbb{R}} ax \underbrace{\int_{y\in\mathbb{R}} \mathsf{p}_{\mathsf{xy}}(x,y)\,\mathsf{d}y}_{\mathsf{p}_{\mathsf{x}}(x)}\,\mathsf{d}x + \int_{y\in\mathbb{R}} by \underbrace{\int_{x\in\mathbb{R}} \mathsf{p}_{\mathsf{xy}}(x,y)\,\mathsf{d}x}_{\mathsf{p}_{\mathsf{y}}(y)}\,\mathsf{d}y + c \underbrace{\int_{y\in\mathbb{R}} \int_{x\in\mathbb{R}} \mathsf{p}_{\mathsf{xy}}(x,y)\,\mathsf{d}x\,\mathsf{d}y}_{1} \\
&= a \underbrace{\int_{x\in\mathbb{R}} x\mathsf{p}_{\mathsf{x}}(x)\,\mathsf{d}x}_{\mathsf{EX}} + b \underbrace{\int_{y\in\mathbb{R}} y\mathsf{p}_{\mathsf{y}}(y)\,\mathsf{d}y}_{\mathsf{EY}} + c \\
&= \boxed{\left(a\mathsf{E}_{\mathsf{x}}\mathsf{X}\right) + \left(b\mathsf{E}_{\mathsf{y}}\mathsf{Y}\right) + c}
\end{aligned}$$

✏

Corollary 1.1. *Let* E *be the* EXPECTATION OPERATOR *over a* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$. *Let* $spLLF$ *be a* VECTOR SPACE OF RANDOM VARIABLES *over* $(\Omega, \mathbb{E}, \mathsf{P})$.

COR

(1). $\mathsf{E}\mathbb{0} = \mathbb{0}$ *and*

(2). $\mathsf{E}(-\mathsf{X}) = -(\mathsf{EX}) \quad \forall \mathsf{X} \in \boldsymbol{L}^2_{\mathbb{F}}$ *and*

(3). $\mathsf{E}(\mathsf{X} - \mathsf{Y}) = \mathsf{EX} - \mathsf{EY} \quad \forall \mathsf{X},\mathsf{Y} \in \boldsymbol{L}^2_{\mathbb{F}}$ *and*

(4). $\mathsf{E}\left(\sum_{n=1}^{N} \alpha_n \mathsf{X}_n\right) = \sum_{n=1}^{N} \alpha_n \left(\mathsf{EX}_n\right) \quad \forall \alpha_n \in \mathbb{F}, \quad \forall \mathsf{X} \in \boldsymbol{L}^2_{\mathbb{F}}$

✎PROOF: These all follow immediately from the fact that E is a *linear operator* and from Theorem O.1 (page 320). ✏

Remark 1.1. Projecting a stochastic process onto a basis often yields valuable insights into the nature of the underlying data. Typical projection operators include the Fourier operator $\tilde{\mathbf{F}}$, Laplace $\mathbf{L}$, and z-transform $\mathbf{Z}$ ...not to mention wavelet operators. But note that any such projection on a random sequence simply produces another random sequence. For example, the Fourier transform $\tilde{\mathbf{F}}\mathsf{x}(n)$ of a random sequence $\mathsf{x}(n)$ is another random sequence.

One way to overcome this difficulty is to simply invoke a *sampling* operator $\mathbf{S}\mathsf{x}(n)$, yielding a deterministic sequence, and then take the Fourier transform of the resulting deterministic sequence. The problem here is that every time you resample the sequence, you will very likely get a different Fourier transform.

[1] Wilks (1963b), page 73 ⟨§3.2 "Mean value of a random variable"⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

Arguably a better approach (and the standard one at that) is to first invoke the expectation operator $\mathsf{Ex}(n)$, also yielding a deterministic sequence.

The good news here is that because E and all the above mentioned operators are *linear*, we can do all the standard arithmetic acrobatics associated with linear algebra operators (next corollary).

Corollary 1.2. *Let* $\mathbf{M}$ *and* $\mathbf{N}$ *be* LINEAR OPERATORS *(Definition O.3 page 320)*.

COR					
1.	$\mathsf{E}(\mathbf{MN})$	$=$	$(\mathsf{E}\mathbf{M})\mathbf{N}$	$\forall \mathsf{E}\in\mathcal{L}(\boldsymbol{Z},\boldsymbol{W}), \mathbf{M}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{N}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(ASSOCIATIVE)
2.	$\mathsf{E}(\mathbf{M}+\mathbf{N})$	$=$	$(\mathsf{E}\mathbf{M})+(\mathsf{E}\mathbf{N})$	$\forall \mathsf{E}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{M}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), \mathbf{N}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(LEFT DISTRIBUTIVE)
3.	$(\mathsf{E}+\mathbf{M})\mathbf{N}$	$=$	$(\mathsf{E}\mathbf{N})+(\mathbf{MN})$	$\forall \mathsf{E}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{M}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{N}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(RIGHT DISTRIBUTIVE)
4.	$\alpha(\mathsf{E}\mathbf{M})$	$=$	$(\alpha\mathsf{E})\mathbf{M}=\mathsf{E}(\alpha\mathbf{M})$	$\forall \mathsf{E}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{M}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), \alpha\in\mathbb{F}$	(HOMOGENEOUS)

✎PROOF: These all follow immediately from the fact that E is a *linear operator* (Theorem 1.1 page 4) and from properities of all linear operators (Theorem O.4 page 323). ✏

Corollary 1.3. *Let* X *be a* RANDOM VARIABLE *on a* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$.

COR				
	$\mathsf{Var}(a\mathsf{X}+b)$	$=$	$a^2\mathsf{Var}(\mathsf{X})$	$\forall a,b\in\mathbb{R}$
	$\mathsf{Var}(\mathsf{X})$	$=$	$\mathsf{E}_{\mathsf{x}}(\mathsf{X}^2)-(\mathsf{E}_{\mathsf{x}}\mathsf{X})^2$	

✎PROOF:

$$
\begin{aligned}
\mathsf{Var}(\mathsf{X}) &\triangleq \mathsf{E}_{\mathsf{x}}\left[(\mathsf{X}-\mathsf{E}_{\mathsf{x}}\mathsf{X})^2\right] && \text{by definition of } \mathsf{Var} && \text{(Definition 1.2 page 3)}\\
&= \mathsf{E}_{\mathsf{x}}\left[\mathsf{X}^2-2\mathsf{X}\mathsf{E}_{\mathsf{x}}\mathsf{X}+(\mathsf{E}_{\mathsf{x}}\mathsf{X})^2\right] && \text{by } \textit{Binomial Theorem} && \\
&= \mathsf{E}_{\mathsf{x}}\mathsf{X}^2-\mathsf{E}_{\mathsf{x}}\left[2\mathsf{X}\mathsf{E}_{\mathsf{x}}\mathsf{X}\right]+\mathsf{E}_{\mathsf{x}}(\mathsf{E}_{\mathsf{x}}\mathsf{X})^2 && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)}\\
&= \mathsf{E}_{\mathsf{x}}\mathsf{X}^2-2(\mathsf{E}_{\mathsf{x}}\mathsf{X})[\mathsf{E}_{\mathsf{x}}\mathsf{X}]+(\mathsf{E}_{\mathsf{x}}\mathsf{X})^2 && && \\
&= \mathsf{E}_{\mathsf{x}}(\mathsf{X}^2)-(\mathsf{E}_{\mathsf{x}}\mathsf{X})^2 && && \\
\mathsf{Var}(a\mathsf{X}+b) &= \mathsf{E}_{\mathsf{x}}(a\mathsf{X}+b)^2-[\mathsf{E}_{\mathsf{x}}(a\mathsf{X}+b)]^2 && && \\
&= \mathsf{E}_{\mathsf{x}}(a^2\mathsf{X}^2+2ab\mathsf{X}+b^2)-[a(\mathsf{E}_{\mathsf{x}}\mathsf{X})+b]^2 && && \\
&= a^2\mathsf{E}_{\mathsf{x}}\mathsf{X}^2+2ab\mathsf{E}_{\mathsf{x}}\mathsf{X}+b^2-\left[a^2[\mathsf{E}_{\mathsf{x}}\mathsf{X}]^2+2ab\mathsf{E}_{\mathsf{x}}\mathsf{X}+b^2\right] && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)}\\
&= a^2\left[\mathsf{E}_{\mathsf{x}}\mathsf{X}^2-(\mathsf{E}_{\mathsf{x}}\mathsf{X})^2\right] && && \\
&\triangleq a^2\mathsf{Var}(\mathsf{X}) && \text{by previous result} &&
\end{aligned}
$$

✏

Figure 1.1: Jensen's inequality

Jensen's inequality is an extremely useful application of *convex*ity (Definition M.9 page 300) to the *expectation* operator. Jensen's inequality is stated in Corollary 1.4 (next) and illustrated in Figure 1.1 (page 5).

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Corollary 1.4 (Jensen's inequality). [2] *Let* f *be a function in* $\mathbb{R}^{\mathbb{R}}$ *and* X *be a* RANDOM VARIABLE *on* $(\Omega, \mathbb{E}, \mathsf{P})$.

COR $\{\mathsf{f}$ *is* CONVEX$\} \quad \Longrightarrow \quad \{\mathsf{f}(\mathsf{EX}) \leq \mathsf{Ef}(\mathsf{X})\}$

✎PROOF:

1. Proof 1: Let $mx + c$ be a "support line" under $\mathsf{f}(x)$ (Figure 1.1 page 5) such that

$$\begin{aligned} mx + c &< \mathsf{f}(x) \quad \text{for } x \neq \mathsf{EX} \\ mx + c &= \mathsf{f}(x) \quad \text{for } x = \mathsf{EX}. \end{aligned}$$

   Then

$$\begin{aligned} \mathsf{f}(\mathsf{EX}) &= m[\mathsf{EX}] + c \\ &= \mathsf{E}[mX + c] \\ &\leq \mathsf{Ef}(\mathsf{X}) \end{aligned}$$

2. Proof 2 (alternate proof):

$$\begin{aligned} \mathsf{f}(\mathsf{EX}) &\triangleq \mathsf{f}\left(\sum_{x\in\mathbb{E}} x\mathsf{P}(x)\right) \\ &\leq \sum_{x\in\mathbb{E}} \mathsf{f}(x)\mathsf{P}(x) \qquad \text{by } \textit{Jensen's inequality} \text{ for convex sets} \qquad \text{(Theorem M.1 page 300)} \end{aligned}$$

✏

Example 1.1. [3] Some examples of *Jensen's Inequality* (Corollary 1.4 page 6) applied to the *expectation operator* are the following:

EX $(\mathsf{EX})^{-1} < \mathsf{E}\left(\mathsf{X}^{-1}\right) \;\Big|\; \mathsf{E}\left(\log \mathsf{X}\right) < \log(\mathsf{EX}) \;\Big|\; e^{-\mathsf{EX}} \leq \mathsf{E}\left[e^{-\mathsf{X}}\right]$

Theorem 1.2 (Law of the Unconscious Statistician). [4]

THM $\mathsf{E}\left[\mathsf{g}(\mathsf{X})\right] = \int_{x\in\mathbb{R}} \mathsf{g}(x)\mathsf{p}_{\mathsf{X}}(x)\,\mathrm{d}x$

1.3 Expectation inequalities

Theorem 1.3 (Markov's inequality). [5] *Let* $\mathsf{X} : \Omega \to [0, \infty)$ *be a non-negative valued* RANDOM VARIABLE *and* $a \in (0, \infty)$. *Then*

THM $\mathsf{P}\{\mathsf{X} \geq a\} \leq \frac{1}{a}\mathsf{EX}$

[2] Shao (2003) page 31 ⟨"1.3 Distributions and Their Characteristics"⟩, Cover and Thomas (1991), page 25, Jensen (1906), pages 179–180

[3] Shao (2003) pages 31–32 ⟨"Example 1.18"⟩, Dekking et al. (2006) page 110 ⟨"8.5 Solutions to the quick exercises"⟩

[4] Suhov et al. (2005) page 145 ⟨(2.69)⟩, Allen (2018) page 490 ⟨18.3.4 The Law of the Unconscious Statistician⟩, Papoulis (1990) page 124 ⟨Fundamental Theorem⟩

[5] Ross (1998), page 395

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$
\begin{aligned}
I &\triangleq \begin{cases} 1 & \text{for} \quad \mathsf{X} \geq a \\ 0 & \text{for} \quad \mathsf{X} < a \end{cases} \\
aI &\leq \mathsf{X} \\
I &\leq \frac{1}{a}\mathsf{X} \\
\mathsf{E}I &\leq \mathsf{E}\left(\frac{1}{a}\mathsf{X}\right)
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{P}\{\mathsf{X} \geq a\} &= 1 \cdot \mathsf{P}\{\mathsf{X} \geq a\} + 0 \cdot \mathsf{P}\{\mathsf{X} < a\} \\
&= \mathsf{E}I \\
&\leq \mathsf{E}\left(\frac{1}{a}\mathsf{X}\right) \\
&= \frac{1}{a}\mathsf{E}\mathsf{X}
\end{aligned}
$$

✏

Theorem 1.4 (Chebyshev's inequality)**.** [6] *Let* X *be a* RANDOM VARIABLE *with mean* μ *and variance* σ^2.

THM $\mathsf{P}\{|\mathsf{X} - \mu| \geq a\} \leq \dfrac{\sigma^2}{a^2}$

✎PROOF:

$$
\begin{aligned}
\mathsf{P}\{|\mathsf{X} - \mu| \geq a\} &= \mathsf{P}\left\{(\mathsf{X} - \mu)^2 \geq a^2\right\} \\
&\leq \frac{1}{a^2}\mathsf{E}(\mathsf{X} - \mu)^2 && \text{by } \textit{Markov's inequality} && \text{(Theorem 1.3 page 6)} \\
&= \frac{\sigma^2}{a^2}
\end{aligned}
$$

✏

Theorem 1.5 (Kolmogorov's inequality)**.** [7] *Let* X *be a* RANDOM VARIABLE *with mean* μ *and variance* σ^2.

THM $\left\{\begin{array}{lll} (A). & (\mathsf{x}_n) \textit{ are } \text{INDEPENDENT} & \textit{and} \\ (B). & \textit{Each } \mathsf{x}_n \textit{ has } \text{ZERO-MEAN} & \textit{and} \\ (C). & \textit{Each } \mathsf{x}_n \textit{ has variance } \sigma^2 & \end{array}\right\} \Longrightarrow \mathsf{P}\left[\left|\sum_{n=1}^{N} \mathsf{x}_n\right| < \lambda \sum_{n=1}^{N} \mathsf{x}_n^2\right] \geq 1 - \frac{1}{\lambda^2}$

1.4 Joint and conditional probability spaces

Sometimes the problem of finding the expected value of a *random variable* X can be simplified by "conditioning X on Y".

Theorem 1.6. *Let* X *and* Y *be* RANDOM VARIABLES*. Then*

THM $\mathsf{E}_{\mathsf{x}}\mathsf{X} = \mathsf{E}_{\mathsf{y}}\mathsf{E}_{\mathsf{x}|\mathsf{y}}(\mathsf{X}|\mathsf{Y})$

[6] Ross (1998), page 396
[7] Wilks (1963b), page 107 ⟨§4.5 "Kolmogorov's inequality"⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✎PROOF:

$$
\begin{aligned}
\mathsf{E}_{\mathsf{y}}\mathsf{E}_{\mathsf{x}|\mathsf{y}}(\mathsf{X}|\mathsf{Y}) &\triangleq \mathsf{E}_{\mathsf{y}}\left[\int_{x\in\mathbb{R}} x\mathsf{p}(\mathsf{X}=x|\mathsf{Y})\,\mathsf{d}x\right]\\
&\triangleq \int_{y\in\mathbb{R}}\left[\int_{x\in\mathbb{R}} x\mathsf{p}(x|\mathsf{Y}=y)\,\mathsf{d}x\right]\mathsf{p}(y)\,\mathsf{d}y\\
&= \int_{y\in\mathbb{R}}\int_{x\in\mathbb{R}} x\mathsf{p}(x|y)\mathsf{p}(y)\,\mathsf{d}x\,\mathsf{d}y\\
&= \int_{x\in\mathbb{R}} x\int_{y\in\mathbb{R}} \mathsf{p}(x,y)\,\mathsf{d}y\,\mathsf{d}x\\
&= \int_{x\in\mathbb{R}} x\mathsf{p}(x)\,\mathsf{d}x\\
&\triangleq \mathsf{E}_{\mathsf{x}}\mathsf{X}
\end{aligned}
$$

✏

1.5 Expectation inner product space

When possible, we like to generalize any given mathematical structure to a more general mathematical structure and then take advantage of the properties of that more general structure. Such a generalization can be done with *random variable*s. Random variables can be viewed as vectors in a vector space. Furthermore, the expectation of the product of two *random variable*s (e.g. $\mathsf{E}(\mathsf{XY})$) can be viewed as an *inner product* in an *inner product space*. Since we have an *inner product* space, we can then immediately use all the properties of *inner product space*s, *norm*ed spaces, vector spaces, metric spaces, and topological spaces.

Theorem 1.7. [8] *Let R be a ring, $(\Omega,\mathbb{E},\mathsf{P})$ be a* PROBABILITY SPACE, E *the expectation operator, and $\boldsymbol{V}=\{\mathsf{X}|\mathsf{X}:\Omega\to R\}$ be the set of all random vectors in* PROBABILITY SPACE $(\Omega,\mathbb{E},\mathsf{P})$.

THM			
	(1).	$\boldsymbol{V}\triangleq\{\mathsf{X}\|\mathsf{X}:\Omega\to R\}$	*is a* VECTOR SPACE.
	(2).	$\langle\mathsf{X}\,\|\,\mathsf{Y}\rangle\triangleq\mathsf{E}(\mathsf{XY}^*)$	*is an* INNER PRODUCT.
	(3).	$\|\mathsf{X}\|\triangleq\sqrt{\mathsf{E}(\mathsf{XX}^*)}$	*is a* NORM.
	(4).	$(\boldsymbol{V},\langle\triangle\,\|\,\triangledown\rangle)$	*is an* INNER PRODUCT SPACE.

✎PROOF:

1. Proof that $\boldsymbol{V}$ is a vector space:

1)	$\forall\mathsf{X},\mathsf{Y},\mathsf{Z}\in\boldsymbol{V}$	$(\mathsf{X}+\mathsf{Y})+\mathsf{Z}=\mathsf{X}+(\mathsf{Y}+\mathsf{Z})$	(+ is associative)
2)	$\forall\mathsf{X},\mathsf{Y}\in\boldsymbol{V}$	$\mathsf{X}+\mathsf{Y}=\mathsf{Y}+\mathsf{X}$	(+ is commutative)
3)	$\exists 0\in\boldsymbol{V}$ such that $\forall\mathsf{X}\in\boldsymbol{V}$	$\mathsf{X}+0=\mathsf{X}$	(+ identity)
4)	$\forall\mathsf{X}\in\boldsymbol{V}\exists\mathsf{Y}\in\boldsymbol{V}$ such that	$\mathsf{X}+\mathsf{Y}=0$	(+ inverse)
5)	$\forall\alpha\in S$ and $\mathsf{X},\mathsf{Y}\in\boldsymbol{V}$	$\alpha\cdot(\mathsf{X}+\mathsf{Y})=(\alpha\cdot\mathsf{X})+(\alpha\cdot\mathsf{Y})$	($\cdot$ distributes over +)
6)	$\forall\alpha,\beta\in S$ and $\mathsf{X}\in\boldsymbol{V}$	$(\alpha+\beta)\cdot\mathsf{X}=(\alpha\cdot\mathsf{X})+(\beta\cdot\mathsf{X})$	($\cdot$ pseudo-distributes over +)
7)	$\forall\alpha,\beta\in S$ and $\mathsf{X}\in\boldsymbol{V}$	$\alpha(\beta\cdot\mathsf{X})=(\alpha\cdot\beta)\cdot\mathsf{X}$	($\cdot$ associates with $\cdot$)
8)	$\forall\mathsf{X}\in\boldsymbol{V}$	$1\cdot\mathsf{X}=\mathsf{X}$	($\cdot$ identity)

[8] Lindquist and Picci (2015) pages 25–26 ⟨2.1 Hilbert Space of Second-Order Random Variables. $\langle\xi\,|\,\eta\rangle=\mathsf{E}\{\xi\bar{\eta}\}$⟩, Caines (1988) page 21 ⟨$Exy=\int_\Omega x(\omega)y(\omega)dP(\omega)$⟩, Caines (2018) page 21 ⟨$Exy=\int_\Omega x(\omega)y(\omega)dP(\omega)$⟩, Moon and Stirling (2000), pages 105–106

COPYRIGHT © 2019 DANIEL J. GREENHOE

2. Proof that $\langle \mathsf{X} \mid \mathsf{Y} \rangle \triangleq \mathsf{E}(\mathsf{X}\mathsf{Y}^*)$ is an *inner product*.

1)	$\mathsf{E}(\mathsf{X}\mathsf{X}^*)$	≥ 0	$\forall \mathsf{X} \in \boldsymbol{V}$	(non-negative)
2)	$\mathsf{E}(\mathsf{X}\mathsf{X}^*)$	$= 0 \iff \mathsf{X} = 0$	$\forall \mathsf{X} \in \boldsymbol{V}$	(non-degenerate)
3)	$\mathsf{E}(\alpha \mathsf{X}\mathsf{Y}^*)$	$= \alpha \mathsf{E}(\mathsf{X}\mathsf{Y}^*)$	$\forall \mathsf{X}, \mathsf{Y} \in \boldsymbol{V},\ \forall \alpha \in \mathbb{C}$	(homogeneous)
4)	$\mathsf{E}[(\mathsf{X}+\mathsf{Y})\mathsf{Z}^*]$	$= \mathsf{E}(\mathsf{X}\mathsf{Z}^*) + \mathsf{E}(Y\mathsf{Z}^*)$	$\forall \mathsf{X}, \mathsf{Y}, \mathsf{Z} \in \boldsymbol{V}$	(additive)
5)	$\mathsf{E}(\mathsf{X}\mathsf{Y}^*)$	$= \mathsf{E}(\mathsf{Y}\mathsf{X}^*)$	$\forall \mathsf{X}, \mathsf{Y} \in \boldsymbol{V}$	(conjugate symmetric).

3. Proof that $\|\mathsf{X}\| \triangleq \sqrt{\mathsf{E}(\mathsf{X}\mathsf{X}^*)}$ is a *norm*: This *norm* is simply induced by the above *inner product*.

4. Proof that $(\boldsymbol{V}, \langle \triangle \mid \triangledown \rangle)$ is an *inner product space*: Because $\boldsymbol{V}$ is a vector space and $\langle \triangle \mid \triangledown \rangle$ is an *inner product*, $(\boldsymbol{V}, \langle \triangle \mid \triangledown \rangle)$ is an *inner product space*.

✏

The next theorem gives some results that follow directly from vector space properties:

Theorem 1.8. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE *with* EXPECTATION *functional* E.

THM

(1). $$\sqrt{\mathsf{E}\left(\sum_{n=1}^{N} \mathsf{X}_n\right)} \leq \sum_{n=1}^{N} \mathsf{E}(\mathsf{X}_n X_n^*)$$ *(*GENERALIZED TRIANGLE INEQUALITY*)*

(2). $$\left|\mathsf{E}(\mathsf{X}\mathsf{Y}^*)\right|^2 \leq \mathsf{E}(\mathsf{X}\mathsf{X}^*)\,\mathsf{E}(Y Y^*)$$ *(*CAUCHY-SCHWARTZ INEQUALITY*)*

(3). $$2\mathsf{E}(\mathsf{X}\mathsf{X}^*) + 2\mathsf{E}(Y Y^*) = \mathsf{E}[(\mathsf{X}+\mathsf{Y})(\mathsf{X}+\mathsf{Y})^*] + \mathsf{E}[(\mathsf{X}-Y)(\mathsf{X}-Y)^*]$$ (PARALLELOGRAM LAW)

✎PROOF:

1. $\left(\mathbb{R}^{\Omega}, \mathsf{E}(\mathsf{x},\mathsf{y})\right)$ is an *inner product space*. Proof: Theorem 1.7 (page 8).

2. Because $\left(\mathbb{R}^{\Omega}, \mathsf{E}(\mathsf{x},\mathsf{y})\right)$ is an *inner product space*, the other properties follow:
   1. *Generalized triangle inequality*: Theorem L.1 page 287
   2. *Cauchy-Schwartz inequality*: Theorem K.2 page 272
   3. *Parallelogram Law*: Theorem K.7 page 279

✏

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 2

RANDOM SEQUENCES

❝ *A likely impossibility is always preferable to an unconvincing possibility.*❞
Aristotle (384 BC – 322 BC) [1]

❝ *We are quite in danger of sending highly trained and highly intelligent young men out into the world with tables of erroneous numbers under their arms, and with a dense fog in the place where their brains ought to be. In this century, of course, they will be working on guided missiles and advising the medical profession on the control of disease, and there is no limit to the extent to which they could impede every sort of national effort.*❞

Ronald A. Fisher, (1890–1962), Statistician, at a lecture in 1958 at Michigan State University [2]

2.1 Definitions

Definition 2.1.

DEF

A ***random sequence*** *is a* SEQUENCE *over a* PROBABILITY SPACE *(Definition A.2 page 167)*.

Definition 2.2. [3] *Let* $\mathsf{x}(n)$ *and* $\mathsf{y}(n)$ *be* RANDOM SEQUENCES.

[1] quote: http://en.wikiquote.org/wiki/Aristotle
image: http://en.wikipedia.org/wiki/Aristotle

[2] quote: Yates and Mather (1963) page 107. image: http://www.genetics.org/content/154/4/1419

[3] Papoulis (1984) page 263 $\langle R_{xy}(m) = E\{\mathsf{x}(m)\mathsf{y}^*(0)\}\rangle$, Wilks (1963b), page 77 ⟨§3.4 "Moments of two-dimensional random variables"⟩, Cadzow (1987) page 341 $\langle r_{xy}(m) = E[\mathsf{x}(m)\mathsf{y}^*(0)]\rangle$, MatLab (2018b) $\langle R_{xy}(m) = E\{x_{n+m}y_n^*\}\rangle$, MatLab (2018a) $\langle R_{xy}(m) = E\{x_{n+m}y_n^*\}\rangle$

DEF

The ***mean***	$\mu_{\mathsf{X}}(n)$	*of* $\mathsf{x}(n)$ *is*	$\mu_{\mathsf{X}}(n)$	$\triangleq$	$\mathsf{E}[\mathsf{x}(n)]$
The ***variance***	$\sigma^2_{\mathsf{X}}(n)$	*of* $\mathsf{x}(n)$ *is*	$\sigma^2_{\mathsf{X}}(n)$	$\triangleq$	$\mathsf{E}\Big(\big[\mathsf{x}(n)-\mu_{\mathsf{X}}(n)\big]^2\Big)$
The ***cross-correlation***	$\mathsf{R}_{\mathsf{xy}}(n,m)$	*of* $\mathsf{x}(n)$ *and* $\mathsf{y}(n)$ *is*	$\mathsf{R}_{\mathsf{xy}}(n,m)$	$\triangleq$	$\mathsf{E}\big[\mathsf{x}(n+m)\mathsf{y}^*(n)\big]$
The ***auto-correlation***	$\mathsf{R}_{\mathsf{xx}}(n,m)$	*of* $\mathsf{x}(n)$ *is*	$\mathsf{R}_{\mathsf{xx}}(n,m)$	$\triangleq$	$\mathsf{R}_{\mathsf{xy}}(n,m)\big\vert_{\mathsf{y}=\mathsf{x}}$

2.2 Properties

Theorem 2.1.

THM

$$\mathsf{R}_{\mathsf{xx}}(n,m) = \mathsf{R}^*_{\mathsf{xx}}(n+m,-m)$$
$$\mathsf{R}_{\mathsf{xy}}(n,m) = \mathsf{R}^*_{\mathsf{yx}}(n+m,-m)$$

✎PROOF:

$\mathsf{R}_{\mathsf{xy}}(n,m) \triangleq \mathsf{E}\big[\mathsf{x}(n+m)\mathsf{y}^*(n)\big]$	by definition of $\mathsf{R}_{\mathsf{xy}}(n,m)$	(Definition 2.2 page 11)
$= \mathsf{E}\big[\mathsf{y}^*(n)\mathsf{x}(n+m)\big]$	by *commutative* property of $(\mathbb{C},+,\cdot,0,1)$	(Definition F.5 page 208)
$= \big(\mathsf{E}\big[\mathsf{y}(n)\mathsf{x}^*(n+m)\big]\big)^*$	by *distributive* property of $*$-**algebras**	(Definition J.3 page 266)
$= \big(\mathsf{E}\big[\mathsf{y}(n+m-m)\mathsf{x}^*(n+m)\big]\big)^*$	by *additive identity* property of $(\mathbb{R},+,\cdot,0,1)$	(Definition F.5 page 208)
$\triangleq \mathsf{R}^*_{\mathsf{yx}}(n+m,-m)$	by definition of $\mathsf{R}_{\mathsf{xy}}(n,m)$	(Definition 2.2 page 11)

$\mathsf{R}_{\mathsf{xx}}(n,m) = \mathsf{R}_{\mathsf{xy}}(n,m)\big\vert_{\mathsf{y}=\mathsf{x}}$	by $\mathsf{y}=\mathsf{x}$ constraint
$= \mathsf{R}^*_{\mathsf{xy}}(n+m,-m)\big\vert_{\mathsf{y}=\mathsf{x}}$	by previous result
$= \mathsf{R}^*_{\mathsf{xx}}(n+m,-m)$	by $\mathsf{y}=\mathsf{x}$ constraint

✏

2.3 Wide Sense Stationary processes

Definition 2.3. *Let* $\mathsf{x}(n)$ *be a* RANDOM SEQUENCE *with* MEAN $\mu_{\mathsf{X}}(n)$ *and* VARIANCE $\sigma^2(n)$ *(Definition 2.2 page 11)*.

DEF

$\mathsf{x}(n)$ *is* ***wide sense stationary (WSS)*** *if*

1. $\mu_{\mathsf{X}}(n)$ *is* CONSTANT *with respect to* n (STATIONARY IN THE 1ST MOMENT) *and*
2. $\sigma^2_{\mathsf{X}}(n)$ *is* CONSTANT *with respect to* n (STATIONARY IN THE 2ND MOMENT)

Definition 2.4. [4] *Let* $\mathsf{x}(n)$ *be a* RANDOM SEQUENCE *with statistics* $\mu_{\mathsf{X}}(n)$, $\sigma^2_{\mathsf{X}}(n)$, $\mathsf{R}_{\mathsf{xx}}(n,m)$, *and* $\mathsf{R}_{\mathsf{xy}}(n,m)$ *(Definition 2.2 page 11)*.

DEF

$\{$ x *and* y *are* WIDE SENSE STATIONARY $\}$ $\Longrightarrow$

(1).	*The* ***mean***	μ_{X}	*of* $\mathsf{x}(n)$ *is*	μ_{X}	$\triangleq$	$\mathsf{E}[\mathsf{x}(0)]$
(2).	*The* ***variance***	σ^2_{X}	*of* $\mathsf{x}(n)$ *is*	σ^2_{X}	$\triangleq$	$\mathsf{E}\Big(\big[\mathsf{x}(0)-\mu_{\mathsf{X}}\big]^2\Big)$
(4).	*The* ***cross-correlation***	$\mathsf{R}_{\mathsf{xy}}(m)$	*of* $\mathsf{x}(n)$ *and* $\mathsf{y}(n)$ *is*	$\mathsf{R}_{\mathsf{xy}}(m)$	$\triangleq$	$\mathsf{E}\big[\mathsf{x}(m)\mathsf{y}^*(0)\big]$
(3).	*The* ***auto-correlation***	$\mathsf{R}_{\mathsf{xx}}(m)$	*of* $\mathsf{x}(n)$ *is*	$\mathsf{R}_{\mathsf{xx}}(m)$	$\triangleq$	$\mathsf{R}_{\mathsf{xy}}(m)\big\vert_{\mathsf{y}=\mathsf{x}}$

[4] Papoulis (1984) page 263 ($\langle$"$R_{xy}(\tau) = E\big\{\mathsf{x}(t+\tau)\mathsf{y}^*(t)\big\}$"$\rangle$, Cadzow (1987) page 341 $\langle r_{xy}(n) = E\big[\mathsf{x}(k+n)\mathsf{y}^*(k)\big]$ (10.41)$\rangle$

COPYRIGHT © 2019 DANIEL J. GREENHOE

Remark 2.1. The $\mathsf{R}_{\mathsf{xy}}(n,m)$ of Definition 2.2 (page 11) and the $\mathsf{R}_{\mathsf{xy}}(m)$ of Definition 2.4 (page 12) (etc.) are examples of *function overload*—that is, functions that use the same mnemonic but are distinguished by different domains. Perhaps a more common example of function overload is the "+" mnemonic. Traditionally it is used with domain of the natural numbers $\mathbb{N}$ as in $3+2$. Later it was extended for domain real numbers $\mathbb{R}$ as in $\sqrt{3}+\sqrt{2}$, or even complex numbers $\mathbb{C}$ as in $\left(\sqrt{3}+i\sqrt{2}\right)+(e+i\pi)$. And it was even more dramatically extended for use with domain $\mathbb{R}^N \times \mathbb{R}^M$ in "linear algebra" as in

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

Remark 2.2. The definition for $\mathsf{R}_{\mathsf{xy}}(m)$ can be defined with the conjugate $*$ on either x or y, or on neither or both; and moreover x may either lead or lag y. In total, there are $2\times 2\times 2 = 8$ different ways to define $\mathsf{R}_{\mathsf{xy}}(m)$.[5] and $\mathsf{R}_{\mathsf{xx}}(m)$ involve complex numbers. This may seem curious when typical ADCs provide real-valued sequences. Note however that complex-valued sequences often come up in signal processing due to two common situations:[6]

1. The presence of an *FFT* operator in the signal processing path
2. communications channel processing involving phase discrimination (e.g. PSK and QAM).

Proposition 2.1. *Let* $\mathsf{y}(n)$ *be a* RANDOM SEQUENCE, $\mathsf{x}(n)$ *a* RANDOM SEQUENCE *with* AUTO-CORRELATION $\mathsf{R}_{\mathsf{xx}}(n,m)$, *and* R_{xy} *the* CROSS-CORRELATION *of* x *and* y.

PRP

$$\left\{\begin{array}{l} \mathsf{x} \text{ and } \mathsf{y} \text{ are} \\ \text{WIDE SENSE STATIONARY} \\ \text{(WSS) (Definition 7.1 page 45)} \end{array}\right\} \Longrightarrow \left\{\begin{array}{lcll} \mathsf{R}_{\mathsf{xx}}(n,m) & = & \mathsf{R}_{\mathsf{xx}}(m) & \forall n\in\mathbb{Z} \\ \mathsf{R}_{\mathsf{xy}}(n,m) & = & \mathsf{R}_{\mathsf{xy}}(m) & \forall n\in\mathbb{Z} \\ \text{(Definition 2.2 page 11)} & & \text{(Definition 2.4 page 12)} & \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{R}_{\mathsf{xy}}(n,m) &\triangleq \mathsf{E}\left[\mathsf{x}[n+m]\mathsf{y}^*[n]\right] && \text{by definition of } \mathsf{R}_{\mathsf{xy}}(n,m) && \text{(Definition 2.2 page 11)} \\
&= \mathsf{E}\left[\mathsf{x}[n-n+m]\mathsf{y}^*[n-n]\right] && \text{by } \textit{wide sense stationary} \text{ hypothesis} \\
&= \mathsf{E}\left[\mathsf{x}[m]\mathsf{y}^*[0]\right] \\
&\triangleq \mathsf{R}_{\mathsf{xy}}(m) && \text{by definition of } \mathsf{R}_{\mathsf{xy}}(m) && \text{(Definition 2.4 page 12)} \\
\mathsf{R}_{\mathsf{xx}}(n,m) &= \mathsf{R}_{\mathsf{xy}}(n,m)\big|_{\mathsf{y}=\mathsf{x}} \\
&= \mathsf{R}_{\mathsf{xy}}(m)\big|_{\mathsf{y}=\mathsf{x}} && \text{by previous result} \\
&= \mathsf{R}_{\mathsf{xx}}(m)
\end{aligned}$$

✏

Figure 2.1: *auto-correlation* $\mathsf{R}_{\mathsf{xx}}(m)$

[5] Greenhoe (2019)

[6] S. Lawrence Marple (1987) page ?, S. Lawrence Marple (2019) pages 48–50 ⟨§"2.12 Extra: Source of Complex-Valued Signals"⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

Corollary 2.1. *Let* $\mathsf{x}(n)$ *be a* RANDOM SEQUENCE *with* AUTO-CORRELATION $\mathsf{R}_{\mathsf{xx}}(n,m)$, $\mathsf{y}(n)$ *a* RANDOM SEQUENCE *with* AUTO-CORRELATION $\mathsf{R}_{\mathsf{yy}}(n,m)$, *and* $\mathsf{R}_{\mathsf{xy}}(n,m)$ *the* CROSS-CORRELATION *of* x *and* y. *Let* $\mathbf{S}$ *be a* SYSTEM *with input* $\mathsf{x}(n)$ *and output* $\mathsf{y}(n)$.

COR

$$\left\{\begin{array}{lll} (A). & \mathsf{x} \text{ is WSS} & \text{and} \\ (B). & \mathsf{y} \text{ is WSS} & \text{and} \\ (C). & \mathbf{S} \text{ is LTI} & \end{array}\right\} \Longrightarrow \left\{\begin{array}{llllll} (1). & \mathsf{R}_{\mathsf{xy}}(m) & = & \mathsf{R}^*_{\mathsf{yx}}(-m) & & \text{and} \\ (2). & \mathsf{R}_{\mathsf{xx}}(m) & = & \mathsf{R}^*_{\mathsf{xx}}(-m) & \text{(CONJUGATE SYMMETRIC)} & \text{and} \\ (3). & \mathbf{R}_{\mathsf{e}}\mathsf{R}_{\mathsf{xx}}(m) & = & \mathbf{R}_{\mathsf{e}}\mathsf{R}_{\mathsf{xx}}(-m) & \text{(SYMMETRIC)} & \text{and} \\ (4). & \mathbf{I}_{\mathsf{m}}\mathsf{R}_{\mathsf{xx}}(m) & = & -\mathbf{I}_{\mathsf{m}}\mathsf{R}_{\mathsf{xx}}(-m) & \text{(ANTI-SYMMETRIC)} & \text{and} \\ (5). & \left|\mathsf{R}_{\mathsf{xx}}(m)\right| & = & \left|\mathsf{R}_{\mathsf{xx}}(-m)\right| & \text{(SYMMETRIC)} & \text{and} \\ (6). & \angle\mathsf{R}_{\mathsf{xx}}(m) & = & -\angle\mathsf{R}_{\mathsf{xx}}(-m) & \text{(ANTI-SYMMETRIC)} & \end{array}\right\}$$

✎PROOF:

$$\begin{array}{rlll} \mathsf{R}_{\mathsf{xy}}(m) & = \mathsf{R}_{\mathsf{xy}}(n,m) & \text{by Proposition 2.1 page 13} & \text{and hypotheses (A),(B)} \\ & = \mathsf{R}^*_{\mathsf{yx}}(n+m,-m) & \text{by Theorem 2.1 page 12} & \text{and hypothesis (B)} \\ & = \mathsf{R}^*_{\mathsf{yx}}(-m) & \text{by Proposition 2.1 page 13} & \text{and hypothesis (A)} \\ \mathsf{R}_{\mathsf{xx}}(m) & = \mathsf{R}_{\mathsf{xx}}(n,m) & \text{by Proposition 2.1 page 13} & \text{and hypothesis (A)} \\ & = \mathsf{R}^*_{\mathsf{xx}}(n+m,-m) & \text{by Theorem 2.1 page 12} & \text{and hypothesis (B)} \\ & = \mathsf{R}^*_{\mathsf{xx}}(-m) & \text{by Proposition 2.1 page 13} & \text{and hypothesis (A)} \end{array}$$

✏

2.4 Spectral density

Definition 2.5. *Let* $\mathsf{x}(n)$ *and* $\mathsf{y}(n)$ *be* WIDE SENSE STATIONARY RANDOM SEQUENCES *with auto-correlation* $\mathsf{R}_{\mathsf{xx}}(m)$ *and cross-correlation* $\mathsf{R}_{\mathsf{xy}}(m)$. *Let* $\mathbf{Z}$ *be the* Z-TRANSFORM OPERATOR *(Definition R.4 page 370)*.

DEF

The ***z-domain cross spectral density*** ***(CSD)*** $\check{\mathsf{S}}_{\mathsf{xy}}(z)$ *of* $\mathbf{x}$ *and* $\mathbf{y}$ *is*

$$\check{\mathsf{S}}_{\mathsf{xy}}(z) \triangleq \mathbf{Z}\mathsf{R}_{\mathsf{xy}}(m) \triangleq \sum_{m\in\mathbb{Z}} \mathsf{R}_{\mathsf{xy}}(m)z^{-m}$$

The ***z-domain power spectral density*** ***(PSD)*** $\check{\mathsf{S}}_{\mathsf{xx}}(z)$ *of* $\mathbf{x}$ *is* $\quad \check{\mathsf{S}}_{\mathsf{xx}}(z) \triangleq \check{\mathsf{S}}_{\mathsf{xy}}(z)\big|_{\mathsf{y}(n)=\mathsf{x}(n)}$

Definition 2.6. *Let* $\mathsf{x}(n)$ *and* $\mathsf{y}(n)$ *be* WIDE SENSE STATIONARY RANDOM SEQUENCES *with auto-correlation* $\mathsf{R}_{\mathsf{xx}}(m)$ *and cross-correlation* $\mathsf{R}_{\mathsf{xy}}(m)$. *Let* $\check{\mathbf{F}}$ *be the* DISCRETE TIME FOURIER TRANSFORM *(DTFT) operator (Definition Q.1 page 359)*.

DEF

The ***auto-spectral density*** $\check{\mathsf{S}}_{\mathsf{xx}}(z)$ *of* $\mathbf{x}$ *is*

$$\tilde{\mathsf{S}}_{\mathsf{xx}}(z) \triangleq \check{\mathbf{F}}\mathsf{R}_{\mathsf{xx}}(m) \triangleq \sum_{m\in\mathbb{Z}} \mathsf{R}_{\mathsf{xx}}(m)e^{-i\omega m}$$

The ***cross spectral density*** ***(CSD)*** $\check{\mathsf{S}}_{\mathsf{xy}}(z)$ *of* $\mathbf{x}$ *and* $\mathbf{y}$ *is*

$$\tilde{\mathsf{S}}_{\mathsf{xy}}(z) \triangleq \check{\mathbf{F}}\mathsf{R}_{\mathsf{xy}}(m) \triangleq \sum_{m\in\mathbb{Z}} \mathsf{R}_{\mathsf{xy}}(m)e^{-i\omega m}$$

The ***auto-spectral density*** *is also called* ***power spectral density (PSD)***.

Theorem 2.2. *Let* $\mathbf{S}$ *be a system with* IMPULSE RESPONSE $\mathsf{h}(n)$, INPUT $\mathsf{x}(n)$, *and* OUTPUT $\mathsf{y}(n)$.

THM

$$\left\{\ \mathsf{x} \text{ and } \mathsf{y} \text{ are WIDE SENSE STATIONARY}\ \right\} \Longrightarrow \left\{\begin{array}{llll} (1). & \check{\mathsf{S}}_{\mathsf{xx}}(z) & = & \check{\mathsf{S}}^*_{\mathsf{xx}}\left(\frac{1}{z^*}\right) \quad \text{and} \\ (2). & \check{\mathsf{S}}_{\mathsf{yx}}(z) & = & \check{\mathsf{S}}^*_{\mathsf{xy}}\left(\frac{1}{z^*}\right) \end{array}\right\}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

✎PROOF:

$$\begin{aligned}
\check{S}_{yx}(z) &\triangleq \mathbf{Z}R_{yx}(m) && \text{by definition of } \check{S}_{xy}(z) && \text{(Definition 2.6 page 14)}\\
&\triangleq \sum_{m\in\mathbb{Z}} R_{yx}(m)z^{-m} && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)}\\
&\triangleq \sum_{m\in\mathbb{Z}} R^*_{xy}(-m)z^{-m} && \text{by Corollary 2.1 page 14} && \\
&= \left[\sum_{m\in\mathbb{Z}} R_{xy}(-m)(z^*)^{-m}\right]^* && \text{by } \textit{antiautomorphic} \text{ property of } {}^*\text{-algebras} && \text{(Definition J.3 page 266)}\\
&= \left[\sum_{-p\in\mathbb{Z}} R_{xy}(p)(z^*)^{p}\right]^* && \text{where } p \triangleq -m && \implies m=-p\\
&= \left[\sum_{p\in\mathbb{Z}} R_{xy}(p)(z^*)^{p}\right]^* && \text{by } \textit{absolutely summable} \text{ property} && \text{(Definition R.2 page 369)}\\
&= \left[\sum_{p\in\mathbb{Z}} R_{xy}(p)\left(\frac{1}{z^*}\right)^{-p}\right]^* && && \\
&= \check{S}^*_{xy}\left(\frac{1}{z^*}\right) && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)}\\
\check{S}_{xx}(z) &= \check{S}_{xy}(z)\big|_{y=x} && && \\
&= \check{S}^*_{yx}(z)\big|_{y=x} && && \\
&= \check{S}^*_{xy}\left(\frac{1}{z^*}\right)\Big|_{y=x} && \text{by (2)—previous result} && \\
&= \check{S}^*_{xx}\left(\frac{1}{z^*}\right) && &&
\end{aligned}$$

✏

Corollary 2.2. *Let* $\mathbf{S}$ *be a system with* IMPULSE RESPONSE $h(n)$, INPUT $x(n)$, *and* OUTPUT $y(n)$.

COR

$$\left\{\begin{array}{ll} (A). & h \textit{ is } \text{LTI} \quad \textit{and} \\ (B). & x \textit{ and } y \textit{ are } \text{WSS} \end{array}\right\} \implies \left\{\begin{array}{llll} (1). & \tilde{S}^*_{xy}(\omega) = \tilde{S}_{yx}(\omega) & \text{(CONJUGATE-SYMMETRIC)} & \textit{and} \\ (2). & \tilde{S}^*_{xx}(\omega) = \tilde{S}_{xx}(\omega) & \text{(CONJUGATE SYMMETRIC)} & \textit{and} \\ (3). & \tilde{S}_{xx}(\omega) \in \mathbb{R} & \text{(REAL-VALUED)} & \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\tilde{S}^*_{xy}(\omega) &= \check{S}^*_{xy}(z)\big|_{z=e^{i\omega}} && \text{by definition of } \textit{DTFT} && \text{(Definition Q.1 page 359)}\\
&= \check{S}^{**}_{yx}\left(\frac{1}{z^*}\right)\Big|_{z=e^{i\omega}} && \text{by Theorem 2.2 page 14} && \\
&= \check{S}_{yx}\left(\frac{1}{z^*}\right)\Big|_{z=e^{i\omega}} && \text{by } \textit{involutory} \text{ property of } *\textbf{-algebras} && \text{(Definition J.3 page 266)}\\
&= \check{S}_{yx}\left(\frac{1}{e^{i\omega *}}\right) && && \\
&= \check{S}_{yx}\left(e^{i\omega}\right) && && \\
&= \tilde{S}_{yx}(\omega) && \text{by definition of } \textit{DTFT} && \text{(Definition Q.1 page 359)}\\
\tilde{S}^*_{xx}(\omega) &= \check{S}^*_{xx}(z)\big|_{z=e^{i\omega}} && \text{by definition of } \textit{DTFT} && \text{(Definition Q.1 page 359)}\\
&= \check{S}^{**}_{xx}\left(\frac{1}{z^*}\right)\Big|_{z=e^{i\omega}} && \text{by Theorem 2.2 page 14} && \\
&= \check{S}_{xx}\left(\frac{1}{z^*}\right)\Big|_{z=e^{i\omega}} && \text{by } \textit{involutory} \text{ property of } *\textbf{-algebras} && \text{(Definition J.3 page 266)}\\
&= \check{S}_{xx}\left(\frac{1}{e^{i\omega *}}\right) && &&
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

$$
\begin{aligned}
&= \check{S}_{xx}\left(e^{i\omega}\right) \\
&= \tilde{S}_{xx}(\omega) && \text{by definition of } DTFT && \text{(Definition Q.1 page 359)} \\
&\implies \tilde{S}_{xx}(\omega) \text{ is } \textit{real-valued} \\
\tilde{S}^*_{xx}(\omega) &= \tilde{S}^*_{xy}(\omega)\big|_{y=x} \\
&= \tilde{S}_{yx}(\omega)\big|_{y=x} && \text{by previous result} \\
&= \tilde{S}_{xx}(\omega)
\end{aligned}
$$

✏

2.5 Spectral Power

The term "*spectral power*" is a bit of an oxymoron because "spectral" deals with leaving the time-domain for the frequency-domain, howbeit the concept of power is solidly founded on the concept of time in that power = energy per time.

However, *Parseval's Theorem* (Proposition H.2 page 232) demonstrates that power in time can also be calculated in frequency. So, it makes some sense to speak of the term "spectral power". Moreover, one way to estimate this power is to average the Fourier Transforms of the product $|x(n)|^2 = x(n)x^*(n)$…that is, to use an estimate of the auto-spectral density $\tilde{S}_{xx}(\omega)$. Thus, an alternate name for *auto-spectral density* is **power spectral density** (PSD).

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 3

CONTINUOUS RANDOM PROCESSES

❝ *A likely impossibility is always preferable to an unconvincing possibility.* ❞
Aristotle (384 BC – 322 BC) [1]

3.1 Definitions

Definition 3.1. [2] *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE.

DEF
The function $\mathsf{x} : \Omega \to \mathbb{R}$ *is a* ***random variable****.*
The function $\mathsf{y} : \mathbb{R} \times \Omega \to \mathbb{R}$ *is a* ***random process****.*

The random process $\mathsf{x}(t, \omega)$, where t commonly represents time and $\omega \in \Omega$ is an outcome of an experiment, can take on more specialized forms depending on whether t and ω are fixed or allowed to vary. These forms are illustrated in Figure 3.1 page 17[3] and Figure 3.2 page 18.

$\mathsf{x}(t, \omega)$	fixed t	variable t
fixed ω	number	time function
variable ω	random variable	random process

Figure 3.1: Specialized forms of a random process $\mathsf{x}(t, \omega)$

[1] quote: http://en.wikiquote.org/wiki/Aristotle
image: http://en.wikipedia.org/wiki/Aristotle
[2] Papoulis (1991), page 63, Papoulis (1991), page 285
[3] Papoulis (1991), pages 285–286

Figure 3.2: Example of a random process $\mathsf{x}(t,\omega)$

Definition 3.2. [4] *Let* $\mathsf{x}(t)$ *and* $\mathsf{y}(t)$ *be random processes.*

DEF						
	The ***mean***	$\mu_{\mathsf{X}}(t)$	*of* $\mathsf{x}(t)$ *is*	$\mu_{\mathsf{X}}(t)$	$\triangleq$	$\mathsf{E}[\mathsf{x}(t)]$
	The ***cross-correlation***	$\mathsf{R}_{\mathsf{xy}}(t)$	*of* $\mathsf{x}(t)$ *and* $\mathsf{y}(t)$ *is*	$\mathsf{R}_{\mathsf{xy}}(t,u)$	$\triangleq$	$\mathsf{E}\left[\mathsf{x}(t)\mathsf{y}^*(u)\right]$
	The ***auto-correlation function***	$\mathsf{R}_{\mathsf{xx}}(t)$	*of* $\mathsf{x}(t)$ *is*	$\mathsf{R}_{\mathsf{xx}}(t,u)$	$\triangleq$	$\mathsf{E}\left[\mathsf{x}(t)\mathsf{x}^*(u)\right]$

Remark 3.1. [5] The equation $\int_{u\in\mathbb{R}} \mathsf{R}_{\mathsf{xx}}(t,u)\mathsf{f}(u)\,\mathsf{d}u$ is a *Fredholm integral equation of the first kind* and $\mathsf{R}_{\mathsf{xx}}(t,u)$ is the *kernel* of the equation.

Theorem 3.1. *Let* $\mathsf{x}(t)$ *and* $\mathsf{y}(t)$ *be random processes with cross-correlation* $\mathsf{R}_{\mathsf{xy}}(t,u)$ *and let* $\mathsf{R}_{\mathsf{xx}}(t,u)$ *be the auto-correlation of* $\mathsf{x}(t)$.

THM				
	$\mathsf{R}_{\mathsf{xx}}(t,u)$	$=$	$\mathsf{R}^*_{\mathsf{xx}}(u,t)$	(CONJUGATE SYMMETRIC)
	$\mathsf{R}_{\mathsf{xy}}(t,u)$	$=$	$\mathsf{R}^*_{\mathsf{yx}}(u,t)$	

✎PROOF:

$$\mathsf{R}_{\mathsf{xx}}(t,u) \triangleq \mathsf{E}\left[\mathsf{x}(t)\mathsf{x}^*(u)\right] = \mathsf{E}\left[\mathsf{x}^*(u)\mathsf{x}(t)\right] = \left(\mathsf{E}\left[\mathsf{x}(u)\mathsf{x}^*(t)\right]\right)^* \triangleq \mathsf{R}^*_{\mathsf{xx}}(u,t)$$
$$\mathsf{R}_{\mathsf{xy}}(t,u) \triangleq \mathsf{E}\left[\mathsf{x}(t)\mathsf{y}^*(u)\right] = \mathsf{E}\left[\mathsf{y}^*(u)\mathsf{x}(t)\right] = \left(\mathsf{E}\left[\mathsf{y}(u)\mathsf{x}^*(t)\right]\right)^* \triangleq \mathsf{R}^*_{\mathsf{yx}}(u,t)$$

✏

[4] Papoulis (1984) page 216 $\langle R_{xy}(t_1,t_2) = E\left\{\mathsf{x}(t_1)\mathsf{y}^*(t_2)\right\}$ (9-35)$\rangle$,

[5] Fredholm (1900), Fredholm (1903), page 365, Michel and Herget (1993), page 97, Keener (1988), page 101

COPYRIGHT © 2019 DANIEL J. GREENHOE

CHAPTER 4

RANDOM PROCESS EIGEN-ANALYSIS

4.1 Definitions

Definition 4.1. *Let* $\mathsf{x}(t)$ *be random processes with* AUTO-CORRELATION *function* *(Definition 3.2 page 18)* $\mathsf{R}_{\mathsf{xx}}(t,u)$.

DEF *The* ***auto-correlation operator*** $\mathbf{R}$ *of* $\mathsf{x}(t)$ *is defined as*

$$\mathbf{R}\mathsf{f} \triangleq \int_{u\in\mathbb{R}} \mathsf{R}_{\mathsf{xx}}(t,u)\mathsf{f}(u)\,\mathrm{d}u$$

Definition 4.2. *Let* $\mathsf{x}(t)$ *be a* RANDOM PROCESS *with* AUTO-CORRELATION $\mathsf{R}_{\mathsf{xx}}(\tau)$ *(Definition 3.2 page 18)*.

DEF *A* RANDOM PROCESS $\mathsf{x}(t)$ *is* ***white*** *if* $\mathsf{R}_{\mathsf{xx}}(\tau) = \delta(\tau)$

If a random process $\mathsf{x}(t)$ is *white* (Definition 4.2 page 19) and the set $\Psi = \{\psi_1(t), \psi_2(t), \ldots, \psi_N(t)\}\}$ is **any** set of orthonormal basis functions, then the innerproducts $\langle n(t) \mid \psi_n(t)\rangle$ and $\langle n(t) \mid \psi_m(t)\rangle$ are *uncorrelated* for $m \neq n$. However, if $\mathsf{x}(t)$ is **colored** (not white), then the innerproducts are not in general uncorrelated. But if the elements of Ψ are chosen to be the eigenfunctions of $\mathbf{R}$ such that $\mathbf{R}\psi_n = \lambda_n\psi_n$, then by Theorem 3.1 (page 18), the set $\{\psi_n(t)\}$ are *orthogonal* and the innerproducts **are** *uncorrelated* even though $\mathsf{x}(t)$ is not white. This criterion is called the Karhunen-Loève criterion for $\mathsf{x}(t)$.

Theorem 4.1. *Let* $\mathbf{R}$ *be an* AUTO-CORRELATION *operator.*

THM

$$\begin{array}{lll} \langle \mathbf{R}\mathsf{x} \mid \mathsf{x}\rangle \geq 0 & \forall \mathsf{x} \in \boldsymbol{X} & (\text{NON-NEGATIVE}) \\ \langle \mathbf{R}\mathsf{x} \mid \mathsf{y}\rangle = \langle \mathsf{x} \mid \mathbf{R}\mathsf{y}\rangle & \forall \mathsf{x},\mathsf{y} \in \boldsymbol{X} & (\text{SELF-ADJOINT}) \end{array}$$

✎PROOF:

1. Proof that $\mathbf{R}$ is *non-negative*:

$$\begin{aligned} \langle \mathbf{R}\mathsf{y} \mid \mathsf{y}\rangle &= \left\langle \int_{u\in\mathbb{R}} \mathsf{R}_{\mathsf{xx}}(t,u)\mathsf{y}(u)\,\mathrm{d}u \mid \mathsf{y}(t)\right\rangle && \text{by definition of } \mathbf{R} && \text{(Definition 4.1 page 19)} \\ &= \left\langle \int_{u\in\mathbb{R}} \mathsf{E}\left[\mathsf{x}(t)\mathsf{x}^*(u)\right]\mathsf{y}(u)\,\mathrm{d}u \mid \mathsf{y}(t)\right\rangle && \text{by definition of } \mathsf{R}_{\mathsf{xx}}(t,u) && \text{(Definition 3.2 page 18)} \\ &= \mathsf{E}\left[\left\langle \int_{u\in\mathbb{R}} \mathsf{x}(t)\mathsf{x}^*(u)\mathsf{y}(u)\,\mathrm{d}u \mid \mathsf{y}(t)\right\rangle\right] && \text{by } \textit{linearity} \text{ of } \langle \triangle \mid \triangledown\rangle \text{ and } \int && \text{(Definition K.1 page 271)} \end{aligned}$$

$$
\begin{aligned}
&= \mathsf{E}\left[\int_{u\in\mathbb{R}} \mathsf{x}^*(u)\mathsf{y}(u)\,du\,\langle \mathsf{x}(t)\,|\,\mathsf{y}(t)\rangle\right] && \text{by } \textit{additivity} \text{ property of } \langle\triangle\,|\,\triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= \mathsf{E}\left[\langle \mathsf{y}(u)\,|\,\mathsf{x}(u)\rangle\,\langle \mathsf{x}(t)\,|\,\mathsf{y}(t)\rangle\right] && \text{by local definition of } \langle\triangle\,|\,\triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= \mathsf{E}\left[\langle \mathsf{x}(u)\,|\,\mathsf{y}(u)\rangle^*\,\langle \mathsf{x}(t)\,|\,\mathsf{y}(t)\rangle\right] && \text{by } \textit{conjugate symmetry} \text{ prop.} && \text{(Definition K.1 page 271)}\\
&= \mathsf{E}|\langle \mathsf{x}(t)\,|\,\mathsf{y}(t)\rangle|^2 && \text{by definition of } |\cdot| && \text{(Definition F.4 page 208)}\\
&\geq 0
\end{aligned}
$$

2. Proof that $\mathbf{R}$ is self-adjoint:

$$
\begin{aligned}
\langle[\mathbf{R}\mathsf{x}](t)\,|\,\mathsf{y}\rangle &= \left\langle \int_{u\in\mathbb{R}} \mathsf{R}_{\mathsf{xx}}(t,u)\mathsf{x}(u)\,du\,|\,\mathsf{y}(t)\right\rangle && \text{by definition of } \mathbf{R} && \text{(Definition 4.1 page 19)}\\
&= \int_{u\in\mathbb{R}} \mathsf{x}(u)\,\langle \mathsf{R}_{\mathsf{xx}}(t,u)\,|\,\mathsf{y}(t)\rangle\,du && \text{by } \textit{additive} \text{ property of } \langle\triangle\,|\,\triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= \int_{u\in\mathbb{R}} \mathsf{x}(u)\,\langle \mathsf{y}(t)\,|\,\mathsf{R}_{\mathsf{xx}}(t,u)\rangle^*\,du && \text{by } \textit{conjugate symmetry} \text{ prop.} && \text{(Definition K.1 page 271)}\\
&= \langle \mathsf{x}(u)\,|\,\langle \mathsf{y}(t)\,|\,\mathsf{R}_{\mathsf{xx}}(t,u)\rangle\rangle && \text{by local definition of } \langle\triangle\,|\,\triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= \left\langle \mathsf{x}(u)\,|\,\int_{t\in\mathbb{R}} \mathsf{y}(t)\mathsf{R}^*_{\mathsf{xx}}(t,u)\,dt\right\rangle && \text{by local definition of } \langle\triangle\,|\,\triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= \left\langle \mathsf{x}(u)\,|\,\int_{t\in\mathbb{R}} \mathsf{y}(t)\mathsf{R}_{\mathsf{xx}}(u,t)\,dt\right\rangle && \text{by property of } \mathsf{R}_{\mathsf{xx}} && \text{(Theorem 3.1 page 18)}\\
&= \left\langle \mathsf{x}(u)\,|\,\underbrace{\mathbf{R}}_{\mathbf{R}^*}\mathsf{y}\right\rangle && \text{by definition of } \mathbf{R} && \text{(Definition 4.1 page 19)}\\
\implies \quad \mathbf{R} = \mathbf{R}^* \quad &\implies \mathbf{R} \text{ is } \textit{self adjoint}
\end{aligned}
$$

✏

4.2 Properties

Theorem 4.2. [1] *Let* $(\!(\lambda_n)\!)_{n\in\mathbb{Z}}$ *be the eigenvalues and* $(\!(\psi_n)\!)_{n\in\mathbb{Z}}$ *be the eigenfunctions of operator* $\mathbf{R}$ *such that* $\mathbf{R}\psi_n = \lambda_n\psi_n$.

THM

1. $\lambda_n \in \mathbb{R}$ — *(eigenvalues of* $\mathbf{R}$ *are* REAL*)*
2. $\lambda_n \neq \lambda_m \implies \langle \psi_n\,|\,\psi_m\rangle = 0$ — *(eigenfunctions associated with distinct eigenvalues are* ORTHOGONAL*)*
3. $\left\|\psi_n(t)\right\|^2 > 0 \implies \lambda_n \geq 0$ — *(eigenvalues are* NON-NEGATIVE*)*
4. $\left\|\psi_n(t)\right\|^2 > 0, \langle \mathbf{R}\mathsf{f}\,|\,\mathsf{f}\rangle > 0 \implies \lambda_n > 0$ — *(if* $\mathbf{R}$ *is* POSITIVE DEFINITE*, then eigenvalues are* POSITIVE*)*

✎PROOF:

1. Proof that eigenvalues are *real-valued*: Because $\mathbf{R}$ is self-adjoint, its eigenvalues are real (Theorem O.18 page 336).

2. eigenfunctions associated with distinct eigenvalues are orthogonal: Because $\mathbf{R}$ is self-adjoint, this property follows (Theorem O.18 page 336).

[1] Keener (1988), pages 114–119

COPYRIGHT © 2019 DANIEL J. GREENHOE

3. Proof that eigenvalues are *non-negative*:

$$\begin{aligned}
0 &\geq \langle \mathbf{R}\psi_n \mid \psi_n \rangle && \text{by definition of non-negative definite}\\
&= \langle \lambda_n \psi_n \mid \psi_n \rangle && \text{by hypothesis}\\
&= \lambda_n \langle \psi_n \mid \psi_n \rangle && \text{by definition of inner-products}\\
&= \lambda_n \left\|\psi_n\right\|^2 && \text{by definition of norm induced by inner-product}
\end{aligned}$$

4. Eigenvalues are *positive* if $\mathbf{R}$ is *positive definite*:

$$\begin{aligned}
0 &> \langle \mathbf{R}\psi_n \mid \psi_n \rangle && \text{by definition of } \textit{positive definite}\\
&= \langle \lambda_n \psi_n \mid \psi_n \rangle && \text{by hypothesis}\\
&= \lambda_n \langle \psi_n \mid \psi_n \rangle && \text{by definition of inner-products}\\
&= \lambda_n \left\|\psi_n\right\|^2 && \text{by definition of norm induced by inner-product}
\end{aligned}$$

✏

Theorem 4.3 (Karhunen-Loève Expansion). [2] *Let* $\mathbf{R}$ *be the* AUTO-CORRELATION OPERATOR *(Definition 4.1 page 19) of a* RANDOM PROCESS $\mathsf{x}(t)$. *Let* $(\!(\lambda_n)\!)_{n\in\mathbb{Z}}$ *be the eigenvalues of* $\mathbf{R}$ *and* $(\!(\psi_n)\!)_{n\in\mathbb{Z}}$ *are the eigenfunctions of* $\mathbf{R}$ *such that* $\mathbf{R}\psi_n = \lambda_n\psi_n$.

THM

$$\underbrace{\left\|\psi_n(t)\right\| = 1}_{\{\psi_n(t)\}\ \textit{are}\ \text{NORMALIZED}} \implies \underbrace{\mathsf{E}\left(\left|\mathsf{x}(t) - \sum_{n\in\mathbb{Z}} \langle \mathsf{x}(t) \mid \psi_n(t)\rangle\, \psi_n(t)\right|^2\right) = 0}_{\text{CONVERGENCE IN PROBABILITY}} \qquad (\{\psi_n(t)\}\ \textit{is a}\ \text{BASIS}\ \textit{for}\ \mathsf{x}(t))$$

✎PROOF:

1. Define $\dot{x}_n \triangleq \langle \mathsf{x}(t) \mid \psi_n(t)\rangle$
2. Define $\mathbf{R}\mathsf{x}(t) \triangleq \int_{u\in\mathbb{R}} \mathsf{R}_{\mathsf{xx}}(t,u)\mathsf{x}(u)\,\mathrm{d}u$
3. lemma: $\mathsf{E}[\mathsf{x}(t)\mathsf{x}(t)] = \sum_{n\in\mathbb{Z}} \lambda_n \left|\psi_n(t)\right|^2$. Proof:

$$\mathsf{E}[\mathsf{x}(t)\mathsf{x}(t)] = \sum_{n\in\mathbb{Z}} \lambda_n \left|\psi_n(t)\right|^2 \qquad \begin{array}{lll} \text{by} & \textit{non-negative}\ \text{property} & \text{(Theorem 4.1 page 19)}\\ \text{and} & \textit{Mercer's Theorem} & \text{(Theorem D.4 page 194)}\end{array}$$

4. lemma:

$$\begin{aligned}
&\mathsf{E}\left[\mathsf{x}(t)\left(\sum_{n\in\mathbb{Z}} \dot{x}_n \psi_n(t)\right)^*\right]\\
&\triangleq \mathsf{E}\left[\mathsf{x}(t)\left(\sum_{n\in\mathbb{Z}} \int_{u\in\mathbb{R}} \mathsf{x}(u)\psi_n^*(u)\,\mathrm{d}u\,\psi_n(t)\right)^*\right] && \text{by definition of } \dot{x} && \text{(definition 1 page 21)}\\
&= \sum_{n\in\mathbb{Z}} \left(\int_{u\in\mathbb{R}} \mathsf{E}\left[\mathsf{x}(t)\mathsf{x}^*(u)\right]\psi_n(u)\,\mathrm{d}u\right)\psi_n^*(t) && \text{by } \textit{linearity} && \text{(Theorem 1.1 page 4)}\\
&\triangleq \sum_{n\in\mathbb{Z}} \left(\int_{u\in\mathbb{R}} \mathsf{R}_{\mathsf{xx}}(t,u)\psi_n(u)\,\mathrm{d}u\right)\psi_n^*(t) && \text{by definition of } \mathsf{R}_{\mathsf{xx}}(t,u) && \text{(Definition 3.2 page 18)}
\end{aligned}$$

[2] Keener (1988), pages 114–119

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

$$\triangleq \sum_{n\in\mathbb{Z}} \left(\mathbf{R}\psi_n(t)\psi_n^*(t)\right) \qquad \text{by definition of } \mathbf{R} \quad \text{(definition 2 page 21)}$$

$$= \sum_{n\in\mathbb{Z}} \lambda_n\psi_n(t)\psi_n^*(t) \qquad \text{by property of } \textit{eigen-system}$$

$$= \sum_{n\in\mathbb{Z}} \lambda_n\left|\psi_n(t)\right|^2$$

5. lemma:

$$\mathsf{E}\left[\sum_{n\in\mathbb{Z}} \dot{x}_n\psi_n(t)\left(\sum_{m\in\mathbb{Z}} \dot{x}_m\psi_m(t)\right)^*\right]$$

$$\triangleq \mathsf{E}\left[\sum_{n\in\mathbb{Z}} \int_{u\in\mathbb{R}} \mathsf{x}(u)\psi_n^*(u)\,\mathrm{d}u\psi_n(t)\left(\sum_{m\in\mathbb{Z}} \int_v \mathsf{x}(v)\psi_m^*(v)\,\mathrm{d}v\psi_m(t)\right)^*\right] \qquad \text{by definition of } \dot{x} \text{ (definition 1 page 21)}$$

$$= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} \int_u \left(\int_v \mathsf{E}\left[\mathsf{x}(u)\mathsf{x}^*(v)\right]\psi_m(v)\,\mathrm{d}v\right)\psi_n^*(u)\,\mathrm{d}u\psi_n(t)\psi_m^*(t) \qquad \text{by } \textit{linearity} \text{ (Theorem 1.1 page 4)}$$

$$= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} \int_u \left(\int_v \mathsf{R}_{\mathsf{xx}}(u,v)\psi_m(v)\,\mathrm{d}v\right)\psi_n^*(u)\,\mathrm{d}u\psi_n(t)\psi_m^*(t) \qquad \text{by definition of } \mathsf{R}_{\mathsf{xx}}(t,u) \text{ (Definition 3.2 page 18)}$$

$$= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} \int_u \left(\mathbf{R}\psi_m(u)\right)\psi_n^*(u)\,\mathrm{d}u\psi_n(t)\psi_m^*(t) \qquad \text{by definition of } \mathbf{R} \text{ (definition 2 page 21)}$$

$$= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} \int_u \left(\lambda_m\psi_m(u)\right)\psi_n^*(u)\,\mathrm{d}u\psi_n(t)\psi_m^*(t) \qquad \text{by property of } \textit{eigen-system}$$

$$= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} \lambda_m\left(\int_{u\in\mathbb{R}} \psi_m(u)\psi_n^*(u)\,\mathrm{d}u\right)\psi_n(t)\psi_m^*(t) \qquad \text{by } \textit{linearity}$$

$$= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} \lambda_m\,\|\psi(t)\|^2\,\bar{\delta}_{mn}\psi_n(t)\psi_m^*(t) \qquad \text{by } \textit{orthogonal} \text{ property (Theorem 4.2 page 20)}$$

$$= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} \lambda_m\bar{\delta}_{mn}\psi_n(t)\psi_m^*(t) \qquad \text{by } \textit{normalized} \text{ hypothesis}$$

$$= \sum_{n\in\mathbb{Z}} \lambda_n\psi_n(t)\psi_n^*(t) \qquad \text{by definition of } \textit{Kronecker delta } \bar{\delta} \quad \text{(Definition K.3 page 283)}$$

$$= \sum_{n\in\mathbb{Z}} \lambda_n\left|\psi_n(t)\right|^2$$

6. Proof that $\{\psi_n(t)\}$ is a *basis* for $\mathsf{x}(t)$:

$$\mathsf{E}\left(\left|\mathsf{x}(t) - \sum_{n\in\mathbb{Z}} \dot{x}_n\psi_n(t)\right|^2\right)$$

$$= \mathsf{E}\left(\left[\mathsf{x}(t) - \sum_{n\in\mathbb{Z}} \dot{x}_n\psi_n(t)\right]\left[\mathsf{x}(t) - \sum_{m\in\mathbb{Z}} \dot{x}_m\psi_m(t)\right]^*\right)$$

$$= \mathsf{E}\left(\mathsf{x}(t)\mathsf{x}^*(t) - \mathsf{x}(t)\left[\sum_{n\in\mathbb{Z}} \dot{x}_n\psi_n(t)\right]^* - \mathsf{x}^*(t)\sum_{n\in\mathbb{Z}} \dot{x}_n\psi_n(t) + \left[\sum_{n\in\mathbb{Z}} \dot{x}_n\psi_n(t)\right]\left[\sum_{m\in\mathbb{Z}} \dot{x}_m\psi_m(t)\right]^*\right)$$

$$= \mathsf{E}\left(\mathsf{x}(t)\mathsf{x}^*(t)\right) - \mathsf{E}\left[\mathsf{x}(t)\left[\sum_{n\in\mathbb{Z}} \dot{x}_n\psi_n(t)\right]^*\right] - \mathsf{E}\left[\mathsf{x}^*(t)\sum_{n\in\mathbb{Z}} \dot{x}_n\psi_n(t)\right] + \mathsf{E}\left[\sum_{n\in\mathbb{Z}} \dot{x}_n\psi_n(t)\left[\sum_{m\in\mathbb{Z}} \dot{x}_m\psi_m(t)\right]^*\right]$$

by *linearity* of E (Theorem 1.1 page 4)

$$= \underbrace{\sum_{n\in\mathbb{Z}} \lambda_n\left|\psi_n(t)\right|^2}_{\text{by (3) lemma}} - \underbrace{\sum_{n\in\mathbb{Z}} \lambda_n\left|\psi_n(t)\right|^2}_{\text{by (4) lemma}} - \underbrace{\left[\sum_{n\in\mathbb{Z}} \lambda_n\left|\psi_n(t)\right|^2\right]^*}_{\text{by (4) lemma}} + \underbrace{\sum_{n\in\mathbb{Z}} \lambda_n\left|\psi_n(t)\right|^2}_{\text{by (5) lemma}}$$

Copyright © 2019 Daniel J. Greenhoe

$$= 0$$

✏

Remark 4.1. The *matrix* **R** is **Toeplitz**. For more information about the properties of *Toeplitz* matrices, see Grenander and Szegö (1958), Widom (1965), Gray (1971), Smylie et al. (1973) page 408 ⟨§"B. Properties of the Toeplitz Matrix"⟩, Grenander and Szegö (1984), Haykin and Kesler (1979), Haykin and Kesler (1983), Böttcher and Silbermann (1999), Gray (2006).

Copyright © 2019 Daniel J. Greenhoe

Part II

Statistical Processing

CHAPTER 5

OPERATIONS ON RANDOM VARIABLES

5.1 Functions of one random variable

Proposition 5.1. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a probability space, and* X *a* RANDOM VARIABLE *with* CUMULATIVE DISTRIBUTION FUNCTION $\mathsf{c}_{\mathsf{X}}(x)$.

PRP

$$\left\{ \begin{array}{l} \mathsf{X} \textit{ is } \text{UNIFORMLY DISTRIBUTED} \\ \textit{(Definition C.1 page 181)} \end{array} \right\} \iff \mathsf{c}_{\mathsf{X}}(x) = \left\{ \begin{array}{ll} 0 & \textit{for } x < 0 \\ x & \textit{for } 0 < x \leq 1 \\ 1 & \textit{otherwise} \end{array} \right\}$$

Theorem 5.1 (Probability integral transform)**.** [1] *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a probability space. Let* X *be a* RANDOM VARIABLE *with* PROBABILITY DENSITY FUNCTION $\mathsf{p}_{\mathsf{X}}(x)$ *and* CUMULATIVE DISTRIBUTION FUNCTION $\mathsf{c}_{\mathsf{X}}(x)$. *Let* Y *be a* RANDOM VARIABLE CUMULATIVE DISTRIBUTION FUNCTION $\mathsf{c}_{\mathsf{Y}}(y)$.

THM

$$\left\{ \begin{array}{ll} \textit{(1).} & \mathsf{Y} = \mathsf{c}_{\mathsf{X}}(\mathsf{X}) \quad \textit{and} \\ \textit{(2).} & \mathsf{p}_{\mathsf{X}}(x) \textit{ is } \text{CONTINUOUS} \end{array} \right\} \implies \left\{ \begin{array}{l} \mathsf{Y} \textit{ is } \text{UNIFORMLY DISTRIBUTED} \\ \textit{(Definition C.1 page 181)} \end{array} \right\}$$

✎PROOF:

$\mathsf{c}_{\mathsf{Y}}(y) \triangleq \mathsf{P}\{\mathsf{Y} \leq y\}$	by definition of *cdf*	(Definition B.2 page 176)
$= \mathsf{P}\left\{\mathsf{c}_{\mathsf{X}}(\mathsf{X}) \leq y\right\}$	by hypothesis (1)	
$= \mathsf{P}\left\{\mathsf{X} \leq \mathsf{c}_{\mathsf{X}}^{-1}(y)\right\}$	by hypothesis (2) and	Proposition A.2 page 170
$\triangleq \mathsf{c}_{\mathsf{X}}\left[\mathsf{c}_{\mathsf{X}}^{-1}(y)\right]$	by definition of *cdf*	(Definition B.2 page 176)
$= y$		
$\implies$ Y is *uniformly distributed*	by	Proposition 5.1 page 27

✏

Theorem 5.2 (Inverse probability integral transform)**.** [2] *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a probability space. Let* X *be a* RANDOM VARIABLE *with* PROBABILITY DENSITY FUNCTION $\mathsf{p}_{\mathsf{X}}(x)$ *and* CUMULATIVE DISTRIBUTION FUNCTION $\mathsf{c}_{\mathsf{X}}(x)$. *Let* Y *be a* RANDOM VARIABLE CUMULATIVE DISTRIBUTION FUNCTION $\mathsf{c}_{\mathsf{Y}}(y)$.

THM

$$\left\{ \begin{array}{ll} \textit{(1).} & \mathsf{Y} = \mathsf{c}_{\mathsf{Z}}^{-1}(\mathsf{X}) \quad \textit{and} \\ \textit{(2).} & \mathsf{Y} \textit{ is } \text{UNIFORMLY DISTRIBUTED} \quad \textit{and} \\ \textit{(3).} & \mathsf{p}_{\mathsf{Z}}(z) \textit{ is } \text{CONTINUOUS} \end{array} \right\} \implies \left\{ \begin{array}{l} \mathsf{p}_{\mathsf{Y}}(y) = \mathsf{p}_{\mathsf{Z}}(y) \\ (\mathsf{Y} \textit{ has distribution } \mathsf{p}_{\mathsf{Z}}(y)) \end{array} \right\}$$

[1] Angus (1994), Roussas (2014) page 232 ⟨Theorem 10⟩, Devroye (1986) page 28 ⟨Theorem 2.1⟩

[2] Devroye (1986) page 28 ⟨Theorem 2.1⟩, Balakrishnan and Lai (2009) page 624 ⟨14.2.1 Introduction⟩

Figure 5.1: $\mathsf{Y} = \mathsf{f}(\mathsf{X})$

✎PROOF:

$\mathsf{c}_{\mathsf{Y}}(y) \triangleq \mathsf{P}\{\mathsf{Y} \le y\}$	by definition of c_{Y}	(Definition B.2 page 176)
$= \mathsf{P}\left\{\mathsf{c}_{\mathsf{Z}}^{-1}(\mathsf{X}) \le y\right\}$	by hypothesis (1)	
$= \mathsf{P}\left\{\mathsf{X} \le \mathsf{c}_{\mathsf{Z}}(y)\right\}$	by hypothesis (3) and	Proposition A.2 page 170
$\triangleq \mathsf{c}_{\mathsf{X}}\left[\mathsf{c}_{\mathsf{Z}}(y)\right]$	by definition of c_{X}	(Definition B.2 page 176)
$= \mathsf{c}_{\mathsf{Z}}(y)$	because $0 \le \mathsf{c}_{\mathsf{Z}}(y) \le 1$ and by	Proposition 5.1 page 27
$\Longrightarrow \quad \mathsf{p}_{\mathsf{Y}}(y) = \mathsf{p}_{\mathsf{Z}}(y)$	(Y has the distribution of Z)	

✏

Definition 5.1. [3] *Let* $\mathsf{f}(x)$ *be a* DIFFERENTIABLE FUNCTION *in* $\mathbb{R}^{\mathbb{R}}$.

DEF *A point* $p \in \mathbb{R}$ *is a* ***critical point*** *of* $\mathsf{f}(x)$ *if*
$$\mathsf{f}'(p) = 0.$$

Theorem 5.3. [4] *Let* X *and* Y *be* RANDOM VARIABLES *in* $\mathbb{R}^{\mathbb{R}}$. *Let* f *be a* DIFFERENTIABLE FUNCTION *in* $\mathbb{R}^{\mathbb{R}}$ *with* N CRITICAL POINTS *(Definition 5.1 page 28). Let the range of* X *be partitioned into* $N+1$ *partitions* $\left\{A_n | n = 1, 2, \ldots, N+1\right\}$ *with partition boundaries set at the* N CRITICAL POINTS *of* $\mathsf{f}(x)$*—as illustrated in Figure 5.1 (page 28). Let* $\mathsf{g}_n(x) \triangleq \mathsf{f}(x)$ *but with domain restricted to* $x \in A_n$.

THM
$$\left\{\begin{array}{ll} (1). & \mathsf{Y} = \mathsf{f}(\mathsf{X}) \quad \textit{and} \\ (2). & \mathsf{f} \textit{ is DIFFERENTIABLE} \end{array}\right\} \Longrightarrow \left\{\mathsf{p}_{\mathsf{Y}}(y) = \sum_{n=1}^{N+1} \frac{\mathsf{p}_{\mathsf{X}}\left(\mathsf{g}_n^{-1}(y)\right)}{\left|\mathsf{f}'\left(\mathsf{g}_n^{-1}(y)\right)\right|}\right\}$$

✎PROOF:

1. The problem with a function $\mathsf{f}(x)$ with at least $N = 1$ critical point is that $\mathsf{f}^{-1}(y)$ is *not invertible*. That is, $\mathsf{f}^{-1}(y)$ has more than one solution (and thus the *relation* $\mathsf{f}^{-1}(y)$ is not a *function*). However, note that in each partition A_n, $\mathsf{f}(x)$ is *invertible* and thus $\mathsf{f}^{-1}(y)$ in that partition has a *unique* solution. Thus, each $\mathsf{g}_n(x)$ *is* invertible in it's domain (and each $\mathsf{g}_n^{-1}(y)$ exists as a function).

[3] Callahan (2010) page 189 ⟨Definition 6.1⟩

[4] Papoulis (1984) pages 95–96 ⟨"Fundamental Theorem"⟩, Papoulis (1990) page 157 ⟨"Fundamental Theorem"⟩, Papoulis (1991), page 93, Haykin (1994) page 235 ⟨0471571768⟩§"4.5 TRANSFORMATIONS OF RANDOM VARIABLES", Proakis (2001), page 30,

COPYRIGHT © 2019 DANIEL J. GREENHOE

2. Using item (1), the remainder of the proof follows …

$$
\begin{aligned}
\mathsf{p}_{\mathsf{Y}}(y) &\triangleq \frac{\mathsf{d}}{\mathsf{d}y}\mathsf{P}\{\mathsf{Y} \le y\} && \text{by definition of } \mathsf{p}_{\mathsf{Y}} \text{ (Definition B.2 page 176)}\\
&= \frac{\mathsf{d}}{\mathsf{d}y}\mathsf{P}\{\mathsf{f}(\mathsf{X}) \le y\} && \text{by hypothesis (1)}\\
&= \frac{\mathsf{d}}{\mathsf{d}y}\sum_{n=1}^{N+1}\mathsf{P}\left\{\mathsf{f}(\mathsf{X}) \le y | \mathsf{X} \in A_n\right\} && \text{by } \textit{sum of products} \text{ (Theorem A.3 page 169)}\\
&= \frac{\mathsf{d}}{\mathsf{d}y}\sum_{n=1}^{N+1}\mathsf{P}\left\{\mathsf{f}(\mathsf{X}) \le y \mid \mathsf{X} \in A_n\right\}\mathsf{P}\left\{\mathsf{X} \in A_n\right\} && \text{by definition of } \mathsf{P}\{X|Y\} \text{ (Definition A.4 page 168)}\\
&= \frac{\mathsf{d}}{\mathsf{d}y}\sum_{n=1}^{N+1}\mathsf{P}\left\{\mathsf{g}_n(\mathsf{X}) \le y \,\middle|\, \mathsf{X} \in A_n\right\}\mathsf{P}\left\{\mathsf{X} \in A_n\right\} && \text{by definition of } \mathsf{g}_n(x)\\
&= \frac{\mathsf{d}}{\mathsf{d}y}\left\{\begin{array}{ll}\displaystyle\sum_{n=1}^{N+1}\mathsf{P}\left\{\mathsf{X} \le \mathsf{g}_n^{-1}(y) \,\middle|\, \mathsf{X} \in A_n\right\}\mathsf{P}\left\{\mathsf{X} \in A_n\right\} & \text{for } \mathsf{f}'(x) \ge 0\\ \displaystyle\sum_{n=1}^{N+1}\mathsf{P}\left\{\mathsf{X} \ge \mathsf{g}_n^{-1}(y) \,\middle|\, \mathsf{X} \in A_n\right\}\mathsf{P}\left\{\mathsf{X} \in A_n\right\} & \text{otherwise}\end{array}\right\} && \text{by item (1)}\\
&= \frac{\mathsf{d}}{\mathsf{d}y}\left\{\begin{array}{ll}\displaystyle\sum_{n=1}^{N+1}\mathsf{P}\left\{\mathsf{X} \le \mathsf{g}_n^{-1}(y) \,\middle|\, \mathsf{X} \in A_n\right\}\mathsf{P}\left\{\mathsf{X} \in A_n\right\} & \text{for } \mathsf{f}'(x) \ge 0\\ \displaystyle\sum_{n=1}^{N+1}\left[1-\mathsf{P}\left\{\mathsf{X} \le \mathsf{g}_n^{-1}(y) \,\middle|\, \mathsf{X} \in A_n\right\}\mathsf{P}\left\{\mathsf{X} \in A_n\right\}\right] & \text{otherwise}\end{array}\right\}\\
&= \frac{\mathsf{d}}{\mathsf{d}y}\left\{\begin{array}{ll}\displaystyle\sum_{n=1}^{N+1}\mathsf{P}\left\{\mathsf{X} \le \mathsf{g}_n^{-1}(y) | \mathsf{X} \in A_n\right\} & \text{for } \mathsf{f}'(x) \ge 0\\ \displaystyle\sum_{n=1}^{N+1}\left[1-\mathsf{P}\left\{\mathsf{X} \le \mathsf{g}_n^{-1}(y) | \mathsf{X} \in A_n\right\}\right] & \text{otherwise}\end{array}\right\} && \text{by definition of } \mathsf{P}\{X|Y\}\\
&= \frac{\mathsf{d}}{\mathsf{d}y}\left\{\begin{array}{ll}\displaystyle\sum_{n=1}^{N+1}\left(\mathsf{P}\left\{\mathsf{X} \le \mathsf{g}_n^{-1}(y)\right\}-\mathsf{P}\left\{\mathsf{X} < \min A_{n-1}\right\}\right) & \text{for } \mathsf{f}'(x) \ge 0\\ \displaystyle\sum_{n=1}^{N+1}\left[1-\left(\mathsf{P}\left\{\mathsf{X} \le \mathsf{g}_n^{-1}(y)\right\}-\mathsf{P}\left\{\mathsf{X} < \min A_{n-1}\right\}\right)\right] & \text{otherwise}\end{array}\right\}\\
&= \frac{\mathsf{d}}{\mathsf{d}y}\left\{\begin{array}{ll}\displaystyle\sum_{n=1}^{N+1}\mathsf{P}\left\{\mathsf{X} \le \mathsf{g}_n^{-1}(y)\right\} & \text{for } \mathsf{f}'(x) \ge 0\\ \displaystyle\sum_{n=1}^{N+1}\left[1-\mathsf{P}\left\{\mathsf{X} \le \mathsf{g}_n^{-1}(y)\right\}\right] & \text{otherwise}\end{array}\right\} && \text{because } \frac{\mathsf{d}}{\mathsf{d}y}\mathsf{P}\{\mathsf{X} < \text{a constant}\} = 0\\
&= \left\{\begin{array}{ll}\displaystyle\sum_{n=1}^{N+1}\frac{\mathsf{d}}{\mathsf{d}y}\mathsf{c}_{\mathsf{X}}\left[\mathsf{g}_n^{-1}(y)\right] & \text{for } \mathsf{f}'(x) \ge 0\\ \displaystyle\sum_{n=1}^{N+1}\frac{\mathsf{d}}{\mathsf{d}y}\left[1-\mathsf{c}_{\mathsf{X}}\left(\mathsf{g}_n^{-1}(y)\right)\right] & \text{otherwise}\end{array}\right\} && \text{by } \textit{linearity} \text{ of } \frac{\mathsf{d}}{\mathsf{d}y} \text{ operator}\\
&= \left\{\begin{array}{ll}\displaystyle\sum_{n=1}^{N+1}\mathsf{p}_{\mathsf{X}}\left[\mathsf{g}_n^{-1}(y)\right]\frac{\mathsf{d}}{\mathsf{d}y}\left[\mathsf{g}_n^{-1}(y)\right] & \text{for } \mathsf{f}'(x) \ge 0\\ \displaystyle\sum_{n=1}^{N+1}\left[-\mathsf{p}_{\mathsf{X}}\left[\mathsf{g}_n^{-1}(y)\right]\frac{\mathsf{d}}{\mathsf{d}y}\left[\mathsf{g}_n^{-1}(y)\right]\right] & \text{otherwise}\end{array}\right\} && \begin{array}{l}\text{by definition of } \mathsf{p}_{\mathsf{X}} \text{ (Definition B.2 page 176)}\\ \text{and the } \textit{chain rule}\end{array}
\end{aligned}
$$

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

Figure 5.2: $\mathsf{Y} = \frac{1}{\mathsf{X}}$

$$= \sum_{n=1}^{N+1} \mathsf{p}_{\mathsf{X}}\left(\mathsf{g}_n^{-1}(y)\right)\left|\frac{d}{dy}\left[\mathsf{g}_n^{-1}(y)\right]\right|$$

$$= \sum_{n=1}^{N+1} \frac{\mathsf{p}_{\mathsf{X}}\left(\mathsf{g}_n^{-1}(y)\right)}{\left|\mathsf{f}'\left(\mathsf{g}_n^{-1}(y)\right)\right|} \qquad \text{by Lemma ?? page ??}$$

✏

Corollary 5.1. [5] *Let* X *and* Y *be* RANDOM VARIABLES *in* $\mathbb{R}^{\mathbb{R}}$. *Let* $a, b \in \mathbb{R}$.

COR

$$\left\{\begin{array}{ll} (1). & \mathsf{Y} = a\mathsf{X} + b \quad \textit{and} \\ (2). & a \neq 0 \end{array}\right\} \Longrightarrow \left\{\mathsf{p}_{\mathsf{Y}}(y) = \frac{1}{|a|}\mathsf{p}_{\mathsf{X}}\left(\frac{y-b}{a}\right)\right\}$$

✎PROOF:

1. Note that $\mathsf{f}(x) = ax + b$ is a *differentiable function* with $N = 0$ *critical points* and $\mathsf{f}'(x) = a$.
2. The inverse of $\mathsf{f}(x)$ is $\mathsf{g}_1(y) = \mathsf{f}^{-1}(y) = \frac{y-b}{a}$.
3. It follows that

$$\mathsf{p}_{\mathsf{Y}}(y) = \sum_{n=1}^{N+1} \frac{\mathsf{p}_{\mathsf{X}}\left(\mathsf{g}_n^{-1}(y)\right)}{\left|\mathsf{f}'\left(\mathsf{g}_n^{-1}(y)\right)\right|} \qquad \text{by Theorem 5.3 (page 28)}$$

$$= \frac{\mathsf{p}_{\mathsf{X}}\left(\mathsf{f}^{-1}(y)\right)}{\left|\mathsf{f}'\left(\mathsf{f}^{-1}(y)\right)\right|} \qquad \text{because } N = 0$$

$$= \frac{\mathsf{p}_{\mathsf{X}}\left(\mathsf{f}^{-1}(y)\right)}{|a|} \qquad \text{by item (1)}$$

$$= \frac{1}{|a|}\mathsf{p}_{\mathsf{X}}\left(\frac{y-b}{a}\right) \qquad \text{by item (2)}$$

✏

[5] Papoulis (1984) page 96 ⟨"Illustrations" 1⟩, Papoulis (1991), page 95, Proakis (2001), page 29

COPYRIGHT © 2019 DANIEL J. GREENHOE

Figure 5.3: $\mathsf{Y} = rvX^2$

Corollary 5.2. [6]

COR $\left\{\mathsf{Y} = \frac{1}{\mathsf{X}}\right\} \implies \left\{\mathsf{p}_\mathsf{Y}(y) = \frac{1}{y^2}\mathsf{p}_\mathsf{X}\left(\frac{1}{y}\right) \quad \textit{for } y > 0\right\}$

✎PROOF:

1. Note that $\mathsf{f}(x) = 1/x$ is a *differentiable function* in $x > 0$ with $N = 0$ *critical points* and $\mathsf{f}'(x) = -1/x^2$.
2. The inverse of $\mathsf{f}(x)$ is $\mathsf{g}_1(y) = \mathsf{f}^{-1}(y) = \frac{1}{y}$.
3. It follows that

$$\begin{aligned}
\mathsf{p}_\mathsf{Y}(y) &= \sum_{n=1}^{N+1} \frac{\mathsf{p}_\mathsf{X}\left(\mathsf{g}_n^{-1}(y)\right)}{\left|\mathsf{f}'\left(\mathsf{g}_n^{-1}(y)\right)\right|} && \text{by Theorem 5.3 (page 28)}\\
&= \frac{\mathsf{p}_\mathsf{X}\left(\mathsf{f}^{-1}(y)\right)}{\left|\mathsf{f}'\left(\mathsf{f}^{-1}(y)\right)\right|} && \text{because } N = 0\\
&= \frac{1}{\left|-1/(1/y)^2\right|}\mathsf{p}_\mathsf{X}\left(\frac{1}{y}\right)\\
&= \frac{1}{y^2}\mathsf{p}_\mathsf{X}\left(\frac{1}{y}\right)
\end{aligned}$$

✏

Corollary 5.3. [7] *Let* X *and* Y *be* RANDOM VARIABLES.

COR $\left\{\mathsf{Y} = \mathsf{X}^2\right\} \implies \left\{\mathsf{p}_\mathsf{Y}(y) = \frac{1}{2\sqrt{y}}\left[\mathsf{p}_\mathsf{X}(-\sqrt{y}) + \mathsf{p}_\mathsf{X}(\sqrt{y})\right]\right\}$

✎PROOF:

1. The roots of $y = x^2$ are $x_1 = -\sqrt{y}$ and $x_2 = +\sqrt{y}$.
2. The derivative of $\mathsf{f}(x) \triangleq y = x^2$ is $\mathsf{f}'(x) = 2x$.

[6] Papoulis (1984) page 97 ⟨Example 5-10⟩, Papoulis (1991), page 94
[7] Papoulis (1984) page 95 ⟨Example 5-9⟩, Devroye (1986) page 27 ⟨Example 4.4⟩, Papoulis (1991), page 95, Proakis (2001), page 29

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Figure 5.4: $Z = \tan \Theta$

3. And so it follows that …

$$
\begin{aligned}
\mathsf{p}_Y(y) &= \sum_{n=1}^{N} \frac{\mathsf{p}_X(x_n)}{|\mathsf{f}'(x_n)|} && \text{by Theorem 5.3 page 28} \\
&= \frac{\mathsf{p}_X(x_1)}{|\mathsf{f}'(x_1)|} + \frac{\mathsf{p}_X(x_2)}{|\mathsf{f}'(x_2)|} && \text{by definition of } \sum \\
&= \frac{\mathsf{p}_X(-\sqrt{y})}{|\mathsf{f}'(-\sqrt{y})|} + \frac{\mathsf{p}_X(\sqrt{y})}{|\mathsf{f}'(\sqrt{y})|} && \text{by item (1)} \\
&= \frac{\mathsf{p}_X(-\sqrt{y})}{2\sqrt{y}} + \frac{\mathsf{p}_X(\sqrt{y})}{2\sqrt{y}} && \text{by item (2)} \\
&= \frac{1}{2\sqrt{y}}\left[\mathsf{p}_X(-\sqrt{y}) + \mathsf{p}_X(\sqrt{y})\right] && \text{by } \textit{linearity} \text{ of } + \text{ operation}
\end{aligned}
$$

✏

Corollary 5.4. [8] *Let* $Z = \tan \Theta$. *Then*

COR $\{Z = \tan \Theta\} \quad \Longrightarrow \quad \left\{ \mathsf{p}_Z(z) = \frac{1}{1+z^2} \sum_{n \in Z} \mathsf{p}_\theta(\operatorname{atan}(z) + n\pi) \right\}$

✎PROOF:

1. The roots of $z = \tan\theta$ are $\{\theta_n = \operatorname{atan} z + n\pi | n \in \mathbb{Z}\}$.
2. The derivative of $z = \tan\theta$ is $\mathsf{f}'(\theta) = \sec^2\theta$.
3. It follows that

$$
\begin{aligned}
\mathsf{p}_Z(z) &= \sum_{n=1}^{N} \frac{\mathsf{p}_\theta(\theta_n)}{|\mathsf{f}'(\theta_n)|} \\
&= \sum_n \frac{\mathsf{p}_\theta(\operatorname{atan} z + n\pi)}{|\mathsf{f}'(\operatorname{atan} z + n\pi)|} \\
&= \sum_n \frac{\mathsf{p}_\theta(\operatorname{atan} z + n\pi)}{|\sec^2(\operatorname{atan} z + n\pi)|} \\
&= \sum_n \cos^2(\operatorname{atan} z + n\pi)\mathsf{p}_\theta(\operatorname{atan} z + n\pi)
\end{aligned}
$$

[8] Papoulis (1991), pages 99–100

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \cos^2(\operatorname{atan} z) \sum_n \mathsf{p}_\theta(\operatorname{atan} z + n\pi)$$

$$= \frac{1}{1+z^2} \sum_n \mathsf{p}_\theta(\operatorname{atan} z + n\pi)$$

✏

5.2 Functions of two random variables

Theorem 5.4. [9] *Let* X, Y, *and* Z *be* RANDOM VARIABLES. *Let* $\star$ *be the* CONVOLUTION *operator* *(Definition P.3 page 352)*.

THM

$$\left\{\begin{array}{lll} (1). & \mathsf{Z} \triangleq \mathsf{X}+\mathsf{Y} & \textit{and} \\ (2). & \mathsf{X} \textit{ and } \mathsf{Y} \textit{ are } \text{INDEPENDENT} & \textit{(Definition A.3 page 167)} \end{array}\right\} \implies \left\{\mathsf{p}_\mathsf{Z}(z) = \mathsf{p}_\mathsf{X}(z) \star \mathsf{p}_\mathsf{Y}(z)\right\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{p}_\mathsf{Z}(z) &\triangleq \frac{\mathsf{d}}{\mathsf{d}z}\mathsf{c}_\mathsf{Z}(z) && \text{by definition of } \mathsf{p}_\mathsf{Z} && \text{(Definition B.2 page 176)} \\
&\triangleq \frac{\mathsf{d}}{\mathsf{d}z}\mathsf{P}\{\mathsf{Z} \le z\} && \text{by definition of } \mathsf{c}_\mathsf{Z} && \text{(Definition B.2 page 176)} \\
&= \frac{\mathsf{d}}{\mathsf{d}z}\mathsf{P}\{\mathsf{X}+\mathsf{Y} \le z\} && \text{by hypothesis (1)} \\
&= \frac{\mathsf{d}}{\mathsf{d}z}\lim_{\varepsilon\to 0}\sum_{n\in\mathbb{Z}} \mathsf{P}\left\{\mathsf{X}+\mathsf{Y} \le z \,|\, y+n\varepsilon < \mathsf{Y} \le y+(n+1)\varepsilon\right\} && \text{by } \textit{sum of products} && \text{(Theorem A.3 page 169)} \\
&= \frac{\mathsf{d}}{\mathsf{d}z}\int_{y\in\mathbb{R}} \mathsf{P}\left\{\mathsf{X}+\mathsf{Y} \le z \mid \mathsf{Y}=y\right\}\mathsf{p}_\mathsf{Y}(y)\,\mathsf{d}y && \text{by definiton of } \mathsf{P}\{\mathsf{X} \mid \mathsf{Y}\} && \text{(Definition A.4 page 168)} \\
&= \frac{\mathsf{d}}{\mathsf{d}z}\int_{y\in\mathbb{R}} \mathsf{P}\left\{\mathsf{X} \le z-y \mid \mathsf{Y}=y\right\}\mathsf{p}_\mathsf{Y}(y)\,\mathsf{d}y \\
&= \frac{\mathsf{d}}{\mathsf{d}z}\int_{y\in\mathbb{R}} \mathsf{P}\{\mathsf{X} \le z-y\}\mathsf{p}_\mathsf{Y}(y)\,\mathsf{d}y && \text{by hypothesis (2)} \\
&\triangleq \frac{\mathsf{d}}{\mathsf{d}z}\int_{y\in\mathbb{R}} \mathsf{c}_\mathsf{X}(z-y)\mathsf{p}_\mathsf{Y}(y)\,\mathsf{d}y && \text{by definition of } \mathsf{c}_\mathsf{X} && \text{(Definition B.2 page 176)} \\
&= \int_{y\in\mathbb{R}} \frac{\mathsf{d}}{\mathsf{d}z}\left[\mathsf{c}_\mathsf{X}(z-y)\mathsf{p}_\mathsf{Y}(y)\right]\mathsf{d}y && \text{by } \textit{linearity} \text{ of } \frac{\mathsf{d}}{\mathsf{d}z} \\
&= \int_{y\in\mathbb{R}} \left[\frac{\mathsf{d}}{\mathsf{d}z}\mathsf{c}_\mathsf{X}(z-y)\right]\mathsf{p}_\mathsf{Y}(y)\,\mathsf{d}y && \text{because } y \text{ is fixed inside the integral} \\
&\triangleq \int_{y} \mathsf{p}_\mathsf{X}(z-y)\mathsf{p}_\mathsf{Y}(y)\,\mathsf{d}y && \text{by definition of } \mathsf{p}_\mathsf{X} && \text{(Definition B.2 page 176)} \\
&= \mathsf{p}_\mathsf{X}(z) \star \mathsf{p}_\mathsf{Y}(z) && \text{by definition of } \star && \text{(Definition P.3 page 352)}
\end{aligned}$$

✏

Theorem 5.5. *Let*

- X_1 *and* X_2 *be random variables with joint distribution* $\mathsf{p}_{\mathsf{X}_1,\mathsf{X}_2}(x_1,x_2)$
- $\mathsf{Y}_1 = \mathsf{f}_1(x_1,x_2)$ *and* $\mathsf{Y}_2 = \mathsf{f}_2(x_1,x_2)$

[9] Papoulis (1990) page 160 ⟨Example 5.16⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Figure 5.5: **Sum** of random variables yields **convolution** of pdfs (Theorem 5.4 page 33)

Then the joint distribution of Y_1 *and* Y_2 *is*

THM

$$\mathsf{p}_{\mathsf{Y}_1,\mathsf{Y}_2}(y_1,y_2) = \frac{\mathsf{p}_{\mathsf{X}_1,\mathsf{X}_2}(x_1,x_2)}{|J(x_1,x_2)|} = \frac{\mathsf{p}_{\mathsf{X}_1,\mathsf{X}_2}(x_1,x_2)}{\begin{vmatrix} \frac{\partial \mathsf{f}_1}{\partial x_1} & \frac{\partial \mathsf{f}_1}{\partial x_2} \\ \frac{\partial \mathsf{f}_2}{\partial x_1} & \frac{\partial \mathsf{f}_2}{\partial x_2} \end{vmatrix}} = \frac{\mathsf{p}_{\mathsf{X}_1,\mathsf{X}_2}(x_1,x_2)}{\frac{\partial \mathsf{f}_1}{\partial x_1}\frac{\partial \mathsf{f}_2}{\partial x_2} - \frac{\partial \mathsf{f}_1}{\partial x_2}\frac{\partial \mathsf{f}_2}{\partial x_1}}$$

Proposition 5.2. *Let* X *and* Y *be random variables with joint distribution* $\mathsf{p}_{\mathsf{XY}}(x,y)$ *and*

$$R^2 \triangleq X^2 + Y^2 \qquad\qquad \Theta \triangleq \operatorname{atan}\frac{\mathsf{Y}}{\mathsf{X}}.$$

Then

PRP

$$\mathsf{p}_{\mathsf{R},\Theta}(r,\theta) = r\ \mathsf{p}_{\mathsf{XY}}(r\cos\theta, r\sin\theta)$$

✎PROOF:

$$\begin{aligned}
\mathsf{p}_{\mathsf{R},\Theta}(r,\theta) &= \frac{\mathsf{p}_{\mathsf{XY}}(x,y)}{|J(x,y)|} = \frac{\mathsf{p}_{\mathsf{XY}}(x,y)}{\begin{vmatrix} \frac{\partial R}{\partial x} & \frac{\partial R}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{vmatrix}} = \frac{\mathsf{p}_{\mathsf{XY}}(x,y)}{\begin{vmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{vmatrix}} \\
&= \frac{\mathsf{p}_{\mathsf{XY}}(x,y)}{\frac{x}{\sqrt{x^2+y^2}}\frac{x}{x^2+y^2} - \frac{y}{\sqrt{x^2+y^2}}\frac{-y}{x^2+y^2}} \\
&= \frac{\mathsf{p}_{\mathsf{XY}}(x,y)}{\frac{x^2+y^2}{(x^2+y^2)^{3/2}}} \\
&= \mathsf{p}_{\mathsf{XY}}(x,y)\frac{(x^2+y^2)^{3/2}}{x^2+y^2} \\
&= \mathsf{p}_{\mathsf{XY}}(r\cos\theta, r\sin\theta)\frac{r^3}{r^2} \\
&= r\ \mathsf{p}_{\mathsf{XY}}(r\cos\theta, r\sin\theta)
\end{aligned}$$

✏

Proposition 5.3. *Let* $\mathsf{X} \sim \mathsf{N}\left(0,\sigma^2\right)$ *and* $\mathsf{Y} \sim \mathsf{N}\left(0,\sigma^2\right)$ *be independent random variables and*

$$R^2 \triangleq X^2 + Y^2 \qquad\qquad \Theta \triangleq \operatorname{atan}\frac{\mathsf{Y}}{\mathsf{X}}.$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

Then

PRP

1. *R and Θ are independent with joint distribution* $\quad \mathsf{p}_{\mathsf{R},\Theta}(r,\theta) = \mathsf{p}_{\mathsf{R}}(r)\mathsf{p}_{\theta}(\theta)$
2. *R has Rayleigh distribution* $\quad \mathsf{p}_{\mathsf{R}}(r) = \frac{r}{\sigma^2}\exp\frac{r^2}{-2\sigma^2}$
3. *Θ has uniform distribution* $\quad \mathsf{p}_{\theta}(\theta) = \frac{1}{2\pi}$

✎PROOF:

$$
\begin{aligned}
\mathsf{p}_{\mathsf{R},\Theta}(r,\theta) &= r\ \mathsf{p}_{\mathsf{XY}}(r\cos\theta, r\sin\theta) && \text{by Proposition 5.2 (page 34)}\\
&= r\ \mathsf{p}_{\mathsf{X}}(r\cos\theta)\ \mathsf{p}_{\mathsf{Y}}(r\sin\theta) && \text{by independence hypothesis}\\
&= r\ \frac{1}{\sqrt{2\pi\sigma^2}}\exp\frac{(r\cos\theta-0)^2}{-2\sigma^2}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\frac{(r\sin\theta-0)^2}{-2\sigma^2}\\
&= \frac{1}{2\pi\sigma^2}\ r\ \exp\frac{r^2(\cos^2\theta+\sin^2\theta)}{-2\sigma^2}\\
&= \frac{1}{2\pi\sigma^2}\ r\ \exp\frac{r^2}{-2\sigma^2}\\
&= \left[\frac{1}{2\pi}\right]\left[\frac{r}{\sigma^2}\exp\frac{r^2}{-2\sigma^2}\right]
\end{aligned}
$$

✏

Proposition 5.4. *Let* X *and* Y *be* RANDOM VARIABLES *with covariance* σ_{xy} *on a* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$.

PRP

$$
\left\{\begin{array}{ll}
(A). & \mathsf{X} \text{ is GAUSSIAN with } \mathsf{N}\left(\mu_{\mathsf{X}},\sigma_{\mathsf{X}}^2\right) \text{ and}\\
(B). & \mathsf{X} \text{ is GAUSSIAN with } \mathsf{N}\left(\mu_{\mathsf{Y}},\sigma_{\mathsf{Y}}^2\right) \text{ and}\\
(C). & \sigma_{\mathsf{xy}} = \mathrm{cov}\,[\mathsf{X},\mathsf{Y}]
\end{array}\right\} \Longrightarrow \left\{\mathsf{P}\{\mathsf{X}>\mathsf{Y}\} = \mathsf{Q}\left(\frac{-\mu_{\mathsf{X}}+\mu_{\mathsf{Y}}}{\sigma_{\mathsf{X}}^2+\sigma_{\mathsf{Y}}^2-2\sigma_{\mathsf{xy}}}\right)\right\}
$$

✎PROOF: Because X and Y are jointly Gaussian, their linear combination $\mathsf{Z} = rvX - \mathsf{Y}$ is also Gaussian. A Gaussian distribution is completely defined by its mean and variance. So, to determine the distribution of Z, we just have to determine the mean and variance of Z.

$$
\begin{aligned}
\mathsf{E}Z &= \mathsf{E}\mathsf{X} - \mathsf{E}Y\\
&= \mu_{\mathsf{X}} - \mu_{\mathsf{Y}}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{var}\,Z &= \mathsf{E}Z^2 - (\mathsf{E}Z)^2\\
&= \mathsf{E}(\mathsf{X}-\mathsf{Y})^2 - (\mathsf{E}\mathsf{X}-\mathsf{E}Y)^2\\
&= \mathsf{E}(\mathsf{X}^2 - 2XY + Y^2) - [(\mathsf{E}\mathsf{X})^2 - 2\mathsf{E}\mathsf{X}\mathsf{E}Y + (\mathsf{E}Y)^2]\\
&= [\mathsf{E}X^2 - (\mathsf{E}\mathsf{X})^2] + [Y^2 - (\mathsf{E}Y)^2] - 2[\mathsf{E}XY - \mathsf{E}\mathsf{X}\mathsf{E}Y]\\
&= \mathrm{var}\,\mathsf{X} + \mathrm{var}\,Y - 2\mathrm{cov}\,[\mathsf{X},\mathsf{Y}]\\
&\triangleq \sigma_{\mathsf{X}}^2 + \sigma_{\mathsf{Y}}^2 - 2\sigma_{\mathsf{xy}}
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{P}\{\mathsf{X}>\mathsf{Y}\} &= \mathsf{P}\{\mathsf{X}-\mathsf{Y}>0\}\\
&= \mathsf{P}\{Z>0\}\\
&= \mathsf{Q}\left(\frac{z-\mathsf{E}Z}{\mathrm{var}\,Z}\right)\Bigg|_{z=0}\\
&= \mathsf{Q}\left(\frac{0-\mu_{\mathsf{X}}+\mu_{\mathsf{Y}}}{\sigma_{\mathsf{X}}^2+\sigma_{\mathsf{Y}}^2-2\sigma_{\mathsf{xy}}}\right)
\end{aligned}
$$

✏

COPYRIGHT © 2019 DANIEL J. GREENHOE

Figure 5.6: The distributions of sums of independent uniformly distributed random variables (Example 5.1 page 36)

Example 5.1. Let $(\!(X_1, X_2, X_3, \ldots)\!)$ be a *sequence* of *independent* (Definition A.3 page 167) *uniformly distributed* random variables. Let $p_N(x)$ be the *probability density function* of $Y \triangleq \sum_{n=1}^{N} X_n$. Some of these distributions are illustrated in Figure 5.6 (page 36). Note that the distributions of the sequence $(\!(p_1, p_2, p_3, \ldots)\!)$ are all *B-splines* (Definition S.2 page 383) and all form a *partition of unity*.

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 6

OPERATORS ON DISCRETE RANDOM SEQUENCES

6.1 LTI operators on random sequences

Theorem 6.1. [1] *Let* $\mathsf{x}(n)$ *be a* RANDOM SEQUENCE *with* MEAN μ_{X} *and* $\mathsf{y}(n)$ *a* RANDOM SEQUENCE *with* MEAN μ_{Y}. *Let* $\mathbf{S}$ *be a system with* IMPULSE RESPONSE $\mathsf{h}(n)$, INPUT $\mathsf{x}(n)$, *and* OUTPUT $\mathsf{y}(n)$.

THM

$$\left\{ \mathbf{S} \text{ is } (LTI) \right\} \Longrightarrow \left\{ \begin{array}{llll} (1). & \mu_{\mathsf{Y}}(n) & = \displaystyle\sum_{k\in\mathbb{Z}} \mathsf{h}(k)\mu_{\mathsf{X}}(n-k) & \triangleq \mathsf{h}(n) \star \mu_{\mathsf{X}}(n) \quad \text{and} \\ (2). & \mathsf{R}_{\mathsf{xy}}(n,m) & = \displaystyle\sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{R}_{\mathsf{xx}}(n-k,m+k) & \quad \text{and} \\ (3). & \mathsf{R}_{\mathsf{yy}}(n,m) & = \displaystyle\sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{R}_{\mathsf{yx}}(n-k,m+k) & \end{array} \right\}$$

✎PROOF:

$$\begin{aligned}
\mu_{\mathsf{Y}}(n) &\triangleq \mathsf{E}\left[\mathsf{y}(n)\right] && \text{by definition of } \mu_{\mathsf{Y}} && \text{(Definition 2.2 page 11)} \\
&= \mathsf{E}\left[\sum_{k\in\mathbb{Z}} \mathsf{h}(k)\mathsf{x}(n-k)\right] && \text{by } LTI \text{ hypothesis} \\
&= \sum_{k\in\mathbb{Z}} \mathsf{h}(k)\mathsf{E}[\mathsf{x}(n-k)] && \text{by } \textit{linear} \text{ property} \\
&= \sum_{k\in\mathbb{Z}} \mathsf{h}(k)\mu_{\mathsf{X}}(n-k) && \text{by definition of } \mu_{\mathsf{X}} && \text{(Definition 2.2 page 11)} \\
&\triangleq \mathsf{h}(n) \star \mu_{\mathsf{X}}(n) && \text{by definition of } \textit{convolution} && \text{(Definition R.3 page 369)}
\end{aligned}$$

$$\begin{aligned}
\mathsf{R}_{\mathsf{xy}}(n,m) &\triangleq \mathsf{E}\left[\mathsf{x}(n+m)\mathsf{y}^*(n)\right] && \text{by definition of } \mathsf{R}_{\mathsf{xy}}(n,m) && \text{(Definition 2.2 page 11)} \\
&= \mathsf{E}\left[\mathsf{x}(n+m)(\mathsf{h}(n) \star \mathsf{x}(n))^*\right] && \text{by } LTI \text{ hypothesis} \\
&\triangleq \mathsf{E}\left[\mathsf{x}(n+m)\left(\sum_{k\in\mathbb{Z}} h(k)\mathsf{x}(n-k)\right)^*\right] && \text{by definition of } \textit{convolution} \star && \text{(Definition R.3 page 369)} \\
&= \mathsf{E}\left[\mathsf{x}(n+m)\sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{x}^*(n-k)\right] && \text{by } \textit{distributive} \text{ property of } *\textbf{-algebras} && \text{(Definition J.3 page 266)}
\end{aligned}$$

[1] Papoulis (1991), page 310

$$= \mathsf{E}\left[\sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{x}(n+m)\mathsf{x}^*(n-k)\right]$$ by *distributive* property of $(\mathbb{C},+,\cdot,0,1)$ (Definition F.5 page 208)

$$= \sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{E}\left[\mathsf{x}(n-k+k+m)\mathsf{x}^*(n-k)\right]$$ by *linear* property of E (Theorem 1.1 page 4)

$$\triangleq \sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{R}_{\mathsf{xx}}(n-k,m+k)$$ by definition of $\mathsf{R}_{\mathsf{xx}}(n,m)$ (Definition 2.2 page 11)

$$\mathsf{R}_{\mathsf{yy}}(n,m) \triangleq \mathsf{E}\left[\mathsf{y}(n+m)\mathsf{y}^*(n)\right]$$ by definition of $\mathsf{R}_{\mathsf{xy}}(n,m)$ (Definition 2.2 page 11)

$$= \mathsf{E}\left[\mathsf{y}(n+m)(\mathsf{h}(n)\star\mathsf{x}(n))^*\right]$$ by *LTI* hypothesis

$$\triangleq \mathsf{E}\left[\mathsf{y}(n+m)\left(\sum_{k\in\mathbb{Z}} \mathsf{h}(k)\mathsf{x}(n-k)\right)^*\right]$$ by definition of *convolution* (Definition R.3 page 369)

$$= \mathsf{E}\left[\mathsf{y}(n+m)\sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{x}^*(n-k)\right]$$ by *distributive* property of $*$-**algebras** (Definition J.3 page 266)

$$= \mathsf{E}\left[\sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{y}(n+m)\mathsf{x}^*(n-k)\right]$$ by *distributive* property of $(\mathbb{C},+,\cdot,0,1)$ (Definition F.5 page 208)

$$= \sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{E}\left[\mathsf{y}(n-k+k+m)\mathsf{x}^*(n-k)\right]$$ by *linear* property of E (Theorem 1.1 page 4)

$$\triangleq \sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{R}_{\mathsf{yx}}(n-k,m+k)$$ by definition of $\mathsf{R}_{\mathsf{xy}}(n,m)$ (Definition 2.2 page 11)

✏

6.2 LTI operators on WSS random sequences

Corollary 6.1. *Let* $\mathbf{S}$ *be the system defined in Theorem 6.1 (page 37).*

COR

$$\left.\begin{array}{ll}(A). & \mathbf{S}\ \textit{is}\ \text{LTI} \quad \textit{and}\\ (B). & \mathsf{x}(n)\ \textit{is}\ \text{WSS}\end{array}\right\} \Longrightarrow \left\{\begin{array}{lll}(1). & \mu_{\mathsf{Y}} = \mu_{\mathsf{X}}\sum_{n\in\mathbb{Z}}\mathsf{h}(k) & \textit{and}\\ (2). & \mathsf{R}_{\mathsf{xy}}(m) = \mathsf{R}_{\mathsf{xx}}(m)\star\mathsf{h}^*(-m) \triangleq \sum_{k\in\mathbb{Z}}\mathsf{h}^*(k)\mathsf{R}_{\mathsf{xx}}(m+k) & \textit{and}\\ (3). & \mathsf{R}_{\mathsf{yy}}(m) = \mathsf{R}_{\mathsf{yx}}(m)\star\mathsf{h}^*(-m) \triangleq \sum_{k\in\mathbb{Z}}\mathsf{h}^*(k)\mathsf{R}_{\mathsf{xy}}(m+k) & \textit{and}\\ (4). & \mathsf{R}_{\mathsf{yy}}(m) = \mathsf{R}^*_{\mathsf{xx}}(m)\star\mathsf{h}(-m)\star\mathsf{h}^*(-m) & \end{array}\right.$$

✎PROOF:

$$\mu_{\mathsf{Y}} = \mu_{\mathsf{Y}}(n)$$ by Proposition 2.1 page 13 and hypothesis (A)

$$= \sum_{n\in\mathbb{Z}} \mathsf{h}(k)\mu_{\mathsf{X}}(n-k)$$ by Theorem 2.1 page 12 and hypothesis (B)

$$= \sum_{n\in\mathbb{Z}} \mathsf{h}(k)\mu_{\mathsf{X}}(0)$$ by Definition 7.1 page 45 and hypothesis (B)

$$= \mu_{\mathsf{X}}(0)\sum_{n\in\mathbb{Z}} \mathsf{h}(k)$$ by *linear* property of $\sum$

$$= \mu_{\mathsf{X}}\sum_{n\in\mathbb{Z}} \mathsf{h}(k)$$ by Proposition 2.1 page 13

[1] Papoulis (1991), page 323

COPYRIGHT © 2019 DANIEL J. GREENHOE

$\mathsf{R}_{\mathsf{xy}}(m) \triangleq \mathsf{R}_{\mathsf{xy}}(0,m)$	by Proposition 2.1 page 13	and hypothesis (A)
$= \sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{R}_{\mathsf{xx}}(0-k,m+k)$	by Theorem 6.1 page 37	and hypothesis (B)
$= \sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{R}_{\mathsf{xx}}(m+k)$	by Proposition 2.1 page 13	and hypothesis (A)
$= \mathsf{h}^*(-m) \star \mathsf{R}_{\mathsf{xx}}(m)$	by Proposition R.2 page 370	
$\mathsf{R}_{\mathsf{yy}}(m) \triangleq \mathsf{R}_{\mathsf{yy}}(0,m)$	by Proposition 2.1 page 13	and hypothesis (A)
$= \sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{R}_{\mathsf{yx}}(n-k,m+k)$	by Theorem 6.1 page 37	and hypothesis (B)
$= \sum_{k\in\mathbb{Z}} \mathsf{h}^*(k)\mathsf{R}_{\mathsf{yx}}(m+k)$	by Proposition 2.1 page 13	and hypothesis (A)
$= \mathsf{h}^*(-m) \star \mathsf{R}_{\mathsf{yx}}(m)$	by Proposition R.2 page 370	
$\mathsf{R}_{\mathsf{yy}}(m) = \mathsf{h}^*(-m) \star \mathsf{R}_{\mathsf{yx}}(m)$	by result (2)	
$= \mathsf{h}^*(-m) \star \mathsf{R}^*_{\mathsf{xy}}(m)$	by Corollary 2.1 page 14	
$= \mathsf{h}^*(-m) \star \left[\mathsf{h}^*(-m) \star \mathsf{R}_{\mathsf{xx}}(m)\right]^*$	by result (1)	
$= \mathsf{h}^*(-m) \star \mathsf{h}(-m) \star \mathsf{R}^*_{\mathsf{xx}}(m)$	by *distributive* property of $*$-**algebras**	(Definition J.3 page 266)

✏

Corollary 6.2. [2] *Let* **S** *be a system with* IMPULSE RESPONSE $\mathsf{h}(n)$, INPUT $\mathsf{x}(n)$, *and* OUTPUT $\mathsf{y}(n)$.

COR

$$\left\{\begin{array}{ll} (A). & \mathsf{h} \text{ is LINEAR TIME INVARIANT} \quad \text{and} \\ (B). & \mathsf{x} \text{ and } \mathsf{y} \text{ are WIDE SENSE STATIONARY} \end{array}\right\} \Longrightarrow \left\{\begin{array}{lll} (1). & \check{\mathsf{S}}_{\mathsf{xy}}(z) = \check{\mathsf{S}}_{\mathsf{xx}}(z)\check{\mathsf{H}}^*\left(\frac{1}{z^*}\right) & \text{and} \\ (2). & \check{\mathsf{S}}_{\mathsf{yy}}(z) = \check{\mathsf{S}}_{\mathsf{yx}}(z)\check{\mathsf{H}}^*\left(\frac{1}{z^*}\right) & \text{and} \\ (3). & \check{\mathsf{S}}_{\mathsf{yy}}(z) = \check{\mathsf{S}}_{\mathsf{xx}}(z)\check{\mathsf{H}}(z)\check{\mathsf{H}}^*\left(\frac{1}{z^*}\right) & \end{array}\right\}$$

✎PROOF: The proof is given in Proposition **??** (page **??**) (1). ✏

Corollary 6.3. *Let* **S** *be a system with* IMPULSE RESPONSE $\mathsf{h}(n)$, INPUT $\mathsf{x}(n)$, *and* OUTPUT $\mathsf{y}(n)$.

COR

$$\left\{\begin{array}{lll} (A). & \mathsf{h} \text{ is} & \text{LTI} \quad \text{and} \\ (B). & \mathsf{x} \text{ and } \mathsf{y} \text{ are} & \text{WSS} \end{array}\right\} \Longrightarrow \left\{\begin{array}{lll} (1). & \tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) = \tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{H}}^*(\omega) & \text{and} \\ (2). & \tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) = \tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\tilde{\mathsf{H}}(\omega) & \text{and} \\ (3). & \tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) = \tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\left|\tilde{\mathsf{H}}(\omega)\right|^2 & \end{array}\right\}$$

✎PROOF: The proof is given in Proposition **??** (page **??**) (1). ✏

6.3 Parallel operators on WSS random sequences

Theorem 6.2. *Let* **S** *be the* SYSTEM *illustrated to the right, where* **T** *is* NOT NECESSARILY LINEAR. *Let*

$$(\!(\mathsf{h}(n))\!) \triangleq \mathbf{H}\bar{\delta}(n) \triangleq \sum_{m\in\mathbb{Z}} \mathsf{h}(m)\bar{\delta}(n-m)$$

be the IMPULSE RESPONSE *of* **H**.

[2] Papoulis (1991), page 323

THM

$$\left\{\begin{array}{lll}\text{(A).} & \mathsf{x}(n) \text{ is} & \textit{WSS} \quad \text{and} \\ \text{(B).} & \mathbf{H} \text{ is} & \textit{LTI}\end{array}\right\} \Longrightarrow \left\{\begin{array}{llll}\text{(1).} & \mathsf{R}_{\mathsf{wy}}(m) & = & \sum_{n\in\mathbb{Z}} \mathsf{h}(n)\mathsf{R}_{\mathsf{xy}}(m-n) \quad (\textit{convolution}) \\ & & \triangleq & \mathsf{h}(m) \star \mathsf{R}_{\mathsf{xy}}(m) \quad \text{and} \\ \text{(2).} & \check{\mathsf{S}}_{\mathsf{wy}}(z) & = & \check{\mathsf{H}}(z)\check{\mathsf{S}}_{\mathsf{xy}}(z) \quad \text{and} \\ \text{(3).} & \tilde{\mathsf{S}}_{\mathsf{wy}}(\omega) & = & \tilde{\mathsf{H}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\end{array}\right\}$$

✎PROOF:

$\mathsf{R}_{\mathsf{wy}}(m) \triangleq \mathsf{E}\left[\mathsf{w}(m)\mathsf{y}^*(0)\right]$	by (A) and definition of R_{wy}	(Definition 2.4 page 12)
$\triangleq \mathsf{E}\left([\mathbf{H}\mathsf{x}](m)\mathsf{y}^*(0)\right)$	by definition of $\mathbf{S}$	
$= \mathbf{H}\mathsf{E}\left(\mathsf{x}(m)\mathsf{y}^*(0)\right)$	by *LTI* hypothesis	(B)
$\triangleq \mathbf{H}\mathsf{R}_{\mathsf{xy}}(m)$	by definition of R_{xy}	(Definition 2.4 page 12)
$= \sum_{n\in\mathbb{Z}} \mathsf{h}(n)\mathsf{R}_{\mathsf{xy}}(m-n)$	by definition of $\mathbf{H}$ *impulse response* $(\!(\mathsf{h}(n))\!)$	
$= \left[\mathsf{h}(m) \star \mathsf{R}_{\mathsf{xy}}(m)\right]$	by definition of *convolution*	(Definition R.3 page 369)
$\check{\mathsf{S}}_{\mathsf{wy}}(z) \triangleq \mathbf{Z}\mathsf{R}_{\mathsf{wy}}(m)$	by definition of $\check{\mathsf{S}}_{\mathsf{wy}}$	(Definition 2.5 page 14)
$= \left[\mathsf{h}(m) \star \mathsf{R}_{\mathsf{xy}}(m)\right]$	by previous result	
$= \check{\mathsf{H}}(z)\check{\mathsf{S}}_{\mathsf{xy}}(z)$	by *Convolution Theorem*	(Theorem R.2 page 372)
$\tilde{\mathsf{S}}_{\mathsf{wy}}(\omega) \triangleq \check{\mathbf{F}}\mathsf{R}_{\mathsf{wy}}(m)$	by definition of $\tilde{\mathsf{S}}_{\mathsf{wy}}$	(Definition 7.3 page 46)
$= \left[\mathsf{h}(m) \star \mathsf{R}_{\mathsf{xy}}(m)\right]$	by previous result	
$= \tilde{\mathsf{H}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)$	by *Convolution Theorem*	(Theorem R.2 page 372)

✏

6.4 Whitening discrete random sequences

Figure 6.1: Poles (×) and zeros (o) of a *minimum phase* filter

Definition 6.1. *Let $\check{\mathsf{H}}(z)$ be the z-transform of the impulse response of a filter. If $\check{\mathsf{H}}(z)$ can be expressed as a rational expression with poles and zeros $r_n e^{i\theta_n}$, then the filter is* ***minimum phase*** *if each $r_n < 1$ (all roots lie inside the unit circle in the complex z-plane).*

See Figure 6.1 page 40.

Note that if $\mathsf{L}(z)$ has a root at $z = re^{i\theta}$, then $\mathsf{L}^*(1/z^*)$ has a root at

$$\frac{1}{z^*} = \frac{1}{\left(re^{i\theta}\right)^*} = \frac{1}{re^{-i\theta}} = \frac{1}{r}e^{i\theta}.$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

That is, if $\mathsf{L}(z)$ has a root inside the unit circle, then $\mathsf{L}^*(1/z^*)$ has a root directly opposite across the unit circle boundary (see Figure 6.2 page 41). A causal stable filter $\check{\mathsf{H}}(z)$ must have all of its poles inside the unit circle. A minimum phase filter is a filter with both its poles and zeros inside the unit circle. One advantage of a minimum phase filter is that its recipricol (zeros become poles and poles become zeros) is also causal and stable.

Figure 6.2: Mirrored roots in complex-z plane

Figure 6.3: Innovations and whitening filters

The next theorem demonstrates a method for "whitening" a *random sequence* $\mathsf{x}(n)$ with a filter constructed from a decomposition of $\mathsf{R}_{\mathsf{xx}}(m)$. The technique is stated precisely in Theorem 6.3 page 41 and illustrated in Figure 6.3 page 41. Both imply two filters with impulse responses $l(n)$ and $\gamma(n)$. Filter $l(n)$ is referred to as the **innovations filter** (because it generates or "innovates" $\mathsf{x}(n)$ from a white noise process $\mathsf{w}(n)$) and $\gamma(n)$ is referred to as the **whitening filter** because it produces a white noise sequence when the input sequence is $\mathsf{x}(n)$.[3]

Theorem 6.3. *Let* $\mathsf{x}(n)$ *be a WSS* RANDOM SEQUENCE *with auto-correlation* $\mathsf{R}_{\mathsf{xx}}(m)$ *and spectral density* $\check{\mathsf{S}}_{\mathsf{xx}}(z)$. ***If*** $\check{\mathsf{S}}_{\mathsf{xx}}(z)$ *has a* ***rational expression****, then the following are true:*

1. *There exists a rational expression* $\mathsf{L}(z)$ *with minimum phase such that*
$$\check{\mathsf{S}}_{\mathsf{xx}}(z) = \mathsf{L}(z)\mathsf{L}^*\left(\frac{1}{z^*}\right).$$
2. *An LTI filter for which the Laplace transform of the impulse response* $\gamma(n)$ *is*
$$\Gamma(z) = \frac{1}{\mathsf{L}(z)}$$
is both causal and stable.
3. *If* $\mathsf{x}(n)$ *is the input to the filter* $\gamma(n)$*, the output* $\mathsf{y}(n)$ *is a* ***white noise sequence*** *such that*
$$\mathsf{S}_{\mathsf{yy}}(z) = 1 \qquad \mathsf{R}_{\mathsf{yy}}(m) = \bar{\delta}(m).$$

[3] Papoulis (1991), pages 401–402

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

✎PROOF:

$$
\begin{aligned}
\mathsf{S}_{\mathsf{ww}}(z) &= \Gamma(z)\Gamma^*\left(\frac{1}{z^*}\right)\check{\mathsf{S}}_{\mathsf{xx}}(z)\\
&= \frac{1}{\mathsf{L}(z)}\frac{1}{\mathsf{L}^*\left(\frac{1}{z^*}\right)}\check{\mathsf{S}}_{\mathsf{xx}}(z)\\
&= \frac{1}{\mathsf{L}(z)}\frac{1}{\mathsf{L}^*\left(\frac{1}{z^*}\right)}\mathsf{L}(z)\mathsf{L}^*\left(\frac{1}{z^*}\right)\\
&= 1
\end{aligned}
$$

✏

COPYRIGHT © 2019 DANIEL J. GREENHOE

CHAPTER 7

OPERATORS ON CONTINUOUS RANDOM SEQUENCES

7.1 LTI Operations on non-stationary random processes

Figure 7.1: Linear system with random process input and output

Theorem 7.1. [1] *Let* $\mathsf{h} : \mathbb{R} \to \mathbb{C}$ *be the impulse response of a linear time-invariant system and Let* $\mathsf{y}(t) = h(t) \star \mathsf{x}(t) \triangleq \int_{u\in\mathbb{R}} h(u)\mathsf{x}(t-u)\,\mathrm{d}u$ *as illustrated in Figure 7.1 page 43. Then*

THM

Correlation functions

$$\begin{aligned}
\mathsf{R}_{\mathsf{xy}}(t,u) &= \mathsf{R}_{\mathsf{xx}}(t,u) \star h^*(u) &&\triangleq \int_v \mathsf{h}^*(v)\mathsf{R}_{\mathsf{xx}}(t,u-v)\,\mathrm{d}v \\
\mathsf{R}_{\mathsf{yy}}(t,u) &= \mathsf{R}_{\mathsf{xy}}(t,u) \star \mathsf{h}(t) &&\triangleq \int_v \mathsf{h}(v)\mathsf{R}_{\mathsf{xy}}(t-v,u)\,\mathrm{d}v \\
\mathsf{R}_{\mathsf{yy}}(t,u) &= \mathsf{R}_{\mathsf{xx}}(t,u) \star \mathsf{h}(t) \star \mathsf{h}^*(u) &&\triangleq \int_w \mathsf{h}^*(w) \int_v \mathsf{h}(v)\mathsf{R}_{\mathsf{xx}}(t-v,u-w)\,\mathrm{d}v\,\mathrm{d}\omega
\end{aligned}$$

Laplace power spectral density functions

$$\begin{aligned}
\check{S}_{\mathsf{xy}}(s,r) &= \check{S}_{\mathsf{xx}}(s,r)\check{\mathsf{h}}^*(r^*) \\
\check{S}_{\mathsf{yy}}(s,r) &= \check{S}_{\mathsf{xy}}(s,r)\check{\mathsf{h}}(s) \\
\check{S}_{\mathsf{yy}}(s,r) &= \check{S}_{\mathsf{xx}}(s,r)\check{\mathsf{h}}(s)\check{\mathsf{h}}^*(r^*)
\end{aligned}$$

Power spectral density functions

$$\begin{aligned}
\mathsf{S}_{\mathsf{xy}}(f,g) &= \mathsf{S}_{\mathsf{xx}}(f,g)\tilde{\mathsf{h}}^*(-g) \\
\mathsf{S}_{\mathsf{yy}}(f,g) &= \mathsf{S}_{\mathsf{xy}}(f,g)\tilde{\mathsf{h}}(\omega) \\
\mathsf{S}_{\mathsf{yy}}(f,g) &= \mathsf{S}_{\mathsf{xx}}(f,g)\tilde{\mathsf{h}}(\omega)\tilde{\mathsf{h}}^*(-g)
\end{aligned}$$

[1] Papoulis (1991), page 312

✎PROOF:

$$
\begin{aligned}
\mathsf{R}_{\mathsf{xy}}(t,u) &\triangleq \mathsf{E}\big[\mathsf{x}(t)\mathsf{y}^*(u)\big] \\
&= \mathsf{E}\left[\mathsf{x}(t)\left(\int_v \mathsf{h}(v)\mathsf{x}(u-v)\,\mathrm{d}v\right)^*\right] \\
&= \mathsf{E}\left[\mathsf{x}(t)\int_v \mathsf{h}^*(v)\mathsf{x}^*(u-v)\,\mathrm{d}v\right] \\
&= \int_v \mathsf{h}^*(v)\mathsf{E}\big[\mathsf{x}(t)\mathsf{x}^*(u-v)\big]\,\mathrm{d}v \\
&= \int_v \mathsf{h}^*(v)\mathsf{R}_{\mathsf{xx}}(t,u-v)\,\mathrm{d}v \\
&\triangleq \mathsf{R}_{\mathsf{xx}}(t,u) \star h^*(u)
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{R}_{\mathsf{yy}}(t,u) &\triangleq \mathsf{E}\big[\mathsf{y}(t)\mathsf{y}^*(u)\big] \\
&= \mathsf{E}\left[\left(\int_v \mathsf{h}(v)\mathsf{x}(t-v)\,\mathrm{d}v\right)\mathsf{y}^*(u)\right] \\
&= \int_v \mathsf{h}(v)\mathsf{E}\big[\mathsf{x}(t-v)\mathsf{y}^*(u)\big]\,\mathrm{d}v \\
&= \int_v \mathsf{h}(v)\mathsf{R}_{\mathsf{xy}}(t-v,u)\,\mathrm{d}v \\
&\triangleq \mathsf{R}_{\mathsf{xy}}(t,u) \star \mathsf{h}(t)
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{R}_{\mathsf{yy}}(t,u) &\triangleq \mathsf{E}\big[\mathsf{y}(t)\mathsf{y}^*(u)\big] \\
&= \mathsf{E}\left[\left(\int_v \mathsf{h}(v)\mathsf{x}(t-v)\,\mathrm{d}v\right)\left(\int_w \mathsf{h}(w)\mathsf{x}(u-w)\,\mathrm{d}\omega\right)^*\right] \\
&= \int_w \mathsf{h}^*(w)\int_v \mathsf{h}(v)\mathsf{E}\big[\mathsf{x}(t-v)\mathsf{x}^*(u-w)\big]\,\mathrm{d}v\,\mathrm{d}\omega \\
&= \int_w \mathsf{h}^*(w)\int_v \mathsf{h}(v)\mathsf{R}_{\mathsf{xx}}(t-v,u-w)\,\mathrm{d}v\,\mathrm{d}\omega \\
&= \int_w \mathsf{h}^*(w)\big[\mathsf{R}_{\mathsf{xx}}(t,u-w) \star \mathsf{h}(t)\big]\,\mathrm{d}\omega \\
&\triangleq \mathsf{R}_{\mathsf{xx}}(t,u) \star \mathsf{h}(t) \star \mathsf{h}^*(u)
\end{aligned}
$$

$$
\begin{aligned}
\check{S}_{\mathsf{xy}}(s,r) &\triangleq \mathbf{L}\mathsf{R}_{\mathsf{xy}}(t,u) \\
&= \mathbf{L}[\mathsf{R}_{\mathsf{xx}}(t,u) \star \mathsf{h}^*(u)] \\
&= \mathbf{L}[\mathsf{R}_{\mathsf{xx}}(t,u)]\mathbf{L}[\mathsf{h}^*(u)] \\
&= \check{S}_{\mathsf{xx}}(s,r)\int_{u\in\mathbb{R}} \mathsf{h}^*(u)e^{-ru}\,\mathrm{d}u \\
&= \check{S}_{\mathsf{xx}}(s,r)\left[\int_{u\in\mathbb{R}} \mathsf{h}(u)e^{-r^*u}\,\mathrm{d}u\right]^* \\
&= \check{S}_{\mathsf{xx}}(s,r)\check{\mathsf{h}}^*(r^*)
\end{aligned}
$$

$$
\begin{aligned}
\check{S}_{\mathsf{yy}}(s,r) &\triangleq \mathbf{L}\mathsf{R}_{\mathsf{yy}}(t,u) \\
&= \mathbf{L}[\mathsf{R}_{\mathsf{xy}}(t,u) \star \mathsf{h}(t)] \\
&= \mathbf{L}[\mathsf{R}_{\mathsf{xy}}(t,u)]\mathbf{L}[\mathsf{h}(t)] \\
&= \check{S}_{\mathsf{xy}}(s,r)\check{\mathsf{h}}(s)
\end{aligned}
$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$
\begin{aligned}
&= \check{S}_{xy}(s,r)\check{h}(s)\\
&= \check{S}_{xx}(s,r)\check{h}^*(r^*)\check{h}(s)\\
&= \check{S}_{xx}(s,r)\check{h}(s)\check{h}^*(r^*)
\end{aligned}
$$

$$
\begin{aligned}
S_{xy}(f,g) &\triangleq \tilde{\mathbf{F}}R_{xy}(t,u)\\
&= \tilde{\mathbf{F}}[R_{xx}(t,u) \star h^*(u)]\\
&= \tilde{\mathbf{F}}[R_{xx}(t,u)]\tilde{\mathbf{F}}[h^*(u)]\\
&= S_{xx}(f,g)\int_{u\in\mathbb{R}} h^*(u)e^{-i2\pi gu}\,du\\
&= S_{xx}(f,g)\left[\int_{u\in\mathbb{R}} h(u)e^{i2\pi gu}\,du\right]^*\\
&= S_{xx}(f,g)\left[\int_{u\in\mathbb{R}} h(u)e^{-i2\pi(-g)u}\,du\right]^*\\
&= S_{xx}(f,g)\tilde{h}^*(-g)
\end{aligned}
$$

$$
\begin{aligned}
S_{yy}(f,g) &\triangleq \tilde{\mathbf{F}}R_{yy}(t,u)\\
&= \tilde{\mathbf{F}}[R_{xy}(t,u) \star h(t)]\\
&= \tilde{\mathbf{F}}[R_{xy}(t,u)]\tilde{\mathbf{F}}[h(t)]\\
&= S_{xy}(f,g)\tilde{h}(\omega)
\end{aligned}
$$

$$
\begin{aligned}
&= S_{xy}(f,g)\tilde{h}(\omega)\\
&= S_{xx}(f,g)\tilde{h}^*(-g)\tilde{h}(\omega)
\end{aligned}
$$

✍

7.2 LTI Operations on WSS random processes

Definition 7.1.

DEF *A random process* $x(t)$ *is* ***wide sense stationary*** *(WSS) if*

(1). $\mu_X(t)$ *is* CONSTANT *with respect to* t (STATIONARY IN THE MEAN) *and*

(2). $R_{xx}(t+\tau,t)$ *is* CONSTANT *with respect to* t (STATIONARY IN CORRELATION)

If a process $x(t)$ is *wide sense stationary*, mean and correlation are often written μ_X and $R_{xx}(\tau)$, respectively. If a pair of processes $x(t)$ and $y(t)$ are *WSS*, then their cross-correlation is commonly written $R_{xy}(\tau)$.

Definition 7.2. *Let* $x(t)$ *and* $y(t)$ *be WSS random processes.*

DEF

$$
\begin{aligned}
\check{S}_{xx}(s) &\triangleq \mathbf{L}R_{xx}(\tau) \triangleq \int_{\tau\in\mathbb{R}} R_{xx}(\tau)e^{-s\tau}\,d\tau\\
\check{S}_{yy}(s) &\triangleq \mathbf{L}R_{yy}(\tau) \triangleq \int_{\tau\in\mathbb{R}} R_{yy}(\tau)e^{-s\tau}\,d\tau\\
\check{S}_{xy}(s) &\triangleq \mathbf{L}R_{xy}(\tau) \triangleq \int_{\tau\in\mathbb{R}} R_{xy}(\tau)e^{-s\tau}\,d\tau\\
\check{S}_{yx}(s) &\triangleq \mathbf{L}R_{yx}(\tau) \triangleq \int_{\tau\in\mathbb{R}} R_{yx}(\tau)e^{-s\tau}\,d\tau
\end{aligned}
$$

Definition 7.3. *Let* $\mathsf{x}(t)$ *and* $\mathsf{y}(t)$ *be WSS random processes.*

DEF

$$\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) \triangleq [\tilde{\mathbf{F}}\mathsf{R}_{\mathsf{xx}}(\tau)](\omega) \triangleq \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{xx}}(\tau)e^{-i\omega\tau}\,\mathrm{d}\tau$$
$$\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) \triangleq [\tilde{\mathbf{F}}\mathsf{R}_{\mathsf{yy}}(\tau)](\omega) \triangleq \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{yy}}(\tau)e^{-i\omega\tau}\,\mathrm{d}\tau$$
$$\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) \triangleq [\tilde{\mathbf{F}}\mathsf{R}_{\mathsf{xy}}(\tau)](\omega) \triangleq \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{xy}}(\tau)e^{-i\omega\tau}\,\mathrm{d}\tau$$
$$\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega) \triangleq [\tilde{\mathbf{F}}\mathsf{R}_{\mathsf{yx}}(\tau)](\omega) \triangleq \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{yx}}(\tau)e^{-i\omega\tau}\,\mathrm{d}\tau$$

Definition 7.4. [2] *Let* $\mathsf{x}(t)$ *be a random variable that is* STATIONARY IN THE MEAN *such that* $\mathsf{E}[\mathsf{x}(t)]$ *is constant with respect to* t.

DEF

$\mathsf{x}(t)$ *is* ***ergodic in the mean*** *if*

$$\underbrace{\mathsf{E}[\mathsf{x}(t)]}_{\text{ENSEMBLE AVERAGE}} = \lim_{\tau\to\infty} \underbrace{\frac{1}{2\tau}\int_{-\tau}^{+\tau} \mathsf{x}(t)\,\mathrm{d}t}_{\text{TIME AVERAGE}}$$

Proposition 7.1.

PRP

$$\{\ \mathsf{x}(t) \text{ is NON-STATIONARY}\ \} \implies \{\ \mathsf{x}(t) \text{ is NOT ERGODIC IN THE MEAN}\ \}$$

✎PROOF: If $\mathsf{x}(t)$ is *non-stationary*, then $\mathsf{E}[\mathsf{x}(t)]$ is not constant with time. But $\lim_{\tau\to\infty}\frac{1}{2\tau}\int_{-\tau}^{+\tau}\mathsf{x}(t)\,\mathrm{d}t$ must be a constant (if it is *convergent*). ✏

Definition 7.5. [3] *Let* $\mathsf{x}(t)$ *be a* WIDE SENSE STATIONARY *random process.*

DEF

(1). *The* ***average power*** $\mathsf{P}_{\mathsf{avg}}$ *of* $\mathsf{x}(t)$ *is* $\quad \mathsf{P}_{\mathsf{avg}}\mathsf{x}(t) \triangleq \lim_{\tau\to\infty}\frac{1}{2\tau}\int_{t\in\mathbb{R}} |\mathsf{x}(t)|^2\,\mathrm{d}t$

(2). *The* ***energy spectral density*** $|\tilde{\mathsf{x}}(\omega)|^2$ *of* $\mathsf{x}(t)$ *is* $\quad |\tilde{\mathsf{x}}(\omega)|^2 \triangleq \left|\tilde{\mathbf{F}}\mathsf{x}(t)\right|^2$

Remark 7.1 (**spectral power**). Why does $\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)$ deserve the name *power spectral density*? This is answered by Theorem 7.2 (next). But to elaborate further, note that $\tilde{\mathsf{S}}_{\mathsf{xx}}$ is the spectral representation of the statistical relationship (the *variance*) between samples of $\mathsf{x}(t)$. For example, if there is no relationship, then $\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) = 1$. But in the case that $\mathsf{x}(t)$ is *ergodic in the mean*, then $\tilde{\mathsf{S}}_{\mathsf{xx}}$ takes on an additional meaning—it describes the "spectral power" present in $\mathsf{x}(t)$. This is demonstrated by the next theorem.

Theorem 7.2. *Let* $\mathsf{x}(t)$ *be a* RANDOM PROCESS.

THM

$$\left\{\begin{array}{ll} (A). & \mathsf{x}(t) \text{ is ERGODIC IN THE MEAN} \quad \textit{and} \\ (B). & \tilde{\mathsf{x}}(\omega) \text{ EXISTS} \end{array}\right\} \implies \left\{\begin{array}{lll} (1). & \tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) = \lim_{\tau\to\infty}\frac{1}{2\tau}\underbrace{|\tilde{\mathsf{x}}(\omega)|^2}_{\text{(ESD)}} & \textit{and} \\ (2). & \mathsf{P}_{\mathsf{avg}}[\mathsf{x}(t)] = \int_{\omega\in\mathbb{R}} \tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\,\mathrm{d}\omega & \end{array}\right\}$$

[2] Papoulis (1984) page 246 ⟨Mean-Ergodic processes⟩, Papoulis (2002) page 523 ⟨12-1 ERGODICITY⟩, KAY (1988) PAGE 58 ⟨3.6 ERGODICITY OF THE AUTOCORRELATION FUNCTION⟩, MANOLAKIS ET AL. (2005) PAGE 106 ⟨ERGODIC RANDOM PROCESSES⟩, KOOPMANS (1995) PAGES 53–61, CADZOW (1987) PAGE 378 ⟨11.13 ERGODIC TIME SERIES⟩, HELSTROM (1991) PAGE 336

[3] ? page 177

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$
\begin{aligned}
\tilde{S}_{xx}(\omega) &\triangleq \int_{\tau\in\mathbb{R}} R_{xx}(\tau)e^{-i\omega\tau}\,d\tau && \text{by definition of } \tilde{S}_{xx}(\omega) && \text{(Definition 7.3 page 46)}\\
&= \int_{\tau\in\mathbb{R}} E\big[x(t+\tau)x^*(t)\big]e^{-i\omega\tau}\,d\tau && \text{by definition of } R_{xx}(t) && \text{(Definition 3.2 page 18)}\\
&= E\left[x^*(t)\int_{\tau\in\mathbb{R}} x(t+\tau)e^{-i\omega\tau}\,d\tau\right] && \text{by } \textit{linearity} \text{ of E operator}\\
&= E\left[x^*(t)\int_{u\in\mathbb{R}} x(u)e^{-i\omega(u-t)}\,du\right] && \text{where } u \triangleq t+\tau \implies \tau = u-t\\
&= E\left[x^*(t)e^{i\omega t}\int_{u\in\mathbb{R}} x(u)e^{-i\omega u}\,du\right]\\
&= E\big[x^*(t)e^{i\omega t}\tilde{x}(\omega)\big] && \text{by definition of } \textit{Fourier Transform} && \text{(Definition P.2 page 349)}\\
&= E\big[x^*(t)e^{i\omega t}\big]\tilde{x}(\omega) && \text{by hypothesis (B)}\\
&= \left[\lim_{\tau\to\infty}\frac{1}{2\tau}\int_{-\tau}^{+\tau} x^*(t)e^{i\omega t}\,dt\right]\tilde{x}(\omega) && \text{by } \textit{ergodic in the mean} \text{ hypothesis} && \text{(Definition 7.4 page 46)}\\
&= \lim_{\tau\to\infty}\frac{1}{2\tau}\left[\int_{t\in\mathbb{R}} x(t)e^{-i\omega t}\,dt\right]^*\tilde{x}(\omega)\\
&= \lim_{\tau\to\infty}\frac{1}{2\tau}\tilde{x}^*(\omega)\tilde{x}(\omega) && \text{by hypothesis (B)}\\
&= \lim_{\tau\to\infty}\frac{1}{2\tau}|\tilde{x}(\omega)|^2
\end{aligned}
$$

$$
\begin{aligned}
\int_{\omega\in\mathbb{R}} \tilde{S}_{xx}(\omega)\,d\omega &= \int_{\omega\in\mathbb{R}} \lim_{\tau\to\infty}\frac{1}{2\tau}|\tilde{x}(\omega)|^2\,d\omega\\
&= \lim_{\tau\to\infty}\frac{1}{2\tau}\int_{\omega\in\mathbb{R}} |\tilde{x}(\omega)|^2\,d\omega\\
&= \lim_{\tau\to\infty}\frac{1}{2\tau}\int_{t\in\mathbb{R}} |x(t)|^2\,dt && \text{by } \textit{Plancheral's formula} && \text{(Theorem P.3 page 351, Theorem H.9 page 234)}\\
&= P_{avg} && \text{by definition of } P_{avg} && \text{(Definition 7.5 page 46)}
\end{aligned}
$$

Thus, $\tilde{S}_{xx}(\omega)$ is the power density of $x(t)$ in the frequency domain. ✏

Figure 7.2: Linear system with WSS random process input and output

Theorem 7.3. [4] *Let* $h : \mathbb{R} \to \mathbb{C}$ *be the impulse response of a linear time-invariant system and let* $y(t) = h(t) \star x(t) \triangleq \int_{u\in\mathbb{R}} h(u)x(t-u)\,du$ *as illustrated in Figure 7.1 page 43. Then*

[4] Papoulis (1991), pages 323–324

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

THM

$$\begin{aligned}
\mathsf{R}_{\mathsf{xy}}(\tau) &= \mathsf{R}_{\mathsf{xx}}(\tau) \star h^*(-\tau) &&\triangleq \int_{u\in\mathbb{R}} h^*(-u)\mathsf{R}_{\mathsf{xx}}(\tau-u)\,\mathrm{d}u\\
\mathsf{R}_{\mathsf{yy}}(\tau) &= \mathsf{R}_{\mathsf{xy}}(\tau) \star h(\tau) &&\triangleq \int_{u\in\mathbb{R}} h(u)\mathsf{R}_{\mathsf{xy}}(\tau-u)\,\mathrm{d}u\\
\mathsf{R}_{\mathsf{yy}}(\tau) &= \mathsf{R}_{\mathsf{xx}}(\tau) \star h(\tau) \star h^*(-\tau) &&\triangleq \int_{v}\int_{u\in\mathbb{R}} h(u-v)h^*(-v)\mathsf{R}_{\mathsf{xx}}(\tau-u-v)\,\mathrm{d}u\,\mathrm{d}v\\
\\
\mathsf{S}_{\mathsf{xy}}(s) &= \mathsf{S}_{\mathsf{xx}}(s)\hat{h}^*(-s^*)\\
\mathsf{S}_{\mathsf{yy}}(s) &= \mathsf{S}_{\mathsf{xy}}(s)\hat{h}(s)\\
\mathsf{S}_{\mathsf{yy}}(s) &= \mathsf{S}_{\mathsf{xx}}(s)\hat{h}(s)\hat{h}^*(-s^*)\\
\\
\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) &= \tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{h}}^*(\omega)\\
\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) &= \tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\tilde{\mathsf{h}}(\omega)\\
\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) &= \tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)|\tilde{\mathsf{h}}(\omega)|^2
\end{aligned}$$

✎ PROOF:

$$\begin{aligned}
\mathsf{R}_{\mathsf{xx}}(\tau) \star h^*(-\tau) &\triangleq \int_{u\in\mathbb{R}} h^*(-u)\mathsf{R}_{\mathsf{xx}}(\tau-u)\,\mathrm{d}u\\
&= \int_{u\in\mathbb{R}} h^*(-u)\mathsf{E}\left[\mathsf{x}(t)\mathsf{x}^*(t-\tau+u)\right]\,\mathrm{d}u\\
&= \mathsf{E}\left[\mathsf{x}(t)\int_{u\in\mathbb{R}} h^*(-u)\mathsf{x}^*(t-\tau+u)\,\mathrm{d}u\right]\\
&= \mathsf{E}\left[\mathsf{x}(t)\int_{u\in\mathbb{R}} h^*(u')\mathsf{x}^*(t-\tau-u')\,\mathrm{d}u'\right]\\
&= \mathsf{E}\left[\mathsf{x}(t)\mathsf{y}^*(t-\tau)\right]\\
&\triangleq \mathsf{R}_{\mathsf{xy}}(\tau)
\end{aligned}$$

$$\begin{aligned}
\mathsf{R}_{\mathsf{xy}}(\tau) \star h(\tau) &\triangleq \int_{u\in\mathbb{R}} h(u)\mathsf{R}_{\mathsf{xy}}(\tau-u)\,\mathrm{d}u\\
&= \int_{u\in\mathbb{R}} h(u)\mathsf{E}\left[\mathsf{x}(t+\tau-u)\mathsf{y}^*(t)\right]\,\mathrm{d}u\\
&= \mathsf{E}\left[\mathsf{y}^*(t)\int_{u\in\mathbb{R}} h(u)\mathsf{x}(t+\tau-u)\,\mathrm{d}u\right]\\
&= \mathsf{E}\left[\mathsf{y}^*(t)\mathsf{y}(t+\tau)\right]\\
&= \mathsf{E}\left[\mathsf{y}(t+\tau)\mathsf{y}^*(t)\right]\\
&\triangleq \mathsf{R}_{\mathsf{yy}}(\tau)
\end{aligned}$$

$$\begin{aligned}
\mathsf{R}_{\mathsf{yy}}(\tau) &= \mathsf{R}_{\mathsf{xy}}(\tau) \star h(\tau)\\
&= \mathsf{R}_{\mathsf{xx}}(\tau) \star h^*(-\tau) \star h(\tau)\\
&= \mathsf{R}_{\mathsf{xx}}(\tau) \star h(\tau) \star h^*(-\tau)
\end{aligned}$$

$$\begin{aligned}
\mathsf{S}_{\mathsf{xy}}(s) &\triangleq \mathbf{L}\mathsf{R}_{\mathsf{xy}}(\tau)\\
&\triangleq \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{xy}}(\tau)e^{-s\tau}\,d\tau\\
&= \int_{\tau\in\mathbb{R}} \left[\mathsf{R}_{\mathsf{xx}}(\tau) \star h^*(-\tau)\right]e^{-s\tau}\,d\tau\\
&= \int_{\tau\in\mathbb{R}} \left[\int_{u\in\mathbb{R}} h^*(-u)\mathsf{R}_{\mathsf{xx}}(\tau-u)\,\mathrm{d}u\right]e^{-s\tau}\,d\tau\\
&= \int_{u\in\mathbb{R}} h^*(-u)\int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{xx}}(\tau-u)e^{-s\tau}\,d\tau\,\mathrm{d}u
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$
\begin{aligned}
&= \int_{u\in\mathbb{R}} h^*(-u) \int_v \mathsf{R}_{\mathsf{xx}}(v) e^{-s(v+u)}\,\mathrm{d}v\,\mathrm{d}u && \text{where } v=\tau-u \iff \tau=v+u\\
&= \int_{u\in\mathbb{R}} h^*(-u) e^{-su}\,\mathrm{d}u \int_v \mathsf{R}_{\mathsf{xx}}(v) e^{-sv}\,\mathrm{d}v\\
&= \int_{u\in\mathbb{R}} h^*(u) e^{-s(-u)}\,\mathrm{d}u \int_v \mathsf{R}_{\mathsf{xx}}(v) e^{-sv}\,\mathrm{d}v\\
&= \left(\int_{u\in\mathbb{R}} h(u) e^{-(-s^*)u}\,\mathrm{d}u\right)^* \int_v \mathsf{R}_{\mathsf{xx}}(v) e^{-sv}\,\mathrm{d}v\\
&\triangleq \hat{h}^*(-s^*)\mathsf{S}_{\mathsf{xx}}(s)
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{S}_{\mathsf{yy}}(s) &\triangleq \mathbf{L}\mathsf{R}_{\mathsf{yy}}(\tau)\\
&\triangleq \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{yy}}(\tau) e^{-s\tau}\,d\tau\\
&= \int_{\tau\in\mathbb{R}} \left[\mathsf{R}_{\mathsf{xy}}(\tau) \star h(\tau)\right] e^{-s\tau}\,d\tau\\
&= \int_{\tau\in\mathbb{R}} \left[\int_{u\in\mathbb{R}} h(u)\mathsf{R}_{\mathsf{xy}}(\tau-u)\,\mathrm{d}u\right] e^{-s\tau}\,d\tau\\
&= \int_{u\in\mathbb{R}} h(u) \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{xy}}(\tau-u) e^{-s\tau}\,d\tau\,\mathrm{d}u\\
&= \int_{u\in\mathbb{R}} h(u) \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{xy}}(v) e^{-s(v+u)}\,d\tau\,\mathrm{d}u && \text{where } v=\tau-u \iff \tau=v+u\\
&= \int_{u\in\mathbb{R}} h(u) e^{-su}\,du \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{xy}}(v) e^{-sv}\,d\tau\\
&\triangleq \hat{h}(s)\mathsf{S}_{\mathsf{xy}}(s)
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{S}_{\mathsf{yy}}(s) &= \hat{h}(s)\mathsf{S}_{\mathsf{xy}}(s)\\
&= \hat{h}(s)\hat{h}^*(-s^*)\mathsf{S}_{\mathsf{xx}}(s)
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) &= \mathsf{S}_{\mathsf{xy}}(s)\big|_{s=j\omega}\\
&= \hat{h}^*(-s^*)\mathsf{S}_{\mathsf{xx}}(s)\big|_{s=j\omega}\\
&= \left(\int_{u\in\mathbb{R}} h(u) e^{-(-s^*)u}\,\mathrm{d}u\right)^* \int_v \mathsf{R}_{\mathsf{xx}}(v) e^{-sv}\,\mathrm{d}v\Big|_{s=j\omega}\\
&= \left(\int_{u\in\mathbb{R}} h(u) e^{-(-j\omega)^*u}\,\mathrm{d}u\right)^* \int_v \mathsf{R}_{\mathsf{xx}}(v) e^{-j\omega v}\,\mathrm{d}v\\
&= \left(\int_{u\in\mathbb{R}} h(u) e^{-j\omega u}\,\mathrm{d}u\right)^* \int_v \mathsf{R}_{\mathsf{xx}}(v) e^{-j\omega v}\,\mathrm{d}v\\
&\triangleq \tilde{\mathsf{h}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) &= \mathsf{S}_{\mathsf{yy}}(s)\big|_{s=j\omega}\\
&= \hat{h}(s)\mathsf{S}_{\mathsf{xy}}(s)\big|_{s=j\omega}\\
&= \int_{u\in\mathbb{R}} h(u) e^{-su}\,du \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{xy}}(v) e^{-sv}\,d\tau\Big|_{s=j\omega}\\
&= \int_{u\in\mathbb{R}} h(u) e^{-j\omega u}\,du \int_{\tau\in\mathbb{R}} \mathsf{R}_{\mathsf{xy}}(v) e^{-j\omega v}\,d\tau\\
&= \tilde{\mathsf{h}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)
\end{aligned}
$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

$$\begin{aligned}\tilde{S}_{yy}(\omega) &= \tilde{h}(\omega)\tilde{S}_{xy}(\omega)\\ &= \tilde{h}(\omega)\tilde{h}^*(\omega)\tilde{S}_{xx}(\omega)\\ &= |\tilde{h}(\omega)|^2\tilde{S}_{xx}(\omega)\end{aligned}$$

✏

Figure 7.3: *auto-correlation* $R_{xx}(\tau)$

Theorem 7.4. *Let* $x : \mathbb{R} \to \mathbb{C}$ *be a WSS random process with auto-correlation* $R_{xx}(\tau)$*. Then* $R_{xx}(\tau)$ *is* ***conjugate symmetric*** *such that (see Figure 7.3 page 50)*

THM

$$\begin{aligned}R_{xx}(\tau) &= R^*_{xx}(-\tau) && \text{(CONJUGATE SYMMETRIC)}\\ \mathbf{R}_e\left[R_{xx}(\tau)\right] &= \mathbf{R}_e\left[R_{xx}(-\tau)\right] && \text{(SYMMETRIC)}\\ \mathbf{I}_m\left[R_{xx}(\tau)\right] &= -\mathbf{I}_m\left[R_{xx}(-\tau)\right] && \text{(ANTI-SYMMETRIC)}\\ |R_{xx}(\tau)| &= |R_{xx}(-\tau)| && \text{(SYMMETRIC)}\\ \angle R_{xx}(\tau) &= \angle R_{xx}(-\tau) && \text{(ANTI-SYMMETRIC).}\end{aligned}$$

✎PROOF:

$$\begin{aligned}R^*_{xx}(\tau) &\triangleq \left(E\left[x(t-\tau)x^*(t)\right]\right)^*\\ &= E\left[x^*(t-\tau)x(t)\right]\\ &= E\left[x(t)x^*(t-\tau)\right]\\ &= E\left[x(u+\tau)x^*(u)\right] && \text{where } u \triangleq t-\tau \iff t = u+\tau\\ &\triangleq R_{xx}(\tau)\end{aligned}$$

$$\begin{aligned}\mathbf{R}_e\left[R_{xx}(\tau)\right] &= \mathbf{R}_e\left[R^*_{xx}(-\tau)\right] &&= \mathbf{R}_e\left[R_{xx}(-\tau)\right]\\ \mathbf{I}_m\left[R_{xx}(\tau)\right] &= \mathbf{I}_m\left[R^*_{xx}(-\tau)\right] &&= -\mathbf{I}_m\left[R_{xx}(-\tau)\right]\\ abs R_{xx}(\tau) &= \left|R^*_{xx}(-\tau)\right| &&= \left|R_{xx}(-\tau)\right|\\ \angle R_{xx}(\tau) &= \angle R^*_{xx}(-\tau) &&= -\angle R_{xx}(-\tau)\end{aligned}$$

✏

7.3 Whitening continuous random sequences

Simple algebraic operations on white noise processes (processes with autocorrelation $R_{xx}(\tau) = \delta(\tau)$) often produce *colored* noise (processes with autocorrelation $R_{xx}(\tau) \neq \delta(\tau)$). Sometimes we would like to process a colored noise process to produce a white noise process. This operation is known as *whitening*. Reasons for why we may want to whiten a noise process include

1. Samples from a white noise process are uncorrelated. If the noise process is Gaussian, then these samples are also independent which often greatly simplifies analysis.

COPYRIGHT © 2019 DANIEL J. GREENHOE

2. Any orthonormal basis can be used to decompose a white noise process. This is not true of a colored noise process. Karhunen–Loève expansion can be used to decompose colored noise. [5]

Definition 7.6. *A **rational expression** $\mathsf{p}(s)$ is a polynomial divided by a polynomial such that*

DEF

$$\mathsf{p}(s) = \frac{\sum_{n=0}^{N} b_n s^n}{\sum_{n=0}^{M} a_n s^n}.$$

*The **zeros** of a rational expression are the roots of its numerator polynomial. The **poles** of a rational expression are the roots of its denominator polynomial.*

Definition 7.7. *Let $\check{\mathsf{h}}(s)$ be the Laplace transform of the impulse response of a filter. If $\check{\mathsf{h}}(s)$ can be expressed as a rational expression with poles and zeros at $a_n + ib_n$, then the filter is **minimum phase** if each $a_n < 0$ (all roots lie in the left hand side of the complex s-plane).*

Note that if $L(s)$ has a root at $s = -a + ib$, then $L^*(-s^*)$ has a root at

$$-s^* = -(-a + ib)^* = -(-a - ib) = a + ib.$$

That is, if $L(s)$ has a root in the left hand plane, then $L^*(-s^*)$ has a root directly opposite across the imaginary axis in the right hand plane (see Figure 7.4 page 51). A causal stable filter $\hat{h}(s)$ must have all of its poles in the left hand plane. A minimum phase filter is a filter with both its poles and zeros in the left hand plane. One advantage of a minimum phase filter is that its recipricol (zeros become poles and poles become zeros) is also causal and stable.

Figure 7.4: Mirrored roots in complex-s plane

Figure 7.5: Innovations and whitening filters

The next theorem demonstrates a method for "whitening" a random process $\mathsf{x}(t)$ with a filter constructed from a decomposition of $\mathsf{R}_{\mathsf{xx}}(\tau)$. The technique is stated precisely in Theorem 7.5 page 52

[5] *Karhunen–Loève expansion*: Section 4.1 page 19

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

and illustrated in Figure 7.5 page 51. Both imply two filters with impulse responses $l(t)$ and $\gamma(t)$. Filter $l(t)$ is referred to as the **innovations filter** (because it generates or "innovates" $\mathsf{x}(t)$ from a white noise process $\mathsf{w}(t)$) and $\gamma(t)$ is referred to as the **whitening filter** because it produces a white noise sequence when the input sequence is $\mathsf{x}(t)$. [6]

Theorem 7.5. *Let $\mathsf{x}(t)$ be a WSS random process with autocorrelation $\mathsf{R}_{\mathsf{xx}}(\tau)$ and spectral density $\mathsf{S}_{\mathsf{xx}}(s)$. **If** $\mathsf{S}_{\mathsf{xx}}(s)$ has a **rational expression**, then the following are true:*

1. *There exists a rational expression $L(s)$ with minimum phase such that*

$$\mathsf{S}_{\mathsf{xx}}(s) = L(s)L^*(-s^*).$$

2. *An LTI filter for which the Laplace transform of the impulse response $\gamma(t)$ is*

$$\Gamma(s) = \frac{1}{L(s)}$$

   is both causal and stable.
3. *If $\mathsf{x}(t)$ is the input to the filter $\gamma(t)$, the output $\mathsf{y}(t)$ is a **white noise sequence** such that*

$$\mathsf{S}_{\mathsf{yy}}(s) = 1 \qquad\qquad \mathsf{R}_{\mathsf{yy}}(\tau) = \delta(\tau).$$

✎PROOF:

$$\begin{aligned}
\mathsf{S}_{\mathsf{ww}}(s) &= \Gamma(s)\Gamma^*(-s^*)\mathsf{S}_{\mathsf{xx}}(s) \\
&= \frac{1}{L(s)}\frac{1}{L^*(-s^*)}\mathsf{S}_{\mathsf{xx}}(s) \\
&= \frac{1}{L(s)}\frac{1}{L^*(-s^*)}L(s)L^*(-s^*) \\
&= 1
\end{aligned}$$

✏

[6] Papoulis (1991), pages 401–402

COPYRIGHT © 2019 DANIEL J. GREENHOE

CHAPTER 8

ADDITIVE NOISE ON RANDOM SEQUENCES

8.1 Additive noise and correlation

Theorem 8.1. *Let* **S** *be the* SYSTEM *illustrated to the right, where* **T** *is* NOT NECESSARILY LINEAR.

THM

$$\left.\begin{array}{lll} \text{(A).} & \mathsf{x}(n) \text{ is} & \textit{WSS} \quad \text{and} \\ \text{(B).} & \mathsf{x}(n) \text{ and } \mathsf{v}(n) \text{ are} & \textit{uncorrelated} \quad \text{and} \\ \text{(C).} & \mathsf{v}(n) \text{ is} & \textit{zero-mean} \end{array}\right\} \Longrightarrow \left\{\begin{array}{llll} \text{(1).} & \mathsf{R}_{\mathsf{yv}}(m) & = & \mathsf{R}_{\mathsf{vv}}(m) \quad \text{and} \\ \text{(2).} & \mathsf{R}_{\mathsf{xy}}(m) & = & \mathsf{R}_{\mathsf{xx}}(m) \quad \text{and} \\ \text{(3).} & \mathsf{R}_{\mathsf{yy}}(m) & = & \mathsf{R}_{\mathsf{xx}}(m) + \mathsf{R}_{\mathsf{vv}}(m) \quad \text{and} \\ \text{(4).} & \mathsf{R}_{\mathsf{xx}}(m) & = & \mathsf{R}_{\mathsf{yy}}(m) + \mathsf{R}_{\mathsf{vv}}(m) - 2\mathbf{R}_{\mathsf{e}}\mathsf{R}_{\mathsf{yv}}(m) \end{array}\right.$$

✎PROOF:

$$\begin{aligned}
\mathsf{R}_{\mathsf{yv}}(m) &\triangleq \mathsf{E}\left[\mathsf{y}(m)\mathsf{v}^*(0)\right] && \text{by (A) and definition of } \mathsf{R}_{\mathsf{yv}} && \text{(Definition 2.4 page 12)} \\
&\triangleq \mathsf{E}\left[(\mathsf{x}(m)+\mathsf{v}(m))\mathsf{v}^*(0)\right] && \text{by definition of } \mathsf{y} \\
&= \mathsf{E}\left[\mathsf{x}(m)\mathsf{v}^*(0)\right] + \mathsf{E}\left[\mathsf{v}(m)\mathsf{v}^*(0)\right] && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)} \\
&= \mathsf{E}\mathsf{x}(m)\mathsf{E}\mathsf{v}^*(0) + \mathsf{E}\left[\mathsf{v}(m)\mathsf{v}^*(0)\right] && \text{by } \textit{uncorrelated} \text{ hypothesis} && \text{(B)} \\
&= \mathsf{E}\mathsf{x}(m)\cancelto{0}{\mathsf{E}\mathsf{v}^*(0)} + \mathsf{E}\left[\mathsf{v}(m)\mathsf{v}^*(0)\right] && \text{by } \textit{zero-mean} \text{ hypothesis} && \text{(C)} \\
&= \mathsf{R}_{\mathsf{vv}}(m) && \text{by definition of } \mathsf{R}_{\mathsf{vv}} && \text{(Definition 2.4 page 12)} \\
\mathsf{R}_{\mathsf{xy}}(m) &\triangleq \mathsf{E}\left[\mathsf{x}(m)\mathsf{y}^*(0)\right] && \text{by (A) and definition of } \mathsf{R}_{\mathsf{xy}} && \text{(Definition 2.4 page 12)} \\
&\triangleq \mathsf{E}\left(\mathsf{x}(m)[\mathsf{x}(0)+\mathsf{v}(0)]^*\right) && \text{by definition of } \mathsf{y} \\
&= \mathsf{E}\left[\mathsf{x}(m)\mathsf{x}^*(0)\right] + \mathsf{E}\left[\mathsf{x}(m)\mathsf{v}^*(0)\right] && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)} \\
&= \mathsf{E}\left[\mathsf{x}(m)\mathsf{x}^*(0)\right] + \mathsf{E}[\mathsf{x}(m)]\mathsf{E}\left[\mathsf{v}^*(0)\right] && \text{by } \textit{uncorrelated} \text{ hypothesis} && \text{(B)} \\
&= \mathsf{E}\left[\mathsf{x}(m)\mathsf{x}^*(0)\right] + \mathsf{E}[\mathsf{x}(m)]\cancelto{0}{\mathsf{E}\left[\mathsf{v}^*(0)\right]} && \text{by } \textit{zero-mean} \text{ hypothesis} && \text{(C)} \\
&= \mathsf{R}_{\mathsf{xx}}(m) && \text{by definition of } \mathsf{R}_{\mathsf{xx}} && \text{(Definition 2.4 page 12)}
\end{aligned}$$

$$\begin{aligned}
\mathsf{R}_{\mathsf{yy}}(m) &\triangleq \mathsf{E}\left[\mathsf{y}(m)\mathsf{y}^*(0)\right] && \text{by (A) and definition of } \mathsf{R}_{\mathsf{yy}} \\
&\triangleq \mathsf{E}\left[(\mathsf{x}(m)+\mathsf{v}(m))(\mathsf{x}(0)+\mathsf{v}(0))^*\right] && \text{by definition of } \mathsf{y} \\
&= \mathsf{E}\left[\mathsf{x}(m)\mathsf{x}^*(0)\right] + \mathsf{E}\left[\mathsf{x}(m)\mathsf{v}^*(0)\right] + \mathsf{E}\left[\mathsf{v}(m)\mathsf{x}^*(0)\right] + \mathsf{E}\left[\mathsf{v}(m)\mathsf{v}^*(0)\right] \\
&= \mathsf{E}\left[\mathsf{x}(m)\mathsf{x}^*(0)\right] + \mathsf{E}\mathsf{x}(m)\mathsf{E}\mathsf{v}^*(0) + \mathsf{E}\mathsf{v}(m)\mathsf{E}\mathsf{x}^*(0) + \mathsf{E}\left[\mathsf{v}(m)\mathsf{v}^*(0)\right] && \text{by } \textit{uncorrelated} \text{ hypothesis (B)} \\
&= \mathsf{E}\left[\mathsf{x}(m)\mathsf{x}^*(0)\right] + \mathsf{E}\mathsf{x}(m)\cancelto{0}{\mathsf{E}\mathsf{v}^*(0)} + \cancelto{0}{\mathsf{E}\mathsf{v}(m)}\mathsf{E}\mathsf{x}^*(0) + \mathsf{E}\left[\mathsf{v}(m)\mathsf{v}^*(0)\right] && \text{by } \textit{zero-mean} \text{ hypothesis (C)}
\end{aligned}$$

$$
\begin{aligned}
&= \mathsf{R}_{\mathsf{xx}}(m) + \mathsf{R}_{\mathsf{vv}}(m) && \text{by definition of } \mathsf{R}_{\mathsf{xx}} \\
\mathsf{R}_{\mathsf{xx}}(m) &\triangleq \mathsf{E}\big[\mathsf{x}(m)\mathsf{x}^*(0)\big] \\
&\triangleq \mathsf{E}\Big(\big[\mathsf{y}(m) - \mathsf{v}(m)\big]\big[\mathsf{y}(0) - \mathsf{v}(0)\big]^*\Big) \\
&= \mathsf{E}\big[\mathsf{y}(m)\mathsf{y}^*(0)\big] - \mathsf{E}\big[\mathsf{y}(m)\mathsf{v}^*(0)\big] - \mathsf{E}\big[\mathsf{v}(m)\mathsf{y}^*(0)\big] + \mathsf{E}\big[\mathsf{v}(m)\mathsf{v}^*(0)\big] \\
&\triangleq \mathsf{R}_{\mathsf{yy}}(m) - \mathsf{R}_{\mathsf{yv}}(m) - \mathsf{R}_{\mathsf{vy}}(m) + \mathsf{R}_{\mathsf{vv}}(m) \\
&= \mathsf{R}_{\mathsf{yy}}(m) + \mathsf{R}_{\mathsf{vv}}(m) - 2\mathbf{R}_{\mathsf{e}}\mathsf{R}_{\mathsf{yv}}(m)
\end{aligned}
$$

✏

Remark 8.1. Because in Theorem 8.1 $\mathsf{y} = \mathsf{x} + \mathsf{v}$ and $\mathsf{R}_{\mathsf{yy}} = \mathsf{R}_{\mathsf{xx}} + \mathsf{R}_{\mathsf{vv}}$, one might assume that R is a kind of *linear operator* (Definition O.3 page 320) and further assume that because $\mathsf{x} = \mathsf{y} - \mathsf{v}$ and $\mathsf{R}_{(-\mathsf{v})(-\mathsf{v})} = \mathsf{R}_{\mathsf{vv}}$, that $\mathsf{R}_{\mathsf{xx}} = \mathsf{R}_{\mathsf{yy}} + \mathsf{R}_{\mathsf{vv}}$. As Theorem 8.1 demonstrates, this is simply **not the case**. The problem here is that y and v are very much *correlated*—in fact y is obviously a *function* of v.

Corollary 8.1. *Let* $\mathbf{S}$ *be the* SYSTEM *illustrated in Theorem 8.1 (page 53).*

COR

$$
\left.\begin{array}{l}\textit{hypotheses of}\\ \textit{Theorem 8.1 (page 53)}\end{array}\right\} \Longrightarrow \left\{\begin{array}{lrcll}
(1). & \check{\mathsf{S}}_{\mathsf{yy}}(z) & = & \check{\mathsf{S}}_{\mathsf{xx}}(z) + \check{\mathsf{S}}_{\mathsf{vv}}(z) & \textit{and} \\
(2). & \check{\mathsf{S}}_{\mathsf{yv}}(z) & = & \check{\mathsf{S}}_{\mathsf{vv}}(z) & \textit{and} \\
(3). & \check{\mathsf{S}}_{\mathsf{yv}}(z) & = & \check{\mathsf{S}}_{\mathsf{yy}}(z) + \check{\mathsf{S}}_{\mathsf{vv}}(z) + \check{\mathsf{S}}_{\mathsf{yv}}(z) + \check{\mathsf{S}}^*_{\mathsf{yv}}(z^*) & \textit{and} \\
(4). & \tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) & = & \tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) + \tilde{\mathsf{S}}_{\mathsf{vv}}(\omega) & \textit{and} \\
(5). & \tilde{\mathsf{S}}_{\mathsf{yv}}(\omega) & = & \tilde{\mathsf{S}}_{\mathsf{vv}}(\omega) & \textit{and} \\
(6). & \tilde{\mathsf{S}}_{\mathsf{yv}}(\omega) & = & \tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) + \tilde{\mathsf{S}}_{\mathsf{vv}}(\omega) + \tilde{\mathsf{S}}_{\mathsf{yv}}(\omega) + \tilde{\mathsf{S}}^*_{\mathsf{yv}}(-\omega) &
\end{array}\right.
$$

✎PROOF:

$$
\begin{aligned}
\check{\mathsf{S}}_{\mathsf{yy}}(z) &\triangleq \mathbf{Z}\mathsf{R}_{\mathsf{yy}}(m) && \text{by definition of } \check{\mathsf{S}}_{\mathsf{xx}} && \text{(Definition 2.5 page 14)} \\
&= \mathbf{Z}\mathsf{R}_{\mathsf{qq}}(m) + \mathbf{Z}\mathsf{R}_{\mathsf{vv}}(m) && \text{by previous result} && (1) \\
&= \check{\mathsf{S}}_{\mathsf{qq}}(z) + \check{\mathsf{S}}_{\mathsf{vv}}(z) && \text{by definition of } \check{\mathsf{S}}_{\mathsf{yy}} && \text{(Definition 7.3 page 46)} \\
\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) &\triangleq \check{\mathbf{F}}\mathsf{R}_{\mathsf{yy}}(m) && \text{by definition of } \tilde{\mathsf{S}}_{\mathsf{yy}} && \text{(Definition 7.3 page 46)} \\
&= \check{\mathbf{F}}\mathsf{R}_{\mathsf{qq}}(m) + \check{\mathbf{F}}\mathsf{R}_{\mathsf{vv}}(m) && \text{by previous result} && (1) \\
&= \tilde{\mathsf{S}}_{\mathsf{qq}}(\omega) + \tilde{\mathsf{S}}_{\mathsf{vv}}(\omega) && \text{by definition of } \tilde{\mathsf{S}}_{\mathsf{yy}} && \text{(Definition 7.3 page 46)}
\end{aligned}
$$

✏

Theorem 8.2. *Let* $\mathbf{S}$ *be the* SYSTEM *illustrated to the right:*

THM

$$
\left\{\begin{array}{lrcll}
(A). & \mathsf{x}(n) \text{ and } \mathsf{r}(n) \text{ are } \textit{wide sense stationary} & & & \text{and} \\
(B). & \mathsf{E}[\mathsf{x}(n)\mathsf{w}(n)] & = & \mathsf{E}\mathsf{x}(n)\mathsf{E}\mathsf{w}(n) \quad (\textit{uncorrelated}) & \text{and} \\
(C). & \mathsf{E}[\mathsf{r}(n)\mathsf{v}(n)] & = & \mathsf{E}\mathsf{r}(n)\mathsf{E}\mathsf{v}(n) \quad (\textit{uncorrelated}) & \text{and} \\
(D). & \mathsf{E}[\mathsf{w}(n)\mathsf{v}(n)] & = & \mathsf{E}\mathsf{w}(n)\mathsf{E}\mathsf{v}(n) \quad (\textit{uncorrelated}) & \text{and} \\
(E). & \mathsf{E}\mathsf{v}(n) = \mathsf{E}\mathsf{w}(n) & = & 0 \quad (\textit{zero-mean}) &
\end{array}\right\} \Longrightarrow \left\{\begin{array}{rcl}
\mathsf{R}_{\mathsf{sy}}(m) & = & \mathsf{R}_{\mathsf{sx}}(m) \\
 & = & \mathsf{R}_{\mathsf{ry}}(m) \\
 & = & \mathsf{R}_{\mathsf{rx}}(m)
\end{array}\right\}
$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

✎PROOF:

$$
\begin{aligned}
R_{sy}(m) &\triangleq E\big[s(m)y^*(0)\big] && \text{by definition of } R_{xy} && \text{(Definition 2.4 page 12)}\\
&\triangleq E\big([r(m)+w(m)][x(0)+v(0)]^*\big) && \text{by definition of } \mathbf{S} && \\
&= E\big[r(m)x^*(0)\big] + E\big[r(m)v^*(0)\big] + E\big[w(m)x^*(0)\big] + E\big[w(m)v^*(0)\big] && && \\
&= E\big[r(m)x^*(0)\big] + Er(m)Ev^*(0) && && \\
&\qquad + Ew(m)Ex^*(0) + Ew(m)Ev^*(0) && \text{by } \textit{uncorrelated} \text{ hypotheses} && \text{(B), (C), and (D)}\\
&= E\big[r(m)x^*(0)\big] + Er(m)\cancelto{0}{Ev^*(0)} && && \\
&\qquad + \cancelto{0}{Ew(m)}Ex^*(0) + \cancelto{0}{Ew(m)}Ev^*(0) && \text{by } \textit{zero-mean} \text{ hypothesis} && \text{(E)}\\
&= R_{rx}(m) && \text{by definition of } R_{rx} && \text{(Definition 2.4 page 12)}\\
R_{sx}(m) &\triangleq E\big[s(m)x^*(0)\big] && \text{by definition of } R_{xy} && \text{(Definition 2.4 page 12)}\\
&\triangleq E\big([r(m)+w(m)]x^*(0)\big) && \text{by definition of } \mathbf{S} && \\
&= E\big[r(m)x^*(0)\big] + Ew(m)Ex^*(0) && \text{by } \textit{uncorrelated} \text{ hypothesis} && \text{(B)}\\
&= E\big[r(m)x^*(0)\big] + \cancelto{0}{Ew(m)}Ex^*(0) && \text{by } \textit{zero-mean} \text{ hypothesis} && \text{(E)}\\
&= R_{rx}(m) && \text{by definition of } R_{rx} && \text{(Definition 2.4 page 12)}\\
R_{ry}(m) &\triangleq E\big[r(m)y^*(0)\big] && \text{by definition of } R_{xy} && \text{(Definition 2.4 page 12)}\\
&\triangleq E\big(r(m)[x(0)+v(0)]^*\big) && \text{by definition of } \mathbf{S} && \\
&= E\big[r(m)x^*(0)\big] + Er(m)Ev^*(0) && \text{by } \textit{uncorrelated} \text{ hypothesis} && \text{(C)}\\
&= E\big[r(m)x^*(0)\big] + Er(m)\cancelto{0}{Ev^*(0)} && \text{by } \textit{zero-mean} \text{ hypothesis} && \text{(E)}\\
&= R_{rx}(m) && \text{by definition of } R_{rx} && \text{(Definition 2.4 page 12)}
\end{aligned}
$$

✏

Corollary 8.2. *Let* $\mathbf{S}$ *be the* SYSTEM *illustrated in Theorem 8.2 (page 54).*

$$
\text{COR}\quad \left\{\begin{array}{l}\textit{hypotheses of}\\ \textit{Theorem 8.2 (page 54)}\end{array}\right\} \Longrightarrow \left\{\begin{array}{llllllll} (1). & \check{S}_{sy}(z) & = & \check{S}_{sx}(z) & = & \check{S}_{ry}(z) & = & \check{S}_{rx}(z) \quad \textit{and}\\ (2). & \tilde{S}_{sy}(\omega) & = & \tilde{S}_{sx}(\omega) & = & \tilde{S}_{ry}(\omega) & = & \tilde{S}_{rx}(\omega)\end{array}\right\}
$$

✎PROOF:

$$
\begin{aligned}
\check{S}_{sy}(\omega) &\triangleq \mathbf{Z}R_{sy}(m) && \text{by definition of } \check{S}_{xy} && \text{(Definition 2.5 page 14)}\\
&= \mathbf{Z}R_{rx}(m) && \text{by previous result} && (1)\\
&= \check{S}_{rx}(z) && \text{by definition of } \check{S}_{xy} && \text{(Definition 2.5 page 14)}\\
\tilde{S}_{sy}(\omega) &\triangleq \check{\mathbf{F}}R_{sy}(m) && \text{by definition of } \tilde{S}_{xy} && \text{(Definition 7.3 page 46)}\\
&= \check{\mathbf{F}}R_{rx}(m) && \text{by previous result} && (1)\\
&= \tilde{S}_{rx}(\omega) && \text{by definition of } \tilde{S}_{xy} && \text{(Definition 7.3 page 46)}
\end{aligned}
$$

✏

8.2 Additive noise and operators

Lemma 8.1. *Let* $\mathbf{S}$ *be the* SYSTEM *illustrated in Figure 8.2 (page 57) (A).*

$$
\text{LEM}\quad \left\{\begin{array}{llllll} (A). & R_{xx}(n_1,m) & = & R_{xx}(n_2,m) & \text{(WSS)} & \textit{and}\\ (B). & E[x(n)v(n)] & = & Ex(n)Ev(n) & \text{(UNCORRELATED)} & \textit{and}\\ (E). & Ev(n) & = & 0 & \text{(ZERO-MEAN)} & \end{array}\right\} \Longrightarrow \left\{\begin{array}{llll} (1). & R_{xy}(m) & = & R_{xq}(m) \quad \textit{and}\\ (2). & \check{S}_{xy}(z) & = & \check{S}_{xq}(z) \quad \textit{and}\\ (3). & \tilde{S}_{xy}(\omega) & = & \tilde{S}_{xq}(\omega)\end{array}\right\}
$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

Figure 8.1: Additive noise with *linear*/**non-linear** operator **T**

✎PROOF:

$\mathsf{R}_{xy}(m) \triangleq \mathsf{E}\big[\mathsf{x}(m)\mathsf{y}^*(0)\big]$		by definition of R_{xy}	(Definition 2.4 page 12)
$\triangleq \mathsf{E}\big[\mathsf{x}(m)(\mathsf{q}(0)+\mathsf{v}(0))^*\big]$		by definition of **S**	(Figure 8.2 page 57) (A)
$= \mathsf{E}\big[\mathsf{x}(m)\mathsf{q}^*(0)+\mathsf{p}(m)\mathsf{v}^*(0)\big]$			
$= \mathsf{E}\big[\mathsf{x}(m)\mathsf{q}^*(0)\big]+\mathsf{E}\big[\mathsf{x}(m)\mathsf{v}^*(0)\big]$			
$= \mathsf{E}\big[\mathsf{x}(m)\mathsf{q}^*(0)\big]+[\mathsf{E}\mathsf{x}(m)]\big[\mathsf{E}\mathsf{v}^*(0)\big]$		by *uncorrelated* hypothesis	(B)
$= \mathsf{E}\big[\mathsf{x}(m)\mathsf{q}^*(0)\big]+\big[\mathsf{E}\mathsf{p}(m)\big]\cancelto{0}{\big[\mathsf{E}\mathsf{v}^*(0)\big]}$		by *zero-mean* hypothesis	(E)
$= \mathsf{R}_{xq}(m)$		by definition of R_{xq}	(Definition 2.4 page 12)
$\check{\mathsf{S}}_{xy}(z) \triangleq \mathbf{Z}\mathsf{R}_{xy}(m)$		by definition of $\check{\mathsf{S}}_{xy}$	(Definition 2.5 page 14)
$= \mathbf{Z}\mathsf{R}_{xq}(m)$		by previous result	(1)
$= \check{\mathsf{S}}_{xq}(z)$		by definition of $\check{\mathsf{S}}_{xq}$	(Definition 2.5 page 14)
$\tilde{\mathsf{S}}_{xy}(\omega) \triangleq \check{\mathbf{F}}\mathsf{R}_{xy}(m)$		by definition of $\tilde{\mathsf{S}}_{xy}$	(Definition 7.3 page 46)
$= \check{\mathbf{F}}\mathsf{R}_{xq}(m)$		by previous result	(1)
$= \tilde{\mathsf{S}}_{xq}(\omega)$		by definition of $\tilde{\mathsf{S}}_{xq}$	(Definition 7.3 page 46)

✏

Lemma 8.2. *Let* **S** *be the* SYSTEM *illustrated in Figure 8.2 (page 57) (B).*

LEM
$$\left\{\begin{array}{lll} (A). & \mathsf{x}(n) \text{ is} & (\text{WSS}) \quad \text{and} \\ (B). & \mathsf{u}(n) \text{ is} & (\text{ZERO-MEAN}) \quad \text{and} \\ (C). & \mathsf{x}(n) \text{ and } \mathsf{u}(n) \text{ are} & (\text{UNCORRELATED}) \end{array}\right\} \Longrightarrow \left\{\begin{array}{llll} (1). & \mathsf{R}_{pq}(m) & = & \mathsf{R}_{xy}(m)+\mathsf{R}_{uy}(m) \quad \text{and} \\ (2). & \check{\mathsf{S}}_{pq}(z) & = & \check{\mathsf{S}}_{xy}(z)+\check{\mathsf{S}}_{uy}(z) \quad \text{and} \\ (3). & \tilde{\mathsf{S}}_{pq}(\omega) & = & \tilde{\mathsf{S}}_{xy}(\omega)+\tilde{\mathsf{S}}_{uy}(\omega) \end{array}\right\}$$

✎PROOF:

$\mathsf{R}_{xy}(m) \triangleq \mathsf{E}\big[\mathsf{x}(m)\mathsf{y}^*(0)\big]$		by definition of R_{xy}	(Definition 2.4 page 12)
$\triangleq \mathsf{E}\big(\big[\mathsf{p}(m)-\mathsf{u}(m)\big]\mathsf{y}^*(0)\big)$		by definition of **S**	
$= \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)-\mathsf{u}(m)\mathsf{y}^*(0)\big]$			
$= \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)\big]-\mathsf{E}\big[\mathsf{u}(m)\mathsf{y}^*(0)\big]$		because E is a *linear operator*	(Theorem 1.1 page 4)
$\triangleq \mathsf{R}_{py}(m)-\mathsf{R}_{uy}(m)$		by definition of R_{xy}	(Definition 2.4 page 12)

✏

Lemma 8.3. *Let* **S** *be the* SYSTEM *illustrated in Figure 8.2 (page 57) (C).*

LEM
$$\left\{\begin{array}{lll} (A). & \mathsf{x}(n) \text{ is} & (\text{WSS}) \quad \text{and} \\ (B). & \mathsf{u}(n) \text{ is} & (\text{ZERO-MEAN}) \quad \text{and} \\ (C). & \mathsf{x}(n) \text{ and } \mathsf{u}(n) \text{ are} & (\text{UNCORRELATED}) \end{array}\right\} \Longrightarrow \left\{\begin{array}{llll} (1). & \mathsf{R}_{xy}(m) & = & \mathsf{R}_{py}(m) \quad \text{and} \\ (2). & \check{\mathsf{S}}_{xy}(z) & = & \check{\mathsf{S}}_{py}(z) \quad \text{and} \\ (3). & \tilde{\mathsf{S}}_{xy}(\omega) & = & \tilde{\mathsf{S}}_{py}(\omega) \end{array}\right\}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$
\begin{aligned}
\mathsf{R}_{\mathsf{xy}}(m) &\triangleq \mathsf{E}\big[\mathsf{x}(m)\mathsf{y}^*(0)\big] && \text{by definition of } \mathsf{R}_{\mathsf{py}} && \text{(Definition 2.4 page 12)}\\
&\triangleq \mathsf{E}\big(\big[\mathsf{p}(m)+\mathsf{u}(m)\big]\mathsf{y}^*(0)\big) && \text{by definition of } \mathbf{S} && \\
&= \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)+\mathsf{u}(m)\mathsf{y}^*(0)\big] && \text{by field properties of } (\mathbb{R},+,\cdot,0,1) && \\
&= \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)\big]+\mathsf{E}\big[\mathsf{u}(m)\mathsf{y}^*(0)\big] && \text{because } \mathsf{E} \text{ is a } \textit{linear operator} && \text{(Theorem 1.1 page 4)}\\
&= \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)\big]+\mathsf{E}[\mathsf{u}(m)]\mathsf{E}\big[\mathsf{y}^*(0)\big] && \text{by } \textit{uncorrelated} \text{ hypothesis} && \text{(C)}\\
&= \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)\big]+\mathsf{E}[\mathsf{u}(m)]\cancelto{0}{\mathsf{E}\big[\mathsf{y}^*(0)\big]} && \text{by } \textit{zero-mean} \text{ hypothesis} && \text{(B)}\\
&\triangleq \mathsf{R}_{\mathsf{py}}(m) && \text{by definition of } \mathsf{R}_{\mathsf{xy}} && \text{(Definition 2.4 page 12)}
\end{aligned}
$$

✏

(A) *Communications additive noise system*
$\tilde{\mathsf{S}}_{\mathsf{pq}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{py}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)+\tilde{\mathsf{S}}_{\mathsf{uy}}(\omega)$
(Theorem 8.3 page 57)

(B) *Measurement additive noise system*
$\tilde{\mathsf{S}}_{\mathsf{pq}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{py}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{xq}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)$
(Theorem 8.4 page 57)

Figure 8.2: *linear* / **non-linear** additive noise systems

Theorem 8.3 (*communications additive noise cross-correlation*).
Let $\mathbf{S}$ *be the* SYSTEM *illustrated in Figure 8.2 page 57 (A).*

THM

$$
\left.\begin{array}{llll}
(A). & \mathsf{x}(n) \textit{ is} & \text{WSS} & \textit{and}\\
(B). & \mathsf{u}(n) \textit{ is} & \text{ZERO-MEAN} & \textit{and}\\
(C). & \mathsf{v}(n) \textit{ is} & \text{ZERO-MEAN} & \textit{and}\\
(D). & \mathsf{x}(n),\mathsf{u}(n),\mathsf{v}(n) \textit{ are} & \text{UNCORRELATED} &
\end{array}\right\}
\Longrightarrow
\left\{\begin{array}{lll}
(1). & \mathsf{R}_{\mathsf{pq}}(m)=\mathsf{R}_{\mathsf{py}}(m)=\mathsf{R}_{\mathsf{xy}}(m)+\mathsf{R}_{\mathsf{uy}}(m) & \textit{and}\\
(2). & \check{\mathsf{S}}_{\mathsf{pq}}(z)=\check{\mathsf{S}}_{\mathsf{py}}(z)=\check{\mathsf{S}}_{\mathsf{xy}}(z)+\check{\mathsf{S}}_{\mathsf{uy}}(z) & \textit{and}\\
(3). & \tilde{\mathsf{S}}_{\mathsf{pq}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{py}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)+\tilde{\mathsf{S}}_{\mathsf{uy}}(\omega) &
\end{array}\right.
$$

✎PROOF:

$$
\begin{aligned}
\mathsf{R}_{\mathsf{pq}}(m) &= \mathsf{R}_{\mathsf{py}}(m) && \text{by Lemma 8.1 page 55} && \\
\mathsf{R}_{\mathsf{pq}}(m) &= \mathsf{R}_{\mathsf{xq}}(m)+\mathsf{R}_{\mathsf{uq}}(m) && \text{by Lemma 8.2 page 56} && \\
\mathsf{R}_{\mathsf{py}}(m) &\triangleq \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)\big] && \text{by definition } \mathsf{R}_{\mathsf{py}} && \text{(Definition 2.4 page 12)}\\
&\triangleq \mathsf{E}\big([\mathsf{x}(m)+\mathsf{u}(m)]\mathsf{y}^*(0)\big) && \text{by definition } \mathbf{S} && \text{(Figure 8.2 page 57) (A)}\\
&= \mathsf{E}\big[\mathsf{x}(m)\mathsf{y}^*(0)+\mathsf{u}(m)\mathsf{y}^*(0)\big] && && \\
&= \mathsf{E}\big[\mathsf{x}(m)\mathsf{y}^*(0)\big]+\mathsf{E}\big[\mathsf{u}(m)\mathsf{y}^*(0)\big] && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)}\\
&= \mathsf{R}_{\mathsf{xy}}(m)+\mathsf{R}_{\mathsf{uy}}(m) && \text{by definitions } \mathsf{R}_{\mathsf{xy}} \text{ and } \mathsf{R}_{\mathsf{uy}} && \text{(Definition 2.4 page 12)}
\end{aligned}
$$

✏

Theorem 8.4 (*measurement additive noise cross-correlation*).
Let $\mathbf{S}$ *be the* SYSTEM *illustrated in Figure 8.2 page 57 (B).*

THM

$$
\left.\begin{array}{llll}
(A). & \mathsf{x}(n) \textit{ is} & \text{WSS} & \textit{and}\\
(B). & \mathsf{u}(n) \textit{ is} & \text{ZERO-MEAN} & \textit{and}\\
(C). & \mathsf{v}(n) \textit{ is} & \text{ZERO-MEAN} & \textit{and}\\
(D). & \mathsf{x}(n),\mathsf{u}(n),\mathsf{v}(n) \textit{ are} & \text{UNCORRELATED} &
\end{array}\right\}
\Longrightarrow
\left\{\begin{array}{lll}
(1). & \mathsf{R}_{\mathsf{pq}}(m)=\mathsf{R}_{\mathsf{py}}(m)=\mathsf{R}_{\mathsf{xq}}(m)=\mathsf{R}_{\mathsf{xy}}(m) & \textit{and}\\
(2). & \check{\mathsf{S}}_{\mathsf{pq}}(z)=\check{\mathsf{S}}_{\mathsf{py}}(z)=\check{\mathsf{S}}_{\mathsf{xq}}(z)=\check{\mathsf{S}}_{\mathsf{xy}}(z) & \textit{and}\\
(3). & \tilde{\mathsf{S}}_{\mathsf{pq}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{py}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{xq}}(\omega)=\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) &
\end{array}\right.
$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✎PROOF:

$$
\begin{aligned}
\mathsf{R}_{\mathsf{pq}}(m) &= \mathsf{R}_{\mathsf{py}}(m) && \text{by Lemma 8.1 page 55} \\
\mathsf{R}_{\mathsf{pq}}(m) &= \mathsf{R}_{\mathsf{xq}}(m) && \text{by Lemma 8.3 page 56} \\
\mathsf{R}_{\mathsf{xy}}(m) &\triangleq \mathsf{E}\big[\mathsf{x}(m)\mathsf{y}^*(0)\big] && \text{by definition } \mathsf{R}_{\mathsf{xy}} \qquad \text{(Definition 2.4 page 12)} \\
&\triangleq \mathsf{E}\big(\big[\mathsf{p}(m)+\mathsf{u}(m)\big]\mathsf{y}^*(0)\big) && \text{by definition } \mathbf{S} \qquad \text{(Figure 8.2 page 57) (B)} \\
&= \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)+\mathsf{u}(m)\mathsf{y}^*(0)\big] \\
&= \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)\big]+\mathsf{E}\big[\mathsf{u}(m)\mathsf{y}^*(0)\big] && \text{by } \textit{linearity} \text{ of } \mathsf{E} \qquad \text{(Theorem 1.1 page 4)} \\
&= \mathsf{E}\big[\mathsf{p}(m)\mathsf{y}^*(0)\big]+\cancelto{0}{\mathsf{E}\big[\mathsf{u}(m)\mathsf{y}^*(0)\big]} && \text{by } \textit{uncorrelated} \text{ hypothesis} \qquad \text{(D)} \\
&= \mathsf{R}_{\mathsf{py}}(m) && \text{by definition of } \mathsf{R}_{\mathsf{py}} \qquad \text{(Definition 2.4 page 12)}
\end{aligned}
$$

✏

8.3 Additive noise and LTI operators

Figure 8.3: Additive noise with **LTI** operator **H**

Corollary 8.3. *Let* **S** *be the* SYSTEM *illustrated in Figure 8.3 (page 58) (A).*

COR

$$
\left\{
\begin{array}{llll}
(A). & \mathsf{x}(n) \textit{ is} & (\text{WSS}) & \textit{and} \\
(B). & \mathsf{u}(n) \textit{ is} & (\text{ZERO-MEAN}) & \textit{and} \\
(C). & \mathsf{x}(n) \textit{ and } \mathsf{u}(n) \textit{ are} & (\text{UNCORRELATED}) & \textit{and} \\
(D). & \mathbf{H} \textit{ is} & (\text{LTI}) &
\end{array}
\right\}
\Longrightarrow
\left\{ \tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) = \tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) \right\}
$$

✎PROOF:

$$
\begin{aligned}
\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) &= \tilde{\mathsf{S}}_{\mathsf{xq}}(\omega) && \text{by Lemma 8.1 page 55} \\
&= \tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) && \text{by Corollary 6.3 page 39}
\end{aligned}
$$

✏

Theorem 8.5. *Let* **S** *be the* SYSTEM *illustrated in Figure 8.3 (page 58) (B).*

THM

$$
\left\{
\begin{array}{llll}
(A). & \mathsf{x}(n) \textit{ is} & (\text{WSS}) & \textit{and} \\
(B). & \mathsf{u}(n) \textit{ is} & (\text{ZERO-MEAN}) & \textit{and} \\
(C). & \mathsf{x}(n) \textit{ and } \mathsf{u}(n) \textit{ are} & (\text{UNCORRELATED}) & \textit{and} \\
(D). & \mathbf{H} \textit{ is} & (\text{LTI}) &
\end{array}
\right\}
\Longrightarrow
\left\{
\begin{array}{llll}
(1). & \mathsf{R}_{\mathsf{yx}}(m) = \mathsf{h}(m) \star \mathsf{R}_{\mathsf{xx}}(m) & \textit{and} \\
(2). & \check{\mathsf{S}}_{\mathsf{yx}}(z) = \check{\mathsf{H}}(z)\check{\mathsf{S}}_{\mathsf{xx}}(z) & \textit{and} \\
(3). & \tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) = \tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) &
\end{array}
\right\}
$$

✎PROOF:

1. definition: Let $(\mathsf{h}(n))$ be the *impulse response* of operator **H** such that
$$\mathbf{H}\bar{\delta}(n) \triangleq \sum_{m\in\mathbb{Z}} \mathsf{h}(m)\bar{\delta}(n-m)$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

2. lemma: $\mathbf{H}\boldsymbol{x}(n) = \sum_{m\in\mathbb{Z}} \mathsf{h}(n)\boldsymbol{x}(m-n) = \mathsf{h}(n) \star \mathsf{R}_{\mathsf{xx}}(n)$.

   Proof: by the *linear time-invariant* hypotheses (D) and definition of *convolution* operator $\star$ (Definition P.3 page 352)

3. Proof that $\mathsf{R}_{\mathsf{yx}}(m) = \mathsf{h}(n) \star \mathsf{R}_{\mathsf{xx}}(m)$:

$$
\begin{aligned}
\mathsf{R}_{\mathsf{yx}}(m) &\triangleq \mathsf{E}\big[\mathsf{y}(m)\mathsf{x}^*(0)\big] && \text{by definition of } \mathsf{R}_{\mathsf{py}} && \text{(Definition 2.4 page 12)}\\
&= \mathsf{E}\big([\mathbf{H}\mathsf{x}(m) + \mathbf{H}\mathsf{u}(m)]\mathsf{x}^*(0)\big) && \text{by } \textit{linear} \text{ hypothesis} && \text{(D)}\\
&= \mathsf{E}\big(\big[\mathbf{H}\mathsf{x}^*(m)\big]\mathsf{x}^*(0) + [\mathbf{H}\mathsf{u}(0)]\mathsf{x}^*(0)\big) && && \\
&= \mathsf{E}\big(\big[\mathbf{H}\mathsf{x}^*(m)\big]\mathsf{x}^*(0)\big) + \mathsf{E}\big([\mathbf{H}\mathsf{u}(0)]\mathsf{x}^*(0)\big) && \text{by } \textit{linearity} \text{ of } \mathbf{E} && \text{(Theorem ?? page ??)}\\
&= \mathbf{H}\mathsf{E}\big[\mathsf{x}(m)\mathsf{x}^*(0)\big] + \mathbf{H}\mathsf{E}\big[\mathsf{u}(m)\mathsf{x}^*(0)\big] && \text{by } \textit{LTI} \text{ hypotheses} && \text{(D)}\\
&= \mathbf{H}\mathsf{E}\big[\mathsf{x}(m)\mathsf{x}^*(0)\big] + \mathbf{H}\mathsf{E}\mathsf{u}(m)\mathsf{E}\mathsf{x}^*(0) && \text{by } \textit{uncorrelated} \text{ hypothesis} && \text{(C)}\\
&= \mathbf{H}\mathsf{E}\big[\mathsf{x}(m)\mathsf{x}^*(0)\big] + \mathbf{H}\cancelto{0}{\mathsf{E}\mathsf{u}(m)}\mathsf{E}\mathsf{x}^*(0) && \text{by } \textit{zero-mean} \text{ hypothesis} && \text{(B)}\\
&= \mathbf{H}\mathsf{R}_{\mathsf{xx}}(m) && \text{by definition of } \mathsf{R}_{\mathsf{xx}} && \text{(Definition 2.4 page 12)}\\
&= \mathsf{h}(m) \star \mathsf{R}_{\mathsf{xx}}(m) && \text{by (2) lemma} &&
\end{aligned}
$$

✏

When $\mathbf{H}$ is *LTI*, what effect does the additive uncorrelated noise sources have on the cross-statistical properties of x and y? Corollary 8.5 (next) demonstrates that, amazingly, under very general conditions, the noise sources have **no effect**.

Corollary 8.4. *Let* $\mathbf{S}$ *be the* SYSTEM *illustrated in Figure 8.3 (page 58) (C).*

COR

$$
\left\{
\begin{array}{llll}
(A). & \mathsf{x}(n) \textit{ is} & (\text{WSS}) & \textit{and}\\
(B). & \mathsf{u}(n) \textit{ is} & (\text{ZERO-MEAN}) & \textit{and}\\
(C). & \mathsf{x}(n) \textit{ and } \mathsf{w}(n) \textit{ are} & (\text{UNCORRELATED}) & \textit{and}\\
(D). & \mathbf{H} \textit{ is} & (\text{LTI}) &
\end{array}
\right\}
\Longrightarrow
\left\{ \tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) = \tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{pp}}(\omega) \right\}
$$

✎PROOF:

$$
\begin{aligned}
\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) &= \tilde{\mathsf{S}}_{\mathsf{py}}(\omega) && \text{by Lemma 8.3 page 56}\\
&= \tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{pp}}(\omega) && \text{by Corollary 6.3 page 39}
\end{aligned}
$$

✏

Figure 8.4:

Corollary 8.5. *Let* $\mathbf{S}$ *be the* SYSTEM *illustrated in Figure 8.4 page 59 (A).*

COR

$$
\left\{
\begin{array}{llll}
(A). & \mathsf{x}(n) \textit{ is} & (\text{WSS}) & \textit{and}\\
(B). & \mathsf{u}(n) \textit{ is} & (\text{ZERO-MEAN}) & \textit{and}\\
(C). & \mathsf{v}(n) \textit{ is} & (\text{ZERO-MEAN}) & \textit{and}\\
(D). & \mathsf{x}(n), \mathsf{u}(n), \mathsf{v}(n) \textit{ are} & (\text{UNCORRELATED}) & \textit{and}\\
(E). & \mathbf{H} \textit{ is} & (\text{LTI}) &
\end{array}
\right\}
\Longrightarrow
\left\{ \tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) = \tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) \right\}
$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✎PROOF:

$$\begin{aligned}\tilde{S}_{yx}(\omega) &= \tilde{S}_{qx}(\omega) && \text{by Lemma 8.1 page 55}\\ &= \tilde{H}(\omega)\tilde{S}_{xx}(\omega) && \text{by Corollary 6.3 page 39}\end{aligned}$$

✏

Corollary 8.6. *Let* **S** *be the* SYSTEM *illustrated in Figure 8.4 page 59 (B).*

COR

$$\left\{\begin{array}{llll} (A). & \mathsf{x}(n) \text{ is} & \text{WSS} & \text{and}\\ (B). & \mathsf{w}(n) \text{ is} & \text{ZERO-MEAN} & \text{and}\\ (C). & \mathsf{v}(n) \text{ is} & \text{ZERO-MEAN} & \text{and}\\ (D). & \mathsf{x}(n), \mathsf{w}(n), \mathsf{v}(n) \text{ are} & \text{UNCORRELATED} & \text{and}\\ (E). & \mathbf{H} \text{ is} & \text{LTI} & \end{array}\right\} \Longrightarrow \left\{ \tilde{S}_{xy}(\omega) = \tilde{H}^*(\omega)\tilde{S}_{pp}(\omega) \right\}$$

✎PROOF:

$$\begin{aligned}\tilde{S}_{yx}(\omega) &= \tilde{S}_{qx}(\omega) && \text{by Lemma 8.1 page 55}\\ &= \tilde{S}_{qp}(\omega) && \text{by Lemma 8.1 page 55}\\ &= \tilde{H}(\omega)\tilde{S}_{pp}(\omega) && \text{by Corollary 6.3 page 39}\end{aligned}$$

✏

8.4 Additive noise and dual operators

Figure 8.5: Dual Additive Noise Systems

Corollary 8.7. *Let* **S** *be the* SYSTEM *illustrated in Figure 8.5 (page 60) (A).*

COR

$$\left\{\begin{array}{llll} (A). & \mathbf{H} \text{ is} & \text{LTI} & \text{and}\\ (B). & \mathsf{x}(n) \text{ is} & \text{WSS} & \text{and}\\ (C). & \mathsf{u} \text{ and } \mathsf{v} \text{ are} & \text{ZERO-MEAN} & \text{and}\\ (D). & \mathsf{x}, \mathsf{u}, \mathsf{v} \text{ are} & \text{UNCORRELATED} & \end{array}\right\} \Longrightarrow \left\{\begin{array}{llll} (1). & \check{S}_{sy}(z) = \check{H}(z)\left[\check{S}_{xy}(z) + \check{S}_{uy}(z)\right] & \text{and}\\ (2). & \tilde{S}_{sy}(\omega) = \tilde{H}(\omega)\left[\tilde{S}_{xy}(\omega) + \tilde{S}_{uy}(\omega)\right] & \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}\check{S}_{sy}(z) &= \check{S}_{rq}(z) && \text{by Corollary 8.2 page 55} && \text{and (B), (C) and (D)}\\ &= \check{H}(z)\check{S}_{pq}(z) && \text{by Theorem 6.2 page 39} && \text{and (A)}\\ &= \check{H}(z)\left[\check{S}_{xq}(z) + \check{S}_{uq}(z)\right] && \text{by Lemma 8.2 page 56}\\ &= \check{H}(z)\left[\check{S}_{xy}(z) + \check{S}_{uy}(z)\right] && \text{by Lemma 8.1 page 55}\\ \tilde{S}_{sy}(\omega) &= \check{S}_{sy}(z)\big|_{z=e^{i\omega}}\\ &= \check{H}(z)\left[\check{S}_{xy}(z) + \check{S}_{uy}(z)\right]\big|_{z=e^{i\omega}} && \text{by previous result} && (1)\\ &= \tilde{H}(\omega)\left[\tilde{S}_{xy}(\omega) + \tilde{S}_{uy}(\omega)\right]\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

Corollary 8.8. *Let* $\mathbf{S}$ *be the* SYSTEM *illustrated in Figure 8.5 (page 60) (B).*

COR

$$\left\{\begin{array}{llll} (A). & \mathbf{H}\ \textit{is} & \text{LTI} & \textit{and} \\ (B). & \mathsf{x}(n)\ \textit{is} & \text{WSS} & \textit{and} \\ (C). & \mathsf{u}\ \textit{and}\ \mathsf{v}\ \textit{are} & \text{ZERO-MEAN} & \textit{and} \\ (D). & \mathsf{p}, \mathsf{u}, \mathsf{v}\ \textit{are} & \text{UNCORRELATED} & \end{array}\right\} \Longrightarrow \left\{\begin{array}{llll} (1). & \check{\mathsf{S}}_{\mathsf{sy}}(z) & = & \check{\mathsf{H}}(z)\check{\mathsf{S}}_{\mathsf{xy}}(z) \quad \textit{and} \\ (2). & \tilde{\mathsf{S}}_{\mathsf{sy}}(\omega) & = & \tilde{\mathsf{H}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\check{\mathsf{S}}_{\mathsf{sy}}(z) &= \check{\mathsf{S}}_{\mathsf{rq}}(z) && \text{by Corollary 8.2 page 55} && \text{and (B), (C) and (D)} \\
&= \check{\mathsf{H}}(z)\check{\mathsf{S}}_{\mathsf{pq}}(z) && \text{by Theorem 6.2 page 39} && \text{and (A)} \\
&= \check{\mathsf{H}}(z)\check{\mathsf{S}}_{\mathsf{xq}}(z) && \text{by Lemma 8.3 page 56} && \\
&= \check{\mathsf{H}}(z)\check{\mathsf{S}}_{\mathsf{xy}}(z) && \text{by Lemma 8.1 page 55} && \\
\tilde{\mathsf{S}}_{\mathsf{sy}}(\omega) &= \check{\mathsf{S}}_{\mathsf{sy}}(z)\big|_{z=e^{i\omega}} && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)} \\
&= \check{\mathsf{H}}(z)\check{\mathsf{S}}_{\mathsf{xy}}(z)\big|_{z=e^{i\omega}} && \text{by previous result} && (1) \\
&= \tilde{\mathsf{H}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega) &&&&
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

Copyright © 2019 Daniel J. Greenhoe

Part III

Statistical Estimation

CHAPTER 9

ESTIMATION OVERVIEW

9.1 Estimation types

Estimation types. Let $\mathsf{x}(t;\theta)$ be a waveform with parameter θ. There are three basic types of estimation of x:

1. *detection*:
   - The waveform $\mathsf{x}(t;\theta_n)$ is known except for the value of parameter θ_n.
   - The parameter θ_n is one of a finite set of values.
   - Estimate θ_n and thereby also estimate $\mathsf{x}(t;\theta)$.
2. *parametric* estimation:
   - The waveform $\mathsf{x}(t;\theta)$ is known except for the value of parameter θ.
   - The parameter θ is one of an infinite set of values.
   - Estimate θ and thereby also estimate $\mathsf{x}(t;\theta)$.
3. *nonparametric* estimation:
   - The waveform $\mathsf{x}(t)$ is unknown and assumed without any parameter θ.
   - Estimate $\mathsf{x}(t)$.

Estimation criterion. Optimization requires a criterion against which the quality of an estimate is measured.[1] The most demanding and general criterion is the *Bayesian* criterion. The Bayesian criterion requires knowledge of the probability distribution functions and the definition of a *cost function*. Other criterion are special cases of the Bayesian criterion such that the cost function is defined in a special way, no cost function is defined, and/or the distribution is not known (Figure 9.2 page 68).

Estimation techniques. Estimation techniques can be classified into five groups (Figure 9.2 page 68):[2]

[1] Mandyam D. Srinath (1996) ⟨013125295X⟩.

[2] Nelles (2001) page 26 ⟨"Fig 2.2 Overview of linear and nonlinear optimization techniques"⟩, Nelles (2001) page 33 ⟨"Fig 2.5 The Bayes method is the most general approach but…"⟩, Nelles (2001) page 63 ⟨"Table 3.3 Relationship between linear recursive and nonlinear optimization techniques"⟩, Nelles (2001) page 66

1. sequential decoding
2. norm minimization
3. gradient search
4. inner product analysis
5. direct search

Sequential decoding is a non-linear estimation family. Perhaps the most famous of these is the Veterbi algorithm which uses a trellis to calculate the estimate. The Verterbi algorithm has been shown to yield an optimal estimate in the maximal likelihood (ML) sense. Norm minimization and gradient search algorithms are all linear algorithms. While this restriction to linear operations often simplifies calculations, it often yields an estimate that is not optimal in the ML sense.

9.2 Estimation criterion

Figure 9.1: Estimation criterion

Definition 9.1. *Let*

(A). $\mathsf{x}(t;\theta)$ *be a random process with unknown parameter* θ

(B). $\mathsf{y}(t)$ *an observed random process which is statistically dependent on* $\mathsf{x}(t;\theta)$

(C). $\mathsf{C}(\theta, \mathsf{p}(x, y))$ *be a cost function.*

Then the following ***estimates*** *are defined as described here:*

DEF

(1). ***Bayesian estimate*** $\hat{\theta}_{\mathsf{B}} \triangleq \arg\min_{\theta} C(\theta, \mathsf{p}(x, y))$

(2). ***Mean square estimate*** *("MS estimate")* $\hat{\theta}_{\mathsf{mms}} \triangleq \arg\min_{\theta} \mathsf{E}\, \|C(\theta, \mathsf{p}(x, y))\|^2$

(3). ***mini-max estimate*** *("MM estimate")* $\hat{\theta}_{\mathsf{mm}} \triangleq \arg\min_{\theta} \max_{\mathsf{p}} C(\theta, \mathsf{p}(x, y))$

(4). ***maximum a-posteriori probability estimate*** *("MAP estimate")* $\hat{\theta}_{\mathsf{map}} \triangleq \arg\max_{\theta} \mathsf{P}\{\mathsf{x}(t;\theta)|\mathsf{y}(t)\}$

(5). ***maximum likelihood estimate*** *("ML estimate")* $\hat{\theta}_{\mathsf{ml}} \triangleq \arg\max_{\theta} \mathsf{P}\{\mathsf{y}(t)|\mathsf{x}(t;\theta)\}$

Copyright © 2019 Daniel J. Greenhoe

Theorem 9.1. *Let* $\mathsf{x}(t;\theta)$ *be a random process with unknown parameter* θ.

THM $\{\mathsf{P}\{\theta\} = \text{CONSTANT}\} \quad \Longrightarrow \quad \{\hat{\theta}_{\mathsf{map}} = \hat{\theta}_{\mathsf{ml}}\}$

✎PROOF:

$$\begin{aligned}
\hat{\theta}_{\mathsf{map}} &\triangleq \arg\max_{\theta} \mathsf{P}\{\mathsf{x}(t;\theta)|\mathsf{y}(t)\} && \text{by definition of } \hat{\theta}_{\mathsf{map}} && \text{(Definition 9.1 page 66)}\\
&\triangleq \arg\max_{\theta} \frac{\mathsf{P}\{\mathsf{x}(t;\theta)\wedge \mathsf{y}(t)\}}{\mathsf{P}\{r(t)\}} && \text{by definition of } \textit{conditional probability} && \text{(Definition A.4 page 168)}\\
&\triangleq \arg\max_{\theta} \frac{\mathsf{P}\{r(t)|\mathsf{x}(t;\theta)\}\mathsf{P}\{\mathsf{x}(t;\theta)\}}{\mathsf{P}\{\mathsf{y}(t)\}} && \text{by definition of } \textit{conditional probability} && \text{(Definition A.4 page 168)}\\
&= \arg\max_{\theta} \mathsf{P}\{\mathsf{y}(t)|\mathsf{x}(t;\theta)\}\mathsf{P}\{\mathsf{x}(t;\theta)\} && \text{because } \mathsf{y}(t) \text{ is independent of } \theta\\
&= \arg\max_{\theta} \mathsf{P}\{\mathsf{y}(t)|\mathsf{x}(t;\theta)\}\\
&\triangleq \hat{\theta}_{\mathsf{ml}} && \text{by definition of } \hat{\theta}_{\mathsf{ml}} && \text{(Definition 9.1 page 66)}
\end{aligned}$$

✏

9.3 Measures of estimator quality

Definition 9.2. [3]

DEF *The* ***mean square error*** $\mathsf{mse}(\hat{\theta})$ *of an estimate* $\hat{\theta}$ *of a parameter* θ *is defined as*

$$\mathsf{mse}(\hat{\theta}) \triangleq \mathsf{E}\left[\left(\hat{\theta}-\theta\right)^2\right]$$

Definition 9.3. [4]

DEF *The* ***normalized rms error*** $\epsilon(\hat{\theta})$ *of an estimate* $\hat{\theta}$ *of a parameter* θ *is defined as*

$$\epsilon(\hat{\theta}) \triangleq \frac{\sqrt{\mathsf{mse}(\hat{\theta})}}{\theta} \triangleq \frac{\sqrt{\mathsf{E}\left[\left(\hat{\theta}-\theta\right)^2\right]}}{\theta}$$

Definition 9.4. [5]

DEF *The* ***mean integrated square error*** $\mathsf{mse}(\hat{\theta})$ *of an estimate* $\hat{\theta}$ *of a parameter* θ *is defined as*

$$\mathsf{mse}(\hat{\theta}) \triangleq \mathsf{E}\int_{\theta\in\mathbb{R}}\left[\left(\hat{\theta}-\theta\right)^2\right]$$

The *mean square error* of $\hat{\theta}$ can be expressed as the sum of two components: the variance of $\hat{\theta}$ and the bias of $\hat{\theta}$ squared (next Theorem). For an example of Theorem 9.2 in action, see the proof for the $\mathsf{mse}(\hat{\mu})$ of the *arithmetic mean estimate* as provided in Theorem 13.1 (page 97).

Theorem 9.2. [6] *Let* $\mathsf{mse}(\hat{\theta})$ *be the* MEAN SQUARE ERROR *(Definition 9.2 page 67) and* $\epsilon(\hat{\theta})$ *the* NORMALIZED

[3] Silverman (1986) page 35 ⟨§"1.3.2 Measures of discrepancy…"⟩, Bendat and Piersol (2010) ⟨§"1.4.3 Error Analysis Criteria"⟩, Bendat and Piersol (1966), page 183§"5.3 Statistical Errors for Parameter Estimates"

[4] Bendat and Piersol (2010) ⟨§"1.4.3 Error Analysis Criteria"⟩

[5] Silverman (1986) page 35 ⟨§"1.3.2 Measures of discrepancy…"⟩, Rosenblatt (1956) page 835 ⟨"integrated mean square error"⟩

[6] Choi (1978) page 76, Kay (1988) page 45 ⟨§"3.3 Estimation Theory"⟩, Stuart and Ord (1991) page 629 ⟨"Minium mean-square-error estimation"⟩, Clarkson (1993) page 51 ⟨§"2.6 Estimation of Moments"⟩, Bendat and Piersol (2010) ⟨§"1.4.3 Error Analysis Criteria"⟩, Bendat and Piersol (1966), page 183§"5.3 Statistical Errors for Parameter Estimates", Bendat and Piersol (1980) page 39 ⟨§"2.4.1 Bias versus Random Errors"⟩

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

RMS ERROR *(Definition 9.3 page 67) of an estimator* $\hat{\theta}$.

THM

$$\mathrm{mse}(\hat{\theta}) = \underbrace{\mathsf{E}\left[\left(\hat{\theta}-\mathsf{E}\hat{\theta}\right)^2\right]}_{\textit{variance of }\hat{\theta}} + \underbrace{\left[\mathsf{E}\hat{\theta}-\theta\right]^2}_{\textit{bias of }\hat{\theta}\textit{ squared}} \qquad \epsilon(\hat{\theta}) = \frac{\sqrt{\mathsf{E}\left[\left(\hat{\theta}-\mathsf{E}\hat{\theta}\right)^2\right]+\left[\mathsf{E}\hat{\theta}-\theta\right]^2}}{\theta}$$

✎PROOF:

$$\begin{aligned}
\mathrm{mse}(\hat{\theta}) &\triangleq \mathsf{E}\left[\left(\hat{\theta}-\theta\right)^2\right] && \text{by definition of mse} && \text{(Definition 9.2 page 67)}\\
&= \mathsf{E}\left[\left(\hat{\theta}\underbrace{-\mathsf{E}\hat{\theta}+\mathsf{E}\hat{\theta}}_{0}-\theta\right)^2\right] && \text{by } \textit{additive identity} \text{ property of } (\mathbb{C},+,\cdot,0,1)\\
&= \mathsf{E}\left[\left(\hat{\theta}-\mathsf{E}\hat{\theta}\right)^2+\underbrace{\left(\mathsf{E}\hat{\theta}-\theta\right)^2}_{\text{constant}}-2\left(\hat{\theta}-\mathsf{E}\hat{\theta}\right)\left(\mathsf{E}\hat{\theta}-\theta\right)\right] && \text{by } \textit{Binomial Theorem}\\
&= \mathsf{E}\left(\hat{\theta}-\mathsf{E}\hat{\theta}\right)^2+\left(\mathsf{E}\hat{\theta}-\theta\right)^2-2\mathsf{E}\left[\hat{\theta}\mathsf{E}\hat{\theta}-\hat{\theta}\theta-\mathsf{E}\hat{\theta}\hat{\theta}+\mathsf{E}\hat{\theta}\theta\right] && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)}\\
&= \mathsf{E}\left(\hat{\theta}-\mathsf{E}\hat{\theta}\right)^2+\left(\mathsf{E}\hat{\theta}-\theta\right)^2-2\underbrace{\left[\mathsf{E}\hat{\theta}\mathsf{E}\hat{\theta}-\mathsf{E}\hat{\theta}\mathsf{E}\theta-\mathsf{E}\hat{\theta}\mathsf{E}\hat{\theta}+\mathsf{E}\hat{\theta}\mathsf{E}\theta\right]}_{0} && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)}\\
&= \mathsf{E}\left(\hat{\theta}-\mathsf{E}\hat{\theta}\right)^2+\left(\mathsf{E}\hat{\theta}-\theta\right)^2
\end{aligned}$$

✏

Definition 9.5. [7]

DEF

An estimate $\hat{\theta}$ *of a parameter* θ *is a* ***minimum variance unbiased estimator*** *(MVUE) if*

(1). $\mathsf{E}\hat{\theta} = 0$ (UNBIASED) *and*

(2). *no other unbiased estimator* $\hat{\phi}$ *has smaller variance* $\mathrm{var}(\hat{\phi})$

9.4 Estimation techniques

Figure 9.2: Estimation techniques

[7] Choi (1978) page 76, Shao (2003) page 161 ⟨§"The UMVUE"⟩, Bolstad (2007) page 164 ⟨§"Minimum Variance Unbiased Estimator"⟩,

COPYRIGHT © 2019 DANIEL J. GREENHOE

9.5 Sequential decoding

Figure 9.3: Viterbi algorithm trellis

It has been shown that the Viterbi algorithm (trellis) produces an optimal estimate in the maximal likelihood (ML) sense. A Verterbi trellis is shown in Figure 9.3 (page 69).

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 10

NORM MINIMIZATION

Figure 10.1: Estimation using gradient of cost function

Norm minimization techniques are very powerful in that an optimum solution can be computed in one step without iteration or recursion. In this section we present two types of norm minimization:[1]

1. minimum mean square estimation (MMSE):
   The MMS estimate is a *stochastic* estimate. To compute the MMS estimate, the we do not need to know the actual data values, but we must know certain system statistics which are the input data autocorrelation and input/output crosscorrelation. The cost function is the expected value of the norm squared error.
2. least square estimation (LSE):
   The LS estimate is a *deterministic* estimate. To compute the LS estimate, we must know the actual data values (although these may be "noisy" measurements). The cost function is the norm squared error.

Solutions to both are given in terms of two matrices:

[1] The Least Squares algorithm is nothing new to mathematics. It was first published by Legendre in 1805, but there is a credible claim by Gauss that he had it as far back as 1795. Gauss, by the way, was also the first to discover the FFT. References: Sorenson (1970) page 63, Plackett (1972), Stigler (1981), Dutka (1995)

Y: Autocorrelation matrix
W: Crosscorrelation matrix.

10.1 Minimum mean square estimation

Definition 10.1. *Let the following vectors, matrices, and functions be defined as follows:*

DEF				
	$\boldsymbol{x}$	$\in$	$\mathbb{C}^m$	*data vector*
	$\boldsymbol{y}$	$\in$	$\mathbb{C}^n$	*processed data vector*
	$\hat{\boldsymbol{y}}$	$\in$	$\mathbb{C}^n$	*processed data estimate vector*
	$\boldsymbol{e}$	$\in$	$\mathbb{C}^n$	*error vector*
	$\boldsymbol{p}$	$\in$	$\mathbb{R}^m$	*parameter vector*
	U	$\in$	C_{mn}	*regression matrix*
	R	$\in$	C_{mm}	*autocorrelation matrix*
	W	$\in$	$\mathbb{C}^m$	*cross-correlation vector*
	C	$:$	$\mathbb{R}^m \to \mathbb{R}^{\vdash}$	*cost function*

Theorem 10.1 (Minimum mean square estimation)**.** *Let*

$$\begin{aligned}
\hat{\boldsymbol{y}}(\boldsymbol{p}) &\triangleq U^H\boldsymbol{p}\\
\boldsymbol{e}(\boldsymbol{p}) &\triangleq \hat{\boldsymbol{y}} - \boldsymbol{y}\\
\mathsf{C}(\boldsymbol{p}) &\triangleq \mathsf{E}\,\|\boldsymbol{e}\|^2 \triangleq \mathsf{E}\left[\boldsymbol{e}^H\boldsymbol{e}\right]\\
\hat{\theta}_{\mathsf{mms}} &\triangleq \arg\min_{\boldsymbol{p}} \mathsf{C}(\boldsymbol{p})\\
R &\triangleq \mathsf{E}\left[UU^H\right]\\
W &\triangleq \mathsf{E}\left[U\boldsymbol{y}\right].
\end{aligned}$$

Then

THM

$$\begin{aligned}
\hat{\theta}_{\mathsf{mms}} &= (\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)\\
\mathsf{C}(\boldsymbol{p}) &= \boldsymbol{p}^H R\boldsymbol{p} - (W^H\boldsymbol{p})^* - W^H\boldsymbol{p} + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y}\\
\nabla_{\boldsymbol{p}}\mathsf{C}(\boldsymbol{p}) &= 2\mathbf{R}_{\mathsf{e}}[Y]\boldsymbol{p} - 2\mathbf{R}_{\mathsf{e}}W\\
\mathsf{C}(\hat{\theta}_{\mathsf{mms}}) &= \begin{cases} \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} + (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) \\ \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} - (\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W) \quad \textit{if } R \textit{ is } \textbf{\textit{real-valued}} \end{cases}
\end{aligned}$$

✎PROOF: See APPENDIX E (page 195) for a Matrix Calculus reference.

1. Proof that cost $\mathsf{C}(\boldsymbol{p}) = \boldsymbol{p}^H R\boldsymbol{p} - (W^H)^*\boldsymbol{p} - W^H\boldsymbol{p} + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y}$:

$$\begin{aligned}
\boxed{\mathsf{C}(\boldsymbol{p})} &\triangleq \mathsf{E}\,\|\boldsymbol{e}\|^2 && \text{by definition of } \textit{cost function}\ \mathsf{C}\\
&\triangleq \mathsf{E}\left[\boldsymbol{e}^H\boldsymbol{e}\right] && \text{by definition of } \textit{norm}\ \|\cdot\|\\
&\triangleq \mathsf{E}\left[(\hat{\boldsymbol{y}}-\boldsymbol{y})^H(\hat{\boldsymbol{y}}-\boldsymbol{y})\right] && \text{by definition of } \textit{error vector}\ \boldsymbol{e}\\
&\triangleq \mathsf{E}\left[\left(U^H\boldsymbol{p}-\boldsymbol{y}\right)^H\left(U^H\boldsymbol{p}-\boldsymbol{y}\right)\right] && \text{by definition of estimate } \hat{\boldsymbol{y}}\\
&= \mathsf{E}\left[\left(\boldsymbol{p}^HU-\boldsymbol{y}^H\right)\left(U^H\boldsymbol{p}-\boldsymbol{y}\right)\right] && \text{by } \textit{distributive} \text{ prop. of } \textit{*-algebras} \quad \text{(Definition J.3 page 266)}\\
&= \mathsf{E}\left[\boldsymbol{p}^HUU^H\boldsymbol{p} - \boldsymbol{p}^HU\boldsymbol{y} - \boldsymbol{y}^HU^H\boldsymbol{p} + \boldsymbol{y}^H\boldsymbol{y}\right] && \text{by } \textit{matrix algebra} \text{ ring property}
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$
\begin{aligned}
&= p^H \mathsf{E}\left[UU^H\right]p - p^H \mathsf{E}\left[U\boldsymbol{y}\right] - \mathsf{E}\left[\boldsymbol{y}^H U^H\right]p + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} && \text{by } \textit{linearity } \mathsf{E} \\
&= p^H \mathsf{E}\left[UU^H\right]p - (\mathsf{E}\left[U\boldsymbol{y}\right]^H p)^H - \mathsf{E}\left[U\boldsymbol{y}\right]^H p + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} \\
&= p^H Rp - (W^H p)^H - W^H p + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} \\
&= p^H Rp - (W^H p)^* - W^H p + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} \\
&= \boxed{p^H Rp - (W^H)^* p - W^H p + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y}} \\
&= p^H Rp - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right]p + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y}
\end{aligned}
$$

(Theorem 1.1 page 4

2. Proof that optimal $p_{\text{opt}} = (\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)$:

$$
\begin{aligned}
\nabla_p \mathsf{C}(p) &= \nabla_p\left[p^H Rp - (W^H)^* p - W^H p + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y}\right] && \text{by previous resut} \\
&= Rp + R^T p - \nabla_p\left[(W^H)^* p + W^H p\right] + 0 && \text{by } \textit{quadratic form} \text{ result (Theorem E.6 page 200)} \\
&= Rp + R^T p - [(W^H)^*]^T - [W^H]^T + 0 && \text{by } \textit{affine equations} \text{ result (Theorem E.3 page 199)} \\
&= Rp + (R^H)^* p - W - W^* && \text{by definition of } \textit{Hermitian Transpose}^H \\
&= Rp + R^* p - W - W^* && \text{because } R \text{ is } \textit{Hermitian symmetric} \\
&= (R + R^*)p - (W + W^*) && \text{by } \textit{ring} \text{ property} \\
&= 2(\mathbf{R}_{\mathsf{e}}Y)p - 2\mathbf{R}_{\mathsf{e}}W && \text{by definition of } \mathbf{R}_{\mathsf{e}} \text{ (Definition J.5 page 267)} \\
\implies p_{\text{opt}} &= (\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) && \text{by setting } \nabla_p \mathsf{C}(p) = 0
\end{aligned}
$$

3. Cost of optimal p_{opt}:

$$
\begin{aligned}
\mathsf{C}(p_{\text{opt}}) &= p_{\text{opt}}^H R p_{\text{opt}} - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right]p_{\text{opt}} + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} && \text{by item (1)} \\
&= [(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)]^H R[(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)] - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right][(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)] + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} && \text{by item (2)} \\
&= (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-H} R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right](\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} \\
&= (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}R^H)^{-1} R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right](\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} \\
&= (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1} R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y}
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{C}(p_{\text{opt}})|_{R\text{ real}} &= (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1} R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} \\
&= (\mathbf{R}_{\mathsf{e}}W^H)R^{-1}RR^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W) + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} \\
&= (\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W) + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} \\
&= \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} - (\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W)
\end{aligned}
$$

In many adaptive filter and equalization applications, the autocorrelation matrix U is simply the m-element random data vector $\boldsymbol{x}(k)$ at time k, as in the *Wiener-Hopf equations* (next).

Corollary 10.1 (Wiener-Hopf equations). [2]

COR

$$
\left\{ U \triangleq \boldsymbol{x}(k) \triangleq \begin{bmatrix} x(k) \\ x(k-1) \\ x(k-2) \\ \vdots \\ x(k-m+1) \end{bmatrix} \right\} \implies \left\{ \begin{aligned} \hat{\theta}_{\text{mms}} &= R^{-1}W \\ \mathsf{C}(\hat{\theta}_{\text{mms}}) &= W^T R^{-1} R R^{-1} W - 2W^T R^{-1} W + \mathsf{E}\boldsymbol{y}^T\boldsymbol{y} \end{aligned} \right\}
$$

[2] Ifeachor and Jervis (1993) pages 547–549 ⟨§"9.3 Basic Wiener filter theory"⟩, Ifeachor and Jervis (2002) pages 651–654 ⟨§"10.3 Basic Wiener filter theory"⟩, Kay (1988) page 51 ⟨§"3.3.3 Random Parameters"⟩

Copyright © 2019 Daniel J. Greenhoe

Figure 10.2: Adative filter example

✎PROOF: This is a special case of the more general case discussed in Theorem 10.1 (page 72). Here, the dimension of U is $m \times 1$ (n=1). As a result, $\boldsymbol{y}$, $\hat{\boldsymbol{y}}$, and $\boldsymbol{e}$ are simply scalar quantities (not vectors). In this special case, we have the following results (Figure 10.2 page 74):

$$\begin{aligned}
\hat{y}(\boldsymbol{p}) &\triangleq \boldsymbol{x}^T \boldsymbol{p} \\
e(\boldsymbol{p}) &\triangleq \hat{y} - y \\
\mathsf{C}(\boldsymbol{p}) &\triangleq \mathsf{E}\,\|\boldsymbol{e}\|^2 \triangleq \mathsf{E}\left[e^2\right] \\
\hat{\theta}_{\text{mms}} &\triangleq \arg\min_{\boldsymbol{p}} \mathsf{C}(\boldsymbol{p}) \\
R &\triangleq \mathsf{E}\left[\boldsymbol{x}\boldsymbol{x}^T\right] \\
W &\triangleq \mathsf{E}[\boldsymbol{x} y] \\
\mathsf{C}(\boldsymbol{p}) &= \boldsymbol{p}^T R \boldsymbol{p} - 2W^T \boldsymbol{p} + \mathsf{E}\left[\boldsymbol{y}^T \boldsymbol{y}\right] \\
\nabla_{\boldsymbol{p}} \mathsf{C}(\boldsymbol{p}) &= 2R\boldsymbol{p} - 2W \\
\mathsf{C}(\hat{\theta}_{\text{mms}})|_{R \text{ real}} &= \mathsf{E}\boldsymbol{y}^T \boldsymbol{y} - W^T R^{-1} W.
\end{aligned}$$

✏

10.2 Least squares

Theorem 10.2 (Least squares)**.** *Let*

$$\begin{aligned}
\hat{\boldsymbol{y}}(\boldsymbol{p}) &\triangleq U^H \boldsymbol{p} \\
\boldsymbol{e}(\boldsymbol{p}) &\triangleq \hat{\boldsymbol{y}} - \boldsymbol{y} \\
\mathsf{C}(\boldsymbol{p}) &\triangleq \|\boldsymbol{e}\|^2 \triangleq \boldsymbol{e}^H \boldsymbol{e} \\
\hat{\theta}_{\text{ls}} &\triangleq \arg\min_{\boldsymbol{p}} \mathsf{C}(\boldsymbol{p}) \\
R &\triangleq U U^H \\
W &\triangleq U \boldsymbol{y}.
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

Then

THM

$$\begin{aligned}
\hat{\theta}_{\mathsf{ls}} &= (\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)\\
\mathsf{C}(\boldsymbol{p}) &= \boldsymbol{p}^H R\boldsymbol{p} - (W^H\boldsymbol{p})^* - W^H\boldsymbol{p} + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y}\\
\nabla_{\boldsymbol{p}}\mathsf{C}(\boldsymbol{p}) &= 2\mathbf{R}_{\mathsf{e}}[Y]\boldsymbol{p} - 2\mathbf{R}_{\mathsf{e}}W\\
\mathsf{C}(\hat{\theta}_{\mathsf{ls}}) &= (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) + \mathsf{E}\boldsymbol{y}^H\boldsymbol{y}\\
\mathsf{C}(\hat{\theta}_{\mathsf{ls}})|_{R\ real} &= \mathsf{E}\boldsymbol{y}^H\boldsymbol{y} - (\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W).
\end{aligned}$$

✎PROOF: See APPENDIX E (page 195) for a Matrix Calculus reference.

$$\begin{aligned}
\mathsf{C}(\boldsymbol{p}) &\triangleq \|\boldsymbol{e}\|^2\\
&= \boldsymbol{e}^H\boldsymbol{e}\\
&= (\hat{\boldsymbol{y}} - \boldsymbol{y})^H(\hat{\boldsymbol{y}} - \boldsymbol{y})\\
&= \left(U^H\boldsymbol{p} - \boldsymbol{y}\right)^H\left(U^H\boldsymbol{p} - \boldsymbol{y}\right)\\
&= \left(\boldsymbol{p}^HU - \boldsymbol{y}^H\right)\left(U^H\boldsymbol{p} - \boldsymbol{y}\right)\\
&= \boldsymbol{p}^HUU^H\boldsymbol{p} - \boldsymbol{p}^HU\boldsymbol{y} - \boldsymbol{y}^HU^H\boldsymbol{p} + \boldsymbol{y}^H\boldsymbol{y}\\
&= \boldsymbol{p}^HR\boldsymbol{p} - (W^H\boldsymbol{p})^H - W^H\boldsymbol{p} + \boldsymbol{y}^H\boldsymbol{y}\\
&= \boldsymbol{p}^HR\boldsymbol{p} - (W^H\boldsymbol{p})^* - W^H\boldsymbol{p} + \boldsymbol{y}^H\boldsymbol{y}\\
&= \boldsymbol{p}^HR\boldsymbol{p} - (W^H)^*\boldsymbol{p} - W^H\boldsymbol{p} + \boldsymbol{y}^H\boldsymbol{y}\\
&= \boldsymbol{p}^HR\boldsymbol{p} - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right]\boldsymbol{p} + \boldsymbol{y}^H\boldsymbol{y}
\end{aligned}$$

$$\begin{aligned}
\nabla_{\boldsymbol{p}}\mathsf{C}(\boldsymbol{p}) &= \nabla_{\boldsymbol{p}}\left[\boldsymbol{p}^HR\boldsymbol{p} - (W^H)^*\boldsymbol{p} - W^H\boldsymbol{p} + \boldsymbol{y}^H\boldsymbol{y}\right]\\
&= R\boldsymbol{p} + R^T\boldsymbol{p} - [(W^H)^*]^T - [W^H]^T + 0\\
&= R\boldsymbol{p} + (R^H)^*\boldsymbol{p} - W - W^*\\
&= R\boldsymbol{p} + R^*\boldsymbol{p} - W - W^*\\
&= (R + R^*)\boldsymbol{p} - (W + W^*)\\
&= 2(\mathbf{R}_{\mathsf{e}}Y)\boldsymbol{p} - 2\mathbf{R}_{\mathsf{e}}W
\end{aligned}$$

$$\boldsymbol{p}_{\mathsf{opt}} = (\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)$$

$$\begin{aligned}
\mathsf{C}(\boldsymbol{p}_{\mathsf{opt}}) &= \boldsymbol{p}_{\mathsf{opt}}^HR\boldsymbol{p}_{\mathsf{opt}} - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right]\boldsymbol{p}_{\mathsf{opt}} + \boldsymbol{y}^H\boldsymbol{y}\\
&= [(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)]^HR[(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)] - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right][(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W)] + \boldsymbol{y}^H\boldsymbol{y}\\
&= (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-H}R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right](\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) + \boldsymbol{y}^H\boldsymbol{y}\\
&= (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}R^H)^{-1}R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2\mathbf{R}_{\mathsf{e}}\left[W^H\right](\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) + \boldsymbol{y}^H\boldsymbol{y}\\
&= (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) + \boldsymbol{y}^H\boldsymbol{y}
\end{aligned}$$

$$\begin{aligned}
\mathsf{C}(\boldsymbol{p}_{\mathsf{opt}})|_{R\ \mathrm{real}} &= (\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}R(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)(\mathbf{R}_{\mathsf{e}}Y)^{-1}(\mathbf{R}_{\mathsf{e}}W) + \boldsymbol{y}^H\boldsymbol{y}\\
&= (\mathbf{R}_{\mathsf{e}}W^H)R^{-1}RR^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W) + \boldsymbol{y}^H\boldsymbol{y}\\
&= (\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W) - 2(\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W) + \boldsymbol{y}^H\boldsymbol{y}\\
&= \boldsymbol{y}^H\boldsymbol{y} - (\mathbf{R}_{\mathsf{e}}W^H)R^{-1}(\mathbf{R}_{\mathsf{e}}W)
\end{aligned}$$

✏

Example 10.1 (Polynomial approximation).
Suppose we **know** the locations $\{(x_n, y_n)|n = 1,2,3,4,5\}$ of 5 data points. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ represent the

locations of these points such that

$$\boldsymbol{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} \qquad \boldsymbol{y} \triangleq \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}$$

Suppose we want to find a second order polynomial

$$cx^2 + bx + a$$

that best approximates these 5 points in the least squares sense. We define the matrix U (known) and vector $\hat{\theta}$ (to be computed) as follows:

$$U^H \triangleq \underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \end{bmatrix}}_{\text{Vandermonde matrix}}{}^{3} \qquad \hat{\theta} \triangleq \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Then, using Theorem 10.2 (page 74), the best coefficients $\hat{\theta}$ for the polynomial are

$$\begin{aligned}
\hat{\theta} &= \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\
&= R^{-1}W \\
&= (UU^H)^{-1}\,(U\boldsymbol{y}) \\
&= \left(\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \end{bmatrix}^H \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \end{bmatrix}\right)^{-1} \left(\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \end{bmatrix}^H \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}\right)
\end{aligned}$$

[3] Horn and Johnson (1990)29

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 11

GRADIENT SEARCH TECHNIQUES

11.1 Gradient search techniques

One of the biggest advantages of using a gradient search technique is that they can be implemented *recursively* as shown in the next equation. The general form of the gradient search parameter estimation techniques is[1]

THM $\boldsymbol{p}_n = \boldsymbol{p}_{n-1} - \eta_{n-1} R \left[\nabla_p \mathsf{C}(\boldsymbol{p}_n)\right]$ where at time n

$\boldsymbol{p}_n$	is the *state*	(vector)
η_n	is the *step size*	(scalar)
Y	is the *direction*	(matrix)
$\nabla_p \mathsf{C}(\boldsymbol{p}_n)$	is the *gradient* of the cost function $\mathsf{C}(\boldsymbol{p}_n)$	(vector)

Two major categories of gradient search techniques are

- steepest descent (includes LMS)
- Newton's method (includes RLS and Kalman filters).

The key difference between the two is that ***steepest descent* uses only first derivative information**, while ***Newton's method* uses both first and second derivative information** making it converge much faster but with significantly higher complexity.

First derivative techniques

Steepest descent. In this algorithm, $R = I$ (identity matrix). First derivative information is contained in $\nabla\mathsf{C}$. Second derivative information, if present, is contained in Y. Thus, steepest descent algorithms do not use second derivative information.

THM $\boldsymbol{p}_n = \boldsymbol{p}_{n-1} - \eta_{n-1} \left[\nabla_p \mathsf{C}(\boldsymbol{p}_n)\right]$

[1] Nelles (2001), page 90

Least Mean Squares (LMS). [2] This is a special case of *steepest descent*. In minimum mean square estimation (Section 10.1 page 72), the cost function $\mathsf{C}(\boldsymbol{p})$ is defined as a *statistical average* of the error vector such that $\mathsf{C}(\boldsymbol{p}) = \mathsf{E}\left[\boldsymbol{e}^H\boldsymbol{e}\right]$. In this case the gradient $\nabla\mathsf{C}$ is difficult to compute. However, the LMS algorithm greatly simplifies the problem by instead defining the cost function as a function of the *instantaneous error* such that

$$\begin{aligned}
\boldsymbol{y} &= y(n)\\
\hat{\boldsymbol{y}} &= \hat{y}(n)\\
\mathsf{C}(\boldsymbol{p}) &= \|e(n)\|^2\\
&= e^2(n)\\
&= (\hat{y}(n) - y(n))^2
\end{aligned}$$

Computing the gradient of this cost function is then just a special case of *least squares estimation* (Section 10.2 page 74). Using LS, we let $U = \boldsymbol{x}^T$ and hence

$$\begin{aligned}
\nabla_{\boldsymbol{p}}\mathsf{C}(\boldsymbol{p}) &= 2U^TU\boldsymbol{p} - 2U^T\boldsymbol{y} && \text{by Theorem 10.2 page 74}\\
&= 2\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{p} - 2\boldsymbol{x}y && \text{by above definitions}\\
&= 2\boldsymbol{x}\hat{y} - 2\boldsymbol{x}y\\
&= 2\boldsymbol{x}(\hat{y} - y)\\
&= 2\boldsymbol{x}e(n)
\end{aligned}$$

The LMS algorithm uses this instantaneous gradient for $\nabla\mathsf{C}$, lets $R = I$, and uses a constant step size η to give

THM $\boldsymbol{p}_n = \boldsymbol{p}_{n-1} - 2\eta\boldsymbol{x}_n e(n)$

Second derivative techniques

Newton's Method. This algorithm uses the *Hessian* matrix H, which is the second derivative of the cost function $\mathsf{C}(\boldsymbol{p})$, and lets $R = H^{-1}$.

$$\begin{aligned}
H_n &\triangleq && \nabla_{\boldsymbol{p}}\nabla_{\boldsymbol{p}}\mathsf{C}(\boldsymbol{p}_n)\\
\boldsymbol{p}_n &= \boldsymbol{p}_{n-1} - \eta_{n-1}H_n^{-1}\left[\nabla_{\boldsymbol{p}}\mathsf{C}(\boldsymbol{p}_n)\right]
\end{aligned}$$

Kalman filtering [3]

$$\begin{aligned}
\gamma(k) &= \frac{1}{x^T(k)P(k-1)\boldsymbol{x}(k) + 1}P(k-1)x(k)\\
P(k) &= (I - \gamma(k)x^T(k))P(k-1) + V\\
e(k) &= y(k) - x^T(k)\hat{\boldsymbol{p}}(k-1)\\
\hat{\boldsymbol{p}}(k) &= \hat{\boldsymbol{p}}(k-1) + \gamma(k)e(k)
\end{aligned}$$

[2] Manolakis et al. (2000), page 526
[3] Nelles (2001), page 66

Copyright © 2019 Daniel J. Greenhoe

Recursive Least Squares (RLS) [4] This algorithm is a special case of either the RLS with forgetting or the Kalman filter.

$$\begin{aligned}
\gamma(k) &= \frac{1}{x^T(k)P(k-1)\boldsymbol{x}(k)+1}P(k-1)x(k)\\
P(k) &= (I-\gamma(k)x^T(k))P(k-1)\\
e(k) &= y(k)-x^T(k)\hat{\boldsymbol{p}}(k-1)\\
\hat{\boldsymbol{p}}(k) &= \hat{\boldsymbol{p}}(k-1)+\gamma(k)e(k)
\end{aligned}$$

11.2 Direct search

A direct search algorithm may be used in cases where the cost function over $\boldsymbol{p}$ has several local minimums, making convergence difficult. Furthermore, direct search algorithms can be very computationally demanding.

[4] Nelles (2001), page 66

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 12

KL-EXPANSION APPLICATION

12.1 Sufficient statistics

Theorem 12.1 (page 81) (next) shows that the finite set $Y \triangleq \{\dot{y}_n | n = 1, 2, \ldots, N\}$ provides just as much information as having the entire $\mathsf{y}(t)$ waveform (an uncountably infinite number of values) with respect to the following cases:

1. the conditional probability of $\mathsf{x}(t;\hat{\theta})$ given $\mathsf{y}(t)$
2. the *MAP estimate* of the information sequence
3. the *ML estimate* of the information sequence.

That is, even with a drastic reduction in the amount of information from uncountably infinite to finite N, no information is lost with respect to the quantities listed above.

This amazing result is very useful in practical system implementation and also for proving other theoretical results (notably estimation and detection theorems which come later in this chapter).

Theorem 12.1 (Sufficient statistic theorem). [1] *Let Ψ be an orthonormal basis for $\mathsf{x}(t;\hat{\theta})$ such that*

$$\begin{aligned} \mathsf{y}(t) &= \mathsf{x}(t;\hat{\theta}) + \mathsf{v}(t) \\ \Psi &= \{\psi_n | n = 1, 2, \ldots, N\} \end{aligned}$$

*Then $Y \triangleq \{\dot{y}_n | n = 1, 2, \ldots, N\}$ is a **sufficient statistic** for $\mathsf{y}(t)$ such that…*

THM

$$\left\{ \mathsf{v}(t) \text{ is AWGN} \right\} \Longrightarrow \left\{ \begin{array}{llll} (1). & \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|\mathsf{y}(t)\right\} & = & \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|Y\right\} \\ (2). & \hat{\theta}_{\mathsf{map}} \triangleq \arg\max_{\hat{\theta}} \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|\mathsf{y}(t)\right\} & = & \arg\max_{\hat{\theta}} \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|Y\right\} \\ (3). & \hat{\theta}_{\mathsf{ml}} \triangleq \arg\max_{\hat{\theta}} \mathsf{P}\left\{\mathsf{y}(t)|\mathsf{x}(t;\hat{\theta})\right\} & = & \arg\max_{\hat{\theta}} \mathsf{P}\left\{Y|\mathsf{x}(t;\hat{\theta})\right\} \end{array} \right\}$$

✎PROOF: Let $\mathsf{v}'(t) \triangleq \mathsf{v}(t) - \sum_{n=1}^{N} \dot{v}_n \psi_n(t)$.

[1] Fisher (1922) page 316 ⟨"Criterion of Sufficiency"⟩

1. The relationship between Y and $\mathsf{v}'(t)$ is given by

$$\begin{aligned}
\mathsf{y}(t) &= \sum_{n=1}^{N} \langle \mathsf{y}(t) \mid \psi_n(t) \rangle \psi_n(t) + \left[\mathsf{y}(t) - \sum_{n=1}^{N} \langle \mathsf{y}(t) \mid \psi_n(t) \rangle \psi_n(t) \right] \\
&= \sum_{n=1}^{N} \langle \mathsf{y}(t) \mid \psi_n(t) \rangle \psi_n(t) + \left[\mathsf{y}(t) - \sum_{n=1}^{N} \langle \mathsf{x}(t) + \mathsf{v}(t) \mid \psi_n(t) \rangle \psi_n(t) \right] \\
&= \sum_{n=1}^{N} \dot{y}_n \psi_n(t) + \left[\mathsf{x}(t) + \mathsf{v}(t) - \sum_{n=1}^{N} \langle \mathsf{x}(t) \mid \psi_n(t) \rangle \psi_n(t) - \sum_{n=1}^{N} \langle \mathsf{v}(t) \mid \psi_n(t) \rangle \psi_n(t) \right] \\
&= \sum_{n=1}^{N} \dot{y}_n \psi_n(t) + \mathsf{x}(t) + \mathsf{v}(t) - \mathsf{x}(t) - \left[\mathsf{v}(t) - \mathsf{v}'(t) \right] \\
&= \sum_{n=1}^{N} \dot{y}_n \psi_n(t) + \mathsf{v}'(t).
\end{aligned}$$

2. lemma: Y and $\mathsf{v}'(t)$ are *uncorrelated*:

$$\begin{aligned}
\mathsf{E}\left[\dot{y}_n \mathsf{v}'(t)\right] &= \mathsf{E}\left[\langle \mathsf{y}(t) \mid \psi_n(t) \rangle \left(\mathsf{v}(t) - \sum_{n=1}^{N} \langle \mathsf{v}(t) \mid \psi_n(t) \rangle \psi_n(t) \right) \right] \\
&= \mathsf{E}\left[\langle \mathsf{x}(t) + \mathsf{v}(t) \mid \psi_n(t) \rangle \left(\mathsf{v}(t) - \sum_{n=1}^{N} \langle \mathsf{v}(t) \mid \psi_n(t) \rangle \psi_n(t) \right) \right] \\
&= \mathsf{E}\left[\left(\langle \mathsf{x}(t) \mid \psi_n(t) \rangle + \langle \mathsf{v}(t) \mid \psi_n(t) \rangle \right) \left(\mathsf{v}(t) - \sum_{n=1}^{N} \langle \mathsf{v}(t) \mid \psi_n(t) \rangle \psi_n(t) \right) \right] \\
&= \mathsf{E}\left[\left(\dot{x}_n + \dot{\upsilon}_n \right) \left(\mathsf{v}(t) - \sum_{n=1}^{N} \dot{\upsilon}_n \psi_n(t) \right) \right] \\
&= \mathsf{E}\left[\dot{x}_n \mathsf{v}(t) - \dot{x}_n \sum_{n=1}^{N} \dot{\upsilon}_n \psi_n(t) + \dot{\upsilon}_n \mathsf{v}(t) - \dot{\upsilon}_n \sum_{n=1}^{N} \dot{\upsilon}_n \psi_n(t) \right] \\
&= \mathsf{E}\left[\dot{x}_n \mathsf{v}(t)\right] - \mathsf{E}\left[\dot{x}_n \sum_{n=1}^{N} \dot{\upsilon}_n \psi_n(t) \right] + \mathsf{E}\left[\langle \mathsf{v}(t) \mid \psi_n(t) \rangle \mathsf{v}(t) \right] - \mathsf{E}\left[\sum_{m=1}^{N} \dot{\upsilon}_n \dot{\upsilon}_m \psi_m(t) \right] \\
&= \dot{x}_n \cancelto{0}{\mathsf{E}\mathsf{v}(t)} - \dot{x}_n \sum_{n=1}^{N} \cancelto{0}{\mathsf{E}\left[\dot{\upsilon}_n\right]} \psi_n(t) + \mathsf{E}\left[\left\langle \mathsf{v}(t)\mathsf{v}(\hat{\theta}) \mid \psi_n(\hat{\theta}) \right\rangle \right] - \sum_{m=1}^{N} \mathsf{E}\left[\dot{\upsilon}_n \dot{\upsilon}_m\right] \psi_m(t) \\
&= 0 - 0 + \left\langle \mathsf{E}\left[\mathsf{v}(t) n(\hat{\theta})\right] \mid \psi_n(\hat{\theta}) \right\rangle - \sum_{m=1}^{N} N_o \bar{\delta}_{mn} \psi_m(t) \qquad \text{(because } \dot{\upsilon}_n \text{ is white)} \\
&= \left\langle N_o \delta(t - \hat{\theta}) \mid \psi_n(\hat{\theta}) \right\rangle - N_o \psi_n(t) \\
&= N_o \psi_n(t) - N_o \psi_n(t) \\
&= 0 \\
&\Longrightarrow \textbf{uncorrelated}
\end{aligned}$$

3. lemma: $\dot{y}_n$ and $\mathsf{v}'(t)$ are *uncorrelated*. Proof:
   By (2) lemma, $\dot{y}_n$ and $\mathsf{v}'(t)$ are *uncorrelated*. By hypothesis, they are *Gaussian*, and thus are also **independent**.

4. Proof that $\mathsf{P}\left\{\mathsf{x}(t;\hat{\theta}) | \mathsf{y}(t)\right\} = \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta}) | \dot{y}_1, \dot{y}_2, \ldots, \dot{y}_N\right\}$:

Copyright © 2019 Daniel J. Greenhoe

$$
\begin{aligned}
\mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|\mathsf{y}(t)\right\} &= \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|\sum_{n=1}^{N}\dot{y}_n\psi_n(t)+\mathsf{v}'(t)\right\} \\
&= \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|Y,\mathsf{v}'(t)\right\} && \text{because } Y \text{ and } \mathsf{v}'(t) \text{ can be extracted by } \langle\cdots\,|\,\psi_n(t)\rangle \\
&= \frac{\mathsf{P}\left\{Y,\mathsf{v}'(t)|\mathsf{x}(t;\hat{\theta})\right\}P\left\{\mathsf{x}(t;\hat{\theta})\right\}}{\mathsf{P}\{Y,\mathsf{v}'(t)\}} \\
&= \frac{\mathsf{P}\left\{Y|\mathsf{x}(t;\hat{\theta})\right\}\mathsf{P}\left\{\mathsf{v}'(t)|\mathsf{x}(t;\hat{\theta})\right\}\mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})\right\}}{\mathsf{P}\{Y\}\mathsf{P}\{\mathsf{v}'(t)\}} && \text{by } \textit{independence} \text{ of } Y \text{ and } \mathsf{v}'(t) \text{ ((3) lemma page 82)} \\
&= \frac{\mathsf{P}\left\{Y|\mathsf{x}(t;\hat{\theta})\right\}\mathsf{P}\left\{\mathsf{v}'(t)\right\}\mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})\right\}}{\mathsf{P}\{Y\}\mathsf{P}\{\mathsf{v}'(t)\}} && \text{by } \textit{independence} \text{ of } \mathsf{x} \text{ and } \mathsf{v} \\
&= \frac{\mathsf{P}\left\{Y|\mathsf{x}(t;\hat{\theta})\right\}\mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})\right\}}{\mathsf{P}\{Y\}} \\
&= \frac{\mathsf{P}\left\{Y,\mathsf{x}(t;\hat{\theta})\right\}}{\mathsf{P}\{Y\}} \\
&= \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|Y\right\} && \text{by definition of } \textit{conditional probability} \text{ (Definition A.4 pag}
\end{aligned}
$$

5. Proof that Y is a sufficient statistic for the *MAP estimate*:

$$
\begin{aligned}
\hat{\theta}_{\mathsf{map}} &\triangleq \arg\max_{\hat{\theta}} \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|\mathsf{y}(t)\right\} && \text{by definition of } \textit{MAP estimate} \\
&= \arg\max_{\hat{\theta}} \mathsf{P}\left\{\mathsf{x}(t;\hat{\theta})|Y\right\} && \text{by item (4)}
\end{aligned}
$$

6. Proof that Y is a *sufficient statistic* for the *ML estimate*:

$$
\begin{aligned}
\hat{\theta}_{\mathsf{ml}} &\triangleq \arg\max_{\hat{\theta}} \mathsf{P}\left\{\mathsf{y}(t)|\mathsf{x}(t;\hat{\theta})\right\} && \text{by definition of } \textit{ML estimate} \\
&= \arg\max_{\hat{\theta}} \mathsf{P}\left\{\sum_{n=1}^{N}\dot{y}_n\psi_n(t)+\mathsf{v}'(t)|\mathsf{x}(t;\hat{\theta})\right\} \\
&= \arg\max_{\hat{\theta}} \mathsf{P}\left\{Y,\mathsf{v}'(t)|\mathsf{x}(t;\hat{\theta})\right\} && \text{because } Y \text{ and } \mathsf{v}'(t) \text{ can be extracted by } \langle\cdots\,|\,\psi_n(t)\rangle \\
&= \arg\max_{\hat{\theta}} \mathsf{P}\left\{Y|\mathsf{x}(t;\hat{\theta})\right\}\mathsf{P}\left\{\mathsf{v}'(t)\right\}\mathsf{x}(t;\hat{\theta}) && \text{by } \textit{independence} \text{ of } Y \text{ and } \mathsf{v}'(t) \\
&= \arg\max_{\hat{\theta}} \mathsf{P}\left\{Y|\mathsf{x}(t;\hat{\theta})\right\}\mathsf{P}\left\{\mathsf{v}'(t)\right\} && \text{by } \textit{independence} \text{ of } \mathsf{x}(t) \text{ and } \mathsf{v}'(t) \\
&= \arg\max_{\hat{\theta}} \mathsf{P}\left\{Y|\mathsf{x}(t;\hat{\theta})\right\} && \text{by } \textit{independence} \text{ of } \mathsf{v}'(t) \text{ and } \hat{\theta}
\end{aligned}
$$

✍

Theorem 12.2. *Let* $\mathbf{C} = \mathbf{C}_{\mathsf{a}}$ *be an additive noise channel.*

THM

$$\underbrace{\mathbf{C} = \mathbf{C}_{\mathsf{a}}}_{\textit{additive noise channel}} \implies \left\{\ \mathsf{E}(\dot{y}_n|\theta) = \dot{x}_n(\theta) + \mathsf{E}\dot{v}_n\ \right\}$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

✎PROOF:

$$\begin{aligned}
\mathsf{E}(\dot{y}_n|\theta) &\triangleq \left(\langle \mathsf{y}(t) \,|\, \psi_n(t)\rangle \,|\theta\right) \\
&= \langle \mathsf{x}(t;\theta) + \mathsf{n}(t) \,|\, \psi_n(t)\rangle \\
&= \langle \mathsf{x}(t;\theta) \,|\, \psi_n(t)\rangle + \langle \mathsf{n}(t) \,|\, \psi_n(t)\rangle \\
&= \left\langle \sum_{k=1}^{N} \dot{x}_k(\theta)\psi_k(t) \,|\, \psi_n(t) \right\rangle + \dot{\upsilon}_n \\
&= \sum_{k=1}^{N} \dot{x}_k(\theta) \langle \psi_k(t) \,|\, \psi_n(t)\rangle + \dot{\upsilon}_n \\
&= \dot{x}_n(\theta) + \dot{\upsilon}_n
\end{aligned}$$

$$\begin{aligned}
\mathsf{E}(\dot{y}_n|\theta) &= \mathsf{E}\left[\dot{x}_n(\theta) + \dot{\upsilon}_n\right] \\
&= \mathsf{E}\dot{x}_n(\theta) + \mathsf{E}\dot{\upsilon}_n \\
&= \dot{x}_n(\theta)
\end{aligned}$$

✏

Figure 12.1: Additive *Gaussian* noise channel Statistics

Theorem 12.3. *Let* $\mathbf{C} = \mathbf{C}_{\mathsf{agn}}$ *be an additive gaussian noise channel with distribution* $\mathsf{n}(t) \sim \mathsf{N}\left(0, \sigma^2\right)$ *for all* t.

THM

$$\underbrace{\mathbf{C} = \mathbf{C}_{\mathsf{agn}}}_{\textit{additive GAUSSIAN channel}} \implies \left\{ \begin{array}{rcll}
\mathsf{E}\dot{\upsilon}_n &=& 0 & \\
\mathsf{E}(\dot{y}_n|\theta) &=& \dot{x}_n(\theta) & \\
\dot{\upsilon}_n &\sim& \mathsf{N}\left(0, \sigma^2\right) & \textit{(noise projections are GAUSSIAN)} \\
\dot{y}_n|\theta &\sim& \mathsf{N}\left(\dot{x}_n(\theta), \sigma^2\right) & \textit{(receiver projections are GAUSSIAN)}
\end{array} \right\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{E}\dot{\upsilon}_n &= \mathsf{E}\,\langle \mathsf{n}(t) \,|\, \psi_n(t)\rangle \\
&= \langle \mathsf{E}\mathsf{n}(t) \,|\, \psi_n(t)\rangle \\
&= \langle 0 \,|\, \psi_n(t)\rangle \\
&= 0
\end{aligned}$$

$$\begin{aligned}
(\dot{y}_n|\theta) &\triangleq \langle \mathsf{y}(t) \,|\, \psi_n(t)\rangle|\theta \\
&= \langle \mathsf{x}(t;\theta) + \mathsf{n}(t) \,|\, \psi_n(t)\rangle \\
&= \langle \mathsf{x}(t;\theta) \,|\, \psi_n(t)\rangle + \langle \mathsf{n}(t) \,|\, \psi_n(t)\rangle \\
&= \left\langle \sum_{k=1}^{N} \dot{x}_k(\theta)\psi_k(t) \,|\, \psi_n(t) \right\rangle + \dot{\upsilon}_n \\
&= \sum_{k=1}^{N} \dot{x}_k(\theta) \langle \psi_k(t) \,|\, \psi_n(t)\rangle + \dot{\upsilon}_n
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \dot{x}_n(\theta) + \dot{\upsilon}_n$$

$$\begin{aligned}
\mathsf{E}(\dot{y}_n|\theta) &= \mathsf{E}\left[\dot{x}_n(\theta) + \dot{\upsilon}_n\right] \\
&= \mathsf{E}\dot{x}_n(\theta) + \mathsf{E}\dot{\upsilon}_n \\
&= \dot{x}_n(\theta)
\end{aligned}$$

The distributions follow because they are linear operations on Gaussian processes. ✏

Theorem 12.4. *Let* $\mathbf{C} = \mathbf{C}_{\mathsf{awn}}$ *be an additive white noise channel.*

THM

$$\underbrace{\mathbf{C} = \mathbf{C}_{\mathsf{awn}}}_{\textit{additive white channel}} \implies \left\{\begin{array}{lcll}
\mathsf{E}\dot{\upsilon}_n & = & 0 & \textit{(noise projection is zero-mean)} \\
\mathsf{E}(\dot{y}_n|\theta) & = & \dot{x}_n(\theta) & \textit{(expected receiver projection = transmitted projection)} \\
\mathrm{cov}\left[\dot{\upsilon}_n, \dot{\upsilon}_m\right] & = & \sigma^2\bar{\delta}_{nm} & \textit{(noise projections are UNCORRELATED)} \\
\mathrm{cov}\left[\dot{y}_n|\theta, \dot{y}_m|\theta\right] & = & \sigma^2\bar{\delta}_{nm} & \textit{(receiver projections are UNCORRELATED)}
\end{array}\right\}$$

✎PROOF: Because the noise is additive (see Theorem 12.2 (page 83))

$$\begin{aligned}
\mathsf{E}\dot{\upsilon}_n &= 0 \\
(\dot{y}_n|\theta) &= \dot{x}_n(\theta) + \dot{\upsilon}_n \\
\mathsf{E}(\dot{y}_n|\theta) &= \dot{x}_n(\theta).
\end{aligned}$$

Because the noise is also *white*,

$$\begin{aligned}
\mathrm{cov}\left[\dot{\upsilon}_m, \dot{\upsilon}_n\right] &= \mathrm{cov}\left[\langle \mathsf{n}(t) \,|\, \psi_m(t)\rangle, \langle \mathsf{n}(t) \,|\, \psi_n(t)\rangle\right] \\
&= \mathsf{E}\left[\langle \mathsf{n}(t) \,|\, \psi_m(t)\rangle \langle \mathsf{n}(t) \,|\, \psi_n(t)\rangle\right] \\
&= \mathsf{E}\left[\langle \mathsf{n}(t) \,|\, \psi_m(t)\rangle \left\langle n(\hat{\theta}) \,|\, \psi_n(\hat{\theta})\right\rangle\right] \\
&= \mathsf{E}\left[\left\langle n(\hat{\theta}) \langle \mathsf{n}(t) \,|\, \psi_m(t)\rangle \,|\, \psi_n(\hat{\theta})\right\rangle\right] \\
&= \mathsf{E}\left[\left\langle \left\langle n(\hat{\theta})\mathsf{n}(t) \,|\, \psi_m(t)\right\rangle \,|\, \psi_n(\hat{\theta})\right\rangle\right] \\
&= \left\langle \left\langle \mathsf{E}\left[n(\hat{\theta})\mathsf{n}(t)\right] \,|\, \psi_m(t)\right\rangle \,|\, \psi_n(\hat{\theta})\right\rangle \\
&= \left\langle \left\langle \sigma^2\delta(t-\hat{\theta}) \,|\, \psi_m(t)\right\rangle \,|\, \psi_n(\hat{\theta})\right\rangle \\
&= \sigma^2 \langle \psi_n(t) \,|\, \psi_m(t)\rangle \\
&= \left\{\begin{array}{ll} \sigma^2 & \text{for } n = m \\ 0 & \text{for } n \neq m. \end{array}\right.
\end{aligned}$$

$$\begin{aligned}
\mathrm{cov}\left[\dot{y}_n|\theta, \dot{y}_m|\theta\right] &= \mathsf{E}\left[\dot{y}_n\dot{y}_m|\theta\right] - [\mathsf{E}\dot{y}_n|\theta][\mathsf{E}\dot{y}_m|\theta] \\
&= \mathsf{E}\left[(\dot{x}_n(\theta) + \dot{\upsilon}_n)(\dot{x}_m(\theta) + \dot{\upsilon}_m)\right] - \dot{x}_n(\theta)\dot{x}_m(\theta) \\
&= \mathsf{E}\left[(\dot{x}_n(\theta) + \dot{\upsilon}_n)(\dot{x}_m(\theta) + \dot{\upsilon}_m)\right] - \dot{x}_n(\theta)\dot{x}_m(\theta) \\
&= \mathsf{E}\left[\dot{x}_n(\theta)\dot{x}_m(\theta) + \dot{x}_n(\theta)\dot{\upsilon}_m + \dot{\upsilon}_n\dot{x}_m(\theta) + \dot{\upsilon}_n\dot{\upsilon}_m\right] - \dot{x}_n(\theta)\dot{x}_m(\theta) \\
&= \dot{x}_n(\theta)\dot{x}_m(\theta) + \dot{x}_n(\theta)\mathsf{E}\left[\dot{\upsilon}_m\right] + \mathsf{E}\left[\dot{\upsilon}_n\right]\dot{x}_m(\theta) + \mathsf{E}\left[\dot{\upsilon}_n\dot{\upsilon}_m\right] - \dot{x}_n(\theta)\dot{x}_m(\theta) \\
&= 0 + \dot{x}_n(\theta)\cdot 0 + 0\cdot\dot{x}_m(\theta) + \mathrm{cov}\left[\dot{\upsilon}_n, \dot{\upsilon}_m\right] + [\mathsf{E}\dot{\upsilon}_n][\mathsf{E}\dot{\upsilon}_m] \\
&= \sigma^2\bar{\delta}_{nm} + 0\cdot 0 \\
&= \left\{\begin{array}{ll} \sigma^2 & \text{for } n = m \\ 0 & \text{for } n \neq m. \end{array}\right.
\end{aligned}$$

✏

Theorem 12.5. *Let* $\mathbf{C} = \mathbf{C}_{\mathsf{awgn}}$ *be an additive gaussian noise channel with distribution* $\mathsf{n}(t) \sim \mathsf{N}\left(0, \sigma^2\right)$ *for all* t.

COPYRIGHT © 2019 DANIEL J. GREENHOE

Figure 12.2: Additive white *Gaussian* noise channel statistics

THM

$$\underbrace{\mathbf{C} = \mathbf{C}_{\mathsf{awgn}}}_{AWGN} \Longrightarrow \begin{cases} \dot{\upsilon}_n \sim \mathsf{N}\left(0, \sigma^2\right) & \textit{(noise projections are GAUSSIAN)} \\ \dot{y}_n|\theta \sim \mathsf{N}\left(\dot{x}_n(\theta), \sigma^2\right) & \textit{(receiver projections are GAUSSIAN)} \\ \operatorname{cov}\left[\dot{\upsilon}_n, \dot{\upsilon}_m\right] = \sigma^2\bar{\delta}_{nm} & \textit{(noise projections are UNCORRELATED)} \\ \operatorname{cov}\left[\dot{y}_n, \dot{y}_m\right] = \sigma^2\bar{\delta}_{nm} & \textit{(receiver projections are UNCORRELATED)} \\ \mathsf{P}\{\dot{\upsilon}_n = a \wedge \dot{\upsilon}_m = b\} = \mathsf{P}\{\dot{\upsilon}_n = a\}\mathsf{P}\{\dot{\upsilon}_m = b\} & \textit{(noise projections are INDEPENDENT)} \\ \mathsf{P}\{\dot{y}_n = a \wedge \dot{y}_m = b\} = \mathsf{P}\{\dot{y}_n = a\}\mathsf{P}\{\dot{y}_m = b\} & \textit{(receiver projections are INDEPENDENT)} \end{cases}$$

✎PROOF: The distributions follow because they are *linear* operations on *Gaussian* processes.

By Theorem 12.4 (page 85) (for AWN channel)

$$\begin{aligned} \mathsf{E}\dot{\upsilon}_n &= 0 \\ \operatorname{cov}\left[\dot{\upsilon}_m, \dot{\upsilon}_n\right] &= \sigma^2\bar{\delta}_{mn} \\ \dot{y}_n &= \dot{x}_n + \dot{\upsilon}_n \\ \mathsf{E}\dot{y}_n &= \dot{x}_n \\ \operatorname{cov}\left[\dot{y}_n, \dot{y}_m\right] &= \sigma^2\bar{\delta}_{mn} \end{aligned}$$

Because the processes are *Gaussian*, *uncorrelatedness* implies *independence*. ✏

12.2 Optimal symbol estimation

The AWGN projection statistics provided by Theorem 12.5 (page 85) help generate the optimal ML-estimates for a number of communication systems. These ML-estimates can be expressed in either of two standard forms:

- **Spectral decompostion**: The optimal estimate is expressed in terms of *projections* of signals onto orthonormal basis functions.
- **Matched signal**: The optimal estimate is expressed in terms of the (noisy) received signal correlated with ("matched" with) the (noiseless) transmitted signal.

Theorem 12.6 (page 86) (next) expresses the general optimal *ML estimate* in both of these forms.

Parameter detection is a special case of parameter estimation. In parameter detection, the estimate is a member of an finite set. In parameter estimation, the estimate is a member of an infinite set (Section 12.2 page 86).

Theorem 12.6 (General ML estimation)**.** *Let Ψ be an orthonormal set spanning* $\mathsf{x}(t)$ *such that*

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$\begin{aligned}
\Psi &\triangleq (\!(\psi_1(t), \psi_2(t), \dots, \psi_n(t))\!) \\
\dot{y}_n &\triangleq \langle \mathsf{y}(t) \mid \psi_n(t) \rangle \\
\dot{x}_n &\triangleq \langle \mathsf{x}(t) \mid \psi_n(t) \rangle \\
\mathsf{y}(t) &= \mathsf{x}(t;\hat{\theta}) + \mathsf{n}(t).
\end{aligned}$$

Then the optimal ML-estimate $\hat{\theta}_{\mathsf{ml}}$ of parameter θ is

THM

$$\begin{aligned}
\hat{\theta}_{\mathsf{ml}} &= \arg\min_{\hat{\theta}} \left[\sum_{n=1}^{N} [\dot{y}_n - \dot{x}_n(\hat{\theta})]^2 \right] && \textit{(spectral decomposition)} \\
&= \arg\max_{\hat{\theta}} \left[2 \left\langle \mathsf{y}(t) \mid \mathsf{x}(t;\hat{\theta}) \right\rangle - \left\| \mathsf{x}(t;\hat{\theta}) \right\|^2 \right] && \textit{(matched signal)}
\end{aligned}$$

✎PROOF:

$$\begin{aligned}
\hat{\theta}_{\mathsf{ml}} &= \arg\max_{\hat{\theta}} \mathsf{P}\left\{ \mathsf{y}(t) | \mathsf{x}(t;\hat{\theta}) \right\} \\
&= \arg\max_{\hat{\theta}} \mathsf{P}\left\{ \dot{y}_1, \dot{y}_2, \dots, \dot{y}_n | \mathsf{x}(t;\hat{\theta}) \right\} && \text{by Theorem 12.1 (page 81)} \\
&= \arg\max_{\hat{\theta}} \prod_{n=1}^{N} \mathsf{P}\left\{ \dot{y}_n | \mathsf{x}(t;\hat{\theta}) \right\} \\
&= \arg\max_{\hat{\theta}} \prod_{n=1}^{N} \mathsf{p}\left[\dot{\mathsf{y}}_{\mathsf{n}} | \mathsf{x}(\mathsf{t};\hat{\theta}) \right] \\
&= \arg\max_{\hat{\theta}} \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{[\dot{y}_n - \dot{x}_n(\hat{\theta})]^2}{-2\sigma^2} && \text{by Theorem 12.5 (page 85)} \\
&= \arg\max_{\hat{\theta}} \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^N \exp \frac{-1}{2\sigma^2} \sum_{n=1}^{N} [\dot{y}_n - \dot{x}_n(\hat{\theta})]^2 \\
&= \arg\max_{\hat{\theta}} \left[-\sum_{n=1}^{N} [\dot{y}_n - \dot{x}_n(\hat{\theta})]^2 \right]
\end{aligned}$$

$$\begin{aligned}
&= \arg\max_{\hat{\theta}} \left[-\lim_{N\to\infty} \sum_{n=1}^{N} [\dot{y}_n - \dot{x}_n(\hat{\theta})]^2 \right] && \text{by Theorem 12.1 (page 81)} \\
&= \arg\max_{\hat{\theta}} \left[-\left\| \mathsf{y}(t) - \mathsf{x}(t;\hat{\theta}) \right\|^2 \right] && \text{by } \textit{Plancheral's formula} \quad \text{(Theorem H.9 page 234)} \\
&= \arg\max_{\hat{\theta}} \left[-\left\| \mathsf{y}(t) \right\|^2 + 2\mathbf{R}_{\mathsf{e}} \left\langle \mathsf{y}(t) \mid \mathsf{x}(t;\hat{\theta}) \right\rangle - \left\| \mathsf{x}(t;\hat{\theta}) \right\|^2 \right] \\
&= \arg\max_{\hat{\theta}} \left[2 \left\langle \mathsf{y}(t) \mid \mathsf{x}(t;\hat{\theta}) \right\rangle - \left\| \mathsf{x}(t;\hat{\theta}) \right\|^2 \right] && \text{because } \mathsf{y}(t) \textit{ independent} \text{ of } \hat{\theta}
\end{aligned}$$

✏

Theorem 12.7 (ML amplitude estimation)**.** [2] *Let* $\mathbf{S}$ *be an additive white gaussian noise system such that*

$$\begin{aligned}
\mathsf{y}(t) &= [\mathbf{C}_{\mathsf{awgn}}\mathsf{x}](t) = \mathsf{x}(t;a) + \mathsf{n}(t) \\
\mathsf{x}(t;a) &\triangleq a\lambda(t).
\end{aligned}$$

[2] Mandyam D. Srinath (1996) pages 158–159

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Then

THM

$$\begin{aligned}
\hat{a}_{\mathsf{ml}} &= \frac{1}{\|\lambda(t)\|^2}\langle \mathsf{y}(t) \mid \lambda(t)\rangle && \textit{(optimal ML-estimate of } a\textit{)}\\
&= \frac{1}{\|\lambda(t)\|^2}\sum_{n=1}^{N}\dot{y}_n\dot{\lambda}_n \\
\mathsf{E}\hat{a}_{\mathsf{ml}} &= a && (\hat{a}_{\mathsf{ml}} \textit{ is } \textbf{\textit{unbiased}})\\
\operatorname{var}\hat{a}_{\mathsf{ml}} &= \frac{\sigma^2}{\|\lambda(t)\|^2} && \textit{(variance of estimate } \hat{a}_{\mathsf{ml}}\textit{)}\\
\operatorname{var}\hat{a}_{\mathsf{ml}} &= \textit{CR lower bound} && (\hat{a}_{\mathsf{ml}} \textit{ is an } \textbf{efficient } \textbf{\textit{estimate}})
\end{aligned}$$

✎PROOF:

1. *ML estimate* in "matched signal" form:

$$\begin{aligned}
\hat{a}_{\mathsf{ml}} &= \arg\max_a \left[2\left\langle \mathsf{y}(t) \mid \mathsf{x}(t;\hat{\theta})\right\rangle - \|\mathsf{x}(t;\phi)\|^2\right] && \text{by Theorem 12.6 (page 86)}\\
&= \arg\max_a \left[2\langle \mathsf{y}(t) \mid a\lambda(t)\rangle - \|a\lambda(t)\|^2\right] && \text{by hypothesis}\\
&= \arg_a\left[\frac{\partial}{\partial a}2a\langle \mathsf{y}(t) \mid \lambda(t)\rangle - \frac{\partial}{\partial a}a^2\|\lambda(t)\|^2 = 0\right]\\
&= \arg_a\left[2\langle \mathsf{y}(t) \mid \lambda(t)\rangle - 2a\|\lambda(t)\|^2 = 0\right]\\
&= \arg_a\left[\langle \mathsf{y}(t) \mid \lambda(t)\rangle = a\|\lambda(t)\|^2\right]\\
&= \frac{1}{\|\lambda(t)\|^2}\langle \mathsf{y}(t) \mid \lambda(t)\rangle
\end{aligned}$$

2. *ML estimate* in "spectral decomposition" form:

$$\begin{aligned}
\hat{a}_{\mathsf{ml}} &= \arg\min_a\left(\sum_{n=1}^{N}\left[\dot{y}_n - \dot{x}_n(a)\right]^2\right) && \text{by Theorem 12.6 (page 86)}\\
&= \arg_a\left(\frac{\partial}{\partial a}\sum_{n=1}^{N}\left[\dot{y}_n - \dot{x}_n(a)\right]^2 = 0\right)\\
&= \arg_a\left(2\sum_{n=1}^{N}\left[\dot{y}_n - \dot{x}_n(a)\right]\frac{\partial}{\partial a}\dot{x}_n(a) = 0\right)\\
&= \arg_a\left(\sum_{n=1}^{N}\left[\dot{y}_n - \langle a\lambda(t) \mid \psi_n(t)\rangle\right]\frac{\partial}{\partial a}\langle a\lambda(t) \mid \psi_n(t)\rangle = 0\right)\\
&= \arg_a\left(\sum_{n=1}^{N}\left[\dot{y}_n - a\langle \lambda(t) \mid \psi_n(t)\rangle\right]\frac{\partial}{\partial a}(a\langle \lambda(t) \mid \psi_n(t)\rangle) = 0\right)\\
&= \arg_a\left(\sum_{n=1}^{N}\left[\dot{y}_n - a\dot{\lambda}_n\right]\langle \lambda(t) \mid \psi_n(t)\rangle = 0\right)\\
&= \arg_a\left(\sum_{n=1}^{N}\left[\dot{y}_n - a\dot{\lambda}_n\right]\dot{\lambda}_n = 0\right)\\
&= \arg_a\left(\sum_{n=1}^{N}\dot{y}_n\dot{\lambda}_n = \sum_{n=1}^{N}a\dot{\lambda}_n^2\right)\\
&= \left(\frac{1}{\sum_{n=1}^{N}\dot{\lambda}_n^2}\right)\sum_{n=1}^{N}\dot{y}_n\dot{\lambda}_n
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \frac{1}{\|\lambda(t)\|^2} \sum_{n=1}^{N} \dot{y}_n \dot{\lambda}_n$$

3. Prove that the estimate $\hat{a}_{\mathsf{ml}}$ is **unbiased**:

$$\begin{aligned}
\mathsf{E}\hat{a}_{\mathsf{ml}} &= \mathsf{E}\frac{1}{\|\lambda(t)\|^2}\int_{t\in\mathbb{R}} \mathsf{y}(t)\lambda(t)\,\mathsf{d}t && \text{by previous result}\\
&= \mathsf{E}\frac{1}{\|\lambda(t)\|^2}\int_{t\in\mathbb{R}} [a\lambda(t)+\mathsf{n}(t)]\lambda(t)\,\mathsf{d}t && \text{by hypothesis}\\
&= \frac{1}{\|\lambda(t)\|^2}\int_{t\in\mathbb{R}} \mathsf{E}[a\lambda(t)+\mathsf{n}(t)]\lambda(t)\,\mathsf{d}t && \text{by linearity of } \int\cdot\,\mathsf{d}t \text{ and } \mathsf{E}\\
&= \frac{1}{\|\lambda(t)\|^2}a\int_{t\in\mathbb{R}} \lambda^2(t)\,\mathsf{d}t && \text{by } \mathsf{E} \text{ operation}\\
&= \frac{1}{\|\lambda(t)\|^2}a\,\|\lambda(t)\|^2 && \text{by definition of } \|\cdot\|^2\\
&= a
\end{aligned}$$

4. Compute the variance of $\hat{a}_{\mathsf{ml}}$:

$$\begin{aligned}
\mathsf{E}\hat{a}_{\mathsf{ml}}^2 &= \mathsf{E}\left[\frac{1}{\|\lambda(t)\|^2}\int_{t\in\mathbb{R}} \mathsf{y}(t)\lambda(t)\,\mathsf{d}t\right]^2\\
&= \mathsf{E}\left[\frac{1}{\|\lambda(t)\|^4}\int_{t\in\mathbb{R}} \mathsf{y}(t)\lambda(t)\,\mathsf{d}t\int_{v} \mathsf{y}(v)\lambda(v)\,\mathsf{d}v\right]\\
&= \mathsf{E}\left[\frac{1}{\|\lambda(t)\|^4}\int_{t\in\mathbb{R}}\int_{v} [a\lambda(t)+\mathsf{n}(t)][a\lambda(v)+\mathsf{n}(v)]\lambda(t)\lambda(v)\,\mathsf{d}v\,\mathsf{d}t\right]\\
&= \mathsf{E}\left[\frac{1}{\|\lambda(t)\|^4}\int_{t\in\mathbb{R}}\int_{v} [a^2\lambda(t)\lambda(v)+a\lambda(t)\mathsf{n}(v)+a\lambda(v)\mathsf{n}(t)+\mathsf{n}(t)\mathsf{n}(v)]\lambda(t)\lambda(v)\,\mathsf{d}v\,\mathsf{d}t\right]\\
&= \left[\frac{1}{\|\lambda(t)\|^4}\int_{t\in\mathbb{R}}\int_{v} [a^2\lambda(t)\lambda(v)+0+0+\sigma^2\delta(t-v)]\lambda(t)\lambda(v)\,\mathsf{d}v\,\mathsf{d}t\right]\\
&= \frac{1}{\|\lambda(t)\|^4}\int_{t\in\mathbb{R}}\int_{v} a^2\lambda^2(t)\lambda^2(v)\,\mathsf{d}v\,\mathsf{d}t + \frac{1}{\|\lambda(t)\|^4}\int_{t\in\mathbb{R}}\int_{v} \sigma^2\delta(t-v)\lambda(t)\lambda(v)\,\mathsf{d}v\,\mathsf{d}t\\
&= \frac{1}{\|\lambda(t)\|^4}a^2\int_{t\in\mathbb{R}} \lambda^2(t)\,\mathsf{d}t\int_{v} \lambda^2(v)\,\mathsf{d}v + \frac{1}{\|\lambda(t)\|^4}\sigma^2\int_{t\in\mathbb{R}} \lambda^2(t)\,\mathsf{d}t\\
&= a^2\frac{1}{\|\lambda(t)\|^4}\,\|\lambda(t)\|^2\,\|\lambda(v)\|^2 + \frac{1}{\|\lambda(t)\|^4}\sigma^2\,\|\lambda(t)\|^2\\
&= a^2 + \frac{\sigma^2}{\|\lambda(t)\|^2}
\end{aligned}$$

$$\begin{aligned}
\operatorname{var}\hat{a}_{\mathsf{ml}} &= \mathsf{E}\hat{a}_{\mathsf{ml}}^2 - (\mathsf{E}\hat{a}_{\mathsf{ml}})^2\\
&= \left(a^2 + \frac{\sigma^2}{\|\lambda(t)\|^2}\right) - \left(a^2\right)\\
&= \frac{\sigma^2}{\|\lambda(t)\|^2}
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

5. Compute the Cramér-Rao Bound:

$$\begin{aligned}
\mathsf{p}\left[\mathsf{y(t)|x(t;a)}\right] &= \mathsf{p}\left[\dot{\mathsf{y}}_1, \dot{\mathsf{y}}_2, \ldots, \dot{\mathsf{y}}_N|\mathsf{x(t;a)}\right] \\
&= \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(\dot{y}_n - a\dot{\lambda}_n)^2}{-2\sigma^2} \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \exp \frac{1}{-2\sigma^2} \sum_{n=1}^{N} (\dot{y}_n - a\dot{\lambda}_n)^2
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial a} \ln \mathsf{p}\left[\mathsf{y(t)|x(t;a)}\right] &= \frac{\partial}{\partial a} \ln \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \exp \frac{1}{-2\sigma^2} \sum_{n=1}^{N} (\dot{y}_n - a\dot{\lambda}_n)^2 \\
&= \frac{\partial}{\partial a} \ln \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N + \frac{\partial}{\partial a} \ln \exp \frac{1}{-2\sigma^2} \sum_{n=1}^{N} (\dot{y}_n - a\dot{\lambda}_n)^2 \\
&= \frac{\partial}{\partial a} \left[\frac{1}{-2\sigma^2} \sum_{n=1}^{N} (\dot{y}_n - a\dot{\lambda}_n)^2\right] \\
&= \frac{1}{-2\sigma^2} \sum_{n=1}^{N} 2(\dot{y}_n - a\dot{\lambda}_n)(-\dot{\lambda}_n) \\
&= \frac{1}{\sigma^2} \sum_{n=1}^{N} \dot{\lambda}_n(\dot{y}_n - a\dot{\lambda}_n)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial a^2} \ln \mathsf{p}\left[\mathsf{y(t)|x(t;a)}\right] &= \frac{\partial}{\partial a}\frac{\partial}{\partial a} \ln \mathsf{p}\left[\mathsf{y(t)|x(t;a)}\right] \\
&= \frac{\partial}{\partial a}\frac{1}{\sigma^2} \sum_{n=1}^{N} \dot{\lambda}_n(\dot{y}_n - a\dot{\lambda}_n) \\
&= \frac{1}{\sigma^2} \sum_{n=1}^{N} \dot{\lambda}_n(-\dot{\lambda}_n) \\
&= \frac{-1}{\sigma^2} \sum_{n=1}^{N} \dot{\lambda}_n^2 \\
&= \frac{-\|\lambda(t)\|^2}{\sigma^2}
\end{aligned}$$

$$\begin{aligned}
\operatorname{var} \hat{a}_{\mathsf{ml}} &\triangleq \mathsf{E}\left[\hat{a}_{\mathsf{ml}} - \mathsf{E}\hat{a}_{\mathsf{ml}}\right]^2 \\
&= \mathsf{E}\left[\hat{a}_{\mathsf{ml}} - a\right]^2 \\
&\geq \frac{-1}{\mathsf{E}\left(\frac{\partial^2}{\partial a^2} \ln \mathsf{p}\left[\mathsf{y(t)|x(t;a)}\right]\right)} \\
&= \frac{-1}{\mathsf{E}\left(\frac{-\|\lambda(t)\|^2}{\sigma^2}\right)} \\
&= \frac{\sigma^2}{\|\lambda(t)\|^2} \qquad \text{(Cramér-Rao lower bound of the variance)}
\end{aligned}$$

6. Proof that $\hat{a}_{\mathsf{ml}}$ is an *efficient* estimate:

   An estimate is *efficient* if $\operatorname{var} \hat{a}_{\mathsf{ml}} =$ CR lower bound. We have already proven this, so $\hat{a}_{\mathsf{ml}}$ is an *efficient*

Copyright © 2019 Daniel J. Greenhoe

estimate.

Also, even without explicitly computing the variance of $\hat{a}_{\mathsf{ml}}$, the variance equals the *Cramér-Rao lower bound* (and hence $\hat{a}_{\mathsf{ml}}$ is an *efficient* estimate) if and only if

$$\hat{a}_{\mathsf{ml}} - a = \left(\frac{-1}{\mathsf{E}\left[\frac{\partial^2}{\partial a^2}\ln \mathsf{p}\left[\mathsf{y(t)|x(t;a)}\right]\right]}\right)\left(\frac{\partial}{\partial a}\ln \mathsf{p}\left[\mathsf{y(t)|x(t;a)}\right]\right)$$

$$\begin{aligned}\left(\frac{-1}{\mathsf{E}\left(\frac{\partial^2}{\partial a^2}\ln \mathsf{p}\left[\mathsf{y(t)|x(t;a)}\right]\right)}\right)\left(\frac{\partial}{\partial a}\ln \mathsf{p}\left[\mathsf{y(t)|x(t;a)}\right]\right) &= \left(\frac{\sigma^2}{\|\lambda(t)\|^2}\right)\left(\frac{1}{\sigma^2}\sum_{n=1}^{N}\dot{\lambda}(\dot{y}-a\dot{\lambda})\right)\\ &= \frac{1}{\|\lambda(t)\|^2}\sum_{n=1}^{N}\dot{\lambda}\dot{y} - \frac{1}{\|\lambda(t)\|^2}\sum_{n=1}^{N}\dot{\lambda}^2\\ &= \hat{a}_{\mathsf{ml}} - a\end{aligned}$$

✏

Theorem 12.8 (ML phase estimation). [3] *Let* **S** *be an additive white gaussian noise system such that*

$$\begin{aligned}\mathsf{y}(t) &= [\mathbf{C}_{\mathsf{awgn}}\mathsf{x}](t) = \mathsf{x}(t;\phi) + \mathsf{n}(t)\\ \mathsf{x}(t;\phi) &= A\cos(2\pi f_c t + \phi).\end{aligned}$$

Then the optimal ML-estimate of parameter ϕ is

THM

$$\hat{\phi}_{\mathsf{ml}} = -\operatorname{atan}\left(\frac{\langle \mathsf{y}(t) \mid \sin(2\pi f_c t)\rangle}{\langle \mathsf{y}(t) \mid \cos(2\pi f_c t)\rangle}\right)$$

✎PROOF:

$$\begin{aligned}\hat{\phi}_{\mathsf{ml}} &= \arg\max_{\phi}\left[2\left\langle \mathsf{y}(t) \mid \mathsf{x}(t;\hat{\theta})\right\rangle - \|\mathsf{x}(t;\phi)\|^2\right] && \text{by Theorem 12.6 (page 86)}\\ &= \arg\max_{\phi}\left[2\left\langle \mathsf{y}(t) \mid \mathsf{x}(t;\phi)\right\rangle\right] && \text{because } \|\mathsf{x}(t;\phi)\| \text{ does not depend on } \phi\\ &= \arg_{\phi}\left[\frac{\partial}{\partial\phi}\left\langle \mathsf{y}(t) \mid \mathsf{x}(t;\phi)\right\rangle = 0\right]\\ &= \arg_{\phi}\left[\left\langle \mathsf{y}(t) \mid \frac{\partial}{\partial\phi}\mathsf{x}(t;\phi)\right\rangle = 0\right] && \text{because } \langle\cdot\mid\cdot\rangle \text{ is a linear operator}\\ &= \arg_{\phi}\left[\left\langle \mathsf{y}(t) \mid \frac{\partial}{\partial\phi}A\cos(2\pi f_c t + \phi)\right\rangle = 0\right]\\ &= \arg_{\phi}\left[\left\langle \mathsf{y}(t) \mid -A\sin(2\pi f_c t + \phi)\right\rangle = 0\right]\\ &= \arg_{\phi}\left[-A\left\langle \mathsf{y}(t) \mid \cos(2\pi f_c t)\sin\phi + \sin(2\pi f_c t)\cos\phi\right\rangle = 0\right]\\ &= \arg_{\phi}\left[\sin\phi\left\langle \mathsf{y}(t) \mid \cos(2\pi f_c t)\right\rangle = -\cos\phi\left\langle \mathsf{y}(t) \mid \sin(2\pi f_c t)\right\rangle\right]\\ &= \arg_{\phi}\left[\frac{\sin\phi}{\cos\phi} = -\frac{\langle \mathsf{y}(t) \mid \sin(2\pi f_c t)\rangle}{\langle \mathsf{y}(t) \mid \cos(2\pi f_c t)\rangle}\right]\\ &= \arg_{\phi}\left[\tan\phi = -\frac{\langle \mathsf{y}(t) \mid \sin(2\pi f_c t)\rangle}{\langle \mathsf{y}(t) \mid \cos(2\pi f_c t)\rangle}\right]\\ &= -\operatorname{atan}\left(\frac{\langle \mathsf{y}(t) \mid \sin(2\pi f_c t)\rangle}{\langle \mathsf{y}(t) \mid \cos(2\pi f_c t)\rangle}\right)\end{aligned}$$

✏

[3] Mandyam D. Srinath (1996) pages 159–160

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Theorem 12.9 (ML estimation of a function of a parameter). [4] *Let* $\mathbf{S}$ *be an additive white gaussian noise system such that*

$$\begin{aligned} \mathsf{y}(t) &= [\mathbf{C}_{\mathsf{awgn}}\mathsf{x}](t) = \mathsf{x}(t;\hat{\theta}) + \mathsf{n}(t) \\ \mathsf{x}(t;\hat{\theta}) &= \mathsf{g}(\hat{\theta}) \end{aligned}$$

where g *is one-to-one and onto (invertible).*

THM

Then the optimal ML-estimate of parameter u *is*
$$\hat{\theta}_{\mathsf{ml}} = \mathsf{g}^{-1}\left(\frac{1}{N}\sum_{n=1}^{N}\dot{y}_n\right).$$

If an ML ESTIMATE $\hat{\theta}_{\mathsf{ml}}$ *is unbiased* ($\mathsf{E}\hat{\theta}_{\mathsf{ml}} = \theta$) *then*
$$\operatorname{var}\hat{\theta}_{\mathsf{ml}} \geq \frac{\sigma^2}{N}\frac{1}{\left[\frac{\partial \mathsf{g}(\theta)}{\partial\theta}\right]^2}.$$

If $\mathsf{g}(\theta) = \theta$ *then* $\hat{\theta}_{\mathsf{ml}}$ *is an* ***efficient*** *estimate such that*
$$\operatorname{var}\hat{\theta}_{\mathsf{ml}} = \frac{\sigma^2}{N}.$$

✎PROOF:

$$\begin{aligned}
\hat{\theta}_{\mathsf{ml}} &= \arg\min_{\hat{\theta}}\left[\sum_{n=1}^{N}[\dot{y}_n - \mathsf{g}(\hat{\theta})]^2\right] \\
&= \arg_{\hat{\theta}}\left[\frac{\partial}{\partial\hat{\theta}}\sum_{n=1}^{N}[\dot{y}_n - \mathsf{g}(\hat{\theta})]^2 = 0\right] \\
&= \arg_{\hat{\theta}}\left[2\sum_{n=1}^{N}[\dot{y}_n - \mathsf{g}(\hat{\theta})]\frac{\partial}{\partial\hat{\theta}}\mathsf{g}(\hat{\theta}) = 0\right] \\
&= \arg_{\hat{\theta}}\left[2\sum_{n=1}^{N}[\dot{y}_n - \mathsf{g}(\hat{\theta})] = 0\right] \\
&= \arg_{\hat{\theta}}\left[\sum_{n=1}^{N}\dot{y}_n = N\mathsf{g}(\hat{\theta})\right] \\
&= \arg_{\hat{\theta}}\left[\mathsf{g}(\hat{\theta}) = \frac{1}{N}\sum_{n=1}^{N}\dot{y}_n\right] \\
&= \arg_{\hat{\theta}}\left[u = \mathsf{g}^{-1}\left(\frac{1}{N}\sum_{n=1}^{N}\dot{y}_n\right)\right] \\
&= \mathsf{g}^{-1}\left(\frac{1}{N}\sum_{n=1}^{N}\dot{y}_n\right)
\end{aligned}$$

If $\hat{\theta}_{\mathsf{ml}}$ is unbiased ($\mathsf{E}\hat{\theta}_{\mathsf{ml}} = \theta$), we can use the *Cramér-Rao bound* to find a lower bound on the variance:

$$\begin{aligned}
\operatorname{var}\hat{\theta}_{\mathsf{ml}} &\triangleq \mathsf{E}\left[\hat{\theta}_{\mathsf{ml}} - \mathsf{E}\hat{\theta}_{\mathsf{ml}}\right]^2 \\
&= \mathsf{E}\left[\hat{\theta}_{\mathsf{ml}} - \theta\right]^2 \\
&\geq \frac{-1}{\mathsf{E}\left(\frac{\partial^2}{\partial\theta^2}\ln \mathsf{p}\left[\mathsf{y}(t)|\mathsf{x}(t;\theta)\right]\right)} \\
&= \frac{-1}{\mathsf{E}\left(\frac{\partial^2}{\partial\theta^2}\ln \mathsf{p}\left[\dot{\mathsf{y}}_1, \dot{\mathsf{y}}_2, \ldots, \dot{\mathsf{y}}_N|\mathsf{x}(t;\theta)\right]\right)}
\end{aligned}$$

[4] Mandyam D. Srinath (1996) pages 142–143

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$
\begin{aligned}
&= \frac{-1}{\mathsf{E}\left(\frac{\partial^2}{\partial\theta^2}\ln\left[\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \exp\frac{-1}{2\sigma^2}\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]^2\right]\right)}\\
&= \frac{-1}{\mathsf{E}\left(\frac{\partial^2}{\partial\theta^2}\ln\left[\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N\right]+\frac{\partial^2}{\partial\theta^2}\ln\left[\exp\frac{-1}{2\sigma^2}\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]^2\right]\right)}\\
&= \frac{-1}{\mathsf{E}\left(\frac{\partial^2}{\partial\theta^2}\left(\frac{-1}{2\sigma^2}\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]^2\right)\right)}\\
&= \frac{2\sigma^2}{\mathsf{E}\left(\frac{\partial}{\partial\theta}\frac{\partial}{\partial\theta}\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]^2\right)}\\
&= \frac{2\sigma^2}{\mathsf{E}\left(-2\frac{\partial}{\partial\theta}\frac{\partial\mathsf{g}(\theta)}{\partial\theta}\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]\right)}\\
&= \frac{-\sigma^2}{\mathsf{E}\left(\frac{\partial\mathsf{g}^2(\theta)}{\partial\theta^2}\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]+\frac{\partial\mathsf{g}(\theta)}{\partial\theta}\frac{\partial}{\partial\theta}\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]\right)}\\
&= \frac{-\sigma^2}{\mathsf{E}\left(\frac{\partial\mathsf{g}^2(\theta)}{\partial\theta^2}\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]-N\frac{\partial\mathsf{g}(\theta)}{\partial\theta}\frac{\partial\mathsf{g}(\theta)}{\partial\theta}\right)}\\
&= \frac{-\sigma^2}{\frac{\partial\mathsf{g}^2(\theta)}{\partial\theta^2}\sum_{n=1}^N\mathsf{E}[\dot{y}_n-\mathsf{g}(\theta)]-N\frac{\partial\mathsf{g}(\theta)}{\partial\theta}\frac{\partial\mathsf{g}(\theta)}{\partial\theta}}\\
&= \frac{-\sigma^2}{-N\frac{\partial\mathsf{g}(\theta)}{\partial\theta}\frac{\partial\mathsf{g}(\theta)}{\partial\theta}}\\
&= \frac{\sigma^2}{N}\frac{1}{\left[\frac{\partial\mathsf{g}(\theta)}{\partial\theta}\right]^2}
\end{aligned}
$$

The inequality becomes equality (an *efficient* estimate) if and only if

$$
\hat{\theta}_{\mathsf{ml}}-\theta=\left(\frac{-1}{\mathsf{E}\left(\frac{\partial^2}{\partial\theta^2}\ln\mathsf{p}\left[\mathsf{y}(\mathsf{t})|\mathsf{x}(\mathsf{t};\theta)\right]\right)}\right)\left(\frac{\partial}{\partial\theta}\ln\mathsf{p}\left[\mathsf{y}(\mathsf{t})|\mathsf{x}(\mathsf{t};\theta)\right]\right).
$$

$$
\begin{aligned}
\left(\frac{-1}{\mathsf{E}\left(\frac{\partial^2}{\partial\theta^2}\ln\mathsf{p}\left[\mathsf{y}(\mathsf{t})|\mathsf{x}(\mathsf{t};\theta)\right]\right)}\right)\left(\frac{\partial}{\partial\theta}\ln\mathsf{p}\left[\mathsf{y}(\mathsf{t})|\mathsf{x}(\mathsf{t};\theta)\right]\right) &= \left(\frac{\sigma^2}{N}\frac{1}{\left[\frac{\partial\mathsf{g}(\theta)}{\partial\theta}\right]^2}\right)\left(\frac{-1}{2\sigma^2}(2)\frac{\partial\mathsf{g}(\theta)}{\partial\theta}\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]\right)\\
&= -\frac{1}{N}\frac{1}{\frac{\partial\mathsf{g}(\theta)}{\partial\theta}}\left(\sum_{n=1}^N[\dot{y}_n-\mathsf{g}(\theta)]\right)\\
&= -\frac{1}{\frac{\partial\mathsf{g}(\theta)}{\partial\theta}}\left(\frac{1}{N}\sum_{n=1}^N\dot{y}_n-\mathsf{g}(\theta)\right)\\
&= -\frac{1}{\frac{\partial\mathsf{g}(\theta)}{\partial\theta}}\left(\hat{\theta}_{\mathsf{ml}}-\mathsf{g}(\theta)\right)\\
&= -(\hat{\theta}_{\mathsf{ml}}-\theta)
\end{aligned}
$$

Copyright © 2019 Daniel J. Greenhoe

12.3 Colored noise

This chapter presented several theorems whose results depended on the noise being white. However if the noise is **colored**, then these results are invalid. But there is still hope for colored noise. Processing colored signals can be accomplished using two techniques:

1. Karhunen-Loève basis functions [5]
2. whitening filter [6]

Karhunen-Loève. If the noise is white, the set $\{\langle \mathsf{y}(t) \mid \psi_n(t)\rangle\}$ is a sufficient statistic regardless of which set $\{\psi_n(t)\}$ of orthonormal basis functions are used. If the noise is colored, and if $\{\psi_n(t)\}$ satisfy the Karhunen-Loève criterion

$$\int_{t_2} \mathsf{R}_{\mathsf{xx}}(t_1, t_2)\psi_n(t_2)\, \mathrm{d}t_2 = \lambda_n \psi_n(t_1)$$

then $\{\langle \mathsf{y}(t) \mid \psi_n(t)\rangle\}$ is still a sufficient statistic.

Whitening filter. The whitening filter makes the received signal $\mathsf{y}(t)$ statistically white (uncorrelated in time). In this case, any orthonormal basis set can be used to generate sufficient statistics.

12.4 Signal matching

Detection methods. There are basically two types of detection methods:

1. signal matching
2. orthonormal decomposition.

Let S be the set of transmitted waveforms and Y be a set of orthonormal basis functions that span S. *Signal matching* computes the innerproducts of a received signal $\mathsf{y}(t)$ with each signal from S. *Orthonormal decomposition* computes the innerproducts of $\mathsf{y}(t)$ with each signal from the set Y.

In the case where $|S|$ is large, often $|Y| << |S|$ making orthonormal decomposition much easier to implement. For example, in a QAM-64 modulation system, signal matching requires $|S| = 64$ innerproduct calculations, while orthonormal decomposition only requires $|Y| = 2$ innerproduct calculations because all 64 signals in S can be spanned by just 2 orthonormal basis functions.

Maximizing SNR. Theorem 12.1 (page 81) shows that the innerproducts of $\mathsf{y}(t)$ with basis functions of Y is *sufficient* for optimal detection. Theorem 12.10 (page 94) (next) shows that a receiver can maximize the SNR of a received signal when signal matching is used.

Theorem 12.10. *Let* $\mathsf{x}(t)$ *be a transmitted signal,* $\mathsf{n}(t)$ *noise, and* $\mathsf{y}(t)$ *the received signal in an AWGN channel. Let the* SIGNAL TO NOISE RATIO *SNR be defined as*

[5] **Karhunen-Loève**: Section 4.1 (page 19)
[6] *Continuous data whitening*: Section 7.3 page 50
Discrete data whitening: Section 6.4 page 40

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$\mathrm{SNR}[\mathsf{y}(t)] \triangleq \frac{|\langle \mathsf{x}(t) \mid \mathsf{x}(t)\rangle|^2}{\mathsf{E}\left[|\langle \mathsf{n}(t) \mid \mathsf{x}(t)\rangle|^2\right]}.$$

THM $\mathrm{SNR}[\mathsf{y}(t)] \le \dfrac{2\,\|\mathsf{x}(t)\|^2}{N_o}$ *and is maximized (equality) when* $\mathsf{x}(t) = a\mathsf{x}(t)$, *where* $a \in \mathbb{R}$.

✎PROOF:

$$\begin{aligned}
\mathrm{SNR}[\mathsf{y}(t)] &\triangleq \frac{|\langle \mathsf{x}(t) \mid \mathsf{x}(t)\rangle|^2}{\mathsf{E}\left[|\langle \mathsf{n}(t) \mid \mathsf{x}(t)\rangle|^2\right]} \\
&= \frac{|\langle \mathsf{x}(t) \mid f(t)\rangle|^2}{\mathsf{E}\left[\left[\int_{t\in\mathbb{R}} \mathsf{n}(t)\mathsf{x}^*(t)\,dt\right]\left[\int_{\hat\theta} n(\hat\theta)f^*(\hat\theta)\,du\right]^*\right]} \\
&= \frac{|\langle \mathsf{x}(t) \mid \mathsf{x}(t)\rangle|^2}{\mathsf{E}\left[\int_{t\in\mathbb{R}}\int_{\hat\theta} \mathsf{n}(t)n^*(\hat\theta)\mathsf{x}^*(t)\mathsf{x}(\hat\theta)\,dtdu\right]} \\
&= \frac{|\langle \mathsf{x}(t) \mid f(t)\rangle|^2}{\int_{t\in\mathbb{R}}\int_{\hat\theta} \mathsf{E}\left[\mathsf{n}(t)n^*(\hat\theta)\right]\mathsf{x}^*(t)\mathsf{x}(\hat\theta)\,dtdu} \\
&= \frac{|\langle \mathsf{x}(t) \mid \mathsf{x}(t)\rangle|^2}{\int_{t\in\mathbb{R}}\int_{\hat\theta} \frac{1}{2}N_o\delta(t-\hat\theta)\mathsf{x}^*(t)\mathsf{x}(\hat\theta)\,dtdu} \\
&= \frac{|\langle \mathsf{x}(t) \mid \mathsf{x}(t)\rangle|^2}{\frac{1}{2}N_o\int_{t\in\mathbb{R}} \mathsf{x}^*(t)\mathsf{x}(t)\,\mathrm{d}t} \\
&= \frac{|\langle \mathsf{x}(t) \mid \mathsf{x}(t)\rangle|^2}{\frac{1}{2}N_o\,\|\mathsf{x}(t)\|^2} \\
&\le \frac{|\|\mathsf{x}(t)\|\;\|\mathsf{x}(t)\||^2}{\frac{1}{2}N_o\,\|\mathsf{x}(t)\|^2} && \text{by } \textit{Cauchy-Schwarz Inequality} \quad \text{(Theorem K.2 page 272)} \\
&= \frac{2\,\|\mathsf{x}(t)\|^2}{N_o}
\end{aligned}$$

The Cauchy-Schwarz Inequality becomes an equality (SNR is maximized) when $\mathsf{x}(t) = a\mathsf{x}(t)$. ✏

Implementation. The innerproduct operations can be implemented using either

1. a correlator or
2. a matched filter.

A correlator is simply an integrator of the form $\langle \mathsf{y}(t) \mid f(t)\rangle = \displaystyle\int_0^T \mathsf{y}(t)f(t)\,\mathrm{d}t$.

A matched filter introduces a function $\mathsf{h}(t)$ such that $\mathsf{h}(t) = \mathsf{x}(T-t)$ (which implies $\mathsf{x}(t) = h(T-t)$) giving

$$\underbrace{\langle \mathsf{y}(t) \mid \mathsf{x}(t)\rangle = \int_0^T \mathsf{y}(t)\mathsf{x}(t)\,\mathrm{d}t}_{\text{correlator}} = \underbrace{\int_0^\infty \mathsf{x}(\tau)h(t-\tau)\,\mathrm{d}\tau\Big|_{t=T} = \mathsf{x}(t)\star\mathsf{h}(t)|_{t=T}}_{\text{matched filter}}.$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

This shows that $\mathsf{h}(t)$ is the impulse response of a filter operation sampled at time T. By Theorem 12.10 (page 94), the optimal impulse response is $\mathsf{h}(T-t) = f(t) = \mathsf{x}(t)$. That is, the optimal $\mathsf{h}(t)$ is just a "flipped" and shifted version of $\mathsf{x}(t)$.

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 13

MOMENT ESTIMATION

13.1 Mean Estimation

Theorem 13.1. *Let* $\hat{\mu} \triangleq \sum_{n=1}^{N} \lambda_n \mathsf{x}_n$ *with* $\sum_{n=1}^{N} \lambda_n = 1$ *be the* ARITHMETIC MEAN *(Definition N.4 page 311)*.

THM

$$\left\{\begin{array}{lll} (A). & (\mathsf{x}_n) \text{ is WIDE SENSE STATIONARY} & \text{and} \\ (B). & \mu \triangleq \mathsf{E}\mathsf{x}_n & \text{and} \\ (C). & (\mathsf{x}_n) \text{ is UNCORRELATED} & \text{and} \\ (D). & \hat{\mu} \triangleq \sum_{n=1}^{N} \lambda_n \mathsf{x}_n \quad \text{(ARITHMETIC MEAN)} & \end{array}\right\} \Longrightarrow \left\{\begin{array}{lll} (1). & \mathsf{E}\hat{\mu} = \mu \quad \text{(UNBIASED)} & \text{and} \\ (2). & \mathrm{var}(\hat{\mu}) = \sigma^2 \sum_{n=1}^{N} \lambda_n^2 & \text{and} \\ (3). & \mathrm{mse}(\hat{\mu}) = \sigma^2 \sum_{n=1}^{N} \lambda_n^2 & \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{E}\hat{\mu} &\triangleq \mathsf{E}\sum_{n\in\mathbb{Z}} \lambda_n \mathsf{x}_n && \text{by definition of } \textit{arithmetic mean} && \text{(Definition N.4 page 311)} \\
&= \sum_{n\in\mathbb{Z}} \lambda_n \mathsf{E}\mathsf{x}_n && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)} \\
&= \mu \sum_{n\in\mathbb{Z}} \lambda_n && \text{by } \textit{WSS} \text{ hypothesis} && \text{(A)} \\
&= \mu && \text{by } \textstyle\sum \lambda_n = 1 \text{ hypothesis} && \text{(Definition N.4 page 311)} \\
\mathrm{var}(\hat{\mu}) &\triangleq \mathsf{E}(\hat{\mu} - \mathsf{E}\hat{\mu})^2 && \text{by definition of } \textit{variance} \\
&= \mathsf{E}(\hat{\mu} - \mu)^2 && \text{by previous result} \\
&= \mathsf{E}\left(\sum_{n=1}^{N} \lambda_n \mathsf{x}_n - \mu\right)^2 && \text{by definition of } \hat{\mu} \\
&= \mathsf{E}\left[\sum_{n=1}^{N} \lambda_n \mathsf{x}_n - \mu \underbrace{\sum_{n=1}^{N} \lambda_n}_{1}\right]^2 && \text{by } \textstyle\sum \lambda_n = 1 \text{ hypothesis} && \text{(Definition N.4 page 311)}
\end{aligned}$$

$$= \mathsf{E}\left[\sum_{n=1}^{N} \lambda_n(\mathsf{x}_n - \mu)\right]^2$$

$$= \mathsf{E}\left[\sum_{n=1}^{N} \lambda_n(\mathsf{x}_n - \mu) \sum_{m=1}^{N} \lambda_m(\mathsf{x}_m - \mu)\right]$$

$$= \sum_{n=1}^{N}\sum_{m=1}^{N} \lambda_n\lambda_m\big(\mathsf{E}\big[(\mathsf{x}_n - \mu)(\mathsf{x}_m - \mu)\big]\big)$$

$$= \sum_{n=1}^{N}\sum_{m=1}^{N} \lambda_n\lambda_m\big(\mathsf{E}\big[\mathsf{x}_n\mathsf{x}_m\big] - \mu\mathsf{E}\big[\mathsf{x}_n\big] - \mu\mathsf{E}\big[\mathsf{x}_m\big] + \mu^2\big)$$

$$= \sum_{n=1}^{N}\sum_{m=1}^{N} \lambda_n\lambda_m\big(\mathsf{E}\big[\mathsf{x}_n\mathsf{x}_m\big] - \mu^2 - \mu^2 + \mu^2\big)$$

$$= \sum_{n=1}^{N}\sum_{m=1}^{N} \lambda_n\lambda_m\big(\mathsf{E}\big[\mathsf{x}_n\mathsf{x}_m\big] - \mu^2\big)$$

$$= \sum_{n=1}^{N} \lambda_n^2\big(\mathsf{E}\big[\mathsf{x}_n^2\big] - \mu^2\big) + \sum_{n=1}^{N}\sum_{m\neq n} \lambda_n\lambda_m\big(\mathsf{E}\big[\mathsf{x}_n\mathsf{x}_m\big] - \mu^2\big)$$

$$= \sum_{n=1}^{N} \lambda_n^2\big(\mathsf{E}\big[\mathsf{x}_n^2\big] - \mu^2\big) + \sum_{n=1}^{N}\sum_{m\neq n} \lambda_n\lambda_m\big(\mathsf{E}\mathsf{x}_n\mathsf{E}\mathsf{x}_m - \mu^2\big)$$

$$= \sum_{n=1}^{N} \lambda_n^2\sigma^2 + \cancelto{0}{\sum_{n=1}^{N}\sum_{m\neq n} \lambda_n\lambda_m\big(\mu\mu - \mu^2\big)} \qquad \text{by } \textit{WSS} \text{ hypothesis} \qquad \text{(A)}$$

$$= \sigma^2 \sum_{n=1}^{N} \lambda_n^2$$

$$\mathsf{mse}(\hat{\mu}) = \mathsf{E}(\hat{\mu} - \mathsf{E}\hat{\mu})^2 + (\mathsf{E}\hat{\mu} - \mu)^2 \qquad \text{by Theorem 9.2 page 67}$$

$$= \sigma^2 \sum_{n=1}^{N} \lambda_n^2 + (\mu - \mu)^2 \qquad \text{by previous results}$$

$$= \sigma^2 \sum_{n=1}^{N} \lambda_n^2$$

✏

Definition 13.1.

DEF *The **average** $\hat{\mu}$ of a length N sequence $(\!|\mathsf{x}_n|\!)_1^N$ is defined as* $\hat{\mu} \triangleq \frac{1}{N}\sum_{n=1}^{N} \mathsf{x}_n$

Corollary 13.1. [1]

COR

$$\left\{\begin{array}{lll} (A). & (\!|\mathsf{x}_n|\!) \textit{ is } \text{WIDE SENSE STATIONARY} & \textit{and} \\ (B). & \mu \triangleq \mathsf{E}\mathsf{x}_n & \textit{and} \\ (C). & (\!|\mathsf{x}_n|\!) \textit{ is } \text{UNCORRELATED} & \textit{and} \\ (D). & \hat{\mu} \triangleq \frac{1}{N}\sum_{n=1}^{N} \mathsf{x}_n \quad (\text{AVERAGE}) & \end{array}\right\} \Longrightarrow \left\{\begin{array}{llll} (1). & \mathsf{E}\hat{\mu} = \mu & (\text{UNBIASED}) & \textit{and} \\ (2). & \mathsf{var}(\hat{\mu}) = \frac{\sigma^2}{N} & & \textit{and} \\ (3). & \mathsf{mse}(\hat{\mu}) = \frac{\sigma^2}{N} & (\text{CONSISTENT}) & \end{array}\right\}$$

✎PROOF: These results follow from Theorem 13.1 (page 97) with $\lambda_n = \frac{1}{N}$. ✏

[1] Kay (1988) page 45 ⟨§"3.3 ESTIMATION THEORY"⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

The *arithmetic mean* estimator $\hat{\mu} \triangleq \sum \lambda_n \mathsf{x}_n$ is *unbiased* and *consistent* for any $\sum \lambda_n = 1$ and yields *mean square error* $\mathsf{mse}(\hat{\mu}) = \sigma^2 \sum \lambda_n^2$ (Theorem 13.1 page 97). But…

1. Said qualitatively: "What is the 'best' sequence $(\!(\lambda_n)\!)$ to use?"
2. Said quantitatively: "What sequence $(\!(\lambda_n)\!)$ yields the smallest $\mathsf{mse}(\hat{\mu})$?"

For example, would fashioning $(\!(\lambda_n)\!)$ to be a scaled version of a standard window function, like the *Hanning window*[2] illustrated below, yield the best $\mathsf{mse}(\hat{\mu})$?

Proposition 13.1 (page 100) answers question (2) stating that the best sequence in terms of minimal mse is the sequence $(\!(\lambda_n)\!) \triangleq \frac{1}{N}(\!(\ldots,1,1,1,\ldots)\!)$, which is the *average* estimator, which yields $\mathsf{mse}(\hat{\mu}) = \frac{\sigma^2}{N}$ (Corollary 13.1 page 98).

That is, it turns out that $\frac{1}{N} \leq \sum \lambda_n^2$ for all possible sequences $(\!(\lambda_n)\!)$. This fact is demonstrated by Lemma 13.1 (next), which in turn follows more or less directly from the ubiquitous *Cauchy-Schwarz Inequality* (Theorem N.6 page 314, Theorem K.2 page 272).

Lemma 13.1.

LEM
$$\left\{\sum_{n=1}^{N} \lambda_n = 1\right\} \quad \Longrightarrow \quad \left\{\frac{1}{N} \leq \sum_{n=1}^{N} \lambda_n^2\right\}$$

✎PROOF:

1. Let the sequence $(\!(a_n)\!)$ be defined as $(\!(a_n)\!) \triangleq (\!(\ldots,1,1,1,\ldots)\!)$
2. Let *inner product* $\langle a_n \mid b_n \rangle$ be defined as $\langle a_n \mid b_n \rangle \triangleq \sum_{n=1}^{N} a_n b_n$
3. Let *norm* $\|a_n\|$ be defined as $\|a_n\| \triangleq \sum_{n=1}^{N} a_n^2$
4. Proof of lemma:

$$\begin{aligned}
\boxed{\frac{1}{N}} &= \frac{1}{N}\left(\sum_{n=1}^{N} \lambda_n\right)^2 && \text{by } \sum_{n=1}^{N} \lambda_n = 1 \text{ hypothesis} \\
&= \frac{1}{N}\left(\sum_{n=1}^{N} a_n \lambda_n\right)^2 && \text{by } (\!(a_n)\!) \triangleq (\!(\ldots,1,1,1,\ldots)\!) \text{ definition} && \text{(definition 1 page 99)} \\
&\boxed{\leq} \frac{1}{N}\left(\sum_{n=1}^{N} a_n^2\right)\left(\sum_{n=1}^{N} \lambda_n^2\right) && \text{by } \textit{Cauchy-Schwartz inequality} && \text{(Theorem N.6 page 314)} \\
&\triangleq \frac{1}{N}\left(\sum_{n=1}^{N} 1^2\right)\left(\sum_{n=1}^{N} \lambda_n^2\right) && \text{by definition of } (\!(a_n)\!) && \text{(definition 1 page 99)} \\
&= \boxed{\sum_{n=1}^{N} \lambda_n^2}
\end{aligned}$$

[2] Abdaheer (2009), page 130

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✏

Proposition 13.1. *Let* mse(*average mean*) *be the mean square error of the* AVERAGE *estimator* *(Corollary 13.1 page 98)* *and* mse(*arithmetic mean*) *be the mean square error of the* ARITHMETIC *estimator* *(Theorem 13.1 page 97)*.

PRP

$$\text{mse}(\textit{average mean}) \le \text{mse}(\textit{arithmetic mean})$$

✎PROOF:

$$\begin{aligned}
\text{mse}(\text{average mean}) &= \sigma^2 \frac{1}{N} && \text{by Corollary 13.1 page 98}\\
&\le \sigma^2 \sum_{n=1}^{N} \lambda_n^2 && \text{by Lemma 13.1 page 99}\\
&= \text{mse}(\text{arithmetic mean}) && \text{by Theorem 13.1 page 97}
\end{aligned}$$

✏

13.2 Variance Estimation

If we know the true *mean* μ of a stationary random process $(\!(\mathsf{x}_n)\!)$, then a reasonable estimate of the variance might be $\frac{1}{N}\sum_{n=1}^{N}(\mathsf{x}_n-\mu)^2$. This estimate has the highly touted property of being *unbiased*:

$$\mathsf{E}\left[\frac{1}{N}\sum_{n=1}^{N}(\mathsf{x}_n-\mu)^2\right] = \frac{1}{N}\sum_{n=1}^{N}\mathsf{E}\left[(\mathsf{x}_n-\mu)^2\right] = \frac{1}{N}\sum_{n=1}^{N}\sigma^2 = \sigma^2$$

Very good. However, in many cases we don't know the **true mean** μ, but rather only have an **estimated mean** $\hat{\mu} \triangleq \frac{1}{N} = \sum_{n=1}^{N}\mathsf{x}_n$. In this case, substituting in the estimated mean for the true mean as in $\hat{\text{var}}_B\big((\!(\mathsf{x}_n)\!)\big)$ (next definition) yields a *biased* variance estimate (next theorem).

Definition 13.2. [3] *Let* $\hat{\mu}$ *be an estimate of the mean of a random sequence* $(\!(\mathsf{x}_n)\!)$.

DEF

The ***sample variance*** $\hat{\text{var}}\big((\!(\mathsf{x}_n)\!)\big)$ *is defined as*

$$\hat{\text{var}}\big((\!(\mathsf{x}_n)\!)\big) \triangleq \frac{1}{N-1}\sum_{n=1}^{N}(\mathsf{x}_n-\hat{\mu})^2$$

The ***biased sample variance*** $\hat{\text{var}}_B\big((\!(\mathsf{x}_n)\!)\big)$ *is here defined as*

$$\hat{\text{var}}_B\big((\!(\mathsf{x}_n)\!)\big) \triangleq \frac{1}{N}\sum_{n=1}^{N}(\mathsf{x}_n-\hat{\mu})^2$$

The factor $\frac{N}{N-1}$ such that $\hat{\text{var}}\big((\!(\mathsf{x}_n)\!)\big) = \frac{N}{N-1}\hat{\text{var}}_B\big((\!(\mathsf{x}_n)\!)\big)$ is known as "*Bessel's correction*". Why such "correction" would be useful at all is demonstrated by Theorem 13.2 (next). Theorem 13.2 demonstrates that the *biased sample variance* $\hat{\text{var}}_B\big((\!(\mathsf{x}_n)\!)\big)$ is *biased*, and multiplication by $\frac{N}{N-1}$ makes it *unbiased*.

[3] Wilks (1963a), page 199 ⟨§"8.2 MEANS AND VARIANCES OF MEAN, VARIANCE,..."⟩, Wilks (1963b), page 199 ⟨§"(b) MEAN AND VARIANCE OF SAMPLE VARIANCE"⟩, Kenney (1947), page 125 ⟨"BESSEL'S CORRECTION"⟩, Bajpai (1967), page 509 ⟨???⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

Theorem 13.2. [4] *Let $\hat{\mu}$ be the* AVERAGE *(Definition N.4 page 311) of a sequence $(\!(\mathsf{x}_n)\!)$. Let $\mu^4 \triangleq \mathsf{E}\left[\left(\mathsf{x}_n - \mu\right)^4\right]$ be the* 4TH CENTRAL MOMENT *of x_n.*

THM

$$\left.\begin{array}{lll} (A). & (\!(\mathsf{x}_n)\!) \text{ is WSS} & \text{and} \\ (B). & \mu \triangleq \mathsf{E}\mathsf{x}_n & \text{and} \\ (C). & (\!(\mathsf{x}_n)\!) \text{ is UNCORRELATED} & \end{array}\right\} \Longrightarrow \left\{\begin{array}{llcll} (1). & \mathsf{E}\hat{\mathsf{var}}_B(\!(\mathsf{x}_n)\!) & = & \frac{N-1}{N}\sigma^2 & \text{(BIASED)} \quad \text{and} \\ (2). & \mathsf{E}\hat{\mathsf{var}}(\!(\mathsf{x}_n)\!) & = & \sigma^2 & \text{(UNBIASED)} \quad \text{and} \\ (3). & \mathsf{var}\left[\hat{\mathsf{var}}(\!(\mathsf{x}_n)\!)\right] & = & \frac{1}{N}\left[\mu^4 - \left(\frac{N-3}{N-1}\right)\sigma^4\right] & \text{(CONSISTENT)} \end{array}\right.$$

✎PROOF:

1. lemma: $\mathsf{E}(\mathsf{x}_n\hat{\mu}) = \frac{1}{N}\sigma^2 + \mu^2$. Proof:

$$\begin{aligned}
\mathsf{E}\left(\mathsf{x}_n\hat{\mu}\right) &\triangleq \mathsf{E}\left(\mathsf{x}_n\frac{1}{N}\sum_{m=1}^{N}\mathsf{x}_m\right) && \text{by definition of } \textit{average} && \text{(Definition N.4 page 311)} \\
&= \mathsf{E}\left(\frac{1}{N}\sum_{m=1}^{N}\mathsf{x}_n\mathsf{x}_m\right) \\
&= \frac{1}{N}\sum_{m=1}^{N}\mathsf{E}\left(\mathsf{x}_n\mathsf{x}_m\right) && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)} \\
&= \frac{1}{N}\left[\mathsf{E}\mathsf{x}_n^2 + \sum_{m\neq n}\mathsf{E}\left(\mathsf{x}_n\mathsf{x}_m\right)\right] \\
&= \frac{1}{N}\left[\mathsf{E}\mathsf{x}_n^2 + \sum_{m\neq n}\left(\mathsf{E}\mathsf{x}_n\right)\left(\mathsf{E}\mathsf{x}_m\right)\right] && \text{by } \textit{uncorrelated} \text{ hypothesis} && \text{(C)} \\
&= \frac{1}{N}\left[\left(\sigma^2+\mu^2\right) + (N-1)\mu^2\right] && \text{by Corollary 1.3 (page 5)} \\
&= \frac{1}{N}\sigma^2 + \mu^2
\end{aligned}$$

2. Proof for (1):

$$\begin{aligned}
\mathsf{E}\hat{\mathsf{var}}_B\left((\!(\mathsf{x}_n)\!)\right) &\triangleq \mathsf{E}\left[\frac{1}{N}\sum_{n=1}^{N}\left(\mathsf{x}_n - \hat{\mu}\right)^2\right] && \text{by definition of } \hat{\mathsf{var}}_B \text{ (Definition 13.2 page 100)} \\
&= \frac{1}{N}\mathsf{E}\left[\sum_{n=1}^{N}\left(\mathsf{x}_n \underbrace{-\mu+\mu}_{0} - \hat{\mu}\right)^2\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\left[\underbrace{\mathsf{E}\left(\mathsf{x}_n-\mu\right)^2}_{\sigma^2} + 2\mathsf{E}\left[(\mathsf{x}_n-\mu)(\mu-\hat{\mu})\right] + \mathsf{E}\left(\mu - \underbrace{\hat{\mu}}_{\mathsf{E}\mu}\right)^2\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\left[\sigma^2 + 2\mathsf{E}\left[\mathsf{x}_n\mu - \mathsf{x}_n\hat{\mu} - \mu^2 + \mu\hat{\mu}\right] + \frac{1}{N}\sigma^2\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\left[\sigma^2 + 2\left[\mu^2 - \mathsf{E}(\mathsf{x}_n\hat{\mu}) - \mu^2 + \mu^2\right] + \frac{1}{N}\sigma^2\right] && \text{by Corollary 13.1 page 98}
\end{aligned}$$

[4] Wilks (1963a), page 199 ⟨§"8.2 MEANS AND VARIANCES OF MEAN, VARIANCE,…"⟩, TUCKER (1965) PAGE 111 ⟨§"8.2 UNBIASED AND CONSISTENT ESTIMATES"⟩, STUART AND ORD (1991) PAGE 609 ⟨§"UNBIASED ESTIMATORS"⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

$$= \frac{1}{N}\sum_{n=1}^{N}\left[\sigma^2 + 2\left[\mu^2 - \left(\mu^2 + \frac{1}{N}\sigma^2\right)\right] + \frac{1}{N}\sigma^2\right] \qquad \text{by } \textit{unbiased} \text{ property of } \hat{\mu}$$ ((1) lemma p

$$= \frac{1}{N}\sum_{n=1}^{N}\left[\sigma^2 - \frac{1}{N}\sigma^2\right]$$

$$= \frac{N-1}{N}\sigma^2$$

3. Proof for (2):

$$\mathsf{E}\hat{\text{var}}\big(\!\!\big(\mathsf{x}_n\big)\!\!\big) \triangleq \mathsf{E}\left[\frac{1}{N-1}\sum_{n=1}^{N}\left(\mathsf{x}_n - \hat{\mu}\right)^2\right] \qquad \text{by definition of } \hat{\text{var}}$$ (Definition 13.2 page 100)

$$= \frac{N}{N-1}\mathsf{E}\left[\frac{1}{N}\sum_{n=1}^{N-1}\left(\mathsf{x}_n - \hat{\mu}\right)^2\right] \qquad \text{by } \textit{linearity} \text{ of } \mathsf{E}$$ (Theorem 1.1 page 4)

$$= \frac{N}{N-1}\left[\frac{N-1}{N}\sigma^2\right] \qquad \text{by } \hat{\text{var}}_B \text{ result}$$ (item (2) page 101)

$$= \sigma^2$$

4. lemma: $\mathsf{E}\left[\left(\hat{\text{var}}\right)^2\right] = \frac{\mu^4}{N} + \frac{(N-1)^2+2}{N(N-1)}\sigma^4$. Proof: No proof here at this time. The assertion is made by Wilks (1963a), page 199 who also without there supplying a proof says, "Carrying out similar mean value operations we find after some reduction that" the result follows.

$$\mathsf{E}\left[\left(\hat{\text{var}}\right)^2\right] \triangleq \mathsf{E}\left[\left(\frac{1}{N-1}\sum_{n=1}^{N}\left(\mathsf{x}_n - \hat{\mu}\right)^2\right)^2\right] \qquad \text{by definition of } \hat{\text{var}}$$ (Definition 13.2 page 100)

$$= \mathsf{E}\left[\left(\frac{1}{N-1}\sum_{n=1}^{N}\left(\mathsf{x}_n - \hat{\mu}\right)^2\right)\left(\frac{1}{N-1}\sum_{m=1}^{N}\left(\mathsf{x}_m - \hat{\mu}\right)^2\right)\right]$$

$$= \left(\frac{1}{N-1}\right)^2\mathsf{E}\left[\sum_{n=1}^{N}\sum_{m=1}^{N}\left(\mathsf{x}_n - \hat{\mu}\right)^2\left(\mathsf{x}_m - \hat{\mu}\right)^2\right]$$

$$= \left(\frac{1}{N-1}\right)^2\mathsf{E}\left[\sum_{n=1}^{N}\sum_{m=1}^{N}\left(\mathsf{x}_n^2 - 2\mathsf{x}_n\hat{\mu} + \hat{\mu}^2\right)\left(\mathsf{x}_m^2 - 2\mathsf{x}_m\hat{\mu} + \hat{\mu}^2\right)\right]$$

$$= \left(\frac{1}{N-1}\right)^2\mathsf{E}\left[\sum_{n=1}^{N}\sum_{m=1}^{N}\left[\begin{array}{c}\left(\mathsf{x}_n^2\mathsf{x}_m^2 - 2\hat{\mu}\mathsf{x}_n^2\mathsf{x}_m + \hat{\mu}^2\mathsf{x}_n^2\right) + \left(-2\hat{\mu}\mathsf{x}_n\mathsf{x}_m^2 + 4\hat{\mu}^2\mathsf{x}_n\mathsf{x}_m - 2\mathsf{x}_n\hat{\mu}^3\right)\\ + \left(\hat{\mu}^2\mathsf{x}_m^2 - 2\hat{\mu}^3\mathsf{x}_m + \hat{\mu}^4\right)\end{array}\right]\right]$$

$$= \left(\frac{1}{N-1}\right)^2\sum_{n=1}^{N}\sum_{m=1}^{N}\left[\begin{array}{c}\mathsf{E}\left[\mathsf{x}_n^2\mathsf{x}_m^2\right] - 2\mathsf{E}\left[\hat{\mu}\mathsf{x}_n^2\mathsf{x}_m\right] - 2\mathsf{E}\left[\hat{\mu}\mathsf{x}_n\mathsf{x}_m^2\right] + \mathsf{E}\left[\hat{\mu}^2\mathsf{x}_n^2\right] + \mathsf{E}\left[\hat{\mu}^2\mathsf{x}_m^2\right]\\ +4\,\mathsf{E}\left[\hat{\mu}^2\mathsf{x}_n\mathsf{x}_m\right] - 2\mathsf{E}\left[\hat{\mu}^3\mathsf{x}_n\right] - 2\mathsf{E}\left[\hat{\mu}^3\mathsf{x}_m\right] + \mathsf{E}\left[\hat{\mu}^4\right]\end{array}\right]$$

$$= \left(\frac{1}{N-1}\right)^2\sum_{n=1}^{N}\sum_{m=1}^{N}\left[\begin{array}{c}\mathsf{E}\left[\mathsf{x}_n^2\mathsf{x}_m^2\right] - 4\mathsf{E}\left[\hat{\mu}\mathsf{x}_n^2\mathsf{x}_m\right] + 2\mathsf{E}\left[\hat{\mu}^2\mathsf{x}_n^2\right]\\ +4\,\mathsf{E}\left[\hat{\mu}^2\mathsf{x}_n\mathsf{x}_m\right] - 4\mathsf{E}\left[\hat{\mu}^3\mathsf{x}_n\right] + \mathsf{E}\left[\hat{\mu}^4\right]\end{array}\right]$$

$$\stackrel{?}{=} \frac{\mu^4}{N} + \frac{(N-1)^2+2}{N(N-1)}\sigma^4$$

Copyright © 2019 Daniel J. Greenhoe

5. Proof for (3):

$$\begin{aligned}
\operatorname{var}\left[\hat{\operatorname{var}}(\mathsf{x}_n)\right] &\triangleq \mathsf{E}\left[\left(\hat{\operatorname{var}} - \mathsf{E}\hat{\operatorname{var}}\right)^2\right] && \text{by definition of } \hat{\operatorname{var}} && \text{(Definition 13.2 page 100)}\\
&= \mathsf{E}\left[\left(\hat{\operatorname{var}} - \sigma^2\right)^2\right] && \text{by (2)} && \text{(item (3) page 102)}\\
&= \mathsf{E}\left[\left(\hat{\operatorname{var}}\right)^2 - 2\sigma^2\hat{\operatorname{var}} + \left(\sigma^2\right)^2\right] && \text{by } \textit{Binomial Theorem} && \\
&= \mathsf{E}\left[\left(\hat{\operatorname{var}}\right)^2\right] - 2\sigma^2\mathsf{E}\left[\hat{\operatorname{var}}\right] + \mathsf{E}\left[\left(\sigma^2\right)^2\right] && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)}\\
&= \mathsf{E}\left[\left(\hat{\operatorname{var}}\right)^2\right] - 2\left(\sigma^2\right)^2 + \left(\sigma^2\right)^2 && \text{by (2)} && \text{(item (3) page 102)}\\
&= \left[\frac{\mu^4}{N} + \left(\frac{(N-1)^2+2}{N(N-1)}\right)\sigma^4\right] - \sigma^4 && \text{by (4) lemma} && \\
&= \frac{1}{N}\left[\mu^4 + \left(\frac{(N^2-N+1)+2-N(N-1)}{N-1}\right)\sigma^4\right] && && \\
&= \frac{1}{N}\left[\mu^4 - \left(\frac{N-3}{N-1}\right)\sigma^4\right] && &&
\end{aligned}$$

✏

13.3 Estimates in terms of moment estimates

Definition 13.3.

DEF

*The **order-k moment estimate** is here defined as*

$$\hat{\mathsf{M}}_k(\mathsf{x}_n) \triangleq \frac{1}{N}\sum_{n=1}^{N} \mathsf{x}_n^k$$

Proposition 13.2.

PRP

$$\begin{aligned}
\hat{\mu}(\mathsf{x}_n) &= \hat{\mathsf{M}}_1\\
\hat{\operatorname{var}}(\mathsf{x}_n) &= \frac{N}{N-1}\hat{\mathsf{M}}_2 - \frac{N}{N-1}\hat{\mathsf{M}}_1^2
\end{aligned}$$

✎PROOF:

$$\begin{aligned}
\hat{\mu}(\mathsf{x}_n) &\triangleq \frac{1}{N}\sum_{n=1}^{N}\mathsf{x}_n && \text{by definition of } \hat{\mu} && \text{(Definition N.4 page 311)}\\
&\triangleq \hat{\mathsf{M}}_1 && \text{by definition of } \hat{\mathsf{M}}_1 && \text{(Definition 13.3 page 103)}
\end{aligned}$$

$$\begin{aligned}
\hat{\operatorname{var}}(\mathsf{x}_n) &\triangleq \frac{1}{N-1}\sum_{n=1}^{N}\left(\mathsf{x}_n - \hat{\mu}\right)^2 && \text{by definition of } \hat{\operatorname{var}} && \text{(Definition 13.2 page 100)}\\
&= \frac{1}{N-1}\sum_{n=1}^{N}\left(\mathsf{x}_n^2 - 2\mathsf{x}_n\hat{\mu} + (\hat{\mu})^2\right) && && \\
&= \frac{1}{N-1}\sum_{n=1}^{N}\mathsf{x}_n^2 - 2\frac{1}{N-1}\sum_{n=1}^{N}\mathsf{x}_n\hat{\mu} + \frac{1}{N-1}\sum_{n=1}^{N}(\hat{\mu})^2 && && \\
&= \frac{N}{N-1}\underbrace{\frac{1}{N}\sum_{n=1}^{N}\mathsf{x}_n^2}_{\hat{\mathsf{M}}_2} - 2\underbrace{\hat{\mu}}_{\hat{\mathsf{M}}_1}\frac{N}{N-1}\underbrace{\frac{1}{N}\sum_{n=1}^{N}\mathsf{x}_n}_{\hat{\mathsf{M}}_1} + \frac{N}{N-1}\underbrace{(\hat{\mu})^2}_{\hat{\mathsf{M}}_1^2} && &&
\end{aligned}$$

$$= \frac{N}{N-1}\hat{\mathsf{M}}_2 - \frac{N}{N-1}\hat{\mathsf{M}}_1^2$$

✏

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 14

CORRELATION ESTIMATION

Definition 14.1. [1]

DEF

*The **windowed auto-correlation estimate** $\hat{\mathsf{R}}_{\mathsf{xx}}(m)$ is defined as*

$$\hat{\mathsf{R}}_{\mathsf{xx}}(m) \triangleq \frac{1}{N} \sum_{n=0}^{N-|m|} \mathsf{x}(n)\mathsf{x}(n+m)$$

Theorem 14.1. [2]

THM

$$\begin{aligned} \mathsf{E}\left[\hat{\mathsf{R}}_{\mathsf{xx}}(m)\right] &= \left(1 - \frac{|m|}{N}\right)\mathsf{R}_{\mathsf{xx}}(m) && \text{(ASYMPTOTICALLY UNBIASED)} \\ \operatorname{var}\left[\hat{\mathsf{R}}_{\mathsf{xx}}(m)\right] &= \frac{1}{N}\sum_{n\in\mathbb{Z}}\left[\mathsf{R}_{\mathsf{xx}}^2(n) + \mathsf{R}_{\mathsf{xx}}(n-m)\mathsf{R}_{\mathsf{xx}}(n+m)\right] && \text{(CONSISTENT)} \end{aligned}$$

[1] Vaseghi (2000) page 271 ⟨§"9.3.3 Energy-Spectral Density and Power-Spectral Density"⟩

[2] Jenkins and Watts (1968), Vaseghi (2000) page 272 ⟨§"9.3.3 Energy-Spectral Density and Power-Spectral Density"⟩

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 15

SPECTRAL ESTIMATION

Quality of spectral estimators[1]

THM				
	Periodogram:	Q	$=$	1
	Welch Method 0% overlap:	Q	$=$	$0.78N\Delta f$
	Welch Method 50% overlap:	Q	$=$	$1.39N\Delta f$
	Bartlett Method:	Q	$=$	$1.11N\Delta f$
	Blackman-Tukey Method:	Q	$=$	$2.34N\Delta f$

[1] Proakis (2002) pages 452–457 ⟨§"8.2.4 Performance Characteristics of Nonparametric Power Spectrum Estimators"⟩, Proakis and Manolakis (1996) pages 916–919 ⟨§"12.2.4 Performance Characteristics of Nonparametric Power Spectrum Estimators"⟩, Rao and Swamy (2018) page 731 ⟨"Table 12.1 Comparison of performance of classical methods"⟩, Salivahanan and Vallavaraj (2001) page 606 ⟨§"12.5 Power Spectrum Estimation: Non-Parametric Methods"⟩, Ifeachor and Jervis (2002) pages 706–707 ⟨§"11.3.7 Comparison of the power spectral density estimation methods"⟩, J.S.Chitode (2009b) page P-100, Abdaheer (2009), page 204

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 16

DENSITY ESTIMATION

References:

- Silverman (1986)

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 17

SYSTEM IDENTIFICATION

❝ *I can therefore gladly admit that falsificationists like myself much prefer an attempt to solve an interesting problem by a bold conjecture, even (and especially) if it so turns out to be false, to any recital of a sequence of irrelevant truisms. We prefer this because we believe that this is the way in which we can learn from our mistakes and that in finding that our conjecture was false we shall have learned much about the truth, and shall have got nearer to the truth.* ❞

Karl R. Popper (1902–1994) [1]

17.1 Estimation techniques

Let **S** be a *system* with *impulse response* $\mathsf{h}(n)$ with with *DTFT* $\tilde{\mathsf{H}}(\omega)$, input $\mathsf{x}(n)$, and output $\mathsf{y}(n)$. Often in the field of "digital signal processing" (DSP), **S** is a "filter" with known $\mathsf{h}(n)$ and $\tilde{\mathsf{H}}(\omega)$ because the filter **S** was designed by a designer who had direct control over $\mathsf{h}(n)$.

However in many other practical situations, **S** is some other system for which $\mathsf{h}(n)$ and $\tilde{\mathsf{H}}(\omega)$ are *not* known…but which we may want to *estimate*. Examples of such **S** is a device on an industrial shaker table, a communication channel, or the entire earth.

Determining $\mathsf{h}(n)$ and/or $\tilde{\mathsf{H}}(\omega)$ is part of an operation called "*system identication*". Determining $\tilde{\mathsf{H}}(\omega)$ in particular is referred to as "*Frequency Response Identification*"[2] or as "*Frequency Response Function*" ("*FRF*") estimation.[3] *FRF* estimation is a challenging problem and one that many have devoted much effort to. This chapter describes some of that effort.

In the early days, people used a rather obvious technique for determining $\tilde{\mathsf{H}}(\omega)$—the humble *sine sweep*. That is, they drove the input with a sine wave with slowly increasing (or decreasing) frequency while measuring the resulting output. This technique, although effective, was "very slow".[4]

---

[1] quote: Popper (1962), page 231, Popper (1963) page 313
image: https://en.wikipedia.org/wiki/File:Karl_Popper.jpg, "no known copyright restrictions"

[2] Shin and Hammond (2008) page 292

[3] Cobb (1988) page 1 ⟨FRF "measurement"⟩

[4] Leuridan et al. (1986)911"Stepped Sine Testing", Cobb (1988) page 1 ⟨Chapter 1—Introduction⟩, Ewins

(A) **communications additive noise model**
The "input signal" is $\mathsf{x}(n)$.

(B) **measurement additive noise model**
The "input signal" is $\mathsf{p}(n)$.

In each model, $\mathsf{x}(n)$ and $\mathsf{y}(n)$ are "known", and $\mathsf{u}(n)$, $\mathsf{v}(n)$, and $\mathsf{w}(n)$ are *not*.
In definition, the two models are **equivalent** under the relation $\mathsf{u}(n) = -\mathsf{w}(n)$.
In practice, they are **different**:

- in (A), x and u would be typically *uncorrelated*;
- in (B), x and $\mathsf{w} = -\mathsf{u}$ are very much *correlated* (x is a function of u).

Figure 17.1: Additive noise systems with *linear*/**non-linear** operator **T**

And there is another problem here—we don't always have control over the input signal. Examples of this include earthquake and volcanic activity analysis.

An alternative to the sine-sweep input is *random sequence* input. All the techniques that follow in this chapter are of this type. A problem with using random sequences directly for estimating $\tilde{\mathsf{H}}(\omega)$ is that the estimate $\hat{\mathsf{H}}(\omega)$ is itself also random. This is not what we want. We want an estimate that we can actually write down on paper or at least plot on paper.

A solution to this is to not use the random sequences directly to estimate $\tilde{\mathsf{H}}(\omega)$, but instead to first use the *expectation* operator E (Definition 1.1 page 3). The expectation operator takes a quantity X that is inherently "random" (with some probability distribution $\mathsf{p}(x)$) and turns it into a deterministic "constant" $\mathsf{E}\mathsf{X}$.

The operator E is also used by the spectral density functions $\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)$ and $\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)$ (Definition 7.3 page 46). And $\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)$ and $\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)$ are what are typically used to calculate an estimate $\hat{\mathsf{H}}(\omega)$.

17.2 Additive noise system models

Consider the additive noise systems illustrated in Figure 18.1 (page 137).

- The illustration on the left is suitable for modeling a communications system where x is the transmitted signal, y is the received signal, u and v are thermal noise, and the "transfer function" **H** is the communications channel (air, water, wires, etc.) that one wishes to estimate.
- The illustration on the right is suitable for modeling a testing system where p is an input test signal (from an industrial shaker or from a naturally occurring signal originating from geophysical activity), w is measurement noise, x is the measured input contaminated by noise, and **H** is the device under test (a piece of equipment, a building, or the entire earth).

Note that the two models are an equivalent system **S** under the relation $\mathsf{u} = -\mathsf{w}$. But although one might expect such a sign difference to wreak mathematical havoc in resulting equations, this is

(1986) pages 125–140 ⟨3.7 Use of different excitation types⟩

Copyright © 2019 Daniel J. Greenhoe

simply not the case here because

$$\tilde{S}_{ww} = \tilde{\mathbf{F}}E[w(m)w*(0)] = \tilde{\mathbf{F}}E\big[(-u(m))\big(-u^*(0)\big)\big] = \tilde{\mathbf{F}}E\big[(u(m))\big(u^*(0)\big)\big] = \tilde{S}_{uu}$$

So the sign difference is not that big of a difference after all. But there are some key differences in practice:

- In the communications model (on the left), the "input signal" is $x(n)$ and the frequency-domain input *signal-to-noise ratio* (*SNR*) is $\tilde{S}_{xx}(\omega)/\tilde{S}_{uu}(\omega)$. In the measurement model (on the right), the "input signal" is $p(n)$ and the frequency-domain input *signal-to-noise ratio* (*SNR*) is $\tilde{S}_{pp}(\omega)/\tilde{S}_{ww}(\omega) = \tilde{S}_{pp}(\omega)/\tilde{S}_{uu}(\omega)$.
- On the left, x and u would be typically *uncorrelated*; on the right, x and $w = -u$ are very much *correlated* (x is a function of u).

17.3 Transfer function estimate definitions and interpretation

As a first attempt at estimating the transfer function $\mathbf{H}$ of $\mathbf{S}$, or at least the magnitude squared of $\mathbf{H}$, we might assume $\mathbf{H}$ to be *LTI*, take a cue from the relation $\tilde{S}_{yy} = \tilde{S}_{xx}|\tilde{H}|^2$ of Corollary 6.3 (page 39), and arrive at a function called "*transmissibility*" (next definition).

Definition 17.1. [5] *Let* $\mathbf{S}$ *be a* SYSTEM *with input* $x(n)$ *and output* $y(n)$.

DEF ***transmissibility*** $\tilde{T}_{xy}(\omega)$ *is defined as* $$\tilde{T}_{xy}(\omega) \triangleq \sqrt{\frac{\tilde{S}_{yy}(\omega)}{\tilde{S}_{xx}(\omega)}}$$

Transmissibility is in essence the ratio of "*spectral power*" (Remark 7.1 page 46) output to *spectral power* input. Note that it is a real-valued function (because $\tilde{S}_{xx}$ and $\tilde{S}_{yy}$ are real-valued). We might suspect that we could attain better estimates of $\mathbf{H}$ by allowing the estimates to be complex-valued. And in fact, all the remaining estimates in this section are in general complex-valued.

And so to start (again), and in the very special (a.k.a unrealistic) case of $\mathbf{S}$ having *zero measurement noise* (*zero measurement error*) ($v = u = w = 0$), $h(n)$ being *linear time invariant* (*LTI*), and input $x(n)$ being *wide sense stationary*…then we can determine (a.k.a "identify") $h(n)$ or $\tilde{H}(\omega)$ exactly by $\tilde{H}(\omega) = \tilde{S}_{yx}(\omega)/\tilde{S}_{xx}(\omega)$ (Corollary 6.3 page 39).

However, in practical situations, there is measurement noise/error. Examples may include "road noise" from a test being performed in a moving vehicle or *quantization noise* from an *analog-to-digital converter* (*ADC*).

If the measurement error is at the output only (and under the assumptions of *LTI* and *WSS*) then $\hat{H}_1$ (next definition) is the ideal estimator in the sense that $\hat{H}_1 = \tilde{H}$ (Corollary 17.4 page 131).

Definition 17.2. [6] *Let* $\mathbf{S}$ *be a* SYSTEM *with input* $x(n)$ *and output* $y(n)$.

[5] Bendat and Piersol (2010) page 469 $\langle |H(f)| = \left[G_{yy}(f)/G_{xx}(f)\right]^{1/2}\rangle$, Yan and Ren (2012) page 204 $\langle (1)\ \left[G_{YY}(s)\right] = [H(s)]\left[G_{FF}(s)\right]\left[H^*(s)\right]^T\rangle$, Goldman (1999) page 179 ⟨Transmissibility …$H'_{ab} = G_{bb}/G_{aa}$ (note: differs by $\sqrt{\cdot}$ from Bendat and Piersol⟩, Zhang et al. (2016), Zhou and Wahab (2018) page 824, https://link.springer.com/chapter/10.1007/978-3-319-54109-9_4

[6] Bendat and Piersol (1993) pages 106–109 ⟨5.1.1 Optimality of Calculations⟩, Bendat and Piersol (2010) page 185 $\langle H_1(f) = G_{xy}(f)/G_{xx}(f)$ (6.37)⟩, Shin and Hammond (2008) page 293 $\langle\ H_1(f) = \tilde{S}_{xy}(f)/\tilde{S}_{xx}(f)$ (9.63); which dif-

DEF *The* ***Least Squares transfer function estimate*** $\hat{\mathsf{H}}_1(\omega)$ *of* **S** *is defined as* $\hat{\mathsf{H}}_1(\omega) \triangleq \dfrac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}$

The estimator $\hat{\mathsf{H}}_1$ is a good start. However in the early 1980s, L. D. Mitchell pointed out that in the presence of input noise, $\hat{\mathsf{H}}_1$ is far from ideal in that it is *biased* with respect to $\tilde{\mathsf{H}}$; in fact, $\hat{\mathsf{H}}_1$ *under estimates* $\tilde{\mathsf{H}}$ (Corollary 17.4 page 131). Mitchell proposed a new estimator $\hat{\mathsf{H}}_2$ (next definition).

This estimator has the special property that when there is input noise but no output noise (and under *LTI*, *WSS*, and *uncorrelated* assumptions), then it is ideal in the sense that $\hat{\mathsf{H}}_2(\omega) = \tilde{\mathsf{H}}(\omega)$ (Corollary 17.4 page 131).

Note also that in the case of both no input and no output noise, then $\hat{\mathsf{H}}_1 = \hat{\mathsf{H}}_2$ (Corollary 6.3 page 39).

Definition 17.3. [7] *Let* **S** *be a* SYSTEM *with input* $\mathsf{x}(n)$ *and output* $\mathsf{y}(n)$.

DEF *The* ***Inverse Method transfer function estimate*** $\hat{\mathsf{H}}_2(\omega)$ *of* **S** *is defined as* $\hat{\mathsf{H}}_2(\omega) \triangleq \dfrac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)}$

Mitchell's $\hat{\mathsf{H}}_2$ contribution "generated a flurry of activity"[8] and soon more $\tilde{\mathsf{H}}$ estimators appeared. So far we have

- $\hat{\mathsf{H}}_1$ which is ideal when there is no input noise but *under estimates* $\tilde{\mathsf{H}}$ when there is (Corollary 17.4 page 131)
- $\hat{\mathsf{H}}_2$ which is ideal when there is no output noise but *over estimates* $\tilde{\mathsf{H}}$ when there is (Corollary 17.4 page 131).

But what about estimators for when there is noise on both input and output? Armed with two estimators that between them account for both input and output noise, an "ad hoc" solution might be to somehow take mean values of $\hat{\mathsf{H}}_1$ and $\hat{\mathsf{H}}_2$ to induce new estimators—this approach summarizes the next three definitions. An arguably more mature approach is to find estimators that are optimal with respect to least squares measures—and this approach summarizes Definition 17.9 – Definition 17.7 (page 117).

Definition 17.4. *Let* **S** *be a* SYSTEM *with input* $\mathsf{x}(n)$ *and output* $\mathsf{y}(n)$.

DEF *The* ***Arithmetic Mean transfer function estimate*** $\hat{\mathsf{H}}_{\mathsf{am}}(\omega)$ *of* **S** *is defined as*

$$\hat{\mathsf{H}}_{\mathsf{am}}(\omega) \triangleq \frac{\left|\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\right|^2 + \tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{2\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)}$$

fers from Definition 17.2, but see APPENDIX ?? page ?? ⟩, Bendat (1978)cited by Cobb(1988)—variance estimate for $\hat{\mathsf{H}}_1$, Allemang et al. (1979) ⟨cited by Shin(2008)⟩, Leuridan et al. (1986) page 910 ⟨*Least Squares Technique*; (8) $[G_{xx}](H) = [G_{xy}]$⟩, Abom (1986)cited by Cobb(1988)—variance estimate for $\hat{\mathsf{H}}_1$, Allemang et al. (1987) pages 54–55 ⟨5.3.1 H_1 Technique; $[H] = [G_{XF}][G_{FF}]^{-1}$ (11)⟩, Cobb (1988) page 2 ⟨[1] $\hat{H}(f) = \hat{G}_{yx}(f)/\hat{G}_{xx}(f)$ (1)⟩, Goyder (1984) page 438 ⟨$H(i\omega) = S_{qp}/S_{pp}$ (3)⟩, Pintelon and Schoukens (2012) page 233 ⟨$\hat{G}(\Omega_k) = S_{yu}(j\omega_k)S_{uu}^{-1}(j\omega_k)$ (7-30)⟩, White et al. (2006) page 678 ⟨$H_1(f) = \hat{S}_{x_m y_m}(f)/\hat{S}_{x_m x_m}(f)$ (1) which differs by conjugate, references Bendat and Piersol⟩,

[7] Shin and Hammond (2008) page 293 ⟨$H_2(f) = \tilde{\mathsf{S}}_{\mathsf{yy}}(f)/\tilde{\mathsf{S}}_{\mathsf{yx}}(f)$ (9.65); which differs from Definition 17.3, but see APPENDIX ?? page ?? ⟩, Bendat and Piersol (2010) page 186 ⟨$H_2(f) = G_{yy}(f)/G_{yx}(f)$ (6.42)⟩, Mitchell (1980) ⟨cited by Cobb(1988)⟩, Mitchell (1982) page 278 ⟨"Define what will be called an inverse method for calculation of a FRF as…"; $H_2(f) = G_{yy}/G_{yx}$ (6); Note this differs with Definition 17.3 by a conjugate, but note that Mitchell seems to follow Bendat (see his [3] and [4]), which would explain this difference (APPENDIX ?? page ??)⟩, Cobb (1988) page 3 ⟨[2] $\hat{H}(f) = \hat{G}_{yy}(f)/\hat{G}_{xy}(f)$ (1)⟩, White et al. (2006) page 678 ⟨$H_2(f) = \hat{S}_{y_m y_m}(f)/\hat{S}_{y_m x_m}(f)$ (2) which differs by conjugate, references Bendat and Piersol⟩

[8] Cobb (1988) page 3

COPYRIGHT © 2019 DANIEL J. GREENHOE

Proposition 17.1. [9] *Let* **S** *be a* SYSTEM *with input* $\mathsf{x}(n)$ *and output* $\mathsf{y}(n)$.

PRP

$$\hat{\mathsf{H}}_{\mathsf{am}}(\omega) = \frac{\hat{\mathsf{H}}_1(\omega) + \hat{\mathsf{H}}_2(\omega)}{2} \quad (\textbf{\textit{arithmetic mean}} \text{ of } \hat{\mathsf{H}}_1 \text{ and } \hat{\mathsf{H}}_2)$$

✎PROOF:

$$\hat{\mathsf{H}}_{\mathsf{am}}(\omega) \triangleq \frac{\left|\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\right|^2 + \tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{2\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)} \quad \text{by definition of } \hat{\mathsf{H}}_{\mathsf{am}} \quad \text{(Definition 17.4 page 114)}$$

$$= \frac{\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)}{2\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)} + \frac{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{2\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)} = \frac{\dfrac{\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} + \dfrac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)}}{2}$$

$$= \frac{\hat{\mathsf{H}}_1(\omega) + \hat{\mathsf{H}}_2(\omega)}{2} \quad \text{by definitions of } \hat{\mathsf{H}}_1 \text{ and } \hat{\mathsf{H}}_2 \quad \text{(Definition 17.2 page 113, Definition 17.3 page 114)}$$

✏

Definition 17.5. *Let* **S** *be a* SYSTEM *with input* $\mathsf{x}(n)$ *and output* $\mathsf{y}(n)$.

DEF

The ***Geometric mean transfer function estimate*** $\hat{\mathsf{H}}_{\mathsf{gm}}(\omega)$ *of* **S** *is defined as*

$$\hat{\mathsf{H}}_{\mathsf{gm}}(\omega) \triangleq \frac{\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)}{\left|\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\right|}\sqrt{\frac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}}$$

Proposition 17.2. [10] *Let* **S** *be a* SYSTEM *with input* $\mathsf{x}(n)$ *and output* $\mathsf{y}(n)$.

PRP

$$\pm\hat{\mathsf{H}}_{\mathsf{gm}}(\omega) = \sqrt{\hat{\mathsf{H}}_1(\omega)\hat{\mathsf{H}}_2(\omega)} \quad (\textbf{\textit{geometric mean}} \text{ of } \hat{\mathsf{H}}_1 \text{ and } \hat{\mathsf{H}}_2)$$

✎PROOF:

$$\pm\hat{\mathsf{H}}_{\mathsf{gm}}(\omega) \triangleq \pm\frac{\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)}{\left|\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\right|}\sqrt{\frac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}} \quad \text{by definition of } \hat{\mathsf{H}}_{\mathsf{gm}} \quad \text{(Definition 17.5 page 115)}$$

$$= \sqrt{\frac{\left[\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)\right]^2}{\left|\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\right|^2}\frac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}} = \sqrt{\frac{\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)}\frac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}} = \sqrt{\frac{\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}\frac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)}}$$

$$= \sqrt{\hat{\mathsf{H}}_1(\omega)\hat{\mathsf{H}}_2(\omega)} \quad \text{by definitions of } \hat{\mathsf{H}}_1 \text{ and } \hat{\mathsf{H}}_2 \quad \text{(Definition 17.2 page 113, Definition 17.3 page 114)}$$

$$= \textbf{Geometric mean} \text{ of } \hat{\mathsf{H}}_1(\omega) \text{ and } \hat{\mathsf{H}}_2(\omega)$$

Note that for a complex number $z \triangleq |z|e^{i\phi}$, $\sqrt{z}$ has two solutions:[11]

$$\sqrt{z} = \sqrt{|z|e^{i\phi}} = \{z_1, z_2\} = \left\{\sqrt{|z|}e^{i(\phi/2)}, \sqrt{|z|}e^{i(\phi/2+\pi)}\right\} = \pm\sqrt{|z|}e^{i(\phi/2)}$$

because $z_1^2 = z$ and $z_2^2 = z$. ✏

Note that the *geometric mean estimator* (Definition 17.5 page 115) and *transmissibility* (Definition 17.1 page 113) are closely related (next).

[9] Mitchell (1982) page 279 ⟨"Frequency Response Calculation: The Average Method"⟩, Zheng et al. (2002) page 918 ⟨"1.3 Arithmetic Mean Estimator H_3"⟩

[10] Zheng et al. (2002) page 918 ⟨"1.4 Geometric Mean Estimator H_4"⟩

[11] Many many thanks to Ben Cleveland for his help with this!!!

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Proposition 17.3. *Let $\phi(\omega)$ be the* PHASE *of $\tilde{S}_{xy}(\omega)$ such that $\tilde{S}_{xy}(\omega) \triangleq |\tilde{S}_{xy}(\omega)|e^{-i\phi(\omega)}$*

PRP

$$\hat{H}_{gm}(\omega) = \tilde{T}_{xy}(\omega)\, e^{-i\phi(\omega)} \quad \left(\begin{array}{lcll} |\hat{H}_{gm}(\omega)| & = & \tilde{T}_{xy}(\omega) & \textit{is the } \textsc{magnitude} \textit{ of } \hat{H}_{gm}(\omega) \quad \textit{and} \\ \angle\hat{H}_{gm}(\omega) & = & -\angle\tilde{S}_{xy}(\omega) & \textit{is the } \textsc{phase} \textit{ of } \hat{H}_{gm}(\omega) \end{array} \right)$$

✎PROOF: Let $\phi(\omega)$ be the *phase* of

$$\begin{aligned}
\hat{H}_{gm}(\omega) &\triangleq \sqrt{\hat{H}_1(\omega)\hat{H}_2(\omega)} && \text{by definition of } \hat{H}_{gm} && \text{(Definition 17.5 page 115)} \\
&\triangleq \sqrt{\frac{\tilde{S}^*_{xy}(\omega)}{\tilde{S}_{xx}(\omega)}\frac{\tilde{S}_{yy}(\omega)}{\tilde{S}_{xy}(\omega)}} && \text{by definitions of } \hat{H}_1 \text{ and } \hat{H}_2 && \text{(Definition 17.2 page 113, Definition 17.3 page 114)} \\
&= \sqrt{\frac{\tilde{S}_{yy}(\omega)}{\tilde{S}_{xx}(\omega)}\frac{\tilde{S}^*_{xy}(\omega)}{\tilde{S}_{xy}(\omega)}} \\
&= \tilde{T}_{xy}(\omega)\sqrt{\frac{\tilde{S}^*_{xy}(\omega)}{\tilde{S}_{xy}(\omega)}} && \text{by definition of } \tilde{T}_{xy} && \text{(Definition 17.1 page 113)} \\
&= \tilde{T}_{xy}(\omega)\sqrt{\frac{|\tilde{S}_{xy}(\omega)|e^{-i\phi(\omega)}}{|\tilde{S}_{xy}(\omega)|e^{i\phi(\omega)}}} && \text{where } \tilde{S}_{xy}(\omega) \triangleq |\tilde{S}_{xy}(\omega)|e^{-i\phi(\omega)} \\
&= \tilde{T}_{xy}(\omega)\sqrt{e^{-i2\phi(\omega)}} \\
&= \tilde{T}_{xy}(\omega)\, e^{-i\phi(\omega)}
\end{aligned}$$

✏

Remark 17.1. *Transmissibility* $\tilde{T}_{xy}$ is a kind of "black sheep" of the system identification function family. All the other members of this family ($\hat{H}_1$, $\hat{H}_2$, $\hat{H}_v$, $\hat{H}_s$) are *complex-valued*, but $\tilde{T}_{xy}$ is only *real-valued*—a seemingly ordinary Joe born into a super-hero family. But Proposition 17.3 suggests that $\tilde{T}_{xy}$ is not simply a "black sheep", but rather a "dark horse" with abilities that can easily be unleashed by slight redefinition. In particular, Proposition 17.3 demonstrates that $\tilde{T}_{xy}$ is the *magnitude* of the geometric mean of $\hat{H}_1$ and $\hat{H}_2$. We can thus justifiably define a **complex transmissibility** function as $\hat{H}_{gm}$...and the magnitude of this *complex transmissibility* function is the *ordinary transmissibility* function of Definition 17.1 (page 113).

REM

complex transmissibility $\tilde{T}'_{xy}(\omega) \triangleq \hat{H}_{gm}(\omega)$

Definition 17.6. *Let* **S** *be a* SYSTEM *with input* $x(n)$ *and output* $y(n)$.

DEF

The ***Harmonic mean transfer function estimate*** $\hat{H}_{hm}(\omega)$ *of* **S** *is defined as*

$$\hat{H}_{hm}(\omega) \triangleq \frac{2\tilde{S}_{yy}(\omega)\tilde{S}^*_{xy}(\omega)}{\tilde{S}_{xx}(\omega)\tilde{S}_{yy}(\omega) + |\tilde{S}_{xy}(\omega)|^2}$$

Proposition 17.4. [12] *Let* **S** *be a* SYSTEM *with input* $x(n)$ *and output* $y(n)$.

PRP

$$\hat{H}_{hm}(\omega) = \frac{2}{\frac{1}{\hat{H}_1(\omega)} + \frac{1}{\hat{H}_2(\omega)}} \qquad (\textit{Harmonic mean of } \hat{H}_1 \textit{ and } \hat{H}_2)$$

[12] Carne and Dohrmann (2006) $\langle H_C = [H_A^{-1} + H_B^{-1}]^{-1} \rangle$

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$\begin{aligned}
\hat{\mathsf{H}}_{\mathsf{hm}}(\omega) &\triangleq \frac{2\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) + \left|\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\right|^2} && \text{by definition of } \hat{\mathsf{H}}_{\mathsf{hm}} && \text{(Definition 17.6 page 116)} \\
&= \frac{2}{\dfrac{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) + \left|\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\right|^2}{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)}} = \frac{2}{\dfrac{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}{\tilde{\mathsf{S}}^*_{\mathsf{xy}}(\omega)} + \dfrac{\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}} \\
&= \frac{2}{\dfrac{1}{\hat{\mathsf{H}}_1(\omega)} + \dfrac{1}{\hat{\mathsf{H}}_2(\omega)}} && \text{by definitions of } \hat{\mathsf{H}}_1 \text{ and } \hat{\mathsf{H}}_2 && \text{(Definition 17.2 page 113, Definition 17.3 page 114)} \\
&= \textbf{Harmonic mean} \text{ of } \hat{\mathsf{H}}_1(\omega) \text{ and } \hat{\mathsf{H}}_2(\omega)
\end{aligned}$$

✏

A bit of review reveals $\hat{\mathsf{H}}_1$ at the low end of the estimation problem, $\hat{\mathsf{H}}_2$ at the high end, and $\hat{\mathsf{H}}_{\mathsf{hm}}$, $\hat{\mathsf{H}}_{\mathsf{gm}}$, and $\hat{\mathsf{H}}_{\mathsf{am}}$ somewhere between. But these three "between" estimates are not shown to be optimal in any sense—they are just conceptually interesting. What we might really like is a family of estimators that

- include $\hat{\mathsf{H}}_1$ and $\hat{\mathsf{H}}_2$ as limiting cases
- include the between cases
- are optimal in some sense

The estimator $\hat{\mathsf{H}}_\kappa(\omega;\kappa)$ is one such estimator (next definition) that

- has $\hat{\mathsf{H}}_1$ and $\hat{\mathsf{H}}_2$ as limiting cases (Theorem 17.1 page 119),
- is optimal in the least squares sense (Theorem 17.6 page 132), and
- allows for a system designer to specify an output-input spectral noise ratio $\kappa(\omega)$ that can vary with frequency ω.

Moreover, $\hat{\mathsf{H}}_\kappa(\omega)$ includes some special cases:

- In the case of constant κ, $\hat{\mathsf{H}}_\kappa$ simplifies to the *Scaling transfer function estimate* $\hat{\mathsf{H}}_{\mathsf{s}}$ (Definition 17.8 page 117).
- In the case of $\kappa = 1$, $\hat{\mathsf{H}}_\kappa$ and $\hat{\mathsf{H}}_{\mathsf{s}}$ simplify to the *Total least squares transfer function estimate* $\hat{\mathsf{H}}_{\mathsf{v}}$ (Definition 17.9 page 118).

Definition 17.7. [13] *Let* $\mathbf{S}$ *be a* SYSTEM *with input* $\mathsf{x}(n)$ *and output* $\mathsf{y}(n)$.

DEF

The ***transfer function estimate*** $\hat{\mathsf{H}}_\kappa(\omega;\kappa)$ *with* ***scaling function*** $\kappa(\omega)$ *is defined as*

$$\hat{\mathsf{H}}_\kappa(\omega;\kappa) \triangleq \frac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) - \kappa(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) + \sqrt{\left[\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega) - \kappa(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)\right]^2 + 4\kappa(\omega)\left|\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)\right|^2}}{2\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)}$$

Definition 17.8. [14] *Let* $\mathbf{S}$ *be a* SYSTEM *with input* $\mathsf{x}(n)$ *and output* $\mathsf{y}(n)$.

DEF

The ***Scaling transfer function estimate*** $\hat{\mathsf{H}}_{\mathsf{s}}(\omega;s)$ *of* $\mathbf{S}$ *with* ***scaling parameter*** $s \in [0 : \infty)$ *is defined as* $\quad \hat{\mathsf{H}}_{\mathsf{s}}(\omega;s) \triangleq \hat{\mathsf{H}}_\kappa(\omega;\kappa) \quad$ *with* $\kappa(\omega) \triangleq s^2$

[13] White et al. (2006) page 679 ⟨(6)⟩, Shin and Hammond (2008) page 293 ⟨(9.67)⟩

[14] Shin and Hammond (2008) page 293 ⟨(9.67) with $\kappa(\omega) = s^2$⟩, White et al. (2006) page 679 ⟨(6) with $\kappa(\omega) = s^2$⟩, Leclere et al. (2014) ⟨(10) $\kappa(f) = 1/s^2$ and x and y swapped⟩, Wicks and Vold (1986) page 898 ⟨has additional s in denominator⟩, Zheng et al. (2002) page 918 ⟨(10), seems to differ⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

Definition 17.9. [15] *Let* $\mathbf{S}$ *be a* SYSTEM *with input* $\mathsf{x}(n)$ *and output* $\mathsf{y}(n)$.

DEF *The* ***Total Least Squares transfer function estimate*** $\hat{\mathsf{H}}_{\mathsf{v}}(\omega)$ *of* $\mathbf{S}$ *is defined as*

$$\hat{\mathsf{H}}_{\mathsf{v}}(\omega) \triangleq \hat{\mathsf{H}}_{\kappa}(\omega;\kappa) \qquad \text{with } \kappa(\omega) = 1$$

The previous estimators all assumed two signals: an input $\mathsf{x}(n)$ and an output $\mathsf{y}(n)$. However, in many practical systems, there is a third signal that is "driving" the system. In 1984 Goyder proposed an estimator (next definition) that is based on three signals.

Definition 17.10 (Three channel estimate)**.** [16] *Let* $\mathbf{S}$ *be a system with input* $\mathsf{x}(n)$*, output* $\mathsf{y}(n)$*, and a driver* $\mathsf{p}(n)$.

DEF *The* ***transfer function estimate*** $\hat{\mathsf{H}}_{\mathsf{c}}(\omega)$ *is defined as*

$$\hat{\mathsf{H}}_{\mathsf{c}}(\omega) \triangleq \frac{\tilde{\mathsf{S}}_{\mathsf{py}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{px}}(\omega)}$$

17.4 Estimator relationships

Lemma 17.1.

LEM

$$\frac{\mathrm{d}}{\mathrm{d}p}\left[\tilde{\mathsf{S}}_{\mathsf{yy}} - p\tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left(\tilde{\mathsf{S}}_{\mathsf{yy}} - p\tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}\right] = \frac{\tilde{\mathsf{S}}_{\mathsf{xx}}\left(p\tilde{\mathsf{S}}_{\mathsf{xx}}-\tilde{\mathsf{S}}_{\mathsf{yy}}\right)+2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2-\tilde{\mathsf{S}}_{\mathsf{xx}}\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{xx}}-\tilde{\mathsf{S}}_{\mathsf{yy}}\right)^2+4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{xx}}-\tilde{\mathsf{S}}_{\mathsf{yy}}\right)^2+4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}$$

$$\frac{\mathrm{d}}{\mathrm{d}p}\left[p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}\right] = \frac{\tilde{\mathsf{S}}_{\mathsf{yy}}\left(p\tilde{\mathsf{S}}_{\mathsf{yy}}-\tilde{\mathsf{S}}_{\mathsf{xx}}\right)+2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2+\tilde{\mathsf{S}}_{\mathsf{yy}}\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{yy}}-\tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2+4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{yy}}-\tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2+4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}$$

✎PROOF:

$$\frac{\mathrm{d}}{\mathrm{d}p}\left[\tilde{\mathsf{S}}_{\mathsf{yy}} - p\tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left(\tilde{\mathsf{S}}_{\mathsf{yy}} - p\tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}\right] = -\tilde{\mathsf{S}}_{\mathsf{xx}} + \frac{-2\tilde{\mathsf{S}}_{\mathsf{xx}}\left(\tilde{\mathsf{S}}_{\mathsf{yy}} - p\tilde{\mathsf{S}}_{\mathsf{xx}}\right) + 4\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}{2\sqrt{\left(\tilde{\mathsf{S}}_{\mathsf{yy}} - p\tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}$$

$$= \frac{4\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2 - 2\tilde{\mathsf{S}}_{\mathsf{xx}}\left(\tilde{\mathsf{S}}_{\mathsf{yy}} - p\tilde{\mathsf{S}}_{\mathsf{xx}}\right) - 2\tilde{\mathsf{S}}_{\mathsf{xx}}\sqrt{\left(\tilde{\mathsf{S}}_{\mathsf{yy}} - p\tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{2\sqrt{\left(\tilde{\mathsf{S}}_{\mathsf{yy}} - p\tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}$$

$$= \frac{\tilde{\mathsf{S}}_{\mathsf{xx}}\left(p\tilde{\mathsf{S}}_{\mathsf{xx}} - \tilde{\mathsf{S}}_{\mathsf{yy}}\right) + 2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2 - \tilde{\mathsf{S}}_{\mathsf{xx}}\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{xx}} - \tilde{\mathsf{S}}_{\mathsf{yy}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{xx}} - \tilde{\mathsf{S}}_{\mathsf{yy}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}$$

$$\frac{\mathrm{d}}{\mathrm{d}p}\left[p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}\right] = \tilde{\mathsf{S}}_{\mathsf{yy}} + \frac{2\tilde{\mathsf{S}}_{\mathsf{yy}}\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right) + 4\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}{2\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}$$

$$= \frac{4\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2 + 2\tilde{\mathsf{S}}_{\mathsf{yy}}\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right) + 2\tilde{\mathsf{S}}_{\mathsf{yy}}\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{2\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}$$

[15] White et al. (2006) page 679 ⟨(6)⟩, Shin and Hammond (2008) page 294 ⟨(9.69)⟩

[16] Shin and Hammond (2008) page 297 ⟨$H_3(f) = S_{ry}(f)/S_{rx}(f)$ (9.78)⟩, Cobb (1988) page 4 ⟨${}^c\hat{H}(f) = \hat{G}_{ys}(f)/\hat{G}_{xs}(f)$ (1.4)⟩, Goyder (1984) page 440 ⟨$H(i\omega) = S_{qz}/S_{pz}$ (5)⟩, Cobb and Mitchell (1990) page 450 ⟨(1)⟩, Pintelon and Schoukens (2012) page 241 ⟨$\hat{G}(\Omega_k) = \hat{G}_{ry}(\Omega_k)\hat{G}_{ru}^{-1}(\Omega_k)$ (7-49)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \frac{\tilde{S}_{yy}(p\tilde{S}_{yy} - \tilde{S}_{xx}) + 2|\tilde{S}_{xy}|^2 + \tilde{S}_{yy}\sqrt{(p\tilde{S}_{yy} - \tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2}}{\sqrt{(p\tilde{S}_{yy} - \tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2}}$$

✏

Lemma 17.2.

LEM

$$\begin{aligned} \tilde{S}_{yy} - p\tilde{S}_{xx} + \sqrt{(\tilde{S}_{yy} - p\tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2} &\geq 0 \\ p\tilde{S}_{yy} - \tilde{S}_{xx} + \sqrt{(p\tilde{S}_{yy} - \tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2} &\geq 0 \end{aligned}$$

✎PROOF:

$$\begin{aligned} & \tilde{S}_{yy} - p\tilde{S}_{xx} + \sqrt{(\tilde{S}_{yy} - p\tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2} && \geq 0 \\ \Longleftrightarrow\ & \sqrt{(p\tilde{S}_{xx} - \tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2} && \geq p\tilde{S}_{xx} - \tilde{S}_{yy} \\ \Longleftrightarrow\ & (p\tilde{S}_{xx} - \tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2 && \geq (p\tilde{S}_{xx} - \tilde{S}_{yy})^2 \\ \Longleftrightarrow\ & 4p|\tilde{S}_{xy}|^2 && \geq 0 \\ \Longleftrightarrow\ & |\tilde{S}_{xy}| && \geq 0 \end{aligned}$$

$$\begin{aligned} & p\tilde{S}_{yy} - \tilde{S}_{xx} + \sqrt{(p\tilde{S}_{yy} - \tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2} && \geq 0 \\ \Longleftrightarrow\ & \sqrt{(\tilde{S}_{xx} - p\tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2} && \geq \tilde{S}_{xx} - p\tilde{S}_{yy} \\ \Longleftrightarrow\ & (\tilde{S}_{xx} - p\tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2 && \geq (\tilde{S}_{xx} - p\tilde{S}_{yy})^2 \\ \Longleftrightarrow\ & 4p|\tilde{S}_{xy}|^2 && \geq 0 \\ \Longleftrightarrow\ & |\tilde{S}_{xy}| && \geq 0 \end{aligned}$$

✏

$\hat{H}_s$ as a function of $p \triangleq s^2$

$\frac{d}{dp}\hat{H}_s$ where $p \triangleq s^2$

Figure 17.2: $\hat{H}_s$ with $\tilde{S}_{xx} = \tilde{S}_{yy} = 1$ and $\tilde{S}_{xy} = \frac{1}{2}$

Theorem 17.1. *Let $\hat{H}_s$ be defined as in Definition 17.8 (page 117).*

THM

$$\begin{aligned} \{s_1 < s_2\} \quad &\Longrightarrow \quad |\hat{H}_s(\omega; s_2)| \leq \hat{H}_s(\omega; s_1) \qquad (\hat{H}_s \textit{ is } \textbf{\textit{monotonically decreasing}} \textit{ in } s) \\ |\hat{H}_1(\omega)| \quad &\leq |\hat{H}_s(\omega; s)| \leq \quad |\hat{H}_2(\omega)| \\ \hat{H}_s(\omega; s)\big|_{s=0} \quad &= \quad \hat{H}_2(\omega) \\ \hat{H}_s(\omega; s)\big|_{s=1} \quad &= \quad \hat{H}_v(\omega) \\ \lim_{s\to\infty} \hat{H}_s(\omega; s) \quad &= \quad \hat{H}_1(\omega) \end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

✎PROOF: I. Proofs for equalities:

$$\hat{\mathsf{H}}_{\mathsf{s}}(\omega;s)\big|_{s=0} \triangleq \left.\frac{\tilde{\mathsf{S}}_{\mathsf{yy}} - s^2\tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left[\tilde{\mathsf{S}}_{\mathsf{yy}} - s^2\tilde{\mathsf{S}}_{\mathsf{xx}}\right]^2 + 4s^2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{2\tilde{\mathsf{S}}_{\mathsf{xy}}}\right|_{s=0} \quad \text{by def. of } \hat{\mathsf{H}}_{\mathsf{s}} \quad \text{(Definition 17.8 page 117)}$$

$$= \frac{\tilde{\mathsf{S}}_{\mathsf{yy}} - 0 + \sqrt{\left[\tilde{\mathsf{S}}_{\mathsf{yy}} - 0\right]^2 + 0}}{2\tilde{\mathsf{S}}_{\mathsf{xy}}} = \frac{\tilde{\mathsf{S}}_{\mathsf{yy}}}{\tilde{\mathsf{S}}_{\mathsf{xy}}}$$

$$\triangleq \hat{\mathsf{H}}_2 \quad \text{by def. of } \hat{\mathsf{H}}_2 \quad \text{(Definition 17.3 page 114)}$$

$$\hat{\mathsf{H}}_{\mathsf{s}}(\omega;s)\big|_{s=1} \triangleq \left.\frac{\tilde{\mathsf{S}}_{\mathsf{yy}} - s^2\tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left[\tilde{\mathsf{S}}_{\mathsf{yy}} - s^2\tilde{\mathsf{S}}_{\mathsf{xx}}\right]^2 + 4s^2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{2\tilde{\mathsf{S}}_{\mathsf{xy}}}\right|_{s=1} \quad \text{by def. of } \hat{\mathsf{H}}_{\mathsf{s}} \quad \text{(Definition 17.8 page 117)}$$

$$= \frac{\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left[\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right]^2 + 4\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{2\tilde{\mathsf{S}}_{\mathsf{xy}}}$$

$$\triangleq \hat{\mathsf{H}}_{\mathsf{v}} \quad \text{by def. of } \hat{\mathsf{H}}_{\mathsf{v}} \quad \text{(Definition 17.9 page 118)}$$

$$\lim_{s\to\infty} \hat{\mathsf{H}}_{\mathsf{s}}(\omega;s) \triangleq \lim_{s\to\infty} \frac{\tilde{\mathsf{S}}_{\mathsf{yy}} - s^2\tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left[\tilde{\mathsf{S}}_{\mathsf{yy}} - s^2\tilde{\mathsf{S}}_{\mathsf{xx}}\right]^2 + 4s^2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{2\tilde{\mathsf{S}}_{\mathsf{xy}}} \quad \text{by def. of } \hat{\mathsf{H}}_{\mathsf{s}} \quad \text{(Definition 17.8 page 117)}$$

$$\triangleq \lim_{p\to 0} \frac{\tilde{\mathsf{S}}_{\mathsf{yy}} - {}^1\!/_p\tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left[\tilde{\mathsf{S}}_{\mathsf{yy}} - {}^1\!/_p\tilde{\mathsf{S}}_{\mathsf{xx}}\right]^2 + 4{}^1\!/_p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{2\tilde{\mathsf{S}}_{\mathsf{xy}}} \quad \text{where } p \triangleq \frac{1}{s^2}$$

$$= \lim_{p\to 0} \frac{p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left[p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right]^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{2p\tilde{\mathsf{S}}_{\mathsf{xy}}} \quad \text{by mult. by } 1 = \frac{p}{p}$$

$$= \lim_{p\to 0} \frac{\frac{\mathrm{d}}{\mathrm{d}p}\left[p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}} + \sqrt{\left[p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right]^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}\right]}{\frac{\mathrm{d}}{\mathrm{d}p}\left[2p\tilde{\mathsf{S}}_{\mathsf{xy}}\right]} \quad \text{by } \textit{l'Hôpital's rule}$$

$$= \lim_{p\to 0} \frac{\dfrac{\tilde{\mathsf{S}}_{\mathsf{yy}}\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right) + 2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2 + \tilde{\mathsf{S}}_{\mathsf{yy}}\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}{\sqrt{\left(p\tilde{\mathsf{S}}_{\mathsf{yy}} - \tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2 + 4p\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}}}{2\tilde{\mathsf{S}}_{\mathsf{xy}}} \quad \text{by Lemma 17.1 page 118}$$

$$= \frac{\tilde{\mathsf{S}}_{\mathsf{yy}}\left(-\tilde{\mathsf{S}}_{\mathsf{xx}}\right) + 2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2 + \tilde{\mathsf{S}}_{\mathsf{yy}}\sqrt{\left(-\tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2}}{2\tilde{\mathsf{S}}_{\mathsf{xy}}\sqrt{\left(-\tilde{\mathsf{S}}_{\mathsf{xx}}\right)^2}}$$

$$= \frac{2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2}{2\tilde{\mathsf{S}}_{\mathsf{xx}}\tilde{\mathsf{S}}_{\mathsf{xy}}} = \frac{\tilde{\mathsf{S}}^*_{\mathsf{xy}}}{\tilde{\mathsf{S}}_{\mathsf{xx}}} \triangleq \hat{\mathsf{H}}_1 \quad \text{by def. of } \hat{\mathsf{H}}_1 \quad \text{(Definition 17.2 page 113)}$$

II. Proof for monoticity:

1. Let $p \triangleq s^2$
2. lemma:

$$\boxed{\left[2\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2 + \tilde{\mathsf{S}}_{\mathsf{xx}}\left(p\tilde{\mathsf{S}}_{\mathsf{xx}} - \tilde{\mathsf{S}}_{\mathsf{yy}}\right)\right]^2}$$
$$= 4\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^4 + 4\tilde{\mathsf{S}}_{\mathsf{xx}}\left|\tilde{\mathsf{S}}_{\mathsf{xy}}\right|^2\left(p\tilde{\mathsf{S}}_{\mathsf{xx}} - \tilde{\mathsf{S}}_{\mathsf{yy}}\right) + \tilde{\mathsf{S}}^2_{\mathsf{xx}}\left(p\tilde{\mathsf{S}}_{\mathsf{xx}} - \tilde{\mathsf{S}}_{\mathsf{yy}}\right)^2$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$\boxed{\le} 4|\tilde{S}_{xy}|^2\tilde{S}_{xx}\tilde{S}_{yy} + 4p\tilde{S}_{xx}|\tilde{S}_{xy}|^2\tilde{S}_{xx} - 4\tilde{S}_{xx}|\tilde{S}_{xy}|^2\tilde{S}_{yy} + \tilde{S}_{xx}^2(p\tilde{S}_{xx} - \tilde{S}_{yy})^2 \quad \left(\begin{array}{l}\text{by } \textit{Cauchy Schwartz inequality} \\ \text{(Theorem K.2 page 272)}\end{array}\right)$$

$$= \cancelto{0}{4\tilde{S}_{xx}\tilde{S}_{yy}|\tilde{S}_{xy}|^2} + 4p\tilde{S}_{xx}|\tilde{S}_{xy}|^2\tilde{S}_{xx} - \cancelto{0}{4\tilde{S}_{xx}\tilde{S}_{yy}|\tilde{S}_{xy}|^2} + \tilde{S}_{xx}^2(p\tilde{S}_{xx} - \tilde{S}_{yy})^2$$

$$= \tilde{S}_{xx}^2\left[(p\tilde{S}_{xx} - \tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2\right]$$

$$= \boxed{\left[\tilde{S}_{xx}\sqrt{(p\tilde{S}_{xx} - \tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2}\right]^2}$$

3. lemma: $2|\tilde{S}_{xy}|^2 + \tilde{S}_{xx}(p\tilde{S}_{xx} - \tilde{S}_{yy}) \le \tilde{S}_{xx}\sqrt{(p\tilde{S}_{xx} - \tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2}$. Proof:

$$2|\tilde{S}_{xy}|^2 + \tilde{S}_{xx}(p\tilde{S}_{xx} - \tilde{S}_{yy}) \le \tilde{S}_{xx}\sqrt{(p\tilde{S}_{xx} - \tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2}$$

$$\iff \left[2|\tilde{S}_{xy}|^2 + \tilde{S}_{xx}(p\tilde{S}_{xx} - \tilde{S}_{yy})\right]^2 \le \left[\tilde{S}_{xx}\sqrt{(p\tilde{S}_{xx} - \tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2}\right]^2 \quad \left(\begin{array}{l}\text{because } f(x) \triangleq x^2 \text{ is} \\ \textit{strictly monotonic increasing}\end{array}\right)$$

The previous inequality is true by (2) lemma, so (3) lemma also true.

4. Proof that $\frac{d}{dp}|\hat{H}_s| \le 0$:

$$\frac{d}{dp}|\hat{H}_s| \triangleq \frac{d}{dp}\left|\frac{\tilde{S}_{yy} - s^2\tilde{S}_{xx} + \sqrt{(\tilde{S}_{yy} - s^2\tilde{S}_{xx})^2 + 4s^2|\tilde{S}_{xy}|^2}}{2\tilde{S}_{xy}}\right| \quad \text{by def. of } \hat{H}_s \text{ (Definition 17.8 page 117)}$$

$$\triangleq \frac{d}{dp}\left[\frac{\tilde{S}_{yy} - p\tilde{S}_{xx} + \sqrt{(\tilde{S}_{yy} - p\tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2}}{2\tilde{S}_{xy}}\right] \quad \text{by definition of } p \text{ (item (1) page 120)}$$

$$= \frac{d}{dp}\left(\frac{\left|\tilde{S}_{yy} - p\tilde{S}_{xx} + \sqrt{(\tilde{S}_{yy} - p\tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2}\right|}{|2\tilde{S}_{xy}|}\right)$$

$$= \frac{d}{dp}\left(\frac{\tilde{S}_{yy} - p\tilde{S}_{xx} + \sqrt{(\tilde{S}_{yy} - p\tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2}}{2|\tilde{S}_{xy}|}\right) \quad \text{by Lemma 17.2 page 119}$$

$$= \frac{\dfrac{\tilde{S}_{xx}(p\tilde{S}_{xx} - \tilde{S}_{yy}) + 2|\tilde{S}_{xy}|^2 - \tilde{S}_{xx}\sqrt{(p\tilde{S}_{xx} - \tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2}}{\sqrt{(\tilde{S}_{yy} - p\tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2}}}{2|\tilde{S}_{xy}|} \quad \text{by Lemma 17.1 page 118}$$

$$= \frac{\tilde{S}_{xx}(p\tilde{S}_{xx} - \tilde{S}_{yy}) + 2|\tilde{S}_{xy}|^2 - \tilde{S}_{xx}\sqrt{(p\tilde{S}_{xx} - \tilde{S}_{yy})^2 + 4p|\tilde{S}_{xy}|^2}}{2|\tilde{S}_{xy}|\sqrt{(\tilde{S}_{yy} - p\tilde{S}_{xx})^2 + 4p|\tilde{S}_{xy}|^2}}$$

$$\le 0 \quad \text{by (3) lemma}$$

✏

Theorem 17.2. *Let* **S** *be a* SYSTEM *with input* $x(n)$ *and output* $y(n)$.

THM $$|\hat{H}_1(\omega)| \le |\hat{H}_{hm}(\omega)| \le |\hat{H}_{gm}(\omega)| \le |\hat{H}_{am}(\omega)| \le |\hat{H}_2(\omega)|$$

✎PROOF:

1. lemma: $\hat{H}_1(\omega) \le \hat{H}_2(\omega)$. Proof:

$$
\begin{aligned}
\left|\hat{H}_1\right| &\triangleq \left|\frac{\tilde{S}_{yx}(\omega)}{\tilde{S}_{xx}(\omega)}\right| && \text{by definition of } \hat{H}_1 && \text{(Definition 17.2 page 113)}\\
&= \left|\frac{\langle y \mid x\rangle}{\|x\|^2}\right| = \frac{|\langle y \mid x\rangle|}{\|x\|^2} \\
&\le \frac{|\langle y \mid x\rangle|}{\|x\|^2}\left|\frac{\|x\|\,\|y\|}{\langle y \mid x\rangle}\right|^2 && \text{by } \textit{Cauchy Schwartz inequality} && \text{Theorem K.2 page 272}\\
&= \frac{\|y\|^2}{|\langle y \mid x\rangle|} = \left|\frac{\|y\|^2}{\langle x \mid y\rangle}\right| = \left|\frac{\tilde{S}_{yy}(\omega)}{\tilde{S}_{xy}(\omega)}\right| \\
&= \left|\hat{H}_2\right| && \text{by definition of } \hat{H}_2 && \text{(Definition 17.3 page 114)}
\end{aligned}
$$

2. remainder of the proof:

$$
\begin{aligned}
\left|\hat{H}_1(\omega)\right| &= \min\left\{\hat{H}_1(\omega), \hat{H}_2(\omega)\right\} && \text{by (1) lemma} && \\
&\le \left|\hat{H}_{hm}(\omega)\right| && \text{by Corollary N.1 page 311} && \text{with } \lambda_1 = \lambda_2 = ½\\
&\le \left|\hat{H}_{gm}(\omega)\right| && \text{by Corollary N.1 page 311} && \text{with } \lambda_1 = \lambda_2 = ½\\
&\le \left|\hat{H}_{am}(\omega)\right| && \text{by Corollary N.1 page 311} && \text{with } \lambda_1 = \lambda_2 = ½\\
&\le \max\left\{\hat{H}_1(\omega), \hat{H}_2(\omega)\right\} && \text{by Corollary N.1 page 311} && \text{with } \lambda_1 = \lambda_2 = ½\\
&= \left|\hat{H}_2(\omega)\right| && \text{by (1) lemma} &&
\end{aligned}
$$

✏

Theorem 17.2 (page 121) compared the magnitudes of several transfer function estimates and demonstrated a simple *linear* relationship. What about phase? The phase of those estimates is even simpler than the magnitude, as demonstrated next.

Proposition 17.5 (Estimator phase)**.** *Let $z \triangleq |z|e^{i\phi}$ be a* COMPLEX *number in the set of complex numbers $\mathbb{C}$. Let $\angle z \triangleq \phi$ be the* PHASE *of z.*

PRP

$$
\begin{aligned}
\angle\hat{H}_1(\omega) &= \angle\hat{H}_{hm}(\omega) = \angle\hat{H}_{gm}(\omega) = \angle\hat{H}_{am}(\omega) = \angle\hat{H}_2(\omega) = \angle\hat{H}_s(\omega) = \angle\hat{H}_v(\omega) = \angle\hat{H}_\kappa(\omega)\\
&= \angle C_{xy}(\omega)\\
&= -\angle\tilde{S}_{xy}(\omega)
\end{aligned}
$$

✎PROOF:

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$\angle\hat{H}_1 \triangleq \angle\frac{\tilde{S}_{yx}}{\tilde{S}_{xx}} \qquad = \qquad = \angle\tilde{S}^*_{xy} = -\angle\tilde{S}_{xy}$$

$\uparrow$ (Definition 17.2 page 113)

$$\angle\hat{H}_{hm} \triangleq \angle\frac{2\tilde{S}_{yy}\tilde{S}^*_{xy}}{\tilde{S}_{xx}\tilde{S}_{yy} + |\tilde{S}_{xy}|^2} \qquad = \qquad = \angle\tilde{S}^*_{xy} = -\angle\tilde{S}_{xy}$$

$\uparrow$ (Definition 17.6 page 116)

$$\angle\hat{H}_{gm} \triangleq \angle\frac{\tilde{S}^*_{xy}}{|\tilde{S}_{xy}|}\sqrt{\frac{\tilde{S}_{yy}}{\tilde{S}_{xx}}} \qquad = \qquad = \angle\tilde{S}^*_{xy} = -\angle\tilde{S}_{xy}$$

$\uparrow$ (Definition 17.5 page 115)

$$\angle\hat{H}_{am} \triangleq \angle\frac{|\tilde{S}_{xy}|^2 + \tilde{S}_{xx}\tilde{S}_{yy}}{2\tilde{S}_{xx}\tilde{S}_{xy}} \qquad = \angle\frac{1}{\tilde{S}_{xy}} = \angle\tilde{S}^*_{xy} = -\angle\tilde{S}_{xy}$$

$\uparrow$ (Definition 17.4 page 114)

$$\angle\hat{H}_2 \triangleq \angle\frac{\tilde{S}_{yy}}{\tilde{S}_{xy}} \qquad = \angle\frac{1}{\tilde{S}_{xy}} = \angle\tilde{S}^*_{xy} = -\angle\tilde{S}_{xy}$$

$\uparrow$ (Definition 17.3 page 114)

$$\angle\hat{H}_s \triangleq \angle\frac{\tilde{S}_{yy} - s^2\tilde{S}_{xx} + \sqrt{[\tilde{S}_{yy} - s^2\tilde{S}_{xx}]^2 + 4s^2|\tilde{S}_{xy}|^2}}{2\tilde{S}_{xy}} \qquad = \angle\frac{1}{\tilde{S}_{xy}} = \angle\tilde{S}^*_{xy} = -\angle\tilde{S}_{xy}$$

$\uparrow$ (Definition 17.8 page 117)

$$\angle\hat{H}_v \triangleq \angle\frac{\tilde{S}_{yy} - \tilde{S}_{xx} + \sqrt{[\tilde{S}_{yy} - \tilde{S}_{xx}]^2 + 4|\tilde{S}_{xy}|^2}}{2\tilde{S}_{xy}} \qquad = \angle\frac{1}{\tilde{S}_{xy}} = \angle\tilde{S}^*_{xy} = -\angle\tilde{S}_{xy}$$

$\uparrow$ (Definition 17.9 page 118)

$$\angle\hat{H}_\kappa \triangleq \angle\frac{\tilde{S}_{yy} - \kappa\tilde{S}_{xx} + \sqrt{[\tilde{S}_{yy} - \kappa\tilde{S}_{xx}]^2 + 4\kappa|\tilde{S}_{xy}|^2}}{2\tilde{S}_{xy}} \qquad = \angle\frac{1}{\tilde{S}_{xy}} = \angle\tilde{S}^*_{xy} = -\angle\tilde{S}_{xy}$$

$\uparrow$ (Definition 17.7 page 117)

$$\angle C_{xy} \triangleq \angle\frac{\tilde{S}^*_{xy}}{\sqrt{\tilde{S}_{xx}\tilde{S}_{yy}}} \qquad = \qquad = \angle\tilde{S}^*_{xy} = -\angle\tilde{S}_{xy}$$

$\uparrow$ (Definition 17.12 page 134)

17.5 Alternate forms

Any standard kit of algebraic tricks should arguably always include the ability to swap the location of a square root between numerator and denominator. If you are of this persuasion, after traveling from the definition of $\hat{H}_s$ on page 117, you won't be dissappointed when arriving at the next proposition (Proposition 17.6 page 123). But it has more use than just allowing you to entertain friends at social occasions. It also makes it very easy to see (using only algebra) what previously employed *l'Hôpital's rule* (using calculus) in the proof of Theorem 17.1—that $\lim_{s\to\infty}\hat{H}_s = \hat{H}_1$.

Proposition 17.6. [17] *Let $\hat{H}_\kappa(\omega;\kappa)$ be defined as in Definition 17.7 (page 117).*

[17] Shin and Hammond (2008) page 293 ⟨(9.67)⟩, Leclere et al. (2014) ⟨(10) $\kappa(f) = 1/s^2$ and x and y swapped⟩

PRP

$$\hat{H}_\kappa(\omega;s) = \frac{2\kappa(\omega)\tilde{S}_{yx}(\omega)}{\kappa(\omega)\tilde{S}_{xx}(\omega)-\tilde{S}_{yy}(\omega)+\sqrt{\left[\kappa(\omega)\tilde{S}_{xx}(\omega)-\tilde{S}_{yy}(\omega)\right]^2+4\kappa(\omega)\left|\tilde{S}_{xy}(\omega)\right|^2}}$$

$$= \frac{2\tilde{S}^*_{xy}}{\tilde{S}_{xx}-\frac{1}{\kappa(\omega)}\tilde{S}_{yy}+\sqrt{\left[\tilde{S}_{xx}-\frac{1}{\kappa(\omega)}\tilde{S}_{yy}\right]^2+\frac{4}{\kappa(\omega)}\left|\tilde{S}_{xy}\right|^2}}$$

✎PROOF: We can transform $\hat{H}_\kappa$ from that found in Definition 17.8 (page 117) to the forms in this proposition by the technique of "*rationalizing the denominator*"[18]

$$\hat{H}_\kappa \triangleq \frac{\tilde{S}_{yy}-\kappa\tilde{S}_{xx}+\sqrt{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}\right]^2+4\kappa\left|\tilde{S}_{xy}\right|^2}}{2\tilde{S}_{xy}} \qquad \text{by definition of } \hat{H}_\kappa \text{ (Definition 17.8 page 117)}$$

$$= \frac{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}+\sqrt{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}\right]^2+4\kappa\left|\tilde{S}_{xy}\right|^2}\right]\overbrace{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}-\sqrt{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}\right]^2+4\kappa\left|\tilde{S}_{xy}\right|^2}\right]}^{\text{"rationalizing factor"}}}{2\tilde{S}_{xy}\underbrace{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}-\sqrt{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}\right]^2+4\kappa\left|\tilde{S}_{xy}\right|^2}\right]}_{\text{"rationalizing factor"}}}$$

$$= \frac{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}\right]^2-\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}\right]^2-4\kappa\left|\tilde{S}_{xy}\right|^2}{2\tilde{S}_{xy}\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}-\sqrt{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}\right]^2+4\kappa\left|\tilde{S}_{xy}\right|^2}\right]} = \frac{-4\kappa\left|\tilde{S}_{xy}\right|^2}{2\tilde{S}_{xy}\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}-\sqrt{\left[\tilde{S}_{yy}-\kappa\tilde{S}_{xx}\right]^2+4\kappa\left|\tilde{S}_{xy}\right|^2}\right]}$$

$$= \boxed{\frac{2\kappa\tilde{S}^*_{xy}}{\kappa\tilde{S}_{xx}-\tilde{S}_{yy}+\sqrt{\left[\kappa\tilde{S}_{xx}-\tilde{S}_{yy}\right]^2+4\kappa\left|\tilde{S}_{xy}\right|^2}}} = \frac{2\frac{\kappa}{\kappa}\tilde{S}^*_{xy}}{\frac{\kappa}{\kappa}\tilde{S}_{xx}-\frac{1}{\kappa}\tilde{S}_{yy}+\sqrt{\frac{1}{s^4}\left[\kappa\tilde{S}_{xx}-\tilde{S}_{yy}\right]^2+\frac{4\kappa}{s^4}\left|\tilde{S}_{xy}\right|^2}}$$

$$= \frac{2\tilde{S}^*_{xy}}{\tilde{S}_{xx}-\frac{1}{\kappa}\tilde{S}_{yy}+\sqrt{\left[\frac{\kappa}{\kappa}\tilde{S}_{xx}-\frac{1}{\kappa}\tilde{S}_{yy}\right]^2+\frac{4}{\kappa}\left|\tilde{S}_{xy}\right|^2}} = \boxed{\frac{2\tilde{S}^*_{xy}}{\tilde{S}_{xx}-\frac{1}{\kappa}\tilde{S}_{yy}+\sqrt{\left[\tilde{S}_{xx}-\frac{1}{\kappa}\tilde{S}_{yy}\right]^2+\frac{4}{\kappa}\left|\tilde{S}_{xy}\right|^2}}}$$

Integrity check for $s=0$ and $s\to\infty$ cases: Let $p \triangleq \kappa$.

$$\lim_{p\to\infty}\hat{H}_\kappa = \lim_{p\to\infty}\frac{2\tilde{S}_{yx}}{\tilde{S}_{xx}-\frac{1}{p}\tilde{S}_{yy}+\sqrt{\left[\tilde{S}_{xx}-\frac{1}{p}\tilde{S}_{yy}\right]^2+\frac{4}{p}\left|\tilde{S}_{xy}\right|^2}} = \frac{2\tilde{S}_{yx}}{\tilde{S}_{xx}+\sqrt{\left[\tilde{S}_{xx}\right]^2}}$$

$$= \frac{\tilde{S}_{yx}}{\tilde{S}_{xx}} \triangleq \hat{H}_1 \qquad \text{by def. of } \hat{H}_1 \text{ (Definition 17.2 page 113)}$$

$$\lim_{p\to 0}\hat{H}_\kappa = \lim_{p\to 0}\frac{2p\tilde{S}_{yx}}{p\tilde{S}_{xx}-\tilde{S}_{yy}+\sqrt{\left[p\tilde{S}_{xx}-\tilde{S}_{yy}\right]^2+4p\left|\tilde{S}_{xy}\right|^2}}$$

$$= \lim_{p\to 0}\frac{\frac{d}{dp}\left(2p\tilde{S}_{yx}\right)}{\frac{d}{dp}\left(p\tilde{S}_{xx}-\tilde{S}_{yy}+\sqrt{\left[p\tilde{S}_{xx}-\tilde{S}_{yy}\right]^2+4p\left|\tilde{S}_{xy}\right|^2}\right)} \qquad \text{by } \textit{l'Hôpital's rule}$$

[18] Slaught and Lennes (1915), page 274 ⟨"197. Rationalizing the Denominator."⟩ https://archive.org/details/elementaryalgebr00slaurich/page/274 Note that the operation in the proof of Proposition 17.6 is being performed essentially in reverse…or rather "rationalizing the numerator".

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \lim_{p\to 0} \frac{2\tilde{S}_{yx}}{\dfrac{\tilde{S}_{xx}\left(p\tilde{S}_{xx}-\tilde{S}_{yy}\right)+2\left|\tilde{S}_{xy}\right|^2+\tilde{S}_{xx}\sqrt{\left(p\tilde{S}_{xx}-\tilde{S}_{yy}\right)^2+4p\left|\tilde{S}_{xy}\right|^2}}{\sqrt{\left(p\tilde{S}_{xx}-\tilde{S}_{yy}\right)^2+4p\left|\tilde{S}_{xy}\right|^2}}} \qquad \text{by Lemma 17.1 page 118}$$

$$= \frac{2\tilde{S}_{yy}\tilde{S}_{yx}}{-\tilde{S}_{xx}\tilde{S}_{yy}+2\left|\tilde{S}_{xy}\right|^2+\tilde{S}_{xx}\tilde{S}_{yy}} = \frac{2\tilde{S}_{yy}\tilde{S}_{yx}}{2\left|\tilde{S}_{xy}\right|^2} = \frac{\tilde{S}_{yy}}{\tilde{S}_{xy}}$$

$$\triangleq \hat{H}_2 \qquad \text{by def. of } \hat{H}_2 \text{ (Definition 17.3 page 114)}$$

✍

17.6 Least squares estimates of non-linear systems

> ❝*The legendary Hungarian mathematician John von Neumann once referred to the the theory of nonequilibrium systems as the "theory of non-elephants," ...Nevertheless, such a theory of non-elephants will be attempted here.*❞
>
> Per Bak, in *"how nature works..."* [19]

(A) *Least squares estimation* for *communications additive noise model* (Corollary 17.1 page 128)

(B) *Least squares estimations* for *measurement additive noise model* (Corollary 17.2 page 129)

Figure 17.3: Least Square estimation (Theorem 17.3 page 126)

Let **S** be the *system* illustrated to the right. **If** there is no measurement noise on the input and output and **if H** is *linear time invariant*, then $\tilde{H} = \tilde{S}_{xy}/\tilde{S}_{xx}$ (Corollary 6.1 page 38). But what if there is output measurement noise? And what if **H** is not *LTI*? What is the best least-squares estimate of $\tilde{H}$? The answer depends on how you define "the best".

The defintion of "best" or "optimal" is given by a *cost function* $C(\hat{H})$. There are several possible cost functions. Definition 17.11 provides some. Theorem 17.3 then demonstrate optimal solutions with respect to these definitions.

Definition 17.11. *Let* **S** *be a system defined as in Figure 17.3 (page 125) (A) or (B). Define the following* COST FUNCTIONS *for spectral* LEAST-SQUARES *estimates:*

DEF

$$C_{rq}(\hat{G}) \triangleq \tilde{F}\,\|r(n)-q(n)\|^2 \triangleq \tilde{F}\,\langle r(n)-q(n)\,|\,r(0)-q(0)\rangle \triangleq \tilde{F}E\left(\left[r(n)-q(n)\right]\left[r(0)-q(0)\right]^*\right)$$

$$C_{sy}(\hat{H}) \triangleq \tilde{F}\,\|s(n)-y(n)\|^2 \triangleq \tilde{F}\,\langle s(n)-y(n)\,|\,s(0)-y(0)\rangle \triangleq \tilde{F}E\left(\left[s(n)-y(n)\right]\left[s(0)-y(0)\right]^*\right)$$

[19] Bak (2013) page 29 ⟨§ Systems in Balance Are Not Complex⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

Lemma 17.3. *Let* $C_{rq}(\hat{G})$ *and* $C_{sy}(\hat{H})$ *be defined as in Definition 17.11 (page 125).*

LEM

$$C_{rq}(\hat{G}) = \tilde{S}_{pp}(\omega)\left|\hat{G}(\omega)\right|^2 - \tilde{S}_{py}(\omega)\hat{G}(\omega) - \tilde{S}^*_{py}(\omega)\hat{G}^*(\omega) + \tilde{S}_{qq}(\omega)$$
$$C_{sy}(\hat{H}) = \tilde{S}_{xx}(\omega)\left|\hat{H}(\omega)\right|^2 - \tilde{S}_{xy}(\omega)\hat{H}(\omega) - \tilde{S}^*_{xy}(\omega)\hat{H}^*(\omega) + \tilde{S}_{yy}(\omega)$$

✎PROOF:

$\boxed{C_{rq}(\hat{G})}$

$$\begin{aligned}
&\triangleq \check{\mathbf{F}}\mathsf{E}\Big(\big[r(n)-q(n)\big]\big[r(0)-q(0)\big]^*\Big) && \text{by definition of } C_{rq} && \text{(Definition 17.11 page 125)}\\
&= \tilde{\mathbf{F}}\big[\mathsf{E}\big[r(n)r^*(0)\big] - \mathsf{E}\big[r(n)q^*(0)\big] - \mathsf{E}\big[q(n)r^*(0)\big] + \mathsf{E}\big[q(n)q^*(0)\big]\big] && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)}\\
&\triangleq \tilde{\mathbf{F}}\big[R_{rr}(m) - R_{rq}(m) - R_{qr}(m) + R_{qq}(m)\big] && \text{by definition of } R_{xy} && \text{(Definition 2.4 page 12)}\\
&\triangleq \big[\tilde{S}_{rr}(\omega) - \tilde{S}_{rq}(\omega) - \tilde{S}_{qr}(\omega) + \tilde{S}_{qq}(\omega)\big] && \text{by definition of } \tilde{S}_{xy} && \text{(Definition 7.3 page 46)}\\
&= \boxed{\tilde{S}_{pp}(\omega)\left|\hat{G}(\omega)\right|^2 - \tilde{S}_{py}(\omega)\hat{G}(\omega) - \tilde{S}^*_{py}(\omega)\hat{G}^*(\omega) + \tilde{S}_{qq}(\omega)} && \text{by (A)–(D) and} && \text{Corollary 8.8 page 61}
\end{aligned}$$

$\boxed{C_{sy}(\hat{H})}$

$$\begin{aligned}
&\triangleq \check{\mathbf{F}}\mathsf{E}\Big(\big[s(n)-y(n)\big]\big[s(0)-y(0)\big]^*\Big) && \text{by definition of } C_{sy} && \text{(Definition 17.11 page 125)}\\
&= \tilde{\mathbf{F}}\big[\mathsf{E}\big[s(n)s^*(0)\big] - \mathsf{E}\big[s(n)y^*(0)\big] - \mathsf{E}\big[y(n)s^*(0)\big] + \mathsf{E}\big[y(n)y^*(0)\big]\big] && \text{by } \textit{linearity} \text{ of } \mathsf{E} && \text{(Theorem 1.1 page 4)}\\
&\triangleq \tilde{\mathbf{F}}\big[R_{ss}(m) - R_{sy}(m) - R_{ys}(m) + R_{yy}(m)\big] && \text{by definition of } R_{xy} && \text{(Definition 2.4 page 12)}\\
&\triangleq \big[\tilde{S}_{ss}(\omega) - \tilde{S}_{sy}(\omega) - \tilde{S}_{ys}(\omega) + \tilde{S}_{yy}(\omega)\big] && \text{by definition of } \tilde{S}_{xy} && \text{(Definition 7.3 page 46)}\\
&= \boxed{\tilde{S}_{xx}(\omega)\left|\hat{H}(\omega)\right|^2 - \tilde{S}_{xy}(\omega)\hat{H}(\omega) - \tilde{S}^*_{xy}(\omega)\hat{H}^*(\omega) + \tilde{S}_{yy}(\omega)} && \text{by (A)–(D) and} && \text{Corollary 8.8 (page 61)}
\end{aligned}$$

✏

Theorem 17.3. *Let* **S** *be the* SYSTEM *illustrated in Figure 17.3 page 125 (A) or (B).*

THM

$$\left.\begin{array}{lll}
(A). & x, u, \textit{ and } v \textit{ are} & \text{WSS} \quad \textit{and}\\
(B). & x, u, \textit{ and } v \textit{ are} & \text{UNCORRELATED} \quad \textit{and}\\
(C). & Eu = Ev = 0 & (\text{ZERO-MEAN}) \quad \textit{and}\\
(D). & \hat{G}_{ls} \textit{ and } \hat{H}_{ls} \textit{ are} & \text{LTI}
\end{array}\right\} \Longrightarrow \left\{\begin{array}{lll}
(1). & \arg\min\limits_{\hat{G}} C_{rq}(\hat{G}) & = \dfrac{\tilde{S}_{yp}(\omega)}{\tilde{S}_{pp}(\omega)}\\[2ex]
(2). & \arg\min\limits_{\hat{H}} C_{sy}(\hat{H}) & = \dfrac{\tilde{S}_{yx}(\omega)}{\tilde{S}_{xx}(\omega)}
\end{array}\right.$$

✎PROOF:

1. Define shorthand function names $\hat{G} \triangleq \hat{G}_{ls}$ and $\hat{H} \triangleq \hat{H}_{ls}$.
2. lemma:

$$\begin{aligned}
0 &= \frac{\partial}{\partial \hat{G}_R} C_{rq}(\hat{G})\\
&= \frac{\partial}{\partial \hat{G}_R}\Big(\tilde{S}_{pp}\left|\hat{G}\right|^2 - \hat{G}\tilde{S}_{py} - \hat{G}^*\tilde{S}^*_{py} + \tilde{S}_{qq}\Big) && \text{by Lemma 17.3 page 126}\\
&= \frac{\partial}{\partial \hat{G}_R}\big(\tilde{S}_{pp}\big[\hat{G}_R^2 + \hat{G}_I^2\big] - (\hat{G}_R + i\hat{G}_I)\tilde{S}_{py} - (\hat{G}_R + i\hat{G}_I)^*\tilde{S}^*_{py} + \tilde{S}_{qq}\big)\\
&= 2\hat{G}_R\tilde{S}_{pp} - \tilde{S}_{py} - \tilde{S}^*_{py} + \cancelto{0}{\frac{\partial}{\partial \hat{G}_R}\tilde{S}_{qq}} && \text{because q does not vary with } \hat{G}\\
&= 2\hat{G}_R\tilde{S}_{pp} - 2\mathbf{R}_e\tilde{S}_{py}\\
&= 2\hat{G}_R\tilde{S}_{pp} - 2\mathbf{R}_e\tilde{S}_{yp} && \text{by Corollary 2.2 page 15}
\end{aligned}$$

$$\Longrightarrow \boxed{\hat{G}_R(\omega) = \frac{\mathbf{R}_e\tilde{S}_{yp}(\omega)}{\tilde{S}_{pp}(\omega)}}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

3. lemma:

$$
\begin{aligned}
0 &= \frac{\partial}{\partial \hat{G}_I} C_{rq}(\hat{G}) \\
&= \frac{\partial}{\partial \hat{G}_I}\left(\tilde{S}_{pp}|\hat{G}|^2 - \hat{G}\tilde{S}_{py} - \hat{G}^*\tilde{S}^*_{py} + \tilde{S}_{qq}\right) && \text{by Lemma 17.3 page 126} \\
&= \frac{\partial}{\partial \hat{G}_I}\left[\tilde{S}_{pp}\left[\hat{G}_R^2 + \hat{G}_I^2\right] - (\hat{G}_R + i\hat{G}_I)\tilde{S}_{py} - (\hat{G}_R - i\hat{G}_I)\tilde{S}^*_{py} + \tilde{S}_{qq}\right] \\
&= 2\hat{G}_I\tilde{S}_{pp} - i\tilde{S}_{py} + i\tilde{S}^*_{py} + \cancelto{0}{\frac{\partial}{\partial \hat{G}_I}\tilde{S}_{qq}} && \text{because q does not vary with } \hat{G} \\
&= 2\hat{G}_I\tilde{S}_{pp} - 2i(i\mathbf{I}_m\tilde{S}_{py}) \\
&= 2\hat{G}_I\tilde{S}_{pp} + 2i(i\mathbf{I}_m\tilde{S}_{yp}) && \text{by Corollary 2.2 page 15} \\
&\Longrightarrow \boxed{\hat{G}_I(\omega) = \frac{\mathbf{I}_m\tilde{S}_{yp}(\omega)}{\tilde{S}_{pp}(\omega)}}
\end{aligned}
$$

4. Proof for $\hat{G} \triangleq \hat{G}_{ls}$ expression:

$$
\begin{aligned}
\boxed{\hat{G}(\omega)} &= \hat{G}_R(\omega) + i\hat{G}_I(\omega) \\
&= \frac{\mathbf{R}_e\tilde{S}_{yp}(\omega)}{\tilde{S}_{pp}(\omega)} + \frac{i\mathbf{I}_m\tilde{S}_{yp}(\omega)}{\tilde{S}_{pp}(\omega)} && \text{by (2) lemma and (3) lemma} \\
&= \boxed{\frac{\tilde{S}_{yp}(\omega)}{\tilde{S}_{pp}(\omega)}}
\end{aligned}
$$

5. lemma:

$$
\begin{aligned}
0 &= \frac{\partial}{\partial \hat{H}_R} C_{sy}(\hat{H}) \\
&= \frac{\partial}{\partial \hat{H}_R}\left(\tilde{S}_{xx}|\hat{H}|^2 - \hat{H}\tilde{S}_{xy} - \hat{H}^*\tilde{S}^*_{xy} + \tilde{S}_{yy}\right) && \text{by Lemma 17.3 page 126} \\
&= \frac{\partial}{\partial \hat{H}_R}\left(\tilde{S}_{xx}\left[\hat{H}_R^2 + \hat{H}_I^2\right] - (\hat{H}_R + i\hat{H}_I)\tilde{S}_{xy} - (\hat{H}_R + i\hat{H}_I)^*\tilde{S}^*_{xy} + \tilde{S}_{yy}\right) \\
&= 2\hat{H}_R\tilde{S}_{xx} - \tilde{S}_{xy} - \tilde{S}^*_{xy} + \cancelto{0}{\frac{\partial}{\partial \hat{H}_R}\tilde{S}_{yy}} && \text{because y does not vary with } \hat{H} \\
&= 2\hat{H}_R\tilde{S}_{xx} - 2\mathbf{R}_e\tilde{S}_{xy} \\
&= 2\hat{H}_R\tilde{S}_{xx} - 2\mathbf{R}_e\tilde{S}_{yx} && \text{by Corollary 2.2 page 15} \\
&\Longrightarrow \boxed{\hat{H}_R(\omega) = \frac{\mathbf{R}_e\tilde{S}_{yx}(\omega)}{\tilde{S}_{xx}(\omega)}}
\end{aligned}
$$

6. lemma:

$$
\begin{aligned}
0 &= \frac{\partial}{\partial \hat{H}_I} C_{sy}(\hat{H}) \\
&= \frac{\partial}{\partial \hat{H}_I}\left(\tilde{S}_{xx}|\hat{H}|^2 - \tilde{S}_{xy}\hat{H} - \tilde{S}^*_{xy}\hat{H}^* + \tilde{S}_{yy}\right) && \text{by Lemma 17.3 page 126} \\
&= \frac{\partial}{\partial \hat{H}_I}\left[\tilde{S}_{xx}\left[\hat{H}_R^2 + \hat{H}_I^2\right] - \tilde{S}_{xy}(\hat{H}_R + i\hat{H}_I) - \tilde{S}^*_{xy}(\hat{H}_R - i\hat{H}_I) + \tilde{S}_{yy}\right] \\
&= 2\hat{H}_I\tilde{S}_{xx} - i\tilde{S}_{xy} + i\tilde{S}^*_{xy} + \cancelto{0}{\frac{\partial}{\partial \hat{H}_R}\tilde{S}_{yy}} && \text{because q does not vary with } \hat{H}
\end{aligned}
$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

$$
\begin{aligned}
&= 2\hat{\mathsf{H}}_I\tilde{\mathsf{S}}_{\mathsf{xx}} - 2i(i\mathbf{I}_{\mathsf{m}}\tilde{\mathsf{S}}_{\mathsf{xy}}) \\
&= 2\hat{\mathsf{H}}_I\tilde{\mathsf{S}}_{\mathsf{xx}} + 2i(i\mathbf{I}_{\mathsf{m}}\tilde{\mathsf{S}}_{\mathsf{yx}}) && \text{by Corollary 2.2 page 15} \\
&= 2\hat{\mathsf{H}}_I\tilde{\mathsf{S}}_{\mathsf{xx}} - 2\mathbf{I}_{\mathsf{m}}\tilde{\mathsf{S}}_{\mathsf{yx}} \\
\implies & \boxed{\hat{\mathsf{H}}_I(\omega) = \frac{\mathbf{I}_{\mathsf{m}}\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}}
\end{aligned}
$$

7. Proof for $\hat{\mathsf{H}} \triangleq \hat{\mathsf{H}}_{\mathsf{ls}}$ expression:

$$
\begin{aligned}
\boxed{\hat{\mathsf{H}}(\omega)} &= \hat{\mathsf{H}}_R(\omega) + i\hat{\mathsf{H}}_I(\omega) \\
&= \frac{\mathbf{R}_{\mathsf{e}}\tilde{\mathsf{S}}_{\mathsf{yp}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} + \frac{i\mathbf{I}_{\mathsf{m}}\tilde{\mathsf{S}}_{\mathsf{yp}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} && \text{by (5) lemma and (6) lemma} \\
&= \frac{\mathbf{R}_{\mathsf{e}}\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} + \frac{i\mathbf{I}_{\mathsf{m}}\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} && \text{by Theorem 8.4 page 57} \\
&= \boxed{\frac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}}
\end{aligned}
$$

✏

Using Theorem 17.3 (previous) we can see that the optimal **least-squares** operators $\hat{\mathsf{G}}_{\mathsf{ls}}$ and $\hat{\mathsf{H}}_{\mathsf{ls}}$ for the **non-linear** operator $\mathbf{T}$ in Figure 17.3 (page 125) (A) and (B) are (next two corollaries)

(1).	$\hat{\mathsf{G}}_{\mathsf{ls}}(\omega)$	$= \dfrac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) + \tilde{\mathsf{S}}_{\mathsf{uu}}(\omega)}$	for (A)—*communication system*
(2).	$\hat{\mathsf{G}}_{\mathsf{ls}}(\omega)$	$= \dfrac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) - \tilde{\mathsf{S}}_{\mathsf{uu}}(\omega)}$	for (B)—*measurement system*
(3).	$\hat{\mathsf{H}}_{\mathsf{ls}}(\omega)$	$= \dfrac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}$	for either (A) or (B)

Corollary 17.1. *Let* $\mathbf{S}$ *be the* SYSTEM *illustrated in Figure 17.3 page 125 (A).*

THM

$$
\left\{ \begin{array}{r} \textit{hypotheses of Theorem 17.3} \\ \textit{page 126} \end{array} \right\} \implies \left\{ \begin{array}{lll} (1). & \arg\min\limits_{\hat{\mathsf{G}}_{\mathsf{ls}}} \mathsf{C}_{\mathsf{rq}}(\hat{\mathsf{G}}_{\mathsf{ls}}) & = \dfrac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) + \tilde{\mathsf{S}}_{\mathsf{uu}}(\omega)} \\ (2). & \arg\min\limits_{\hat{\mathsf{H}}_{\mathsf{ls}}} \mathsf{C}_{\mathsf{sy}}(\hat{\mathsf{H}}_{\mathsf{ls}}) & = \dfrac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} \end{array} \right\}
$$

✎PROOF:

$$
\begin{aligned}
\hat{\mathsf{G}}_{\mathsf{ls}} &= \frac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{pp}}(\omega)} && \text{by Theorem 17.3 page 126} \\
&= \frac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) + \tilde{\mathsf{S}}_{\mathsf{uu}}(\omega)} && \text{by Theorem 8.1 page 53} \\
\hat{\mathsf{H}}_{\mathsf{ls}} &= \frac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} && \text{by Theorem 17.3 page 126}
\end{aligned}
$$

✏

COPYRIGHT © 2019 DANIEL J. GREENHOE

(A) **communications LTI additive noise model** (B) **measurement LTI additive noise model**

Figure 17.4: Additive noise systems with **LTI** operator **H**

Corollary 17.2. *Let* **S** *be the* SYSTEM *illustrated in Figure 17.3 page 125 (B).*

THM

$$\left\{ \begin{array}{r} \textit{hypotheses of Theorem 17.3} \\ \textit{page 126} \end{array} \right\} \Longrightarrow \left\{ \begin{array}{lll} (1). & \arg\min\limits_{\hat{G}_{ls}} C_{rq}(\hat{G}_{ls}) & = \dfrac{\tilde{S}_{yx}(\omega)}{\tilde{S}_{xx}(\omega) - \tilde{S}_{uu}(\omega)} \\ (2). & \arg\min\limits_{\hat{H}_{ls}} C_{sy}(\hat{H}_{ls}) & = \dfrac{\tilde{S}_{yx}(\omega)}{\tilde{S}_{xx}(\omega)} \end{array} \right\}$$

✎PROOF:

$$\begin{aligned} \hat{G} &= \frac{\tilde{S}_{yx}(\omega)}{\tilde{S}_{pp}(\omega)} && \text{by Theorem 17.3 page 126} \\ &= \frac{\tilde{S}_{yx}(\omega)}{\tilde{S}_{xx}(\omega) - \tilde{S}_{uu}(\omega)} && \text{by Theorem 8.1 page 53} \\ \hat{H} &= \frac{\tilde{S}_{yx}(\omega)}{\tilde{S}_{xx}(\omega)} && \text{by Theorem 17.3 page 126} \end{aligned}$$

✏

It follows immediately from Corollary 17.1 (page 128) and Corollary 17.2 (page 129) that, in the special case of no input noise ($u(n) = 0$), $\hat{H}_1$ is the optimal least-squares estimate of $\tilde{H}$ (next corollary).

Corollary 17.3. [20] *Let* **S** *be the* SYSTEM *illustrated in Figure 17.3 page 125 (A) or (B).*

COR

$$\left\{ \begin{array}{ll} (1). & \textit{hypotheses of Theorem 17.3} \quad \textit{and} \\ (2). & u(n) = 0 \end{array} \right\} \Longrightarrow \left\{ \hat{G}_{ls}(\omega) = \hat{H}_{ls}(\omega) = \hat{H}_1(\omega) \right\}$$

17.7 Least squares estimates of linear systems

The previous section did assume the estimates $\hat{H}_1$ and $\hat{H}_2$ to be *linear time invariant* (*LTI*), but it did *not* assume that the system transfer function **T** itself to be *LTI*. But making the LTI assumption on **H** yields some interesting and insightful results, such as those in this section.

Theorem 17.4 (Estimating H in communication additive noise system)**.** *Let* **S** *be the* SYSTEM *illustrated in Figure 17.4 page 129 (A).*

[20] Bendat and Piersol (1980) pages 98–100 ⟨5.1.1 Optimal Character of Calculations; note: proof minimizing $\tilde{S}_{vv}$ but yields same result⟩, Bendat and Piersol (1993) pages 106–109 ⟨5.1.1 Optimality of Calculations⟩, Bendat and Piersol (2010) pages 187–190 ⟨6.1.4 Optimum Frequency Response Functions⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

THM

$$\left\{\begin{array}{lll} (A). & \mathbf{H} \textit{ is} & \text{LINEAR TIME INVARIANT } (\text{LTI}) \textit{ and} \\ (B). & \mathsf{x}(n) \textit{ is} & \text{WIDE-SENSE STATIONARY } (\text{WSS}) \textit{ and} \\ (C). & \mathsf{x}(n), \mathsf{u}(n), \textit{and } \mathsf{v}(n) \textit{ are} & \text{UNCORRELATED} \end{array}\right\}$$

$$\Longrightarrow \left\{\begin{array}{lll} (1). & \hat{\mathsf{H}}_1(\omega) = \tilde{\mathsf{H}}(\omega) & \textit{and} \\ (2). & \hat{\mathsf{H}}_2(\omega) = \dfrac{\tilde{\mathsf{S}}_{\mathsf{vv}}(\omega)}{\tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} + \tilde{\mathsf{H}}(\omega)\left[1 + \dfrac{\tilde{\mathsf{S}}_{\mathsf{uu}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}\right] & \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\hat{\mathsf{H}}_1(\omega) &\triangleq \frac{\tilde{\mathsf{S}}_{\mathsf{yx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} && \text{by definition of } \hat{\mathsf{H}}_1 \quad \text{(Definition 17.2 page 113)} \\
&= \frac{\tilde{\mathsf{H}}(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} && \text{by Corollary 8.5 page 59} \\
&= \tilde{\mathsf{H}}(\omega) \\
\hat{\mathsf{H}}_2(\omega) &\triangleq \frac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xy}}(\omega)} && \text{by definition of } \hat{\mathsf{H}}_2 \quad \text{(Definition 17.3 page 114)} \\
&= \frac{\tilde{\mathsf{S}}_{\mathsf{yy}}(\omega)}{\tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} && \text{by Corollary 8.5 page 59} \\
&= \frac{\tilde{\mathsf{S}}_{\mathsf{vv}}(\omega) + \tilde{\mathsf{S}}_{\mathsf{qq}}(\omega)}{\tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} && \text{by Theorem 8.1 page 53} \\
&= \frac{\tilde{\mathsf{S}}_{\mathsf{vv}}(\omega) + \tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{H}}(\omega)\tilde{\mathsf{S}}_{\mathsf{pp}}(\omega)}{\tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} && \text{by Corollary 6.3 page 39} \\
&= \frac{\tilde{\mathsf{S}}_{\mathsf{vv}}(\omega)}{\tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} + \frac{\tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{H}}(\omega)\left[\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega) + \tilde{\mathsf{S}}_{\mathsf{uu}}(\omega)\right]}{\tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} \\
&= \frac{\tilde{\mathsf{S}}_{\mathsf{vv}}(\omega)}{\tilde{\mathsf{H}}^*(\omega)\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)} + \tilde{\mathsf{H}}(\omega)\left[1 + \frac{\tilde{\mathsf{S}}_{\mathsf{uu}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{xx}}(\omega)}\right]
\end{aligned}$$

✏

Theorem 17.5 (Estimating H in measurement additive noise system). [21] *Let* **S** *be the* SYSTEM *illustrated in Figure 17.4 page 129 (B).*

THM

$$\left\{\begin{array}{lll} (A). & \mathbf{H} \textit{ is} & \text{LINEAR TIME INVARIANT } \textit{and} \\ (B). & \mathsf{x}(n) \textit{ is} & \text{WIDE-SENSE STATIONARY } \textit{and} \\ (C). & \mathsf{x}(n), \mathsf{u}(n), \textit{and } \mathsf{v}(n) \textit{ are} & \text{UNCORRELATED} \end{array}\right\}$$

$$\Longrightarrow \left\{\begin{array}{lll} (1). & \hat{\mathsf{H}}_1(\omega) = \tilde{\mathsf{H}}(\omega)\left[\dfrac{1}{1 + \dfrac{\tilde{\mathsf{S}}_{\mathsf{ww}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{pp}}(\omega)}}\right] & (\text{UNDER-ESTIMATED}) \quad \textit{and} \\ (2). & \hat{\mathsf{H}}_2(\omega) = \tilde{\mathsf{H}}(\omega)\left[1 + \dfrac{\tilde{\mathsf{S}}_{\mathsf{vv}}(\omega)}{\tilde{\mathsf{S}}_{\mathsf{qq}}(\omega)}\right] & (\text{OVER-ESTIMATED}) \end{array}\right\}$$

[21] Shin and Hammond (2008) page 294 $\langle H_1(f) = H(f)$ (9.70); $H_2(f) = H(f)\big(1 + S_{n_y n_y}(f)/S_{yy}(f)\big)$ (9.71)$\rangle$, Shin and Hammond (2008) page 294 $\langle H_1(f) = H(f)/\big(1 + S_{n_x n_x}/S_{xx}(f)\big)$ (9.72); $H_2(f) = H(f)$ (9.73)$\rangle$, Mitchell (1982) page 277 $\langle H_1(f) = H_0(f)/\big(1 + G_{nn}/G_{uu}\big)\rangle$ Mitchell (1982) page 278 $\langle H_2(f) = H_0(f)\big(1 + G_{mm}/G_{vv}\big)\rangle$

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$\begin{aligned}
\hat{H}_1(\omega) &\triangleq \frac{\tilde{S}_{yx}(\omega)}{\tilde{S}_{xx}(\omega)} && \text{by definition of } \hat{H}_1 && \text{(Definition 17.2 page 113)}\\
&= \frac{\tilde{S}_{pp}(\omega)\tilde{H}(\omega)}{\tilde{S}_{xx}(\omega)} && \text{by Corollary 8.4 page 59} && \\
&= \frac{\tilde{S}_{pp}(\omega)\tilde{H}(\omega)}{\tilde{S}_{pp}(\omega)+\tilde{S}_{ww}(\omega)} && \text{by hypothesis (A)} && \text{and Corollary 6.3 page 39}\\
&= \tilde{H}(\omega)\left[\frac{1}{1+\frac{\tilde{S}_{ww}(\omega)}{\tilde{S}_{pp}(\omega)}}\right] && && \\
\hat{H}_2(\omega) &\triangleq \frac{\tilde{S}_{yy}(\omega)}{\tilde{S}_{xy}(\omega)} && \text{by definition of } \hat{H}_2 && \text{(Definition 17.3 page 114)}\\
&= \frac{\tilde{S}_{qq}(\omega)+\tilde{S}_{vv}(\omega)}{\tilde{S}_{xy}(\omega)} && \text{by hypothesis (C)} && \text{and Corollary 8.1 page 54}\\
&= \frac{\tilde{S}_{qq}(\omega)+\tilde{S}_{vv}(\omega)}{\tilde{S}_{xq}(\omega)} && \text{by hypothesis (C)} && \text{and Theorem 8.4 page 57}\\
&= \frac{\tilde{S}_{qq}(\omega)+\tilde{S}_{vv}(\omega)}{\tilde{S}_{pq}(\omega)} && \text{by hypothesis (C)} && \text{and Lemma 8.3 page 56}\\
&= \frac{\tilde{S}_{qq}(\omega)+\tilde{S}_{vv}(\omega)}{\tilde{S}_{qq}(\omega)/\tilde{H}(\omega)} && \text{by } \textit{LTI} \text{ hypothesis (A)} && \text{and Corollary 6.3 page 39}\\
&= \tilde{H}(\omega)\left[1+\frac{\tilde{S}_{vv}(\omega)}{\tilde{S}_{qq}(\omega)}\right] && \text{by hypotheses (A) and (B)} && \text{and Corollary 6.3 page 39}
\end{aligned}$$

✏

Corollary 17.4. *Let* **S** *be the* SYSTEM *illustrated in Figure 17.4 (page 129).*

COR

$$\left\{\begin{array}{ll}(A). & \textit{hypotheses of Theorem 17.5 and}\\ (B). & \mathsf{u}(n)=\mathsf{u}(n)=0 \quad (\text{NO INPUT NOISE})\end{array}\right\} \implies \left\{\hat{H}_1(\omega) = \tilde{H}(\omega) \quad (\text{UNBIASED})\right\}$$

$$\left\{\begin{array}{ll}(A). & \textit{hypotheses of Theorem 17.5 and}\\ (B). & \mathsf{v}(n)=0 \quad (\text{NO OUTPUT NOISE})\end{array}\right\} \implies \left\{\hat{H}_2(\omega) = \tilde{H}(\omega) \quad (\text{UNBIASED})\right\}$$

Lemma 17.4. *Let* **S** *be the* SYSTEM *illustrated in Figure 17.4 (page 129).*

LEM

$$\left\{\textit{There exists } \kappa(\omega) \textit{ such that } \tilde{S}_{vv}(\omega)=\kappa(\omega)\tilde{S}_{uu}(\omega)\right\}$$
$$\implies \left\{\tilde{S}_{uu}(\omega) = \frac{\left|\hat{H}(\omega)\right|^2\tilde{S}_{xx}(\omega)-\hat{H}(\omega)\tilde{S}_{xy}(\omega)-\hat{H}^*(\omega)\tilde{S}^*_{xy}(\omega)+\tilde{S}_{yy}(\omega)}{\kappa(\omega)+\left|\hat{H}(\omega)\right|^2}\right\}$$

✎PROOF:

1. Development based on results of previous chapters:

$$\begin{aligned}
\tilde{S}_{vv} &= \tilde{S}_{yy}-\tilde{S}_{qq} && \text{by Corollary 8.1 page 54}\\
&= \tilde{S}_{yy}-\tilde{S}_{pq}\hat{H} && \text{by Corollary 6.3 page 39}\\
&= \tilde{S}_{yy}-\tilde{S}_{xy}\hat{H} && \text{by Theorem 8.4 page 57}\\
\tilde{S}_{uu} &= \tilde{S}_{xx}-\tilde{S}_{pp} && \text{by Corollary 8.1 page 54}
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

$$
\begin{aligned}
&= \tilde{S}_{xx} - \frac{\tilde{S}_{qp}}{\hat{H}} && \text{by Corollary 6.3 page 39}\\
&= \tilde{S}_{xx} - \frac{\tilde{S}_{yx}}{\hat{H}} && \text{by Theorem 8.4 page 57}\\
\tilde{S}_{uu}\left[\left|\hat{H}\right|^2 + \kappa\right] &= \left|\hat{H}\right|^2\tilde{S}_{uu} + \kappa\tilde{S}_{uu}\\
&\triangleq \tilde{S}_{uu}\left|\hat{H}\right|^2 + \tilde{S}_{vv} && \text{by definition of } \kappa(\omega)\\
&= \left|\hat{H}\right|^2\left[\tilde{S}_{xx} - \frac{\tilde{S}_{yx}}{\hat{H}}\right] + \left[\tilde{S}_{yy} - \tilde{S}_{xy}\hat{H}\right]\\
&= \left|\hat{H}\right|^2\tilde{S}_{xx} - \hat{H}^*\tilde{S}_{yx} - \tilde{S}_{xy}\hat{H} + \tilde{S}_{yy}\\
&\Longrightarrow \boxed{\tilde{S}_{uu}(\omega) = \frac{\left|\hat{H}(\omega)\right|^2\tilde{S}_{xx}(\omega) - \hat{H}(\omega)\tilde{S}_{xy}(\omega) - \hat{H}^*(\omega)\tilde{S}^*_{xy}(\omega) + \tilde{S}_{yy}(\omega)}{\kappa(\omega) + \left|\hat{H}(\omega)\right|^2}}
\end{aligned}
$$

2. Development of Wicks and Vold (Wicks and Vold (1986)):

$$
\begin{aligned}
\tilde{Y} - \tilde{V} = \tilde{Q} = \hat{H}\tilde{P} &= \hat{H}(\tilde{X} - \tilde{U}) && \text{by definition of } \hat{H}\\
\hat{H}\tilde{U} - \tilde{V} &= \hat{H}\tilde{X} - \tilde{Y} && \text{by } \textit{left distributive} \text{ prop.} \qquad \text{(Theorem O.4 page 323}\\
E\left(\left[\hat{H}\tilde{U} - \tilde{V}\right]\left[\hat{H}\tilde{U} - \tilde{V}\right]^*\right) &= E\left(\left[\hat{H}\tilde{X} - \tilde{Y}\right]\left[\hat{H}\tilde{X} - \tilde{Y}\right]^*\right)\\
\left|\hat{H}\right|^2\tilde{S}_{uu} - \hat{H}\cancelto{0}{\tilde{S}_{uv}} - \hat{H}^*\cancelto{0}{\tilde{S}_{vu}} + \tilde{S}_{vv} &= \left|\hat{H}\right|^2\tilde{S}_{xx} - \hat{H}\tilde{S}_{xy} - \hat{H}^*\tilde{S}^*_{xy} + \tilde{S}_{yy} && \text{because } u \text{ and } v \text{ are } \textit{uncorrelated}\\
\left|\hat{H}\right|^2\tilde{S}_{uu} + \kappa\tilde{S}_{uu} &= \left|\hat{H}\right|^2\tilde{S}_{xx} - \hat{H}\tilde{S}_{xy} - \hat{H}^*\tilde{S}^*_{xy} + \tilde{S}_{yy} && \text{by hypothesis}
\end{aligned}
$$

✏

Theorem 17.6. [22] *Let* **S** *be the* SYSTEM *illustrated in Figure 17.4 (page 129). Let* $\hat{H}_\kappa(\omega)$ *be the transfer function estimate defined in Definition 17.7 (page 117).*

THM

$$
\left\{\begin{array}{ll}
(1). & \textit{There exists } \kappa(\omega) \textit{ such that}\\
 & \tilde{S}_{vv}(\omega) = \kappa(\omega)\tilde{S}_{uu}(\omega) \qquad \textit{and}\\
(2). & C(\hat{H}_{ls}) = \tilde{S}_{uu}(\omega)
\end{array}\right\}
\Longrightarrow
\left\{\begin{array}{l}
\arg\min_{\hat{H}} C(\hat{H}) = \hat{H}_\kappa(\omega)\\
(\hat{H}_\kappa \textit{ is the "optimal" estimator for minimizing system noise})
\end{array}\right\}
$$

✎PROOF:

1. Let $F \triangleq \left|\hat{H}(\omega)\right|^2\tilde{S}_{xx}(\omega) - \hat{H}(\omega)\tilde{S}_{xy}(\omega) - \hat{H}^*(\omega)\tilde{S}^*_{xy}(\omega) + \tilde{S}_{yy}(\omega)$ (numerator in Lemma 17.4) and
   $G \triangleq \kappa(\omega) + \left|\hat{H}(\omega)\right|^2$ (denominator in Lemma 17.4)

2. lemma $\left(\frac{\partial}{\partial\hat{H}_R}\tilde{S}_{uu}\right)$:

$$
\begin{aligned}
\boxed{0} &= \frac{1}{2}G^2 0 = \frac{1}{2}G^2\frac{\partial}{\partial\hat{H}_R}\tilde{S}_{uu} && \text{set } \frac{\partial}{\partial\hat{H}_R}\tilde{S}_{uu} = 0 \text{ to find optimum } \hat{H}_R\\
&= \frac{1}{2}G^2\frac{\partial}{\partial\hat{H}_R}\frac{F}{G} && \text{by Lemma 17.4 page 131}\\
&= \frac{1}{2}G^2\frac{(F'G - G'F)}{G^2} && \text{by } \textit{Quotient Rule}\\
&= \frac{1}{2}(F'G - G'F)
\end{aligned}
$$

[22] Wicks and Vold (1986) page 898 ⟨has additional s in denominator⟩, Shin and Hammond (2008) page 293 ⟨(9.67)⟩, White et al. (2006) page 679 ⟨(6)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \frac{1}{2}\left[2\hat{H}_R\tilde{S}_{xx} - \tilde{S}_{xy} - \tilde{S}_{xy}^*\right]G - \frac{1}{2}2\hat{H}_R F \qquad \text{by definition of } F, G \qquad \text{(item (1) page 132)}$$

$$= \boxed{\hat{H}_R\tilde{S}_{xx}G - G\mathbf{R}_e\tilde{S}_{xy} - \hat{H}_R F}$$

3. lemma $\left(\frac{\partial}{\partial\hat{H}_I}\tilde{S}_{uu}\right)$:

$$\boxed{0} = \frac{1}{2}G^2 0 = \frac{1}{2}G^2\frac{\partial}{\partial\hat{H}_I}\tilde{S}_{uu} \qquad \text{set } \frac{\partial}{\partial\hat{H}_I}\tilde{S}_{uu} = 0 \text{ to find optimum } \hat{H}_I$$

$$= \frac{1}{2}G^2 0 = \frac{1}{2}G^2\frac{\partial}{\partial\hat{H}_I}\frac{F}{G} \qquad \text{by Lemma 17.4 page 131}$$

$$= \frac{1}{2}G^2\frac{(F'G - G'F)}{G^2} \qquad \text{by } \textit{Quotient Rule}$$

$$= \frac{1}{2}(F'G - G'F)$$

$$= \frac{1}{2}\left[2\hat{H}_I\tilde{S}_{xx} - i\tilde{S}_{xy} + i\tilde{S}_{xy}^*\right]G - \frac{1}{2}2\hat{H}_I F \qquad \text{by definition of } F, G \qquad \text{(item (1) page 132)}$$

$$= \boxed{\hat{H}_I\tilde{S}_{xx}G + G\mathbf{I}_m\tilde{S}_{xy} - \hat{H}_I F}$$

4. Solve for $\hat{H}$ …

$$0 = 0 + i0 = \frac{1}{2}G^2 0 + \frac{1}{2}G^2 0 = \frac{1}{2}G^2\frac{\partial}{\partial\hat{H}_R}\tilde{S}_{uu} + i\frac{1}{2}G^2\frac{\partial}{\partial\hat{H}_I}\tilde{S}_{uu}$$

$$= \left[\hat{H}_R\tilde{S}_{xx}G - G\mathbf{R}_e\tilde{S}_{xy} - \hat{H}_R F\right] + i\left[\hat{H}_I\tilde{S}_{xx}G + G\mathbf{I}_m\tilde{S}_{xy} - \hat{H}_I F\right] \qquad \text{by (2) lemma and (3) lemma}$$

$$= \hat{H}\tilde{S}_{xx}G - \tilde{S}_{xy}^*G - \hat{H}F \quad \text{because } \mathbf{R}_e(z) + i\mathbf{I}_m(z) = z \text{ and } \mathbf{R}_e(z) - i\mathbf{I}_m(z) = z^*$$

$$= \hat{H}\tilde{S}_{xx}G - \tilde{S}_{yx}G - \hat{H}F \qquad \text{by Corollary 2.2 page 15}$$

$$= \hat{H}\tilde{S}_{xx}\left(\kappa + |\hat{H}|^2\right) - \tilde{S}_{yx}\left(\kappa + |\hat{H}|^2\right) - \hat{H}\left(|\hat{H}|^2\tilde{S}_{xx} - \hat{H}\tilde{S}_{xy} - \hat{H}^*\tilde{S}_{xy}^* + \tilde{S}_{yy}\right) \qquad \text{by } F, G \text{ defs.}$$

$$= \hat{H}\tilde{S}_{xx}\left(\kappa + \cancel{|\hat{H}|^2}\right) - \tilde{S}_{yx}\left(\kappa + |\hat{H}|^2\right) - \hat{H}\left(\cancel{|\hat{H}|^2\tilde{S}_{xx}} - \hat{H}\tilde{S}_{xy} - \hat{H}^*\tilde{S}_{xy}^* + \tilde{S}_{yy}\right)$$

$$= \kappa\hat{H}\tilde{S}_{xx} - \tilde{S}_{yx}\left(\kappa + |\hat{H}|^2\right) + \left(\hat{H}^2\tilde{S}_{xy} + |\hat{H}|^2\tilde{S}_{xy}^* - \hat{H}\tilde{S}_{yy}\right)$$

$$= \kappa\hat{H}\tilde{S}_{xx} - \kappa\tilde{S}_{yx} - \cancel{\tilde{S}_{yx}|\hat{H}|^2} + \left(\hat{H}^2\tilde{S}_{xy} + \cancel{|\hat{H}|^2\tilde{S}_{xy}^*} - \hat{H}\tilde{S}_{yy}\right)$$

$$= \hat{H}^2\tilde{S}_{xy} + \hat{H}\left[\kappa\tilde{S}_{xx} - \tilde{S}_{yy}\right] - \kappa\tilde{S}_{xy}^*$$

$$\Longrightarrow \boxed{\hat{H} = \frac{\left(\tilde{S}_{yy} - \kappa\tilde{S}_{xx}\right) \pm \sqrt{\left(\tilde{S}_{yy} - \kappa\tilde{S}_{xx}\right)^2 + 4\kappa\left|\tilde{S}_{xy}\right|^2}}{2\tilde{S}_{xy}}} \qquad \text{by } \textit{Quadratic Equation}$$

✍

17.8 Coherence

17.8.1 Application

Coherence has two basic purposes:

1. The *coherence* of x and y is a measure of how closely x and y are statistically related. That is, it is an indication of how much x and y "cohere" or "stick" together

2. The *coherence* of x and y is a measure of how reliable are the estimates $\hat{H}_1$ and $\hat{H}_2$ (Definition 17.2 page 113, Definition 17.3 page 114). If the coherence is 0.70 or above, then we can have high confidence that the estimates $\hat{H}_1$ and $\hat{H}_2$ are "good" estimates.[23]

17.8.2 Definitions

Definition 17.12. [24] *Let* **S** *be a* SYSTEM *with input* $x(n)$ *and output* $y(n)$.

DEF

The ***complex coherence*** *function is defined as* $C_{xy}(\omega) \triangleq \dfrac{\tilde{S}^*_{xy}(\omega)}{\sqrt{\tilde{S}_{xx}(\omega)\tilde{S}_{yy}(\omega)}}$

The ***ordinary coherence*** *function is defined as* $\gamma^2_{xy}(\omega) \triangleq \dfrac{\left|\tilde{S}_{xy}(\omega)\right|^2}{\tilde{S}_{xx}(\omega)\tilde{S}_{yy}(\omega)}$

Proposition 17.7.

PRP

$$\gamma^2_{xy}(\omega) = \frac{\hat{H}_1(\omega)}{\hat{H}_2(\omega)}$$

✎PROOF:

$$\boxed{\gamma^2_{xy}(\omega)} \triangleq \frac{\left|\tilde{S}_{xy}\right|^2}{\tilde{S}_{xx}\tilde{S}_{yy}} \qquad \text{by definition of } \gamma^2_{xy} \qquad \text{(Definition 17.12 page 134)}$$

$$= \frac{\tilde{S}^*_{xy}/\tilde{S}_{xx}}{\tilde{S}_{yy}/\tilde{S}_{xy}} \triangleq \boxed{\frac{\hat{H}_1}{\hat{H}_2}} \qquad \text{by definitions of } \hat{H}_1 \text{ and } \hat{H}_2 \qquad \text{(Definition 17.2 page 113, Definition 17.3 page 114)}$$

✏

Remark 17.2. Note that the *complex transmissibility* $\tilde{T}'_{xy}$ of Remark 17.1 provides a nice mathematical symmetry (always a good sign of good direction) with *coherence* in the system identification family tree. In particular, note that the following:

REM

$$C_{xy} \triangleq \sqrt{\frac{\hat{H}^*_1}{\hat{H}_2}} \text{ whereas } \tilde{T}'_{xy} \triangleq \sqrt{\hat{H}_1\hat{H}_2}$$

✎PROOF:

$$\sqrt{\frac{\hat{H}^*_1(\omega)}{\hat{H}_2(\omega)}} \qquad \text{by definition of } \hat{H}_{gm} \qquad \text{(Definition 17.5 page 115)}$$

✏

17.8.3 A warning

Estimators yield, as the name implies, estimates. These estimates in general contain some error.

[23] Liang and Lee (2015) pages 363–365 ⟨7.4.2COHERECE FUNCTION⟩

[24] Chen et al. (2012) page 4699(1), (2), Liang and Lee (2015) pages 363–365 ⟨7.4.2 Coherence function⟩, Ewins (1986) page 131 ⟨$\gamma^2 = H_1(\omega)/H_2(\omega)$ (3.8)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

Example 17.1 (The K=1 Welch estimate of coherence). Suppose we have two *uncorrelated* stationary sequences $\mathsf{x}(n)$ and $\mathsf{y}(n)$. Then, there CSD $\mathsf{S}_{\mathsf{xy}}(\omega)$ should be 0 because

$$\begin{aligned}
\mathsf{S}_{\mathsf{xy}}(\omega) &\triangleq \breve{\mathbf{F}}\mathsf{E}\mathsf{R}_{\mathsf{xy}}(m)\\
&= \breve{\mathbf{F}}\mathsf{E}[x(n)y[n+m]]\\
&= \breve{\mathbf{F}}\left[\mathsf{E}_{\mathsf{x}}(n)\right]\left[\mathsf{E}_{\mathsf{y}}[n+m]\right]\\
&= \breve{\mathbf{F}}[0][0]\\
&= 0
\end{aligned}$$

This will give a coherence of 0 also:

$$C(\omega) = \frac{\mathsf{S}_{\mathsf{xy}}}{\sqrt{\mathsf{S}_{\mathsf{xx}}\mathsf{S}_{\mathsf{yy}}}} = 0$$

However, the Welch estimate with $K = 1$ will yield

$$\begin{aligned}
|C(\omega)| &= \left|\frac{\mathsf{S}_{\mathsf{xy}}}{\sqrt{\mathsf{S}_{\mathsf{xx}}\mathsf{S}_{\mathsf{yy}}}}\right|\\
&= \left|\frac{(\tilde{\mathbf{F}}x)(\tilde{\mathbf{F}}y)^*}{\sqrt{\left|\tilde{\mathbf{F}}x\right|^2\left|\tilde{\mathbf{F}}y\right|^2}}\right|\\
&= 1
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 18

ESTIMATING NOISE

Estimating noise in a system is difficult and many estimation methods are possible.

- Thong et al. (2001)
- Zheng et al. (2002)
- Kim and Kamel (2004)
- Kamel and Sim (2004)

(A) **communications LTI additive noise model** (B) **measurement LTI additive noise model**

Figure 18.1: Additive noise systems with **LTI** operator **H**

Copyright © 2019 Daniel J. Greenhoe

Part IV

Statistical Detection

CHAPTER 19

COMMUNICATION CHANNELS

19.1 System model

Figure 19.1: Communication system model

A communication system is an operator. $\mathbf{S}$ over an information sequence u that generates an estimated information sequence $\hat{s}u$. The system operator factors into a receive operator $\mathbf{R}$, a channel operator $\mathbf{C}$, and a transmit operator $\mathbf{T}$ such that

$$\mathbf{S} = \mathbf{RCT}.$$

The transmit operator operates on an information sequence u to generate a channel signal $\mathsf{s}(t;\boldsymbol{u})$. The channel operator operates on the transmitted signal $\mathsf{s}(t;\boldsymbol{u})$ to generate the received signal $\mathsf{r}(t;\boldsymbol{u})$. The receive operator operates on the received signal $\mathsf{r}(t;\boldsymbol{u})$ to generate the estimate $\hat{s}u$ (see Figure 19.1 (page 141)).

Definition 19.1. *Let U be the set of all sequences u and let*

DEF

$$\begin{array}{llll} \mathbf{S} & : U & \to U & \textit{(system operator)} \\ \mathbf{T} & : U & \to \mathbb{R}^\infty & \textit{(transmit operator)} \\ \mathbf{C} & : \mathbb{R}^\infty & \to \mathbb{R}^\infty & \textit{(channel operator)} \\ \mathbf{R} & : \mathbb{R}^\infty & \to U & \textit{(receive operator)} \end{array}$$

*be operators. A **digital communication system** is the operation $\mathbf{S}$ on the set of information sequences U such that $\mathbf{S} \triangleq \mathbf{RCT}$.*

Communication systems can be continuous or discrete valued in time and/or amplitude:

$s(t) = a(t)\psi(t)$	continuous time t	discrete time t
continuous amplitude $a(t)$	analog communications	discrete-time communications
discrete amplitude $a(t)$	—	digital communications

In this document, we normally take the approach that

1. **C** is stochastic
2. There is no structural constraint on **R**.
3. **R** is optimum with respect to the ML-criterion.

These characteristics are explained more fully below.

19.1.1 Channel operator

Real-world physical channels perform a number of operations on a signal. Often these operations are closely modeled by a channel operator **C**. Properties that characterize a particular channel operator associated with some physical channel include

- linear or non-linear
- time-invariant or time-variant
- memoryless or non-memoryless
- deterministic or stochastic.

Examples of physical channels include free space, air, water, soil, copper wire, and fiber optic cable. Information is carried through a channel using some physical process. These processes include:

Process	Example
electromagnetic waves	free space, air
acoustic waves	water, soil
electric field potential (voltage)	wire
light	fiber optic cable
quantum mechanics	

19.1.2 Receive operator

Let **I** be the *identity operator*.[1] Ideally, **R** is selected such that $\mathbf{RCT} = \mathbf{I}$. In this case we say that **R** is the *left inverse*[2] of **CT** and denote this left inverse by **C**. One example of a system where this inverse exists is the noiseless ISI system. While this is quite useful for mathematical analysis and system design, **C** does not actually exist for any real-world system.

When **C** does not exist, the "ideal" **R** is one that is optimum

1. with respect to some *criterion* (or cost function)
2. and sometimes under some structural *constraint*.

[1] **I** is the *identity operator* if for any operator **X**, $\mathbf{XI} = \mathbf{IX} = \mathbf{X}$.

[2] $\mathbf{X}^{-1}X$ is the *left inverse* of **X** if $\mathbf{X}^{-1}X\mathbf{X} = \mathbf{I}$.
$\mathbf{X}^{-1}X$ is the *right inverse* of **X** if $\mathbf{X}\mathbf{X}^{-1}X = \mathbf{I}$.
$\mathbf{X}^{-1}X$ is the *inverse* of **X** if $\mathbf{X}^{-1}X\mathbf{X} = \mathbf{X}\mathbf{X}^{-1}X = \mathbf{I}$.

Copyright © 2019 Daniel J. Greenhoe

When a structural constraint is imposed on $\mathbf{R}$, the solution is called *structured*; otherwise, it is called *non-structured*.[3] A common example of a structured approach is the use of a transversal filter (FIR filter in DSP) in which optimal coefficients are found for the filter. A structured $\mathbf{R}$ is only optimal with respect to the imposed constraint. Even though $\mathbf{R}$ may be optimal with respect to this structure, $\mathbf{R}$ may not be optimal in general; that is, there may be another structure that would lead to a "better" solution. In a non-structured approach, $\mathbf{R}$ is free to take any form whatsoever (practical or impractical) and therefore leads to the best of the best solutions.

The nature of $\mathbf{R}$ depends heavily on the nature of $\mathbf{C}$. If $\mathbf{C}$ does not exist, then the ideal $\mathbf{R}$ is one that is optimal with respect to some criterion. If $\mathbf{C}$ is deterministic, then appropriate optimization criterion may include

- least square error (LSE) criterion
- minimum absolute error criterion
- minimum peak distortion criterion.

If $\mathbf{C}$ is stochastic then appropriate optimization criterion may include

Bayes:	pdf known and cost function defined
Maximum aposteriori probability (MAP):	pdf known and uniform cost function
Maximum likelihood (ML):	pdf known and no prior probability information
mini-max:	pdf not known but a cost function is defined
Neyman-Pearson:	pdf not known and no cost function defined.

Making $\mathbf{R}$ optimum with respect to one of these criterion leads to an *estimate* $\hat{s}u = \mathbf{RCT}\boldsymbol{u}$ that is also optimum with respect to the same criterion (Definition 9.1 page 66).

19.2 Optimization in the case of additional operations

Figure 19.2: Theorem of reversibility

Often in communication systems, an additional operator $\mathbf{X}$ is inserted such that (see Figure 19.2 (page 143))

$$\mathbf{S} = \mathbf{RXCT}.$$

An example of such an operator $\mathbf{X}$ is a receive filter. Is it still possible to find an $\mathbf{R}$ that will perform as well as the case where $\mathbf{X}$ is not inserted? In general, the answer is "no". For example, if $\mathbf{X}r = 0$, then all received information is lost and obviously there is no $\mathbf{R}$ that can recover from this event. However, in the case where the right inverse $\mathbf{X}^{-1}X$ of $\mathbf{X}$ exists, then the answer to the question is "yes" and an optimum $\mathbf{R}$ still exists. That is, it doesn't matter if an $\mathbf{X}$ is inserted into system as long as $\mathbf{X}$ is invertible. This is stated formally in the next theorem.

Theorem 19.1 (Theorem of Reversibility). [4] *Let*

[3] Trees (2001) page 12
[4] Trees (2001) pages 289–290

- $\hat{\theta}_{=}\mathbf{RCT}\boldsymbol{u}$ *be the optimum estimate of* $\boldsymbol{u}$
- $\mathbf{X}$ *be an operator with right inverse* $\mathbf{X}^{-1}X$.

Then there exists some $\mathbf{R}'$ *such that*

THM $\hat{\theta}_{=}\mathbf{R}'\mathbf{XCT}\boldsymbol{u}$.

✎PROOF: Let $\mathbf{R}' = \mathbf{RX}^{-1}X$. Then

$$\mathbf{R}'\mathbf{XCT}\boldsymbol{u} = \mathbf{RX}^{-1}X\mathbf{CT}\boldsymbol{u} = \mathbf{RCT}\boldsymbol{u} = \hat{\theta}.$$

✏

19.3 Channel Statistics

The receiver needs to make a decision as to what sequence $(\!(u)\!)$ the transmitter has sent. This decision should be optimal in some sense. Very often the optimization criterion is chosen to be the *maximal likelihood (ML)* criterion. The information that the receiver can use to make an optimal decision is the received signal $\mathsf{r}(t)$.

If the symbols in $\mathsf{r}(t)$ are statistically *independent*, then the optimal estimate of the current symbol depends only on the current symbol period of $\mathsf{r}(t)$. Using other symbol periods of $\mathsf{r}(t)$ has absolutely no additional benefit. Note that the AWGN channel is *memoryless*; that is, the way the channel treats the current symbol has nothing to do with the way it has treated any other symbol. Therefore, if the symbols sent by the transmitter into the channel are independent, the symbols coming out of the channel are also independent.

However, also note that the symbols sent by the transmitter are often very intentionally not independent; but rather a strong relationship between symbols is intentionally introduced. This relationship is called *channel coding*. With proper channel coding, it is theoretically possible to reduce the probability of communication error to any arbitrarily small value as long as the channel is operating below its *channel capacity*.

This chapter assumes that the received symbols are statistically independent; and therefore optimal decisions at the receiver for the current symbol are made only from the current symbol period of $\mathsf{r}(t)$.

The received signal $\mathsf{r}(t)$ over a single symbol period contains an uncountably infinite number of points. That is a lot. It would be nice if the receiver did not have to look at all those uncountably infinite number of points when making an optimal decision. And in fact the receiver does indeed not have to. As it turns out, a single finite set of *statistics* $\{\dot{r}_1, \dot{r}_2, \ldots, \dot{r}_N\}$ is sufficient for the receiver to make an optimal decision as to which value the transmitter sent.

Definition 19.2. *Let* $\mathbf{C}$ *be an additive noise channel*

COPYRIGHT © 2019 DANIEL J. GREENHOE

CHAPTER 20

OPTIMAL SYMBOL DETECTION

20.1 ML Estimation

Theorem 20.1. *In an AWGN channel with received signal* $\mathsf{r}(t) = s(t;\phi) + n(t)$ *Let*

- $\mathsf{r}(t) = s(t;\phi) + n(t)$ *be the received signal in an AWGN channel*
- $n(t)$ *a Gaussian white noise process*
- $\mathsf{s}(t;\phi)$ *the transmitted signal such that*

$$s(t;\phi) = \sum_{n\in\mathbb{Z}} a_n \lambda(t-nT)\cos(2\pi f_c t + \theta_n + \phi).$$

Then the optimal ML estimate of ϕ *is either of the two equivalent expressions*

THM

$$\begin{aligned}\hat{\phi}_{\mathsf{ml}} &= -\operatorname{atan}\left[\frac{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\sin(2\pi f_c t + \theta_n)\,\mathsf{d}t}{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\cos(2\pi f_c t + \theta_n)\,\mathsf{d}t}\right] \\ &= \arg_\phi\left(\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\left[\lambda(t-nT)\sin(2\pi f_c t + \theta_n + \phi)\right]\,\mathsf{d}t = 0\right).\end{aligned}$$

✎PROOF:

$$\begin{aligned}\hat{\phi}_{\mathsf{ml}} &= \arg_\phi\left(2\int_{t\in\mathbb{R}} \mathsf{r}(t)\left[\frac{\partial}{\partial\phi}s(t;\phi)\right]\,\mathsf{d}t = \frac{\partial}{\partial\phi}\int_{t\in\mathbb{R}} s^2(t;\phi)\,\mathsf{d}t\right) \qquad \text{by Theorem 12.6 page 86} \\ &= \arg_\phi\left(2\int_{t\in\mathbb{R}} \mathsf{r}(t)\left[\frac{\partial}{\partial\phi}s(t;\phi)\right]\,\mathsf{d}t = \frac{\partial}{\partial\phi}\|s(t;\phi)\|^2\,\mathsf{d}t\right) \\ &= \arg_\phi\left(2\int_{t\in\mathbb{R}} \mathsf{r}(t)\left[\frac{\partial}{\partial\phi}s(t;\phi)\right]\,\mathsf{d}t = 0\right) \\ &= \arg_\phi\left(\int_{t\in\mathbb{R}} \mathsf{r}(t)\left[\frac{\partial}{\partial\phi}\sum_{n\in\mathbb{Z}} a_n\lambda(t-nT)\cos(2\pi f_c t + \theta_n + \phi)\right]\,\mathsf{d}t = 0\right)\end{aligned}$$

$$= \arg_\phi\left(-\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\left[\lambda(t-nT)\sin(2\pi f_c t+\theta_n+\phi)\right]\,\mathrm{d}t = 0\right)$$

$$= \arg_\phi\left(\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\left[\sin(2\pi f_c t+\theta_n)\cos(\phi)+\sin(\phi)\cos(2\pi f_c t+\theta_n)\right]\,\mathrm{d}t = 0\right)$$

$$= \arg_\phi\left(\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\sin(\phi)\cos(2\pi f_c t+\theta_n)\,\mathrm{d}t = -\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\sin(2\pi f_c t+\theta_n)\cos(\phi)\,\mathrm{d}t\right)$$

$$= \arg_\phi\left(\sin(\phi)\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\cos(2\pi f_c t+\theta_n)\,\mathrm{d}t = -\cos(\phi)\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\sin(2\pi f_c t+\theta_n)\,\mathrm{d}t\right)$$

$$= \arg_\phi\left(\frac{\sin(\phi)}{\cos(\phi)} = -\frac{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\sin(2\pi f_c t+\theta_n)\,\mathrm{d}t}{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\cos(2\pi f_c t+\theta_n)\,\mathrm{d}t}\right)$$

$$= \arg_\phi\left(\tan(\phi) = -\frac{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\sin(2\pi f_c t+\theta_n)\,\mathrm{d}t}{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\cos(2\pi f_c t+\theta_n)\,\mathrm{d}t}\right)$$

$$= \arg_\phi\left(\phi = -\operatorname{atan}\left(\frac{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\sin(2\pi f_c t+\theta_n)\,\mathrm{d}t}{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\cos(2\pi f_c t+\theta_n)\,\mathrm{d}t}\right)\right)$$

$$= -\operatorname{atan}\left(\frac{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\sin(2\pi f_c t+\theta_n)\,\mathrm{d}t}{\sum_{n\in\mathbb{Z}} a_n \int_{t\in\mathbb{R}} \mathsf{r}(t)\lambda(t-nT)\cos(2\pi f_c t+\theta_n)\,\mathrm{d}t}\right)$$

✏

20.2 Generalized coherent modulation

Figure 20.1: Distributions of orthonormal components

Theorem 20.2. *Let*

- $(V, \langle\cdot\,|\,\cdot\rangle, S)$ *be a modulation space*
- $\Psi \triangleq \{\psi_n(t) : n = 1,2,\ldots,N\}$ *be a set of orthonormal functions that span* S
- $\dot r_n \triangleq \langle \mathsf{r}(t)\,|\,\psi_n(t)\rangle$
- $R \triangleq \{\dot r_n : n = 1,2,\ldots,N\}$
- $\dot s_n(m) \triangleq \langle s(t;m)\,|\,\psi_n(t)\rangle$

and let V *be partitioned into* ***decision regions***

$$\{D_m : m = 1,2,\ldots,|S|\}$$

such that

$$\mathsf{r}(t) \in D_{\hat m} \iff \hat m = \arg\max_m \mathsf{P}\{s(t;m)|\mathsf{r}(t)\}\,.$$

Copyright © 2019 Daniel J. Greenhoe

*Then the **probability of detection error** is*

THM

$$\mathsf{P}\{error\} = 1 - \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \sum_m \mathsf{P}\{m\ sent\} \int_{\boldsymbol{r}\in D_m} \exp\frac{-1}{2\sigma^2}\sum_{n=1}^N [\dot{r}_n - \dot{s}_n(m)]^2\, d\boldsymbol{r}.$$

✎PROOF:

$$\begin{aligned}
\mathsf{P}\{\text{error}\} &= 1 - \mathsf{P}\{\text{no error}\}\\
&= 1 - \sum_m \mathsf{P}\left\{(m \text{ sent}) \wedge (\hat{m} = m \text{ detected})\right\}\\
&= 1 - \sum_m \mathsf{P}\left\{(\hat{m} = m \text{ detected})|(m \text{ sent})\right\} \mathsf{P}\{m \text{ sent}\}\\
&= 1 - \sum_m \mathsf{P}\{m \text{ sent}\}\, \mathsf{P}\{\boldsymbol{r}|(m \text{ sent})\}\\
&= 1 - \sum_m \mathsf{P}\{m \text{ sent}\} \int_{\boldsymbol{r}\in D_m} \mathsf{p}\,[\boldsymbol{r}|(\mathsf{m} \text{ sent})]\, d\boldsymbol{r}\\
&= 1 - \sum_m \mathsf{P}\{m \text{ sent}\} \int_{\boldsymbol{r}\in D_m} \prod_n \mathsf{p}\left[\dot{\mathsf{r}}_\mathsf{n}|\mathsf{m}\right]\, d\boldsymbol{r}\\
&= 1 - \sum_m \mathsf{P}\{m \text{ sent}\} \int_{\boldsymbol{r}\in D_m} \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp\frac{-[\dot{r}_n - \mathsf{E}\dot{r}_n]^2}{2\sigma^2}\, d\boldsymbol{r}\\
&= 1 - \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \sum_m \mathsf{P}\{m \text{ sent}\} \int_{\boldsymbol{r}\in D_m} \exp\frac{-1}{2\sigma^2}\sum_{n=1}^N [\dot{r}_n - \dot{s}_n(m)]^2\, d\boldsymbol{r}
\end{aligned}$$

✏

20.3 Frequency Shift Keying (FSK)

Theorem 20.3. *In an FSK modulation space, the optimal ML estimator of m is*

THM

$$\hat{m} = \arg\max_m \dot{r}_m.$$

✎PROOF:

$$\begin{aligned}
\hat{m} &= \arg\max_m \mathsf{P}\{\mathsf{r}(t)|s(t;m)\}\\
&= \arg\min_m \sum_{n=1}^N [\dot{r}_n - \dot{s}_n(m)]^2 && \text{by Theorem 12.6 (page 86)}\\
&= \arg\min_m \sum_{n=1}^N [\dot{r}_n^2 - 2\dot{r}_n\dot{s}_n(m) + \dot{s}_n^2(m)]\\
&= \arg\min_m \sum_{n=1}^N [-2\dot{r}_n\dot{s}_n(m) + \dot{s}_n^2(m)] && \dot{r}_n^2 \text{ is independent of } m\\
&= \arg\min_m \sum_{n=1}^N [-2\dot{r}_n a\bar{\delta}_{mn} + a^2\bar{\delta}_{mn}]\\
&= \arg\min_m [-2a\dot{r}_m + a^2]
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

$$= \arg\min_m[-\dot{r}_m] \qquad a \text{ and } 2 \text{ independent of } m$$
$$= \arg\max_m[\dot{r}_m]$$

✏

Theorem 20.4. *If an FSK modulation space let*

$$\begin{array}{rcl|lcl|l} z_2 & \triangleq & \dot{r}_1(1) - \dot{r}_2(1) & z_2 > 0 & \Longrightarrow & \dot{r}_1 > \dot{r}_2 & m = 1 \\ z_3 & \triangleq & \dot{r}_1(1) - \dot{r}_3(1) & z_3 > 0 & \Longrightarrow & \dot{r}_1 > \dot{r}_3 & m = 1 \\ & \vdots & & & & & \\ z_M & \triangleq & \dot{r}_1(1) - \dot{r}_M(1) & z_M > 0 & \Longrightarrow & \dot{r}_1 > \dot{r}_M & m = 1 \end{array}$$

Then the ***probability of detection error*** *is*

THM $$\mathsf{P}\{error\} = 1 - \frac{M-1}{M}\int_0^\infty\int_0^\infty\cdots\int_0^\infty \mathsf{p}(z_2, z_3, \ldots, z_M)\, dz_2 dz_3 \cdots dz_M$$ *where*

$$\mathsf{p}(z_2, z_3, \ldots, z_M) = \frac{1}{(2\pi)^{\frac{M-1}{2}}\sqrt{\det R}} \exp -\frac{1}{2}\begin{bmatrix} z_2 - \dot{s} \\ z_3 - \dot{s} \\ \vdots \\ z_M - \dot{s} \end{bmatrix}^T R^{-1} \begin{bmatrix} z_2 - \dot{s} \\ z_3 - \dot{s} \\ \vdots \\ z_M - \dot{s} \end{bmatrix}$$

and

$$R = \begin{bmatrix} \operatorname{cov}\left[z_2, z_2\right] & \operatorname{cov}\left[z_2, z_3\right] & \cdots & \operatorname{cov}\left[z_2, z_M\right] \\ \operatorname{cov}\left[z_3, z_2\right] & \operatorname{cov}\left[z_3, z_3\right] & \cdots & \operatorname{cov}\left[z_3, z_M\right] \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{cov}\left[z_M, z_2\right] & \operatorname{cov}\left[z_M, z_3\right] & \cdots & \operatorname{cov}\left[z_M, z_M\right] \end{bmatrix} = N_o \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{bmatrix}$$

The inverse matrix R^{-1} is equivalent to (????)

$$R^{-1} \stackrel{?}{=} \frac{1}{MN_o}\begin{bmatrix} M-1 & -1 & \cdots & -1 \\ -1 & M-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & M-1 \end{bmatrix}$$

✎PROOF:

$$\begin{aligned} \mathsf{E}z_k &= \mathsf{E}\left[\dot{r}_{11} - \dot{r}_{1k}\right] \\ &= \mathsf{E}\dot{r}_{11} - \mathsf{E}\dot{r}_{1k} \\ &= \dot{s} - 0 \\ &= \dot{s} \end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$
\begin{aligned}
\operatorname{cov}\left[z_m, z_n\right] &= \mathsf{E}\left[z_m z_n\right] - [\mathsf{E}z_m][\mathsf{E}z_n] \\
&= \mathsf{E}\left[(\dot r_{11} - \dot r_{1m})(\dot r_{11} - \dot r_{1n})\right] - \dot s^2 \\
&= \mathsf{E}\left[\dot r_{11}^2 - \dot r_{11}\dot r_{1n} - \dot r_{1m}\dot r_{11}\dot r_{1m}\dot r_{1n}\right] - \dot s^2 \\
&= [\operatorname{var}\dot r_{11} + (\mathsf{E}\dot r_{11})^2] - \mathsf{E}\left[\dot r_{11}\right]\mathsf{E}\left[\dot r_{1n}\right] - \mathsf{E}\left[\dot r_{1m}\right]\mathsf{E}\left[\dot r_{11}\right] + [\operatorname{cov}\left[\dot r_{1m}, \dot r_{1n}\right] + (\mathsf{E}\dot r_{1m})(\mathsf{E}\dot r_{1n})] - \dot s^2 \\
&= [\operatorname{var}\dot r_{11} + \dot s^2] - a\cdot 0 - 0\cdot a + [\operatorname{cov}\left[\dot r_{1m}, \dot r_{1n}\right] + 0\cdot 0] - \dot s^2 \\
&= \operatorname{var}\dot r_{11} + \operatorname{cov}\left[\dot r_{1m}, \dot r_{1n}\right] \\
&= N_o + \operatorname{cov}\left[\dot r_{1m}, \dot r_{1n}\right] \\
&= \begin{cases} 2N_o & \text{for } m = n \\ N_o & \text{for } m \neq n. \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
P\{\text{error}\} &= 1 - P\{\text{no error}\} \\
&= 1 - \sum_{m=1}^{M} P\{\text{m transmitted}) \wedge (\forall k \neq m, \dot r_m > \dot r_k\} \\
&= 1 - (M-1)P\{1 \text{ transmitted}) \wedge (\dot r_{11} > \dot r_{12}) \wedge (\dot r_{11} > \dot r_{13}) \wedge \cdots \wedge (\dot r_{11} > \dot r_{1M})\} \\
&= 1 - (M-1)P\{(\dot r_{11} - \dot r_{12} > 0) \wedge (\dot r_{11} - \dot r_{13} > 0) \wedge \cdots \wedge (\dot r_{11} - \dot r_{1M} > 0)|1 \text{ transmitted})\}P\{1 \text{ transmitted})\} \\
&= 1 - \frac{M-1}{M}P\{(z_2 > 0) \wedge (z_3 > 0) \wedge \cdots \wedge (z_M > 0)|1 \text{ transmitted})\} \\
&= 1 - \frac{M-1}{M}\int_0^\infty \int_0^\infty \cdots \int_0^\infty \mathsf{p}(z_2, z_3, \ldots, z_M)\, dz_2 dz_3 \cdots dz_M.
\end{aligned}
$$

✍

20.4 Quadrature Amplitude Modulation (QAM)

20.4.1 Receiver statistics

Figure 20.2: Distributions of QAM components

Theorem 20.5. *Let $(V, \langle\cdot\,|\,\cdot\rangle$ be a QAM modulation space such that*

$$
\begin{aligned}
\mathsf{r}(t) &= \mathsf{s}(t; m) + \mathsf{n}(t) \\
\dot r_c &\triangleq \langle \mathsf{r}(t)\,|\,\psi_c(t)\rangle \\
\dot r_s &\triangleq \langle \mathsf{r}(t)\,|\,\psi_s(t)\rangle\,.
\end{aligned}
$$

CC BY-NC-ND

Then $(\dot{r}_c|m)$ *and* $(\dot{r}_s|m)$ *are* ***independent*** *and have* ***marginal distributions***

$$\begin{aligned}(\dot{r}_c|m) &\sim & \mathsf{N}\left(a_m,\sigma^2\right) &= \mathsf{N}\left(r_m\cos\theta_m,\sigma^2\right)\\ (\dot{r}_s|m) &\sim & \mathsf{N}\left(b_m,\sigma^2\right) &= \mathsf{N}\left(r_m\sin\theta_m,\sigma^2\right).\end{aligned}$$

✎PROOF: See Theorem 12.5 (page 85) page 85. ✏

20.4.2 Detection

Theorem 20.6. *Let* $(V,\langle\cdot\,|\,\cdot\rangle,S)$ *be a QAM modulation space with*

$$\begin{aligned}\mathsf{r}(t) &= \mathsf{s}(t;m)+\mathsf{n}(t)\\ \dot{r}_c &\triangleq & \langle\mathsf{r}(t)\,|\,\psi_c(t)\rangle\\ \dot{r}_s &\triangleq & \langle\mathsf{r}(t)\,|\,\psi_s(t)\rangle\,.\end{aligned}$$

Then $\{\dot{r}_c,\dot{r}_s\}$ *are sufficient statistics for optimal ML detection and the optimal ML estimate of* m *is*

$$\hat{u}_{\mathsf{ml}}[m] = \arg\min_m\left[(\dot{r}_c-a_m)^2+(\dot{r}_s-b_m)^2\right].$$

✎PROOF:

$$\begin{aligned}\hat{u}_{\mathsf{ml}}[m] &= \arg\max_m \mathsf{P}\left\{\mathsf{r}(t)|s(t;m)\right\} && \text{by Definition 9.1 (page 66)}\\ &= \arg\min_m \sum_{n=1}^{N}[\dot{r}_n-\dot{s}_n(m)]^2 && \text{by Theorem 12.6 (page 86)}\\ &= \arg\min_m\left[(\dot{r}_c-a_m)^2+(\dot{r}_s-b_m)^2\right]\end{aligned}$$

✏

20.4.3 Probability of error

Figure 20.3: QAM-16 cosstellation and decision regions

Theorem 20.7. *In a QAM-16 constellation as shown in Figure 20.3 (page 150), the probability of error is*

$$\mathsf{P}\left\{error\right\} = \frac{9}{4}Q^2\left(\frac{\dot{s}_{21}-\dot{s}_{11}}{2N_o}\right).$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF: Let

$$d \triangleq \dot{s}_{21} - \dot{s}_{11}.$$

Then

$$\begin{aligned}
\mathsf{P}\{\text{error}\} &= \sum_{m=1}^{M} \mathsf{P}\left\{[s(t;m) \text{ transmitted }] \wedge [(\dot{r}_1, \dot{r}_2) \notin D_m]\right\} \\
&= \sum_{m=1}^{M} \mathsf{P}\left\{[(\dot{r}_1, \dot{r}_2) \notin D_m] | [s(t;m) \text{ transmitted }]\right\} \mathsf{P}\left\{[s(t;m) \text{ transmitted }]\right\} \\
&= \frac{1}{M} \sum_{m=1}^{M} \mathsf{P}\left\{[(\dot{r}_1, \dot{r}_2) \notin D_m] | [s(t;m) \text{ transmitted }]\right\} \\
&= \frac{1}{M}\left[4\mathsf{P}\left\{(\dot{r}_1, \dot{r}_2) \notin D_1 | s_1(t)\right\} + 8\mathsf{P}\left\{(\dot{r}_1, \dot{r}_2) \notin D_2 | s_2(t)\right\} + 4\mathsf{P}\left\{(\dot{r}_1, \dot{r}_2) \notin D_6 | s_6(t)\right\}\right] \\
&= \frac{1}{M}\left[4\int\int_{(x,y)\notin D_1} \mathrm{p}_{xy|1}(x,y)\,\mathrm{d}x\,\mathrm{d}y + 8\int\int_{(x,y)\notin D_2} \mathrm{p}_{xy|2}(x,y)\,\mathrm{d}x\,\mathrm{d}y + \right. \\
&\left. 4\int\int_{(x,y)\notin D_6} \mathrm{p}_{xy|6}(x,y)\,\mathrm{d}x\,\mathrm{d}y\right] \\
&= \frac{1}{M}\left[4\int\int_{(x,y)\notin D_1} \mathrm{p}_{x|1}(x)\mathrm{p}_{y|1}(y)\,\mathrm{d}x\,\mathrm{d}y + 8\int\int_{(x,y)\notin D_2} \mathrm{p}_{x|2}(x)\mathrm{p}_{y|2}(y)\,\mathrm{d}x\,\mathrm{d}y + \right. \\
&\left. 4\int\int_{(x,y)\notin D_6} \mathrm{p}_{x|6}(x)\mathrm{p}_{y|6}(y)\,\mathrm{d}x\,\mathrm{d}y\right] \\
&= \frac{1}{M}\left[4Q\left(\frac{d}{2N_o}\right)Q\left(\frac{d}{2N_o}\right) + 8Q\left(\frac{d}{2N_o}\right)2Q\left(\frac{d}{2N_o}\right) + 4\cdot 2Q\left(\frac{d}{2N_o}\right)2Q\left(\frac{d}{2N_o}\right)\right] \\
&= \frac{9}{4}Q^2\left(\frac{d}{2N_o}\right)
\end{aligned}$$

✏

20.5 Phase Shift Keying (PSK)

20.5.1 Receiver statistics

Figure 20.4: Distributions of PSK components

Theorem 20.8. *Let*

$$\begin{aligned}
\dot{r}_c &\triangleq && \langle \mathsf{r}(t) \mid \psi_c(t) \rangle \\
\dot{r}_s &\triangleq && \langle \mathsf{r}(t) \mid \psi_s(t) \rangle \\
\theta_m &\triangleq && \operatorname{atan}\left[\frac{\dot{r}_s(m)}{\dot{r}_c(m)}\right].
\end{aligned}$$

The statistics $(\dot{r}_c|m)$ *and* $(\dot{r}_s|m)$ *are* ***independent*** *with marginal distributions*

$$\begin{aligned}
(\dot{r}_c|m) &\sim && \mathsf{N}\left(r\cos\theta_m, \sigma^2\right) \\
(\dot{r}_s|m) &\sim && \mathsf{N}\left(r\sin\theta_m, \sigma^2\right) \\
\mathrm{p}_{\theta_m}(\theta|m) &= \int_0^\infty x\mathrm{p}_{\dot{r}_c}(x|m)\mathrm{p}_{\dot{r}_s}(x\tan\theta|m)dx.
\end{aligned}$$

✎PROOF:

Indepence and marginal distributions of $\dot{r}_1(m)$ and $\dot{r}_2(m)$ follow directly from Theorem 12.5 (page 85) (page 85).

Let $X \triangleq \dot{r}_1(m)$, $Y \triangleq \dot{r}_2(m)$ and $\Theta \triangleq \theta_m$. Then[1]

$$\begin{aligned}
\mathsf{p}_\theta(\theta)d\theta &\triangleq \mathsf{P}\{\theta < \Theta \le \theta + d\theta\} \\
&= \mathsf{P}\left\{\theta < \operatorname{atan}\frac{Y}{X} \le \theta + d\theta\right\} \\
&= \mathsf{P}\left\{\tan(\theta) < \frac{Y}{X} \le \tan(\theta + d\theta)\right\} \\
&= \mathsf{P}\left\{\tan(\theta) < \frac{Y}{X} \le \tan\theta + (1+\tan^2\theta)\,\mathsf{d}\theta\right\} \\
&= \int_0^\infty \mathsf{P}\left\{\left[\tan\theta < \frac{Y}{X} \le \tan\theta + (1+\tan^2\theta)\,\mathsf{d}\theta\right] \wedge \left[(x < X \le x + \mathsf{d}x)\right]\right\} \\
&= \int_0^\infty \mathsf{P}\left\{\tan\theta < \frac{Y}{x} \le \tan\theta + (1+\tan^2\theta)\,\mathsf{d}\theta \;\middle|\; x < X \le x + dx\right\} \mathsf{P}\{x < X \le x + \mathsf{d}x\} \\
&= \int_0^\infty \mathsf{P}\left\{x\tan\theta < Y \le x\tan\theta + x(1+\tan^2\theta)\,\mathsf{d}\theta) \;\middle|\; X = x\right\} \mathsf{p}_\mathsf{X}(x)\,\mathsf{d}x \\
&= \int_0^\infty [\mathsf{p}_\mathsf{Y}(x\tan\theta)x(1+\tan^2\theta)]\mathsf{p}_\mathsf{X}(x)\,\mathsf{d}x\,\mathsf{d}\theta \\
&= (1+\tan^2\theta)\int_0^\infty x\mathsf{p}_\mathsf{Y}(x\tan\theta)\mathsf{p}_\mathsf{X}(x)\,\mathsf{d}x\,\mathsf{d}\theta \\
\Longrightarrow \\
\mathsf{p}_\theta(\theta)d\theta &= (1+\tan^2\theta)\int_0^\infty x\mathsf{p}_\mathsf{Y}(x\tan\theta)\mathsf{p}_\mathsf{X}(x)\,\mathsf{d}x
\end{aligned}$$

✏

[1] A similar example is in Papoulis (1991), page 138

COPYRIGHT © 2019 DANIEL J. GREENHOE

20.5.2 Detection

Theorem 20.9. *Let $(V, \langle\cdot\,|\,\cdot\rangle, S)$ be a PSK modulation space with*

$$\begin{aligned} \mathsf{r}(t) &= \mathsf{s}(t;m) + \mathsf{n}(t) \\ \dot{r}_c &\triangleq \langle \mathsf{r}(t)\,|\,\psi_c(t)\rangle \\ \dot{r}_s &\triangleq \langle \mathsf{r}(t)\,|\,\psi_s(t)\rangle\,. \end{aligned}$$

Then $\{\dot{r}_c, \dot{r}_s\}$ are sufficient statistics for optimal ML detection and the optimal ML estimate of m is

$$\hat{u}_{\mathsf{ml}}[m] = \arg\min_m \left[(\dot{r}_1 - r\cos\theta_m)^2 + (\dot{r}_2 - r\sin\theta_m)^2\right].$$

✎PROOF:

$$\begin{aligned} \hat{u}_{\mathsf{ml}}[m] &= \arg\max_m \mathsf{P}\{\mathsf{r}(t)|s(t;m)\} && \text{by Definition 9.1 (page 66)} \\ &= \arg\min_m \sum_{n=1}^{N} [\dot{r}_n - \dot{s}_n(m)]^2 && \text{by Theorem 12.6 (page 86)} \\ &= \arg\min_m \left[(\dot{r}_1 - r\cos\theta_m)^2 + (\dot{r}_2 - r\sin\theta_m)^2\right]. \end{aligned}$$

✏

20.5.3 Probability of error

Figure 20.5: PSK-8 Decision regions

Theorem 20.10. *The probability of error using PSK modulation is*

$$\mathsf{P}\{error\} = M\left[1 - \int_{\frac{2\pi}{M}\left(m-\frac{3}{2}\right)}^{\frac{2\pi}{M}\left(m-\frac{1}{2}\right)} \mathrm{p}_{\theta_1}(\theta)\, d\theta\right].$$

✎PROOF: See Figure 20.5 (page 153).

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

$$\begin{aligned}
\mathsf{P}\{\text{error}\} &= \sum_{m=1}^{M} \mathsf{P}\left\{\text{error}|s(t;m) \text{ was transmitted}\right\} \\
&= M\mathsf{P}\left\{\text{error}|s_1(t) \text{ was transmitted}\right\} \\
&= M\left[1 - \int_{\frac{2\pi}{M}\left(m-\frac{3}{2}\right)}^{\frac{2\pi}{M}\left(m-\frac{1}{2}\right)} \mathrm{p}_{\theta_1}(\theta)\, d\theta\right].
\end{aligned}$$

✏

20.6 Pulse Amplitude Modulation (PAM)

20.6.1 Receiver statistics

Figure 20.6: Distribution of PAM component

Theorem 20.11. *Let* $(V, \langle\cdot\,|\,\cdot\rangle$ *be a PAM modulation space such that*

$$\begin{aligned}
\mathsf{r}(t) &= \mathsf{s}(t;m) + \mathsf{n}(t) \\
\dot{r}_c &\triangleq \langle \mathsf{r}(t)\,|\,\psi_c(t)\rangle \\
\dot{r}_s &\triangleq \langle \mathsf{r}(t)\,|\,\psi_{\mathsf{s}}(t)\rangle\,.
\end{aligned}$$

Then $(\dot{r}|m)$ *has* ***distribution***

$$\dot{r}(m) \sim \mathsf{N}\left(a_m, \sigma^2\right).$$

✎PROOF: This follows directly from Theorem 12.5 (page 85) (page 85). ✏

20.6.2 Detection

Theorem 20.12. *Let* $(V, \langle\cdot\,|\,\cdot\rangle, S)$ *be a PAM modulation space with*

$$\begin{aligned}
\mathsf{r}(t) &= \mathsf{s}(t;m) + \mathsf{n}(t) \\
\dot{r} &\triangleq \langle \mathsf{r}(t)\,|\,\psi(t)\rangle\,.
\end{aligned}$$

Then $\dot{r}$ *is a sufficient statistic for the optimal ML detection of* m *and the optimal ML estimate of* m *is*

$$\hat{u}_{\mathsf{ml}}[m] = \arg\min_m |\dot{r} - a_m|.$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$\begin{aligned}
\hat{u}_{\mathsf{ml}}[m] &= \arg\max_m \mathsf{P}\left\{\mathsf{r}(t)|a_m\right\} && \text{by Definition 9.1 (page 66)}\\
&= \arg\min_m \sum_{n=1}^{N}[\dot{r}_n - \dot{s}_n(m)]^2 && \text{by Theorem 12.6 (page 86)}\\
&= \arg\min_m [\dot{r} - \dot{s}(m)]^2\\
&= \arg\min_m |\dot{r} - \dot{s}(m)|
\end{aligned}$$

✏

20.6.3 Probability of error

Theorem 20.13. *The probability of detection error in a PAM modulation space is*

$$\mathsf{P}\{error\} = 2\frac{M-1}{M}\mathsf{Q}\left[\frac{a_2 - a_1}{2\sqrt{N_o}}\right].$$

✎PROOF: Let $d \triangleq a_2 - a_1$ and $\sigma \triangleq \sqrt{\operatorname{var}\dot{r}} = \sqrt{N_o}$. Also, let the decision regions D_m be as illustrated in Figure 20.7 (page 155). Then

$$\begin{aligned}
\mathsf{P}\{error\} &= \sum_{m=1}^{M}\mathsf{P}\left\{s(t;m)\text{ sent }\wedge r \notin D_m\right\}\\
&= \sum_{m=1}^{M}\mathsf{P}\left\{\dot{r} \notin D_m|\mathsf{s}(t;m)\text{ sent }\right\}\mathsf{P}\left\{s(t;m)\text{ sent }\right\}\\
&= \sum_{m=1}^{M}\mathsf{P}\left\{\dot{r}_m \notin D_m\right\}\frac{1}{M}\\
&= \frac{1}{M}\left(\mathsf{Q}\left[\frac{d}{2\sigma}\right] + 2\mathsf{Q}\left[\frac{d}{2\sigma}\right] + \ldots 2\mathsf{Q}\left[\frac{d}{2\sigma}\right] + \mathsf{Q}\left[\frac{d}{2\sigma}\right]\right)\\
&= 2\frac{M-1}{M}\mathsf{Q}\left[\frac{d}{2\sigma}\right]\\
&= 2\frac{M-1}{M}\mathsf{Q}\left[\frac{\dot{s}_2 - \dot{s}_1}{2\sqrt{N_o}}\right]
\end{aligned}$$

✏

Figure 20.7: 4-ary PAM in AWGN channel

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

CHAPTER 21

NETWORK DETECTION

21.1 Detection

For detection, we need

1. Cost function: for hard decisions, its range must be linearly ordered. For soft decisions, it can be a lattice.

2. system joint and marginal probabilities (for Bayesian detection)

21.2 Bayesian Estimation

Definition 21.1.

DEF				
	H	$\triangleq$	$\{h_1, h_2, h_3, \ldots\}$	*set of hypotheses*
	$\boldsymbol{D}$	$\triangleq$	$\{D_1, D_2, D_3, \ldots\}$	*partition—decision regions*
	X	$\triangleq$	$\{X_1, X_2, X_3, \ldots\}$	*set of sensor inputs*

$$
\begin{aligned}
cost(h; P) &= \min_{\boldsymbol{D}} \sum_i \mathsf{P}\left\{\left[X \in D_i\right] \wedge \left[H \neq h_i\right]\right\} \\
&= \min_{\boldsymbol{D}} \sum_i \mathsf{P}\left\{X \in D_i \;\middle|\; H \neq h_i\right\} \mathsf{P}\left\{H \neq h_i\right\} \\
&= \min_{\boldsymbol{D}} \sum_i \sum_{j \neq i} \left[1 - \mathsf{P}\left\{X \in D_i \;\middle|\; H = h_i\right\}\right] \sum_{j \neq i} \left[1 - \mathsf{P}\left\{H = h_i\right\}\right]
\end{aligned}
$$

$$
\hat{h} = \arg_h cost(h; P)
$$

21.3 Joint Gaussian Model

Assume convexity …

$$
\begin{aligned}
\boldsymbol{D} &= \arg_{\boldsymbol{D}} \min_{\boldsymbol{D}} cost(h;P) \\
&= \arg_{\boldsymbol{D}} \left\{ \frac{\partial}{\partial \boldsymbol{D}} \sum_i \int_{D_i} \mathsf{p}(\boldsymbol{x}|H \neq h_i) \underbrace{\mathsf{p}(H \neq h_i)}_{c} d\boldsymbol{x} = 0 \right\} \\
&= \arg_{\boldsymbol{D}} \left\{ \frac{\partial}{\partial \boldsymbol{D}} c \sum_i \int_{D_i} \mathsf{p}(\boldsymbol{x}|H \neq h_i) d\boldsymbol{x} = 0 \right\} \\
&= \arg_{\boldsymbol{D}} \left\{ \frac{\partial}{\partial \boldsymbol{D}} \sum_i \left[1 - \sum_{j \neq i} \int_{D_i} \mathsf{p}(\boldsymbol{x}|H = h_i) d\boldsymbol{x} \right] = 0 \right\} \\
&= \arg_{\boldsymbol{D}} \left\{ \frac{\partial}{\partial \boldsymbol{D}} \sum_i \left[1 - \sum_{j \neq i} \int_{D_i} \frac{1}{\sqrt{(2\pi)^n |\boldsymbol{M}|}} \exp -\frac{1}{2} (\boldsymbol{x} - \mathsf{E}\boldsymbol{x})^T \boldsymbol{M}^{-1} (\boldsymbol{x} - \mathsf{E}\boldsymbol{x}) d\boldsymbol{z} \right] = 0 \right\} \\
&= \arg_{\boldsymbol{D}} \left\{ \sum_i \left[1 - \sum_{j \neq i} \frac{\partial}{\partial \boldsymbol{D}} \int_{D_i} \frac{1}{\sqrt{(2\pi)^n |\boldsymbol{M}|}} \exp -\frac{1}{2} (\boldsymbol{x} - \mathsf{E}\boldsymbol{x})^T \boldsymbol{M}^{-1} (\boldsymbol{x} - \mathsf{E}\boldsymbol{x}) d\boldsymbol{z} \right] = 0 \right\} \\
&= \arg_{\boldsymbol{D}} \left\{ \sum_i \left[1 - \sum_{j \neq i} \begin{bmatrix} \frac{\partial}{\partial D_1} \\ \frac{\partial}{\partial D_2} \\ \vdots \ \frac{\partial}{\partial D_n} \end{bmatrix} \int_{D_i} \frac{1}{\sqrt{(2\pi)^n |\boldsymbol{M}|}} \exp -\frac{1}{2} (\boldsymbol{x} - \mathsf{E}\boldsymbol{x})^T \boldsymbol{M}^{-1} (\boldsymbol{x} - \mathsf{E}\boldsymbol{x}) d\boldsymbol{z} \right] = 0 \right\} \\
&= \arg_{\boldsymbol{D}} \left\{ \sum_i \left[1 - \sum_{j \neq i} \underbrace{\begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}}_{\text{Jacobian matrix}} \right] \right\}
\end{aligned}
$$

For two variable Gaussian …

$$
\begin{aligned}
cost &= \min_{\boldsymbol{D}} \sum_i \int_{D_i} \mathsf{p}(\boldsymbol{x}|H \neq h_i) \underbrace{\mathsf{p}(H \neq h_i)}_{c} d\boldsymbol{x} \\
&= \min_{\boldsymbol{D}} c \sum_i \int_{D_i} \mathsf{p}(\boldsymbol{x}|H \neq h_i) d\boldsymbol{x} \\
&= \min_{\boldsymbol{D}} c \sum_i \left[1 - \sum_{j \neq i} \int_{D_i} \mathsf{p}(\boldsymbol{x}|H = h_i) d\boldsymbol{x} \right]
\end{aligned}
$$

Copyright © 2019 Daniel J. Greenhoe

$$= \min_{\boldsymbol{D}} c \sum_i \left[1 - \sum_{j \neq i} \int_{D_i} \frac{1}{2\pi\sqrt{|M|}} \exp\left(\frac{z_1^2 \mathsf{E}[z_2 z_2] - 2 z_1 z_2 \mathsf{E}[z_1 z_2] + z_2^2 \mathsf{E}[z_1 z_1]}{-2|M|} \right) d\boldsymbol{z} \right]$$

21.4 2 hypothesis, 2 sensor detection

Theorem 21.1 (centralized case)**.** *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a probability space. Let* $D \subsetneq \mathbb{E}$ *be the* DECISION REGION *indicating hypothesis* $H = h_1$*. Let* $\pi_0 \triangleq \mathsf{P}\{H = h_0\}$ *and* $\pi_1 \triangleq \mathsf{P}\{H = h_1\}$*.*

THM

$$\begin{aligned} D &= \arg\min_D \left[\underbrace{\mathsf{P}\left\{(x,y) \in D | H = h_0\right\} \pi_0}_{\text{error for } H = h_0} + \underbrace{\mathsf{P}\left\{(x,y) \in D^{\mathsf{c}} | H = h_1\right\} \pi_1}_{\text{error for } H = h_1} \right] \\ &= \arg\min_D \left[\underbrace{\pi_0 \int_D \mathrm{p}_0(x,y)\ \mathrm{d}x\ \mathrm{d}y}_{\text{error for } H = h_0} + \underbrace{\pi_1 \int_D \mathrm{p}_1(x,y)\ \mathrm{d}x\ \mathrm{d}y}_{\text{error for } H = h_1} \right] \end{aligned}$$

✎PROOF:

$$\begin{aligned} D &= \arg\min_D \left[\mathsf{P}\{\text{error}\}\right] && \text{by definition of decision region } D \\ &= \arg\min_D \left[\mathsf{P}\{\text{error} \wedge H = h_0\} + \mathsf{P}\{\text{error} \wedge H = h_1\}\right] \\ &= \arg\min_D \left[\mathsf{P}\left\{\text{error} | H = h_0\right\} \pi_0 + \mathsf{P}\left\{\text{error} | H = h_1\right\} \pi_1\right] \\ &= \arg\min_D \left[\mathsf{P}\left\{(x,y) \in D | H = h_0\right\} \pi_0 + \mathsf{P}\left\{(x,y) \in D^{\mathsf{c}} | H = h_1\right\} \pi_1\right] \\ &= \arg\min_D \left[\pi_0 \int_D \mathrm{p}_0(x,y)\ \mathrm{d}x\ \mathrm{d}y + \pi_1 \int_D \mathrm{p}_1(x,y)\ \mathrm{d}x\ \mathrm{d}y\right] \end{aligned}$$

✎

Example 21.1. In the centralized case, the decision regions D in the xy-plane can be any arbitrary shape, as illustrated to the right.

Definition 21.2.

DEF

Let $\mathbf{P}_x$ *and* $\mathbf{P}_y$ *be* ***set projection operators*** *such that*

$$\begin{aligned} D_x &\triangleq \mathbf{P}_x D \\ D_y &\triangleq \mathbf{P}_y D \end{aligned}$$

Proposition 21.1. *Let* $+$ *represent* MINKOWSKI ADDITION

PRP

$$\mathbf{D} = \mathbf{D}_x + \mathbf{D}_y$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Theorem 21.2 (distributed AND case)**.** *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a probability space. Let* $D \subsetneq \mathbb{E}$ *be the* DECISION REGION *indicating hypothesis* $H = h_1$*. Let* $\pi_0 \triangleq \mathsf{P}\{H = h_0\}$ *and* $\pi_1 \triangleq \mathsf{P}\{H = h_1\}$*. Let* $E \triangleq D^c$*.*

THM

$$D = \arg\min_D \begin{pmatrix} \mathsf{P}\{ & x \in E, & y \in E & \}\{H = h_1\}\pi_1 & + \\ \mathsf{P}\{ & x \in E, & y \in D & \}\{H = h_1\}\pi_1 & + \\ \mathsf{P}\{ & x \in D, & y \in E & \}\{H = h_1\}\pi_1 & + \\ \mathsf{P}\{ & x \in D, & y \in D & \}\{H = h_0\}\pi_0 & \end{pmatrix}$$

✎PROOF:

x	y	H	$x \wedge y$	
0	0	0	0	
0	1	0	0	
1	0	0	0	
1	1	0	1	error
0	0	1	0	error
0	1	1	0	error
1	0	1	0	error
1	1	1	1	

$$\begin{aligned} D &= \arg\min_D [\mathsf{P}\{\text{error}\}] && \text{by definition of decision region } D \\ &= \arg\min_D \left[\mathsf{P}\{\text{error} \wedge H = h_0\} + \mathsf{P}\{\text{error} \wedge H = h_1\}\right] \\ &= \arg\min_D \left[\mathsf{P}\left\{\text{error}|H = h_0\right\}\pi_0 + \mathsf{P}\left\{\text{error}|H = h_1\right\}\pi_1\right] \\ &= \arg\min_D \begin{pmatrix} \mathsf{P}\{ & x \in E_x, & y \in E_y & \}\{H = h_1\}\pi_1 & + \\ \mathsf{P}\{ & x \in D_x, & y \in E_y & \}\{H = h_1\}\pi_1 & + \\ \mathsf{P}\{ & x \in E_x, & y \in D_y & \}\{H = h_1\}\pi_1 & + \\ \mathsf{P}\{ & x \in D_x, & y \in D_y & \}\{H = h_0\}\pi_0 & \end{pmatrix} \end{aligned}$$

✏

Example 21.2. In the distributed AND case, the decision regions D in the xy-plane are only simple rectangular shapes, as illustrated to the right.

Proposition 21.2.

PRP

In general, distributed AND *detection is suboptimal.*

✎PROOF: Because only rectangular decision regions are possible, detection is suboptimal. ✏

Theorem 21.3. [1]

THM

For the distributed AND *detection*

$$D_x = \left\{x|\pi_0 \int_{D_y} \mathrm{p}_0(x,y)\ \mathrm{d}x\ \mathrm{d}y \le \pi_1 \int_{D_y} \mathrm{p}_1(x,y)\ \mathrm{d}x\ \mathrm{d}y\right\}$$

[1] Willett et al. (2000), page 3268

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$D_x = \left\{x | y \in D_y \quad \Longrightarrow \quad \mathsf{P}\left\{(x,y)\,|H = h_0\right\}\pi_0 \le \mathsf{P}\left\{(x,y)\,|H = h_1\right\}\pi_1\right\}$$
$$= \left\{x | \pi_0 \int_{D_y} \mathrm{p}_0(x,y)\ \mathrm{d}x\ \mathrm{d}y \le \pi_1 \int_{D_y} \mathrm{p}_1(x,y)\ \mathrm{d}x\ \mathrm{d}y\right\}$$

✏

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

Part V

Appendices

APPENDIX A

PROBABILITY SPACE

Figure A.1: Lattice of mathematical spaces

❝ *It is not certain that everything is certain.*❞
Blaise Pascal (1623–1662), mathematician [1]

[1] quote: http://en.wikiquote.org/wiki/Blaise_Pascal
image: http://en.wikipedia.org/wiki/Image:Blaise_pascal.jpg

A.1 Probability functions

Definition A.1. [2] *Let* $(X, \vee, \wedge, \neg, 0, 1; \leq)$ *be a* LATTICE WITH NEGATION.

DEF

The function P *is a* ***probability function*** *if*

(1).	$\mathsf{P}(1)$	$=$	1		(NORMALIZED)	*and*
(2).	$\mathsf{P}(x)$	$\geq$	0	$\forall x{\in}X$	(NONNEGATIVE)	*and*
(3).	$\bigwedge_{n=1}^{\infty} x_n = 0 \implies \mathsf{P}\left(\bigvee_{n=1}^{\infty} x_n\right)$	$=$	$\sum_{n=1}^{\infty} \mathsf{P}(x_n)$	$\forall x_n{\in}X$	(σ-ADDITIVE)	.

Remark A.1. The advantage of this definition is that P is a *measure*, and hence all the power of measure theory is subsequently at one's disposal in using P. However, it has often been argued that the requirement of *σ-additivity* is unnecessary for a probability function. Even as early as 1930, de Finetti argued against it, in what became a kind of polite running debate with Fréchet.[3] In fact, Kolmogorov himself provided some argument against *σ-additivity* when referring to the closely related *Axiom of Continuity* saying, "Since the new axiom is essential for infinite fields of probability only, it is almost impossible to elucidate its empirical meaning…For, in describing any observable random process we can obtain only finite fields of probability.…" But in its support he added, "This limitation has been found expedient in researches of the most diverse sort."[4]

There are several other definitions of probability that only require *additivity* rather than *σ-additivity*. On a *Boolean lattice*, the **traditional probability** function is defined as[5]

(1).	$\mathsf{P}(1)$	$=$	1		(*normalized*)	and
(2).	$\mathsf{P}(x)$	$\geq$	0	$\forall x{\in}X$	(*nonnegative*)	and
(3).	$x \wedge y = 0 \implies \mathsf{P}(x \vee y)$	$=$	$\mathsf{P}(x) + \mathsf{P}(y)$	$\forall x,y{\in}X$	(*additive*)	.

This definition implies (on a *Boolean lattice*) that

(a).	$\mathsf{P}(0)$	$=$	0		(*nondegenerate*)	and
(b).	$\mathsf{P}(x)$	$\leq$	1	$\forall x{\in}X$	(*upper bounded*)	and
(c).	$\mathsf{P}(x)$	$=$	$1 - \mathsf{P}(x^{\perp})$	$\forall x{\in}X$		and
(d).	$\mathsf{P}(x \vee y)$	$\leq$	$\mathsf{P}(x) + \mathsf{P}(y)$	$\forall x,y{\in}X$	(*subadditive*)	and
(e).	$\mathsf{P}(x \vee y)$	$=$	$\mathsf{P}(x) + \mathsf{P}(y) - \mathsf{P}(x \wedge y)$	$\forall x,y{\in}X$		and
(f).	$x \leq y \implies \mathsf{P}(x)$	$\leq$	$\mathsf{P}(y)$	$\forall x,y{\in}X$	(*monotone*)	.

On a *distributive pseudocomplemented lattice*, the **generalized probability** function has been defined as[6]

(1).	$\mathsf{P}(0)$	$=$	0		(*nondegenerate*)	and
(2).	$\mathsf{P}(1)$	$=$	1		(*normalized*)	and
(3).	$0 \leq \mathsf{P}(1)$	$\leq$	1			and
(4).	$\mathsf{P}(x \vee y)$	$=$	$\mathsf{P}(x) + \mathsf{P}(y) - \mathsf{P}(x \wedge y)$	$\forall x,y{\in}X$		.

On an *orthomodular lattice*, or a *finite modular lattice*, the **quantum probability** function is defined as[7]

(1).	$\mathsf{P}(0)$	$=$	0		(*nondegenerate*)	and
(2).	$\mathsf{P}(1)$	$=$	1		(*normalized*)	and
(3).	$x \perp y \implies \mathsf{P}(x \vee y)$	$=$	$\mathsf{P}(x) + \mathsf{P}(y)$	$\forall x,y{\in}X$	(*additive*)	.

However, for lattices that are not *distributive*, *modular*, or *orthomodular*, none of these definitions

[2] Billingsley (1995) pages 22–23 ⟨Probability Measures⟩, Kolmogorov (1933a), Kolmogorov (1933b), page 16 ⟨*field of probability*⟩, Pap (1995) pages 8–9 ⟨Definition 2.3(13)⟩, Kalmbach (1986) page 27

[3] de Finetti (1930a), Fréchet (1930a), de Finetti (1930b), Fréchet (1930b), de Finetti (1930c), Cifarelli and Regazzini (1996) pages 258–260

[4] Kolmogorov (1933b), page 15

[5] Papoulis (1991) pages 21–22, Kolmogorov (1933b), page 2 ⟨§1. Axioms I–V⟩

[6] Narens (2014) page 118, Narens (2007)

[7] Greechie (1971) page 126 ⟨DEFINITIONS⟩, Narens (2014) page 118

COPYRIGHT © 2019 DANIEL J. GREENHOE

work out so well. Take for example the O_6 *lattice* with the "very reasonable" probability function given in Example **??** (page **??**). This probability space $(\mathrm{O}_6, \mathsf{P})$ fails to be any of the 4 probability functions defined in this Remark. It fails to be a *measure-theoretic* or *traditional probability* function because

$$a \wedge b = 0 \qquad \text{but} \qquad \mathsf{P}(a \vee b) = \mathsf{P}(1) = 1 \neq \tfrac{1}{3} + \tfrac{1}{2} = \mathsf{P}(a) + \mathsf{P}(b)\,.$$

It fails to be a *generalized probability* function because

$$\mathsf{P}(a \vee b) = \mathsf{P}(1) = 1 \neq \tfrac{1}{3} + \tfrac{1}{2} - 0 = \mathsf{P}(a) + \mathsf{P}(b) - \mathsf{P}(0) = \mathsf{P}(a) + \mathsf{P}(b) - \mathsf{P}(a \wedge b)\,.$$

It fails to be an *quantum probability* function because

$$a \perp b = 0 \qquad \text{but} \qquad \mathsf{P}(a \vee b) = \mathsf{P}(1) = 1 \neq \tfrac{1}{3} + \tfrac{1}{2} = \mathsf{P}(a) + \mathsf{P}(b)\,.$$

In each of these cases, the function P fails to be *additive*. The solution of Definition A.1 (page 166) is simply to "switch off" *additivity* when the lattice is not *distributive*. This method is a little "crude", but at least it allows us to define probability on a very wide class of lattices, while retaining compatibility with the *Boolean* case.

A.2 Probability Space

In mathematics, a *space* is simply a set and in the most general definition, nothing else. However, normally for a space to actually be useful, some additional structure is added. One of the most general additional structures is a *topology*; and a space together with a topology is called a *topological space*. A topological space imposes additional structure on a space in the form of subsets and guarantees that these subsets are closed under such fundamental operations as set *union* and set *intersection*. With the additional structure available in a topological space, we are able to analyze such basic concepts as continuity, convergence, and connectivity.

However for a great number of mathematical applications, we need to *measure* mathematical objects—the most general measurement being measures on subsets of some set. Examples of measurement in mathematics include integration and probability. Before measurement can be effectively performed on a set, the set must be equipped with a subset structure. In analysis, arguably the most fundamental subset structure is the humble *topology* (Definition **??** page **??**). However, a simple topology does not provide sufficient structure for effective measurement. For example, often we would not only like to measure some subset A, but also its complement A^{c}. A topology is not closed under the complement operation. So instead of a topology only, we equip the space with a more powerful (and thus less general) structure called a *σ-algebra* (*sigma-algebra*) (Definition **??** page **??**). A *σ-algebra* is a subset structure that is closed under set complement. A set together with a *σ-algebra* is called a *measurable space*. And a set together with a *σ-algebra* and a *measure* on that *σ-algebra* is called a *measure space* (Definition **??** page **??**).

The next definition presents a very important measure space—the *probability space*.

Definition A.2.

DEF

The triple $(\Omega, \mathbb{E}, \mathsf{P})$ *is a* ***probability space*** *if*

(1). Ω *is a* SET *and*

(2). $\mathbb{E}$ *is a* σ-ALGEBRA *on* Ω *(Definition* **??** *page* **??***) and*

(3). $\mathsf{P} : \mathbb{E} \to [0,1]$ *is a* MEASURE *on* $\mathbb{E}$ *(Definition* **??** *page* **??***) .*

If $\boldsymbol{S} \triangleq (\Omega, \mathbb{E}, \mathsf{P})$ *is a* PROBABILITY SPACE *then* x *is an* ***outcome*** *in* $\boldsymbol{S}$ *if* $x \in \Omega$, A *is an* ***event*** *in* $\boldsymbol{S}$ *if* $A \in \mathbb{E}$, *and* $\mathsf{P}A$ *is the* ***probability*** *of* A *in* $\boldsymbol{S}$ *if* A *is an* EVENT *in* $\boldsymbol{S}$.

Definition A.3. [8] *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE *(Definition A.2 page 167)*.

[8] Papoulis (1990) page 52 ⟨Independent Events⟩

DEF *Two* EVENTS *A and B in $\mathbb{E}$ are* ***independent*** *if*

$$\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$$

Definition A.4. [9] *Let $(\Omega, \mathbb{E}, \mathsf{P})$ be a* PROBABILITY SPACE *(Definition A.2 page 167). Let x and y be* EVENTS *in L.*

DEF *The* ***conditional probability of*** *x* ***given*** *y is defined as*

$$\mathsf{P}(A|B) \triangleq \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}$$

A.3 Properties

Proposition A.1.

PRP $\underbrace{(\Omega, \mathbb{E}, \mathsf{P}) \text{ is a PROBABILITY SPACE} \implies (\Omega, \mathbb{E}, \mathsf{P}) \text{ is a MEASURE SPACE}}_{\text{(every probability space is a measure space)}}$

Theorem A.1. [10] *Let $(\Omega, \mathbb{E}, \mathsf{P})$ be a* PROBABILITY SPACE *(Definition A.2 page 167).*

THM

(1).	0	$\leq$	$\mathsf{P}(x) \leq 1$	$\forall x \in X$	(BOUNDED)	*and*
(2).	$\mathsf{P}(x)$	$=$	$1 - \mathsf{P}(x^\perp)$	$\forall x \in X$	(PARTITION OF UNITY)	*and*
(3).	$x \leq y$	$\implies$	$\mathsf{P}(y^\perp) \leq \mathsf{P}(x^\perp)$	$\forall x,y \in X$	(ANTITONE)	

✎PROOF:

1. Proof for $0 \leq \mathsf{P}(x) \leq 1$:

$$\begin{aligned}
\boxed{0} &= \mathsf{P}(0) && \text{by by } \textit{nondegenerate} \text{ property of } \mathsf{P} \text{ (Definition A.2 page 167)} \\
&\leq \boxed{\mathsf{P}(x)} && \text{because } 0 \leq x \text{ and } \textit{monotone} \text{ property of } \mathsf{P} \\
&\leq \mathsf{P}(1) && \text{because } x \leq 1 \text{ and } \textit{monotone} \text{ property of } \mathsf{P} \\
&= \boxed{1} && \text{by } \textit{normalized} \text{ property of } \mathsf{P}
\end{aligned}$$

2. Proof for $\mathsf{P}(x) = 1 - \mathsf{P}(x^\perp)$:
   (a) Proof that P is *additive* (Definition A.2 page 167) over $\{0, x, x^\perp\} \subseteq X$:
      i. $\{0, x, x^\perp\}$ is *distributive*.
      ii. $x \wedge x^\perp = 0$ for all $x \in X$ by the *non-contradiction* property of *orthocomplemented lattices*.
      iii. Therefore, by Definition A.2, P is *additive* over $\{0, x, x^\perp\}$.
   (b) Then …

$$\begin{aligned}
1 - \mathsf{P}(x^\perp) &= \mathsf{P}(1) - \mathsf{P}(x^\perp) && \text{by } \textit{normalized} \text{ property of } \mathsf{P} && \text{(Definition A.2 page 167)} \\
&= \mathsf{P}(x \vee x^\perp) - \mathsf{P}(x^\perp) && \text{by } \textit{excluded middle} \text{ property of ortho. lat.} \\
&= \mathsf{P}(x) + \mathsf{P}(x^\perp) - \mathsf{P}(x^\perp) && \text{by } \textit{additive} \text{ property of } (\Omega, \mathbb{E}, \mathsf{P}) && \text{(item (2a) page 168)} \\
&= \mathsf{P}(x) && \text{by field property of } (\mathbb{R}, +, \cdot, 0, 1)
\end{aligned}$$

3. Proof for $x \leq y \implies \mathsf{P}(y^\perp) \leq \mathsf{P}(x^\perp)$:

$$\begin{aligned}
x \leq y &\implies y^\perp \leq x^\perp && \text{by } \textit{antitone} \text{ property of } \textit{orthocomplemented lattices} \\
&\implies \mathsf{P}(y^\perp) \leq \mathsf{P}(x^\perp) && \text{by } \textit{monotone} \text{ property of } \mathsf{P} && \text{(Definition A.2 page 167)}
\end{aligned}$$

[9] Papoulis (1990) page 45 ⟨2-3 Conditional Probability and Indenpendence⟩
[10] property (1): Papoulis (1991) page 21 ⟨(2-11)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

Theorem A.2. [11] *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE *(Definition A.2 page 167)*.

THM

$$\left.\begin{array}{l} \boldsymbol{L} \text{ is BOOLEAN} \\ \text{(Definition ?? page ??)} \end{array}\right\} \Longrightarrow \left\{\begin{array}{lllll} 1. & \mathsf{P}(x \vee y) & = & \mathsf{P}(x) + \mathsf{P}(y) - \mathsf{P}(x \wedge y) & \forall x,y \in X \quad \text{and} \\ 2. & \mathsf{P}(x \vee y) & \leq & \mathsf{P}(x) + \mathsf{P}(y) & \forall x,y \in X \quad \text{(BOOLE'S INEQUALITY)} \end{array}\right.$$

✎PROOF:

1. lemma: Proof that $\mathsf{P}((\neg x) \wedge y) = \mathsf{P}(y) - \mathsf{P}(x \wedge y)$:

$$\begin{aligned} \mathsf{P}(y) - \mathsf{P}(xy) &= \mathsf{P}(1 \wedge y) - \mathsf{P}(xy) && \text{by definition of } 1 \text{ and } \wedge \\ &= \mathsf{P}\big[(x \vee x^{\perp})y\big] - \mathsf{P}(xy) && \text{by } \textit{excluded middle} \text{ property of } \textit{Boolean lattices} \\ &= \mathsf{P}(xy \vee x^{\perp}y) - \mathsf{P}(xy) && \text{by } \textit{distributive} \text{ property of } \textit{Boolean lattices} \\ &= \mathsf{P}(xy) + \mathsf{P}(x^{\perp}y) - \mathsf{P}(xy) && \text{because } (xy)(x^{\perp}y) = 0 \text{ and by } \textit{additive} \text{ property} \\ &= \mathsf{P}(x^{\perp}y) \end{aligned}$$

2. Proof that $\mathsf{P}(x \vee y) = \mathsf{P}(x) + \mathsf{P}(y) - \mathsf{P}(x \wedge y)$:

$$\begin{aligned} \mathsf{P}(x \vee y) &= \mathsf{P}(x \vee x^{\perp}y) && \text{by property of } \textit{Boolean lattices} \\ &= \mathsf{P}(x) + \mathsf{P}(x^{\perp}y) && \text{because } (x)(x^{\perp}y) = 0 \text{ and by } \textit{additive} \text{ property} \\ &= \mathsf{P}(x) + \mathsf{P}(y) - \mathsf{P}(x \wedge y) && \text{by item (1) (page 169)} \end{aligned}$$

Theorem A.3 (sum of products)**.** *Let* $(X, \vee, \wedge, 0, 1; \leq)$ *be a* BOUNDED LATTICE, $(\Omega, \mathbb{E}, \mathsf{P})$ *a* PROBABILITY SPACE *(Definition A.2 page 167)*, *and* $\{y, x_1, x_2, x_3, \ldots\}$ *a subset of* X.

THM

$$\left\{\begin{array}{lll} 1. & \boldsymbol{L} \text{ is DISTRIBUTIVE} & \text{and} \\ 2. & \{x_1, x_2, \ldots\} \text{ is a PARTITION of } y & \end{array}\right\} \Longrightarrow \left\{\begin{array}{lllll} 1. & \mathsf{P}(y) & = & \sum_n \mathsf{P}(x_n) & \text{and} \\ 2. & \mathsf{P}(y) & = & \sum_n \mathsf{P}(y \wedge x_n) & \text{and} \\ 3. & \mathsf{P}(z \wedge y) & = & \sum_n \mathsf{P}(z \wedge x_n) & \end{array}\right\}$$

✎PROOF:

1. Proof that P is *additive* (Definition A.2 page 167) on $(\Omega, \mathbb{E}, \mathsf{P})$:

   (a) Proof that $(yx_n) \wedge (yx_m) = 0$ for $n \neq m$:

$$\begin{aligned} (yx_n) \wedge (yx_m) &= y(x_n x_m) && \text{by definition of } \wedge \\ &= y \wedge 0 && \text{by } \textit{mutually exclusive} \text{ property of } \textit{partitions} \\ &= 0 && \text{by } \textit{lower bounded} \text{ property of } \textit{bounded lattices} \end{aligned}$$

   (b) Proof that $\boldsymbol{L}$ is *distributive*: by *distributive* hypothesis

2. Proof that $\mathsf{P}(y) = \sum_n \mathsf{P}(x_n)$

$$\begin{aligned} \mathsf{P}(y) &= \mathsf{P}\big(yx_1 \vee yx_2 \vee \cdots \vee yx_n\big) \\ &= \sum_n \mathsf{P}\big(yx_n\big) && \text{by item (1) and } \textit{additive} \text{ property} && \text{(Definition A.2 page 167)} \\ &= \sum_n \mathsf{P}\big(y|x_n\big)\mathsf{P}(x_n) && \text{by } \textit{conditional probability} && \text{(Definition A.4 page 168)} \end{aligned}$$

[11] Papoulis (1991) page 21 ⟨(2-13)⟩, Feller (1970) pages 22–23 ⟨(7.4),(7.6)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

EX

$$\begin{aligned} \mathsf{P}(\varnothing) &= 0 \\ \mathsf{P}(A) &= (1-p)^2 \\ \mathsf{P}(B) &= 2p(1-p) \\ \mathsf{P}(C) &= p^2 \end{aligned} \qquad \begin{aligned} \mathsf{P}(\Omega) &= 1 \\ \mathsf{P}(A \cup B) &= (1-p)^2 + 2p(1-p) \\ \mathsf{P}(A \cup C) &= 2p^2 - 2p + 1 \\ \mathsf{P}(B \cup C) &= -p^2 + 2p \end{aligned}$$

Figure A.2: *Double coin toss* (Example A.2 page 170)

✏

As described in Definition A.2 (page 167), every *probability space* $(\Omega, \mathbb{E}, \mathsf{P})$ contains a probability *measure* $\mathsf{P} : \mathbb{E} \to [0,1]$. This probability *measure* has some basic properties as described in Theorem A.4 (next).

Theorem A.4. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE. *Let* B *be a set and* $\{B_n | n = 1, 2, \ldots, N\}$ *a set of sets.*

THM

$$\left\{ \begin{array}{l} \{B_n | n = 1, 2, \ldots, N\} \text{ is a} \\ \text{PARTITION of } B. \end{array} \right\} \implies \left\{ \begin{array}{lll} (1). & \mathsf{P}(B) = \sum_{n=1}^{N} \mathsf{P}(B_n) & \forall B \in \mathbb{E} \quad \text{and} \\ (2). & \mathsf{P}(AB) = \sum_{n=1}^{N} \mathsf{P}(AB_n) & \forall A, B \in \mathbb{E} \end{array} \right\}$$

✎PROOF: P is a *measure* and by Definition **??** (page **??**). ✏

Proposition A.2. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a probability space, and* X *a* RANDOM VARIABLE *(Definition B.1 page 176) with* PROBABILITY DENSITY FUNCTION $\mathsf{p}_{\mathsf{X}}(x)$ *and* CUMULATIVE DISTRIBUTION FUNCTION $\mathsf{c}_{\mathsf{X}}(x)$.

PRP

(1). $\mathsf{c}_{\mathsf{X}}(x)$ *is* MONOTONE *and*
(2). $\mathsf{p}_{\mathsf{X}}(x)$ *is* CONTINUOUS $\implies$ $\mathsf{c}_{\mathsf{X}}(x)$ *is* STRICTLY MONOTONE *and*
(3). $\mathsf{p}_{\mathsf{X}}(x)$ *is* CONTINUOUS $\implies$ $\mathsf{c}_{\mathsf{X}}(x)$ *is* INVERTIBLE

A.4 Examples

Example A.1 (single coin toss). Let ☺ represent "heads" and ● represent "tails" in a coin toss. Let $0 < p < 1$ be the probability of a head. A *probability space* $(\Omega, \mathbb{E}, \mathsf{P})$ for a single coin toss is as follows:

EX

$$\Omega = \{☺, ●\}$$

$$\mathbb{E} = \left\{ \underbrace{\varnothing}_{\text{neither}}, \underbrace{\{☺\}}_{\text{heads}}, \underbrace{\{●\}}_{\text{tails}}, \underbrace{\{☺, ●\}}_{\text{heads or tails}} \right\}$$

$$\mathsf{P}(X) = \left\{ \begin{array}{lll} 0 & \text{for } X = \varnothing & \text{(neither heads nor tails)} \\ p & \text{for } X = \{☺\} & \text{(heads)} \\ 1-p & \text{for } X = \{●\} & \text{(tails)} \\ 1 & \text{for } X = \{☺, ●\} & \text{(either heads or tails)} \end{array} \right\}$$

$$\begin{aligned} \mathsf{P}(\{☺, ●\}) &= 1 \\ \mathsf{P}(\{☺\}) &= p \\ \mathsf{P}(\{●\}) &= 1-p \\ \mathsf{P}(\varnothing) &= 0 \end{aligned}$$

Example A.2 (Double coin toss). Let ☺ represent "heads" and ● represent "tails" in a double coin toss in which each toss is *independent* (Definition A.3 page 167) of the other. Let $0 < p < 1$ be the probability of a head. The *probability space* $(\Omega, \mathbb{E}, \mathsf{P})$ is illustrated in Figure A.2 (page 170).

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$\mathsf{P}(\Omega) = 1$	by *normalized* property of P	(Definition A.1 page 166)
$\mathsf{P}(C) = \mathsf{P}\{☺☺\}$	by definition of C	
$= \mathsf{P}(☺)\mathsf{P}(☺)$	by definition of *independence*	(Definition A.3 page 167)
$= p^2$	by definition of p	
$\mathsf{P}(A) = \mathsf{P}\{●●\}$	by definition of A	
$= \mathsf{P}(●)\mathsf{P}(●)$	by definition of *independence*	(Definition A.3 page 167)
$= \{1 - \mathsf{P}(☺)\}\{1 - \mathsf{P}(☺)\}$	by *antitone* property of P	(Theorem A.1 page 168)
$= (1-p)^2$	by definition of p	
$\mathsf{P}(B) = \mathsf{P}\{(●☺), (☺●)\}$	by definition of B	
$= \mathsf{P}\{●☺\} + \mathsf{P}\{☺●\}$	by *additive* property of P	(Definition A.1 page 166)
$= \mathsf{P}\{●\}\mathsf{P}\{☺\} + \mathsf{P}\{☺\}\mathsf{P}\{●\}$	by definition of *independence*	(Definition A.3 page 167)
$= (1-p)p + p(1-p)$	by *antitone* property of P	(Theorem A.1 page 168) and definition of p
$= -2p^2 + p + 1$		
$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$	by Theorem A.2	
$= \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(\varnothing)$		
$= (1-p)^2 + (-2p^2 + p + 1) + 0$	by previous results	
$= -p^2 - p + 1$		
$\mathsf{P}(\varnothing) = 0$	by *nondegenerate* property of P	(Definition A.1 page 166)

✏

Figure A.3: *even/odd die probability space* (Example A.3 page 171)

Example A.3 (even/odd die toss). The *probability space* for an **even/odd die toss**, with $0 < p < 1$ being the probability of the die toss being odd, is illustrated in Figure A.3 (page 171).

✎PROOF:

$\mathsf{P}(\Omega) = 1$	by *normalized* property of P	(Definition A.1 page 166)
$\mathsf{P}(C) = \mathsf{P}\{☺☺\}$	by definition of C	
$= \mathsf{P}(☺)\mathsf{P}(☺)$	by definition of *independence*	(Definition A.3 page 167)
$= p^2$	by definition of p	
$\mathsf{P}(A) = \mathsf{P}\{⚀, ⚂, ⚄\}$	by definition of A	
$= p$	by definition of p	
$\mathsf{P}(B) = \mathsf{P}\{⚁, ⚃, ⚅\}$	by definition of B	
$= \mathsf{P}\{⚀, ⚂, ⚄\}^c$	by definition of set complement c	
$= \mathsf{P}A^c$	by definition of A	

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

$$
\begin{aligned}
&= \mathsf{P}(\neg A) && \text{by definition of } \neg \\
&= 1 - \mathsf{P}(A) && \text{by Theorem A.1 page 168} \\
&= 1 - p && \text{by definition of } p \\
\mathsf{P}(\varnothing) &= 0 && \text{by } \textit{nondegenerate} \text{ property of } \mathsf{P} \qquad \text{(Definition A.1 page 166)}
\end{aligned}
$$

✏

The two previous *even/odd die* example (Example A.5 page 172) is in essence the same as the *single coin toss* (Example A.1 page 170). The next offers a little more complexity.

EX

$$
\begin{aligned}
\mathsf{P}(\varnothing) &= 0 & \mathsf{P}(\Omega) &= 1 \\
\mathsf{P}(A) &= (1-p)^2 & \mathsf{P}(A \cup B) &= (1-p)^2 + 2p(1-p) \\
\mathsf{P}(B) &= 2p(1-p) & \mathsf{P}(A \cup C) &= 2p^2 - 2p + 1 \\
\mathsf{P}(C) &= p^2 & \mathsf{P}(B \cup C) &= -p^2 + 2p
\end{aligned}
$$

Figure A.4: 3-4-2 die example (Example A.4 page 172)

Example A.4. Suppose we have a "fair" die and we are primarily interested in the events of the first three { ⚀, ⚁, ⚂ }, the next two, { ⚃, ⚄ } and the final one { ⚅ }. The resulting *probability space* is illustrated in Figure A.4 (page 172).

The two previous examples (Example A.5 page 172, Example A.4 page 172) illustrate a *probabilty spaces* in which the events are *mutually exclusive*. The (next) illustrates one where events are *not*.

Example A.5. Suppose we have a "fair" die and we are primarily interested in the events of the first four ({⚀,⚁,⚂,⚃}) (that is, whether one roll of the die will produce a value in the set { ⚀, ⚁, ⚂, ⚃ }) and the last three ({⚃,⚄,⚅}) However, these events do not by themselves form a σ-*algebra*. Rather under the $\cap$ and $\cup$ operations, these two events generate a total of eight possible events that together form a σ-*algebra*. The resulting *probability space* is illustrated in Figure A.5 (page 173).

But why go through all the trouble of requiring a σ-algebra? Having a σ-algebra in place ensures that anything we might possibly want to measure *can* be measured. It makes sure all possible combinations are taken into account. And why go through the additional trouble of requiring a measure space? With a measure space available, expressing the measure over a complex set is often greatly simplified because the measure space provides nice algebraic properties (namely the σ-*additive* property). Example A.6 (next) illustrates how a rather complex σ-algebra (64 elements) can be compactly represented in a measure space.

Example A.6. Suppose we have a "fair" dice and we are interested in measuring over the power set of events (largest possible algebra—$2^6 = 64$ events). This leads to the probability space $(\Omega, \mathbb{E}, \mathsf{P})$ where

EX

$$
\begin{aligned}
\Omega &= \{⚀,⚁,⚂,⚃,⚄,⚅\} \\
\mathbb{E} &= \mathcal{P}(\Omega) && \text{(the power-set of } \Omega) \\
\mathsf{P}(e) &= \tfrac{1}{6}|e| && (\tfrac{1}{6} \text{ times the number of possible outcomes in event } e)
\end{aligned}
$$

Example A.7 (Gaussian distribution on $\mathbb{R}$). Let $\boldsymbol{B}$ be the *Borel algebra on* $\mathbb{R}$. Let $\boldsymbol{L} \triangleq (\boldsymbol{B}, \subseteq)$ be the lattice formed by the elements of $\boldsymbol{B}$—this lattice is a *Boolean algebra*. Let

$$\mathsf{P}(A) \triangleq \frac{1}{\sqrt{2\pi\sigma^2}} \int_A e^{\frac{x^2}{2\sigma^2}}\, dx \text{ for } A \subseteq \boldsymbol{B}$$

and where $\int$ is the *Lebesgue integral* (Definition **??** page **??**). Then $(\boldsymbol{L}, \mathsf{P})$ is a **probability space**.

Copyright © 2019 Daniel J. Greenhoe

Figure A.5: First 4 / last 3 die example (Example A.5 page 172)

Example A.8 (Gaussian noise). Let $X \sim \mathsf{N}\left(0, \sigma^2\right)$ be a random variable with Gaussian distribution. We can construct the following probability space $(\Omega, \mathbb{E}, \mathsf{P})$:

EX

$$\begin{aligned}
\Omega &= \mathbb{R} \\
\mathbb{E} &= \{\varnothing, \Omega\} \cup \{(a,b) | a, b \in \mathbb{R}, a < b\} \\
\mathsf{P}x &= \begin{cases} 0 & \text{for } x = \varnothing \\ 1 & \text{for } x = \Omega \\ \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b e^{\frac{x^2}{2\sigma^2}}\, \mathrm{d}x & \text{otherwise} \end{cases}
\end{aligned}$$

Example A.9. The set of outcomes Ω can also be a set of waveforms:

EX

$\Omega =$

$$\begin{aligned}
\mathbb{E} &= \mathcal{P}(\Omega) \\
\mathsf{P}(e) &= \tfrac{1}{7}|e|
\end{aligned}$$

A.5 Probability subspaces

Example A.10. Suppose a random process is capable of producing three values $\Omega \triangleq \{x, y, z\}$. There are five *algebras of sets* on Ω and therefore five probability spaces $(\Omega, \mathbb{E}_n, \mathsf{P})$ on Ω with the five values of $\mathbb{E}_n$ listed below: [12]

[12] *algebra of sets*: Definition **??** page **??**

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

Figure A.6: Euclidean 3-dimensional space partitioned as a power lattice

$$
\begin{aligned}
\mathbb{E}_1 &= \{ \quad \varnothing, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad X \quad \} \\
\mathbb{E}_2 &= \{ \quad \varnothing, \quad \{x\}, \qquad\qquad\qquad\qquad\qquad \{y,z\}, \quad X \quad \} \\
\mathbb{E}_3 &= \{ \quad \varnothing, \qquad\quad \{y\}, \qquad\qquad\qquad \{x,z\}, \qquad\quad X \quad \} \\
\mathbb{E}_4 &= \{ \quad \varnothing, \qquad\qquad\quad \{z\}, \quad \{x,y\}, \qquad\qquad\qquad X \quad \} \\
\mathbb{E}_5 &= \{ \quad \varnothing, \quad \{x\}, \quad \{y\}, \quad \{z\}, \quad \{x,y\}, \quad \{x,z\}, \quad \{y,z\}, \quad X \quad \}
\end{aligned}
$$

Suppose the samples of Ω are generated by a physical process such that they are all equally likely to occur. Then by varying the *σ-algebra* $\mathbb{E}_n$ effectively varies the probability distribution of the probability space $(\Omega, \mathbb{E}_n, \mathsf{P})$. This is illustrated by the figure to the left.

Copyright © 2019 Daniel J. Greenhoe

APPENDIX B

PROBABILITY DENSITY FUNCTIONS

❝While writing my book I had an argument with Feller. He asserted that everyone said "random variable" and I asserted that everyone said "chance variable." We obviously had to use the same name in our books, so we decided the issue by a stochastic procedure. That is, we tossed for it and he won.❞

Joseph Leonard Doob (1910–2004), pioneer of and key contributor to mathematical probability[1]

B.1 Random variables

The concept of the *random variable* is widely used in probability and random processes. Before discussing what a *random variable is*, note two things that a *random variable* is *not* (next remark).

Remark B.1. [2] As pointed out by others, the term "random variable" is a "misnomer":

REM

- A *random variable* is **not random**.
- A *random variable* is **not a variable**.

What is it then? It is a *function* (next definition). In particular, it is a function that maps from an underlying stochastic process into $\mathbb{R}$. Any "*randomness*" (whatever that means) it may *appear* to have comes from the stochastic process it is mapping *from*. But the function itself (the *random*

[1] quote: Snell (1997), page 307, Snell (2005), page 251
image: http://www.dartmouth.edu/~chance/Doob/conversation.html

[2] Miller (2006) page 130, Feldman and Valdez-Flores (2010) page 4 ⟨"The name "random variable" is actually a misnomer, since it is not random and not a variable.…the *random variable* simply maps each point (outcome) in the sample space to a number on the real line…Technically, the space into which the *random variable* maps the sample space may be more general than the real line…"⟩, Curry and Feldman (2010) page 4, Trivedi (2016) page 2.1 ⟨"The term "random variable" is actually a misnomer, since a *random variable* X is really a function whose domain is the sample space S, and whose range is the set of all real numbers, $\mathbb{R}$."⟩

variable itself) is very deterministic and well-defined. What gives it the appearance of being random is that the outcome ω of the experiment appears to be random to the observer. So the *random variable* $\mathsf{X}(\omega)$ is simply a function of an underlying mechanism that appears to be random.

Definition B.1. [3] *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE *(Definition A.2 page 167)*.

DEF

A ***random variable*** X *is any function in* $\mathbb{R}^{\Omega}$.

B.2 Probability distributions

The probability information about σ-algebra $\mathbb{E}$ in a *probability space* (Definition A.2 page 167) is completely specified by *measure* P. However, sometimes it is more convenient to express this same *measure* information in terms of the *probability density function* or the *cummulative distribution function* of the *probability space*.

Definition B.2. [4] *Let* X *be a* RANDOM VARIABLE *(Definition B.1 page 176) on a* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$.

DEF

X *has* ***cummulative distribution function*** ***(cdf)*** $\mathsf{c}_{\mathsf{X}}(x)$ *if*

$$\mathsf{c}_{\mathsf{X}}(x) \triangleq \mathsf{P}\{x \in \mathbb{E} | \mathsf{X} < x\}$$

X ***probability density function*** ***(pdf)*** $\mathsf{p}_{\mathsf{X}}(x)$ *if*

$$\mathsf{p}_{\mathsf{X}}(x) \triangleq \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{X}}(x) \triangleq \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{P}\{x \in \mathbb{E} | \mathsf{X} < x\}$$

Remark B.2. Suppose X be a *random variable* on a *probability space* $(\Omega, \mathbb{E}, \mathsf{P})$. Note that

- Both X and $\mathbb{E}$ are *functions*.
- But X is a function that maps from Ω to $\mathbb{R}$,
- whereas P is a function that maps from $\mathbb{E}$ to $\mathbb{R}$.

Definition B.3. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE *(Definition A.2 page 167) and* X *and* $\mathsf{Y} : \Omega \to \mathbb{R}$ *random variables. Then a* ***joint probability density function*** $\mathsf{p}_{\mathsf{X}} : \mathbb{E} \times \Omega \to [0 : 1]$ *and a* ***joint cumulative distribution function*** $\mathsf{c}_{\mathsf{X}} : \mathbb{E} \times \Omega \to [0 : 1]$ *are defined as*

DEF

$$\mathsf{c}_{\mathsf{XY}}(x, y) \triangleq \mathsf{P}\{\mathsf{X} \le x | \mathsf{Y} \le y\} \quad \text{(JOINT CUMULATIVE DISTRIBUTION FUNCTION)}$$

$$\mathsf{p}_{\mathsf{XY}}(x, y) \triangleq \frac{\mathsf{d}}{\mathsf{d}y}\frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{XY}}(x, y) \quad \text{(JOINT PROBABILITY DENSITY FUNCTION)}$$

Definition B.4. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE *(Definition A.2 page 167) and* X *a random variable. Then a* ***conditional probability density function*** $\mathsf{p}_{\mathsf{X}} : \mathbb{E} \times \Omega \to [0 : 1]$ *and a* ***conditional cumulative distribution function*** $\mathsf{c}_{\mathsf{X}} : \mathbb{E} \times \Omega \to [0 : 1]$ *are defined as*

DEF

$$\mathsf{c}_{\mathsf{X}}(x|y) \triangleq \mathsf{P}\{\mathsf{X} \le x | \mathsf{Y} = y\} \quad \text{(CONDITIONAL CUMULATIVE DISTRIBUTION FUNCTION—CDF)}$$

$$\mathsf{p}_{\mathsf{X}}(x|y) \triangleq \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{X}}(x|y) \quad \text{(CONDITIONAL PROBABILITY DENSITY FUNCTION—PDF)}$$

B.3 Properties

Definition B.2 (page 176) defines the pdf and cdf of a *probability space* $(\Omega, \mathbb{E}, \mathsf{P})$ in terms of *measure* P. Conversely, the probability *measure* $\mathsf{P}\{a \le \mathsf{X} < b\}$ of an event $\{a \le \mathsf{X} < b\}$ can be expressed in terms of either the pdf or cdf.

[3] Papoulis (1991), page 63

[4] von der Linden et al. (2014) page 93 ⟨Definitions 7.1, 7.2⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

Proposition B.1. *Let* X *a* RANDOM VARIABLE *with* PDF p_{X} *and* CDF c_{X} *(Definition B.2 page 176) on the* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$ *(Definition A.2 page 167)*.

PRP

$$\left\{\begin{array}{ll} (1). & \mathsf{c}_{\mathsf{X}}(x) \text{ and } \mathsf{c}_{\mathsf{Y}}(y) \text{ are CONTINUOUS} \quad \text{OR} \\ (2). & \mathsf{p}_{\mathsf{X}}(x) \text{ and } \mathsf{p}_{\mathsf{Y}}(y) \text{ are CONTINUOUS} \end{array}\right\}$$
$$\Longrightarrow \left\{\begin{array}{rcl} \mathsf{p}_{\mathsf{X}}(x) & = & \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathsf{P}\{x \le \mathsf{X} < x+\varepsilon\} \\ \mathsf{p}_{\mathsf{XY}}(x,y) & = & \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathsf{P}\{x \le \mathsf{X} < x+\varepsilon \wedge y \le \mathsf{Y} < y+\varepsilon\} \end{array}\right\}$$

✎PROOF:

$$\begin{aligned} \mathsf{p}_{\mathsf{X}}(x) &\triangleq \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{X}}(x) && \text{by definition of } \mathsf{p}_{\mathsf{X}} && \text{(Definition B.2 page 176)} \\ &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathsf{P}\{x \in \mathbb{R} | x \le \mathsf{X} < x+\varepsilon\} && \text{by definition of } \frac{\mathsf{d}}{\mathsf{d}x} && \text{(Definition ?? page ??)} \end{aligned}$$

✏

Theorem B.1. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a probability space,* X *be a random variable, and* (a, b) *a real interval.*

THM

$$\left\{\begin{array}{ll} (1). & \mathsf{c}_{\mathsf{X}}(x) \text{ is CONTINUOUS} \quad \text{OR} \\ (2). & \mathsf{p}_{\mathsf{X}}(x) \text{ is CONTINUOUS} \end{array}\right\} \Longrightarrow \left\{ \mathsf{P}\{a < \mathsf{X} \le b\} = \mathsf{c}_{\mathsf{X}}(b) - \mathsf{c}_{\mathsf{X}}(a) = \int_a^b \mathsf{p}_{\mathsf{X}}(x)\,\mathsf{d}x \right\}$$

✎PROOF:

$$\begin{aligned} \mathsf{P}\{a < \mathsf{X} \le b\} &= \mathsf{P}\{\mathsf{X} \le b\} - \mathsf{P}\{\mathsf{X} < a\} && \text{by } \textit{sum of products} && \text{(Theorem A.3 page 169)} \\ &= \mathsf{P}\{\mathsf{X} \le b\} - \mathsf{P}\{\mathsf{X} \le a\} && \text{by } \textit{continuity} \text{ hypothesis} \\ &\triangleq \mathsf{c}_{\mathsf{X}}(b) - \mathsf{c}_{\mathsf{X}}(a) && \text{by definition of } \mathsf{c}_{\mathsf{X}} && \text{(Definition B.2 page 176)} \end{aligned}$$

$$\begin{aligned} \int_a^b \mathsf{p}_{\mathsf{X}}(x)\,\mathsf{d}x &\triangleq \int_a^b \left[\frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{X}}(x)\right]\mathsf{d}x && \text{by definition of } \mathsf{p}_{\mathsf{X}} && \text{(Definition B.2 page 176)} \\ &= \mathsf{c}_{\mathsf{X}}(x)\big|_{x=b} - \mathsf{c}_{\mathsf{X}}(x)\big|_{x=a} && \text{by } \textit{Fundamental theorem of calculus} \\ &= \mathsf{c}_{\mathsf{X}}(b) - \mathsf{c}_{\mathsf{X}}(a) \end{aligned}$$

✏

Theorem B.2. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE, X *be a* RANDOM VARIABLE, *and* $(a : b)$ *a* REAL INTERVAL.

THM

$$\mathsf{P}\{a \le \mathsf{X} < b\} = \int_a^b \mathsf{p}_{\mathsf{X}}(x)\,\mathsf{d}x = \int_{-\infty}^b \mathsf{c}_{\mathsf{X}}(x)\,\mathsf{d}x - \int_{-\infty}^a \mathsf{c}_{\mathsf{X}}(x)\,\mathsf{d}x$$

The properties of the pdf follow closely the properties of *measure* P.

Theorem B.3. [5]

THM

$$\left\{\begin{array}{ll} (A). & \mathsf{c}_{\mathsf{X}}(x) \text{ is CONTINUOUS} \quad \text{OR} \\ (B). & \mathsf{p}_{\mathsf{X}}(x) \text{ is CONTINUOUS} \end{array}\right\} \Longrightarrow \left\{\begin{array}{lrcl} (1). & \mathsf{p}_{\mathsf{X}|\mathsf{Y}}(x|y) & = & \dfrac{\mathsf{p}_{\mathsf{XY}}(x,y)}{\mathsf{p}_{\mathsf{Y}}(y)} \quad \text{and} \\ (2). & \mathsf{p}_{\mathsf{X}}(x) & = & \displaystyle\int_{y\in\mathbb{R}} \mathsf{p}_{\mathsf{XY}}(x,y)\,\mathsf{d}y \end{array}\right\}$$

[5] Papoulis (1990) page 158 ⟨Auxiliary Variable⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✎PROOF:

$$
\begin{aligned}
\mathsf{p}_{\mathsf{X}|\mathsf{Y}}(x|y) &\triangleq \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{X}|\mathsf{Y}}(x|y) && \text{by definition of } \mathsf{c}_{\mathsf{X}} && \text{(Definition A.4 page 168)}\\
&\triangleq \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathsf{P}\{x \le \mathsf{X} < x+\varepsilon|\mathsf{Y}=y\} && \text{by definition of } \frac{\mathsf{d}}{\mathsf{d}x} && \text{(Definition ?? page ??)}\\
&\triangleq \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\frac{\mathsf{P}\{(x \le \mathsf{X} < x+\varepsilon)\wedge(\mathsf{Y}=y)\}}{\mathsf{P}\{\mathsf{Y}=y\}} && \text{by definition of } \mathsf{P}\{A|B\} && \text{(Definition A.4 page 168)}\\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\frac{\mathsf{P}\{(x \le \mathsf{X} < x+\varepsilon)\wedge(y \le \mathsf{Y} < y+\varepsilon)\}}{\mathsf{P}\{y \le \mathsf{Y} < y+\varepsilon\}} && \text{by } \textit{continuity} \text{ hypothesis}\\
&= \frac{\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathsf{P}\{(x \le \mathsf{X} < x+\varepsilon)\wedge(y \le \mathsf{Y} < y+\varepsilon)\}}{\lim_{\varepsilon\to 0}\mathsf{P}\{y \le \mathsf{Y} < y+\varepsilon\}} && \text{by property of } \lim_{\varepsilon\to 0}\\
&= \frac{\mathsf{p}_{\mathsf{XY}}(x,y)}{\mathsf{p}_{\mathsf{Y}}(y)} && \text{by Proposition B.1 page 177}
\end{aligned}
$$

$$
\begin{aligned}
\int_{y\in\mathbb{R}}\mathsf{p}_{\mathsf{XY}}(x,y)\,\mathsf{d}y &\triangleq \int_{y\in\mathbb{R}}\left[\frac{\mathsf{d}}{\mathsf{d}y}\frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{XY}}(x,y)\right]\mathsf{d}y && \text{by definition of } \mathsf{p}_{\mathsf{X}} && \text{(Definition B.2 page 176)}\\
&= \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{XY}}(x,y)\\
&\triangleq \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{y\in\mathbb{R}}\mathsf{P}\{x \le \mathsf{X} < x+\varepsilon, y \le \mathsf{Y} < y+\varepsilon\}\,\mathsf{d}y && \text{by definition of } \frac{\mathsf{d}}{\mathsf{d}x} && \text{(Definition ?? page ??)}\\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathsf{P}\{x \le \mathsf{X} < x+\varepsilon\}\\
&= \mathsf{p}_{\mathsf{X}}(x) && \text{by Proposition B.1 page 177}
\end{aligned}
$$

✏

Theorem B.4.

THM

$$
\begin{aligned}
\mathsf{c}_{\mathsf{X}}(\sup\mathbb{R}) &= 1\\
\mathsf{c}_{\mathsf{X}}(\inf\mathbb{R}) &= 0
\end{aligned}
$$

✎PROOF:

$$
\begin{aligned}
\mathsf{c}_{\mathsf{X}}(\sup\mathbb{R}) &\triangleq \mathsf{P}\{\mathsf{X} \le \sup\mathbb{R}\} && \text{by definition of } \mathsf{c}_{\mathsf{X}} && \text{(Definition B.2 page 176)}\\
&= 1\\
\mathsf{c}_{\mathsf{X}}(\inf\mathbb{R}) &\triangleq \mathsf{P}\{\mathsf{X} \le \inf\mathbb{R}\} && \text{by definition of } \mathsf{c}_{\mathsf{X}} && \text{(Definition B.2 page 176)}\\
&= 0
\end{aligned}
$$

✏

The properties of the pdf follow closely the properties of measure P.

Theorem B.5.

THM

$$
\mathsf{c}_{\mathsf{X}|\mathsf{Y}}(x|y) = \frac{\frac{\mathsf{d}}{\mathsf{d}y}\mathsf{c}_{\mathsf{XY}}(x,y)}{\mathsf{p}_{\mathsf{Y}}(y)} \qquad \mathsf{p}_{\mathsf{X}|\mathsf{Y}}(x|y) = \frac{\mathsf{p}_{\mathsf{XY}}(x,y)}{\mathsf{p}_{\mathsf{Y}}(y)}
$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$
\begin{aligned}
\mathsf{c}_{\mathsf{X}|\mathsf{Y}}(x|y) &\triangleq \mathsf{P}\{\mathsf{X} \le x \mid \mathsf{Y} = y\} && \text{by definition of } \mathsf{c}_{\mathsf{X}|\mathsf{Y}} && \text{(Definition B.4 page 176)}\\
&\triangleq \frac{\mathsf{P}\{\mathsf{X} \le x|\mathsf{Y} = y\}}{\mathsf{P}\{\mathsf{Y} = y\}} && \text{by definition of } \mathsf{P}\{\mathsf{X}|\mathsf{Y}\} && \text{(Definition A.4 page 168)}\\
&= \lim_{\varepsilon\to 0} \frac{\mathsf{P}\{\mathsf{X} \le x|y < \mathsf{Y} \le y+\varepsilon\}}{\mathsf{P}\{y < \mathsf{Y} \le y+\varepsilon\}}\\
&= \lim_{\varepsilon\to 0} \frac{[\mathsf{P}\{\mathsf{X} \le x|\mathsf{Y} \le y+\varepsilon\} - \mathsf{P}\{\mathsf{X} \le x|\mathsf{Y} \le y\}]/\varepsilon}{[\mathsf{P}\{\mathsf{Y} \le y+\varepsilon\} - \mathsf{P}\{\mathsf{Y} \le y\}]/\varepsilon}\\
&\triangleq \lim_{\varepsilon\to 0} \frac{\left[\mathsf{c}_{\mathsf{XY}}(x, y+\varepsilon) - \mathsf{c}_{\mathsf{XY}}(x, y)\right]/\varepsilon}{\left[\mathsf{c}_{\mathsf{Y}}(y+\varepsilon) - \mathsf{c}_{\mathsf{Y}}(y)\right]/\varepsilon} && \text{by definition of } \mathsf{c}_{\mathsf{XY}} && \text{(Definition B.3 page 176)}\\
&\triangleq \frac{\frac{\mathsf{d}}{\mathsf{d}y}\mathsf{c}_{\mathsf{XY}}(x, y)}{\frac{\mathsf{d}}{\mathsf{d}y}\mathsf{c}_{\mathsf{Y}}(y)} && \text{by definition of } \frac{\mathsf{d}}{\mathsf{d}y}\mathsf{f}(y)\\
&\triangleq \frac{\frac{\mathsf{d}}{\mathsf{d}y}\mathsf{c}_{\mathsf{XY}}(x, y)}{\mathsf{p}_{\mathsf{Y}}(y)} && \text{by definition of } \mathsf{p}_{\mathsf{Y}} && \text{(Definition B.2 page 176)}\\
&= \frac{\frac{\mathsf{d}}{\mathsf{d}y}\mathsf{c}_{\mathsf{XY}}(x, y)}{\mathsf{p}_{\mathsf{Y}}(y)} && \text{because } y \text{ is fixed}
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{p}_{\mathsf{X}|\mathsf{Y}}(x|y) &\triangleq \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{X}|\mathsf{Y}}(x|y) && \text{by definition of } \mathsf{p}_{\mathsf{X}|\mathsf{Y}} && \text{(Definition B.4 page 176)}\\
&= \frac{\mathsf{d}}{\mathsf{d}x}\frac{\frac{\mathsf{d}}{\mathsf{d}y}\mathsf{c}_{\mathsf{XY}}(x, y)}{\mathsf{p}_{\mathsf{Y}}(y)} && \text{by previous result}\\
&= \frac{\frac{\mathsf{d}}{\mathsf{d}x}\frac{\mathsf{d}}{\mathsf{d}y}\mathsf{c}_{\mathsf{XY}}(x, y)}{\mathsf{p}_{\mathsf{Y}}(y)} && \text{because } \mathsf{p}_{\mathsf{Y}}(y) \text{ is not a function of } x\\
&\triangleq \frac{\mathsf{p}_{\mathsf{XY}}(x, y)}{\mathsf{p}_{\mathsf{Y}}(y)} && \text{by definition of } \mathsf{p}_{\mathsf{XY}}(x, y) && \text{(Definition B.3 page 176)}
\end{aligned}
$$

✏

Theorem B.6. *Let $(\Omega, \mathbb{E}, \mathsf{P})$ be a probability space.*

THM		
	$\int_{x\in\mathbb{R}} \mathsf{p}_{\mathsf{X}}(x)\,\mathsf{d}x = 1$	$\int_{x\in\mathbb{R}} \mathsf{p}_{\mathsf{X}|\mathsf{Y}}(x|y)\,\mathsf{d}x = 1$
	$\int_{y\in\mathbb{R}} \mathsf{p}_{\mathsf{XY}}(x, y)\,\mathsf{d}y = \mathsf{p}_{\mathsf{X}}(x) \quad \forall x\in\Omega$	$\int_{x\in\mathbb{R}}\int_{y\in\mathbb{R}} \mathsf{p}_{\mathsf{XY}}(x, y)\,\mathsf{d}y\,\mathsf{d}x = 1$

✎PROOF:

$$
\begin{aligned}
\int_{\mathbb{R}} \mathsf{p}_{\mathsf{X}}(x)\,\mathsf{d}x &= \mathsf{c}_{\mathsf{X}}(\sup\mathbb{R}) - \mathsf{c}_{\mathsf{X}}(\inf\mathbb{R}) && \text{by Theorem B.1 page 177}\\
&= 1 - 0\\
&= 1 && \text{because } 0 \text{ is the additive identity element in } (\mathbb{R}, +, \cdot, 0, 1)\\
\int_{x\in\mathbb{R}} \mathsf{p}_{\mathsf{X}|\mathsf{Y}}(x|y)\,\mathsf{d}x &\triangleq \int_{x\in\mathbb{R}} \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{c}_{\mathsf{X}|\mathsf{Y}}(x|y)\,\mathsf{d}x && \text{by definition of } \mathsf{p}_{\mathsf{X}|\mathsf{Y}}(x|y) \text{ (Definition B.4 page 176)}\\
&= \mathsf{c}_{\mathsf{X}|\mathsf{Y}}(\sup\mathbb{R}|y) - \mathsf{c}_{\mathsf{X}|\mathsf{Y}}(\inf\mathbb{R}|y) && \text{by } \textit{Fundamental theorem of calculus}
\end{aligned}
$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

$$
\begin{aligned}
&= 1 - 0 \\
&= 1 && \text{because } 0 \text{ is the additive identity element in } (\mathbb{R}, +, \cdot, 0, 1) \\
\int_{y\in\mathbb{R}} \mathsf{p}_{\mathsf{XY}}(x,y)\,\mathsf{d}y &= \int_{y\in\mathbb{R}} \mathsf{p}_{\mathsf{YX}}(y,x)\,\mathsf{d}y \\
&= \int_{y\in\mathbb{R}} \mathsf{p}_{\mathsf{Y|X}}(y|x)\mathsf{p}_{\mathsf{X}}(x)\,\mathsf{d}y && \text{by Theorem B.5 page 178} \\
&= \mathsf{p}_{\mathsf{X}}(x)\int_{y\in\mathbb{R}} \mathsf{p}_{\mathsf{Y|X}}(y|x)\,\mathsf{d}y && \text{because } \mathsf{p}_{\mathsf{X}}(x) \text{ is not a function of } y \\
&= \mathsf{p}_{\mathsf{X}}(x)\cdot 1 && \text{by previous result} \\
&= \mathsf{p}_{\mathsf{X}}(x) && \text{because } 1 \text{ is the multiplicative identity element in } (\mathbb{R}, +, \cdot, 0, 1)
\end{aligned}
$$

$$
\begin{aligned}
\int_{x\in\mathbb{R}}\int_{y\in\mathbb{R}} \mathsf{p}_{\mathsf{XY}}(x,y)\,\mathsf{d}y\,\mathsf{d}x &= \int_{x\in\mathbb{R}} \mathsf{p}_{\mathsf{X}}(x)\,\mathsf{d}x && \text{by previous result} \\
&= 1 && \text{by previous result}
\end{aligned}
$$

✏

Copyright © 2019 Daniel J. Greenhoe

APPENDIX C

SOME PROBABILITY DENSITY FUNCTIONS

C.1 Discrete distributions

Example C.1. [1] Suppose we throw two "fair" dice and want to know the probabilities of their sum. Let X represent the sum of the face values of the two dice. The resulting probability distribution is illustrated in Figure C.1 (page 182) and has probability space as follows:

EX

$$\begin{aligned} \Omega &= \{ \text{⚀⚀}, \text{⚀⚁}, \text{⚀⚂}, \ldots, \text{⚅⚅} \} \\ \mathbb{E} &= \left\{ 2^{\{X=n \mid n=2,3,\ldots,10,11,\text{ or } 12\}} \right\} \\ \mathsf{P}(e) &= \frac{1}{36} \mid e \mid \end{aligned}$$

C.2 Continuous distributions

C.2.1 Uniform distribution

Definition C.1. *The* ***uniform distribution*** $\mathsf{p}_{\mathsf{X}}(x)$ *is defined as*

DEF

$$\mathsf{p}_{\mathsf{X}}(x) \triangleq \left\{ \begin{array}{ll} 1 & \textit{for}\, 0 < x \leq 1 \\ 0 & \textit{otherwise} \end{array} \right\}$$

Note that although "simple" in form, in light of *Wold's Theorem*, the value of the *uniform distribution* should *not* be taken lightly.

[1] Osgood (2002)

Figure C.1: Probability distribution for two dice (see Example C.1 page 181)

C.2.2 Gaussian distribution

> *❝Tout le monde y croit cependant, me disait un jour M. Lippmann, car les expérimentateurs s'irnaginent que c'est un théorème de mathématiques, et les mathématiciens que c'est un fait expérimental.❞*

> *❝Everyone believes in it [(the normal distribution)] however, said to me one day Mr. Lippmann, because the experimenters imagine that it is a theorem of mathematics, and mathematicians that it is an experimental fact.❞*

Bernard A. Lippmann as told by Henri Poincaré [2]

Figure C.2: Gaussian pdf with $\mu = 0$ and $\sigma \in [0.1, 2]$.

Definition C.2.

[2] quote: Poincaré (1912), page 171
translation: assisted by Google Translate
image:

Copyright © 2019 Daniel J. Greenhoe

DEF

*The **Gaussian distribution** (or **normal distribution**) has pdf*

$$\mathsf{p}_{\mathsf{X}}(x) \triangleq \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

A random variable X *with this distribution is denoted*

$$\mathsf{X} \sim \mathsf{N}\left(\mu, \sigma^2\right)$$

The function $\mathsf{Q}(x)$ *is defined as the area under a Gaussian PDF with zero mean and variance equal to one from* x *to infinity such that*

$$\mathsf{Q}(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{\frac{-u^2}{2}} \, \mathsf{d}u$$

C.2.3 Gamma distribution

Definition C.3. [3] *Let* $b \in \mathbb{R}$. *The **gamma function*** $\Gamma(b)$ *is*

DEF

$$\Gamma(b) \triangleq \int_0^\infty x^{b-1} e^{-x} \, \mathsf{d}x$$

Proposition C.1. [4] *Let* $b \in \mathbb{R}$ *and* $n \in \mathbb{N}$.

PRP

$$\begin{aligned} \Gamma(b) &= (b-1)\Gamma(b-1) \\ \Gamma(n) &= (n-1)! \end{aligned}$$

✎PROOF: Let

$$\begin{aligned} u &= x^{b-1} & \mathsf{d}u &= (b-1)x^{b-1} \, \mathsf{d}x \\ \mathsf{d}v &= e^{-x} \, \mathsf{d}x & v &= -e^{-x} \end{aligned}$$

$$\begin{aligned} \Gamma(b) &\triangleq \int_0^\infty x^{b-1} e^{-x} \, \mathsf{d}x \\ &= \int_{x=0}^\infty u \, \mathsf{d}v \\ &= uv|_{x=0}^\infty - \int_{x=0}^\infty v \, \mathsf{d}u \\ &= -x^{b-1}e^{-x}|_{x=0}^\infty + (b-1)\int_{x=0}^\infty e^{-x}x^{b-1} \, \mathsf{d}x \\ &= (-0+0) + (b-1)\Gamma(b-1) \end{aligned}$$

Note that

$$\Gamma(1) = \int_0^\infty x^{1-1}e^{-x} \, \mathsf{d}x = \int_0^\infty e^{-x} \, \mathsf{d}x = -e^{-x}|_0^\infty = -0+1 = 1$$

$$\begin{aligned} \Gamma(n) &= (n-1)\Gamma(n-1) \\ &= (n-1)(n-2)\Gamma(n-2) \\ &= (n-1)(n-2)(n-3)\Gamma(n-3) \\ &\vdots \\ &= (n-1)(n-2)(n-3)\cdots(1)\Gamma(1) \\ &= (n-1)(n-2)(n-3)\cdots(1) \\ &\triangleq (n-1)! \end{aligned}$$

[3] Papoulis (1991), page 79, Ross (1998), page 222
[4] Ross (1998), page 223

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✏

Definition C.4. *A **Gamma distribution** (b, λ) has pdf*

DEF
$$\mathsf{p}_{\mathsf{X}}(x) \triangleq \frac{\lambda}{\Gamma(b)} e^{-\lambda x}(\lambda x)^{b-1}$$

Theorem C.1. [5] *Let* X *and* Y *be* RANDOM VARIABLES *on a* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$.

THM
$$\left\{\begin{array}{lll} (A). & \mathsf{X} \text{ and } \mathsf{Y} \text{ are INDEPENDENT} & \text{and} \\ (B). & \mathsf{X} \text{ has GAMMA DISTRIBUTION } (a, \lambda) & \text{and} \\ (C). & \mathsf{Y} \text{ has GAMMA DISTRIBUTION } (b, \lambda) & \text{and} \\ (D). & \mathsf{Z} \triangleq \mathsf{X} + \mathsf{Y} & \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \mathsf{Z} \text{ has Gamma distribution} \\ (a+b, \lambda). \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{p}_{\mathsf{Z}}(z) &= \mathsf{p}_{\mathsf{X}}(z) \star \mathsf{p}_{\mathsf{Y}}(z) \\
&= \int_{u\in\mathbb{R}} \mathsf{p}_{\mathsf{X}}(u)\mathsf{p}_{\mathsf{Y}}(z-u)\,\mathrm{d}u && \text{by definition of } \textit{convolution} \quad \text{(Definition P.3 page 352)} \\
&= \int_0^z \frac{1}{\Gamma(a)}\lambda e^{-\lambda u}(\lambda u)^{a-1}\frac{1}{\Gamma(b)}\lambda e^{-\lambda(z-u)}(\lambda(z-u))^{b-1}\,\mathrm{d}u \\
&= \frac{1}{\Gamma(a)}\frac{1}{\Gamma(b)}e^{-\lambda z}\lambda^{1+1+a-1+b-1}\int_0^z u^{a-1}(z-u)^{b-1}\,\mathrm{d}u \\
&= \frac{1}{\Gamma(a)}\frac{1}{\Gamma(b)}\lambda e^{-\lambda z}\lambda^{a+b-1}\int_0^1 (vz)^{a-1}(z-vz)^{b-1}z\,\mathrm{d}v \\
&= \frac{1}{\Gamma(a)}\frac{1}{\Gamma(b)}\lambda e^{-\lambda z}\lambda^{a+b-1}z^{a-1+b-1+1}\int_0^1 v^{a-1}(1-v)^{b-1}\,\mathrm{d}v \\
&= \left[\frac{1}{\Gamma(a)}\frac{1}{\Gamma(b)}\int_0^1 v^{a-1}(1-v)^{b-1}\,\mathrm{d}v\right]\lambda e^{-\lambda z}(\lambda z)^{a+b-1} \\
&= C\lambda e^{-\lambda z}(\lambda z)^{a+b-1} && \text{where } C \text{ is some constant} \\
&= \frac{\lambda}{\Gamma(a+b)}e^{-\lambda z}(\lambda z)^{a+b-1} && C \text{ must be the value that makes } \int_z \mathsf{p}_{\mathsf{Z}}(z) = 1
\end{aligned}$$

$\Longrightarrow \mathsf{p}_{\mathsf{Z}}(z)$ is a $(a+b, \lambda)$ Gamma distribution

✏

C.2.4 Chi-squared distributions

Definition C.5. [6] *Let* $\mathsf{p}(x)$ *be a* PROBABILITY DENSITY FUNCTION *on a* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$.

DEF
$\mathsf{p}(x)$ *is a **chi-square distribution** if*
$$\mathsf{p}(x) \triangleq \left\{\begin{array}{ll} 0 & \text{if } x < 0 \\ \dfrac{1}{\sqrt{2\pi\sigma^2 x}}\exp\left[-\dfrac{x}{2\sigma^2}\right] & \text{if } x \geq 0 \end{array}\right\} \qquad \text{for } \sigma > 0$$

Theorem C.2. [7]

[5] Ross (1998), page 266
[6] Proakis (2001), page 41, Papoulis (1990) page 219 ⟨7-4 Special Distributions of Statistics, (7-78)⟩
[7] Ross (1998), page 267

COPYRIGHT © 2019 DANIEL J. GREENHOE

THM *The following distributions are equivalent:*

(1). *chi-squared distribution* *and*

(2). *distribution of* X^2 *where* $\mathsf{X} \sim \mathsf{N}\left(0, \sigma^2\right)$ *and*

(3). *Gamma distribution* $\left(\frac{1}{2}, \frac{1}{2\sigma^2}\right)$

✎PROOF:

1. Proof that X^2 has chi-squared distribution:

$$\begin{aligned}
\mathsf{p}_{\mathsf{Y}}(y) &= \frac{1}{2\sqrt{y}}\left[\mathsf{p}_{\mathsf{X}}(-\sqrt{y}) + \mathsf{p}_{\mathsf{X}}(\sqrt{y})\right] && \text{by Corollary 5.3 page 31} \\
&= \frac{1}{2\sqrt{y}}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{(-\sqrt{y}-0)^2}{2\sigma^2} + \frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{(+\sqrt{y}-0)^2}{2\sigma^2}\right] \\
&= \frac{1}{2\sqrt{y}}\left[2\frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{y}{2\sigma^2}+\right] \\
&= \frac{1}{\sqrt{2\pi\sigma^2 y}}\exp-\frac{y}{2\sigma^2}
\end{aligned}$$

2. Proof that chi-distribution is a Gamma distribution (b, λ):

$$\begin{aligned}
b &\triangleq \frac{1}{2} \\
\lambda &\triangleq \frac{1}{2\sigma^2}
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\sqrt{2\pi\sigma^2 y}}\exp-\frac{y}{2\sigma^2} &= \frac{1}{\sqrt{\pi}}\lambda^{1/2}\lambda^{1/2}(\lambda y)^{-1/2}e^{-\lambda y} \\
&= \frac{\lambda}{\sqrt{\pi}}(\lambda y)^{b-1}e^{-\lambda y}
\end{aligned}$$

✏

Definition C.6. [8] *The* ***Chi-squared distribution with*** n ***degrees of freedom*** *has pdf*

DEF
$$\mathsf{p}_{\mathsf{Y}}(y) \triangleq \begin{cases} 0 & : y < 0 \\ \dfrac{1}{2\sigma^2\Gamma(n/2)}\left(\dfrac{y}{2\sigma^2}\right)^{\frac{n}{2}-1}\exp-\dfrac{y}{2\sigma^2} & : y \geq 0 \end{cases}$$

Theorem C.3. [9] *The following distributions are equivalent:*

1. *chi-squared distribution with* n *degrees of freedom*
2. *the distribution of* $\sum_{k=1}^{n} X_k^2$ *where* $\left\{X_k | X_k \sim \mathsf{N}\left(0, \sigma^2\right), k = 1, 2, \ldots, n\right\}$ *are independent random variables.*
3. *Gamma distribution* $\left(\frac{n}{2}, \frac{1}{2\sigma^2}\right)$.

[8] Proakis (2001), page 41
[9] Ross (1998), page 267

✎PROOF:

1. Prove chi-squared distribution with n degrees of freedom is the Gamma distribution $\left(\frac{n}{2}, \frac{1}{2\sigma^2}\right)$:

$$\begin{aligned}
\lambda &\triangleq \frac{1}{2\sigma^2} \\
b &\triangleq \frac{1}{2} \\
\frac{1}{2\sigma^2\Gamma(n/2)}\left(\frac{y}{2\sigma^2}\right)^{\frac{n}{2}-1}\exp-\frac{y}{2\sigma^2} &= \frac{\lambda}{\Gamma(nb)}(\lambda y)^{nb-1}\exp-\lambda y
\end{aligned}$$

2. Prove $\sum_{k=1}^{n} X^2$ is Gamma $\left(\frac{n}{2}, \frac{1}{2\sigma^2}\right)$:

   (a) By Theorem C.2, X_k has Gamma distribution $\left(\frac{1}{2}, \frac{1}{2\sigma^2}\right)$.

   (b) By Theorem C.1, $\sum_{k=1}^{n} X_k^2$ has distribution $\left(\frac{n}{2}, \frac{1}{2\sigma^2}\right)$.

✏

Definition C.7. [10] *A **noncentral chi-square distribution** (μ, σ^2) has pdf*

DEF

$$\mathsf{p}_\mathsf{Y}(y) = \frac{1}{\sqrt{2\pi\sigma^2 y}}\exp\frac{y+\mu^2}{-2\sigma^2}\cosh\frac{\mu\sqrt{y}}{\sigma^2}$$

Theorem C.4.

THM

The following distributions are equivalent:

(1). NON-CENTRAL CHI-SQUARED DISTRIBUTION (μ, σ^2)

(2). *distribution of* X^2 *where* $\mathsf{X} \sim \mathsf{N}(\mu, \sigma^2)$

✎PROOF:

1. Proof that $\mathsf{Y} = X^2$ has a non-central chi-squared distribution:

$$\begin{aligned}
\mathsf{p}_\mathsf{Y}(y) &= \frac{1}{2\sqrt{y}}\left[\mathsf{p}_\mathsf{X}(-\sqrt{y}) + \mathsf{p}_\mathsf{X}(\sqrt{y})\right] \qquad \text{by Corollary 5.3 page 31} \\
&= \frac{1}{2\sqrt{y}}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{(-\sqrt{y}-\mu)^2}{2\sigma^2} + \frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{(+\sqrt{y}-\mu)^2}{2\sigma^2}\right] \\
&= \frac{1}{2\sqrt{y}}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{y+\mu^2}{2\sigma^2}\exp\frac{-2\mu\sqrt{y}}{2\sigma^2} + \frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{y+\mu^2}{2\sigma^2}\exp\frac{2\mu\sqrt{y}}{2\sigma^2}\right] \\
&= \frac{1}{\sqrt{2\pi\sigma^2 y}}\exp-\frac{y+\mu^2}{2\sigma^2}\frac{1}{2}\left[\exp\frac{2\mu\sqrt{y}}{2\sigma^2} + \exp\frac{-2\mu\sqrt{y}}{2\sigma^2}\right] \\
&= \frac{1}{\sqrt{2\pi\sigma^2 y}}\exp\frac{y+\mu^2}{-2\sigma^2}\cosh\frac{\mu\sqrt{y}}{\sigma^2}
\end{aligned}$$

✏

[10] Proakis (2001), page 42

COPYRIGHT © 2019 DANIEL J. GREENHOE

Figure C.3: *Joint Gaussian distributions* $\mathsf{p}_{\mathsf{XY}}(x, y)$ with varying correlations

Definition C.8. [11] *The* α***th-order modified Bessel function of the first kind*** $I_\alpha(x)$ *is*

DEF
$$I_\alpha(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(\alpha+k+1)} \left(\frac{x}{2}\right)^{\alpha+2k}$$

Definition C.9. [12] *The* ***noncentral chi-square with*** n***-degrees of freedom*** *distribution has pdf*

DEF
$$\mathsf{p}_{\mathsf{Y}}(y) = \frac{1}{2\sigma^2}\left(\frac{y}{s^2}\right)^{\frac{n-2}{4}} \exp\frac{y+s^2}{-2\sigma^2} I_{n/2-1}\left(\sqrt{y}\frac{s}{\sigma^2}\right) \qquad \text{where } s^2 \triangleq \sum_{k=1}^{n} \mu_k^2$$

C.2.5 Radial distributions

Definition C.10. [13] *The* ***Rayleigh distribution*** *is the pdf*

DEF
$$\mathsf{p}_{\mathsf{R}}(r) = \begin{cases} 0 & \text{for} \quad r < 0 \\ \frac{r}{\sigma^2}\exp -\frac{r^2}{2\sigma^2} & \text{for} \quad r \geq 0 \end{cases}$$

Note that by Proposition 5.3, this distribution is equivalent to the distribution of $R = \sqrt{X^2+Y^2}$ where X and Y are independent random variables each with distribution $\mathsf{N}\left(0, \sigma^2\right)$.

Definition C.11. [14] *The* ***Rice distribution*** *is the pdf*

DEF
$$\mathsf{p}_{\mathsf{R}}(r) = \begin{cases} 0 & \text{for} \quad r < 0 \\ \frac{r}{\sigma^2}\exp\frac{r^2+s^2}{-2\sigma^2} I_o\left(\frac{rs}{\sigma^2}\right) & \text{for} \quad r \geq 0 \end{cases}$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

C.3 Joint Gaussian distributions

Definition C.12 (Joint Gaussian pdf). [15]

DEF

$$\mathsf{p}(x_1, x_2, \ldots, x_n) \triangleq \frac{1}{\sqrt{(2\pi)^n |\boldsymbol{M}|}} \exp -\frac{1}{2}(\boldsymbol{x} - \mathsf{E}\boldsymbol{x})^T \boldsymbol{M}^{-1}(\boldsymbol{x} - \mathsf{E}\boldsymbol{x}) \qquad \textit{(Gaussian joint pdf)}$$

$$\boldsymbol{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\mathsf{Z}_k \triangleq X_k - \mathsf{E}X_k \qquad \textit{(zero mean random variables)}$$

$$\boldsymbol{M} \triangleq \begin{bmatrix} \mathsf{E}[\mathsf{Z}_1\mathsf{Z}_1] & \mathsf{E}[\mathsf{Z}_1\mathsf{Z}_2] & \cdots & \mathsf{E}[\mathsf{Z}_1\mathsf{Z}_n] \\ \mathsf{E}[\mathsf{Z}_2\mathsf{Z}_1] & \mathsf{E}[\mathsf{Z}_2\mathsf{Z}_2] & & \mathsf{E}[\mathsf{Z}_2\mathsf{Z}_n] \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{E}[\mathsf{Z}_n\mathsf{Z}_1] & \mathsf{E}[\mathsf{Z}_n\mathsf{Z}_2] & \cdots & \mathsf{E}[\mathsf{Z}_n\mathsf{Z}_n] \end{bmatrix} \qquad \textit{(correlation matrix)}$$

Figure C.4: Gaussian pdf with $\mu = 0$ and $\sigma \in [0.1, 2]$.

Example C.2 (1 variable joint Gaussian pdf). The **Gaussian distribution** (or **normal distribution** has pdf

EX

$$\mathsf{p}_{\mathsf{x}}(x) \triangleq \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

[11] Proakis (2001), page 43
[12] Proakis (2001), page 43
[13] Proakis (2001), page 44
[14] Proakis (2001), page 46
[15] Anderson (1984) page 21 ⟨Theorem 2.3.1⟩, Anderson (1958), page 14 ⟨§"2.3 The Multivariate Normal Distribution"⟩, Proakis (2001), page 49, Moon and Stirling (2000), page 34

Copyright © 2019 Daniel J. Greenhoe

$$
\begin{aligned}
t &= \arg_t \min_t \left[\frac{1}{2}\int_t^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} + \frac{1}{2}\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\eta)^2}{2\sigma^2}}\right] \\
&= \arg_t \left\{\frac{\partial}{\partial t}\left[\frac{1}{2}\int_t^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} + \frac{1}{2}\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\eta)^2}{2\sigma^2}}\right] = 0\right\} \\
&= \arg_t \left\{\frac{1}{2\sqrt{2\pi\sigma^2}}\left[\frac{\partial}{\partial t}\int_t^{\infty} e^{\frac{-(x-\mu)^2}{2\sigma^2}} + \frac{\partial}{\partial t}\int_{-\infty}^{t} e^{\frac{-(x-\eta)^2}{2\sigma^2}}\right] = 0\right\} \\
&= \arg_t \left\{\left[\left(e^{\frac{-(\infty-\mu)^2}{2\sigma^2}} 0 - e^{\frac{-(t-\mu)^2}{2\sigma^2}} 1\right) + \left(e^{\frac{-(t-\mu)^2}{2\sigma^2}} 1 - e^{\frac{-(\infty-\mu)^2}{2\sigma^2}} 0\right)\right] = 0\right\} \\
&= \arg_t \left\{\left[e^{\frac{-(t-\eta)^2}{2\sigma^2}} - e^{\frac{-(t-\mu)^2}{2\sigma^2}}\right] = 0\right\} \\
&= \arg_t \left\{(t-\eta)^2 = (t-\mu)^2\right\} \\
&= \frac{\mu+\eta}{2}
\end{aligned}
$$

Example C.3 (2 variable joint Gaussian pdf).

EX

$$
\begin{aligned}
z_1 &\triangleq x_1 - \mathsf{E}x_1 \\
z_2 &\triangleq x_2 - \mathsf{E}x_2 \\
|M| &\triangleq \left|\mathsf{E}[z_1 z_1]\mathsf{E}[z_2 z_2] - \mathsf{E}[z_1 z_2]\mathsf{E}[z_1 z_2]\right| \\
\mathsf{p}(x_1, x_2) &\triangleq \frac{1}{2\pi\sqrt{|M|}} \exp\left(\frac{z_1^2\mathsf{E}[z_2 z_2] - 2z_1 z_2 \mathsf{E}[z_1 z_2] + z_2^2 \mathsf{E}[z_1 z_1]}{-2|M|}\right)
\end{aligned}
$$

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

APPENDIX D

SPECTRAL THEORY

D.1 Operator Spectrum

Definition D.1. [1] *Let* $\mathbf{A} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ *be an operator over the linear spaces* $\boldsymbol{X} = (X, F, \oplus, \otimes)$ *and* $\boldsymbol{Y} \triangleq (Y, F, \oplus, \otimes)$. *Let* $\mathcal{N}(\mathbf{A})$ *be the* NULL SPACE *of* $\mathbf{A}$.

DEF

An ***eigenvalue*** *of* $\mathbf{A}$ *is any value* λ *such that there exists* $\boldsymbol{x}$ *such that* $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$.
The ***eigenspace*** $\boldsymbol{H}_\lambda$ *of* $\mathbf{A}$ *at eigenvalue* λ *is* $\mathcal{N}(\mathbf{A} - \lambda\mathbf{I})$.
An ***eigenvector*** *of* $\mathbf{A}$ *associated with eigenvalue* λ *is any element of* $\mathcal{N}(\mathbf{A} - \lambda\mathbf{I})$.

Example D.1. [2] Let $\mathbf{D}$ be the differntial operator.

EX

The set $\left\{e^{\lambda x} | \lambda \in \mathbb{C}\right\}$ are the eigenvectors of $\mathbf{D}$.

$$\begin{aligned}
\rho(\mathbf{D}) &= \varnothing && (\mathbf{D} \text{ has no non-spectral points whatsoever}) \\
\sigma_{\mathsf{p}}(\mathbf{D}) &= \sigma(\mathbf{D}) && (\text{the spectrum of } \mathbf{D} \text{ is all eigenvalues}) \\
\sigma_{\mathsf{c}}(\mathbf{D}) &= \varnothing && (\mathbf{D} \text{ has no continuous spectrum}) \\
\sigma_{\mathsf{r}}(\mathbf{D}) &= \varnothing && (\mathbf{D} \text{ has no resolvent spectrum})
\end{aligned}$$

✎PROOF:

$$\begin{aligned}
(\mathbf{D} - \lambda\mathbf{I})e^{\lambda x} &= \mathbf{D}e^{\lambda x} - \lambda\mathbf{I}e^{\lambda x} \\
&= \lambda e^{\lambda x} - \lambda e^{\lambda x} \\
&= 0 && \forall \lambda \in \mathbb{C}
\end{aligned}$$

This theorem and proof needs more work and investigation to prove/disprove its claims. ✏

Definition D.2. [3] *Let* $\mathbf{A} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ *be an operator over the linear spaces* $\boldsymbol{X} = (X, F, \oplus, \otimes)$ *and* $\boldsymbol{Y} \triangleq (Y, F, \oplus, \otimes)$.

---

[1] Bollobás (1999), page 168, Descartes (1637), Descartes (1954), Cayley (1858), Hilbert (1904), page 67, Hilbert (1912),

[2] Pedersen (2000), page 79

[3] Michel and Herget (1993), page 439

quantity		$\mathcal{N}(\mathbf{A}-\lambda\mathbf{I}) = \{\mathbb{0}\}$ ($\boldsymbol{x} = \mathbb{0}$ is the only solution)	$\overline{\mathcal{R}(\mathbf{A}-\lambda\mathbf{I})} = \boldsymbol{X}$ (dense)	$(\mathbf{A}-\lambda\mathbf{I})^{-1} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$ (continuous/bounded)
$\rho(\mathbf{A})$	(resolvent set)	1	1	1
$\sigma_{\mathsf{p}}(\mathbf{A})$	(point spectrum)	0		
$\sigma_{\mathsf{r}}(\mathbf{A})$	(residual spectrum)	1	0	
$\sigma_{\mathsf{c}}(\mathbf{A})$	(continuous spectrum)	1	1	0

Table D.1: Spectrum of an operator $\mathbf{A}$

DEF

*The **resolvent set** $\rho(\mathbf{A})$ of operator $\mathbf{A}$ is defined as*

$$\rho(\mathbf{A}) \triangleq \left\{\lambda \in F \,\middle|\, \begin{array}{lll} 1. & \mathcal{N}(\mathbf{A}-\lambda\mathbf{I}) = \{\mathbb{0}\} & \textit{(no non-zero eigenvectors)} \quad \textit{and} \\ 2. & \overline{\mathcal{R}(\mathbf{A}-\lambda\mathbf{I})} = \boldsymbol{X} & \textit{(the range is dense in } \boldsymbol{X}\textit{).} \quad \textit{and} \\ 3. & (\mathbf{A}-\lambda\mathbf{I})^{-1} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{Y}) & \textit{(inverse is continuous/bounded).} \end{array}\right\}$$

*The **spectrum** $\sigma(\mathbf{A})$ of operator $\mathbf{A}$ is defined as*

$$\sigma(\mathbf{A}) \triangleq F \setminus \rho(\mathbf{A}).$$

Definition D.3. [4] *Let $\mathbf{A} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$ be an operator over the linear spaces $\boldsymbol{X} = (X, F, \oplus, \otimes)$ and $\boldsymbol{Y} \triangleq (Y, F, \oplus, \otimes)$.*

DEF

*The **point spectrum** $\sigma_{\mathsf{p}}(\mathbf{A})$ of operator $\mathbf{A}$ is defined as*

$$\sigma_{\mathsf{p}}(\mathbf{A}) \triangleq \left\{\lambda \in F \,\middle|\, 1.\ \mathcal{N}(\mathbf{A}-\lambda\mathbf{I}) \supseteq \{\mathbb{0}\} \quad \textit{(has non-zero eigenvector)}\right\}$$

*The **residual spectrum** $\sigma_{\mathsf{r}}(\mathbf{A})$ of operator $\mathbf{A}$ is defined as*

$$\sigma_{\mathsf{r}}(\mathbf{A}) \triangleq \left\{\lambda \in F \,\middle|\, \begin{array}{lll} 1. & \mathcal{N}(\mathbf{A}-\lambda\mathbf{I}) = \{\mathbb{0}\} & \textit{(no non-zero eigenvectors)} \quad \textit{and} \\ 2. & \overline{\mathcal{R}(\mathbf{A}-\lambda\mathbf{I})} \neq \boldsymbol{X} & \textit{(not dense in } \boldsymbol{X}\textit{—has gaps).} \end{array}\right\}$$

*The **continuous spectrum** $\sigma_{\mathsf{c}}(\mathbf{A})$ of operator $\mathbf{A}$ is defined as*

$$\sigma_{\mathsf{c}}(\mathbf{A}) \triangleq \left\{\lambda \in F \,\middle|\, \begin{array}{lll} 1. & \mathcal{N}(\mathbf{A}-\lambda\mathbf{I}) = \{\mathbb{0}\} & \textit{(no non-zero eigenvectors)} \quad \textit{and} \\ 2. & \overline{\mathcal{R}(\mathbf{A}-\lambda\mathbf{I})} = \boldsymbol{X} & \textit{(dense in } \boldsymbol{X}\textit{).} \quad \textit{and} \\ 3. & (\mathbf{A}-\lambda\mathbf{I})^{-1} \notin \mathcal{B}(\boldsymbol{X},\boldsymbol{Y}) & \textit{(not continuous / not bounded)} \end{array}\right\}$$

The spectral components' definitions are illustrated in the figure to the left and summarized in Table D.1 (page 192). Let a family of operators $\mathbf{B}(\lambda)$ be defined with respect to an operator $\mathbf{A}$ such that $\mathbf{B}(\lambda) \triangleq (\mathbf{A} - \lambda\mathbf{I})$. Normally, we might expect a "normal" or "regular" or even "mundane" operator $\mathbf{B}(\lambda)$ to have the properties

1. $\mathbf{B}(\lambda)\boldsymbol{x} = 0$ if and only if $\boldsymbol{x} = \mathbb{0}$
2. $\mathbf{B}(\lambda)\boldsymbol{x}$ spans virtually all of $\boldsymbol{X}$ as we vary $\boldsymbol{x}$
3. $\mathbf{B}^{-1}(\lambda)$ is continuous.

After all, these are the properties that we would have if $\mathbf{B}(\lambda)$ were simply an affine operator in the

[4] Bollobás (1999), page 168, Hilbert (1906) pages 169–172

Copyright © 2019 Daniel J. Greenhoe

field of real numbers— such as $[\mathbf{B}(\lambda)](x) \triangleq [\lambda](x) = \lambda x$ which is 0 if and only if $x = 0$, has range $\mathcal{R}(\lambda) = \mathbb{R}$, and its inverse $\lambda^{-1}x$ is continuous.

If for some λ the operator $\mathbf{B}(\lambda)$ does have all these "regular" properties, then that λ part of the *resolvent set* of $\mathbf{A}$ and λ is called *regular*. However if for some λ the operator $\mathbf{B}(\lambda)$ fails any of these conditions, then that λ part of the *spectrum* of $\mathbf{A}$. And which conditions it fails determines which component of the spectrum it is in.

Theorem D.1. [5] *Let* $\mathbf{A} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ *be an operator.*

THM

$$\sigma(\mathbf{A}) = \sigma_{\mathsf{p}}(\mathbf{A}) \cup \sigma_{\mathsf{c}}(\mathbf{A}) \cup \sigma_{\mathsf{r}}(\mathbf{A})$$

Theorem D.2 (Spectral Theorem)**.** [6] *Let* $\mathbf{N} \in Y^X$ *be an operator.*

THM

$$\left.\begin{array}{ll} (A). & \underbrace{\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^*}_{\mathbf{N}\ \textit{is}\ \text{NORMAL}} \\ (B). & \mathbf{N}\ \textit{is}\ \text{COMPACT} \end{array}\right\} \implies \left\{\begin{array}{lrcl} (1). & \mathbf{N} & = & \sum_n \lambda_n \mathbf{P}_n \\ (2). & \sum_n \mathbf{P}_n & = & \mathbf{I} \\ (3). & \mathbf{P}_n\mathbf{P}_m & = & \bar{\delta}_{n-m}\mathbf{P}_n \\ (4). & dim(\boldsymbol{H}_n) & < & \infty \\ (5). & \left|\left\{\lambda_n | \lambda_n \neq 0\right\}\right| & \textit{is} & \text{COUNTABLY INFINITE} \end{array}\right.$$

where

$$\begin{array}{rcll} (\!(\lambda_n)\!)_{n\in\mathbb{Z}} & \triangleq & \sigma_{\mathsf{p}}(\mathbf{N}) & \textit{(eigenvalues of } \mathbf{N}\textit{)} \\ \boldsymbol{H}_n & \triangleq & \mathcal{N}(\mathbf{N} - \lambda_n\mathbf{I}) & \textit{(}\lambda_n \textit{ is the eigenspace of } \mathbf{N} \textit{ at } \lambda_n \textit{ in } \boldsymbol{Y}\textit{)} \\ \boldsymbol{H}_n & = & \mathbf{P}_n\boldsymbol{Y} & \textit{(}\mathbf{P}_n \textit{ is the projection operator that generates } \boldsymbol{H}_n\textit{)} \end{array}$$

D.2 Fredholm kernels

Definition D.4. [7]

DEF

A ***Fredholm operator*** $\mathbf{K}$ *is defined as*

$$[\mathbf{K}\mathsf{f}](t) \triangleq \underbrace{\int_a^b \underbrace{\kappa(t,s)}_{\textit{kernel}}\mathsf{f}(s)\,\mathrm{d}s}_{\textit{Fredholm integral equation of the first kind}^{8}} \qquad \forall \mathsf{f} \in \boldsymbol{L}_2([a,b])$$

Example D.2. Examples of Fredholm operators include

1.	Fourier Transform	$[\tilde{\mathbf{F}}\mathsf{x}](f)$	$=$	$\int_t \mathsf{x}(t)e^{-i2\pi ft}\,\mathrm{d}t$	$\kappa(t,f)$	$=$	$e^{-i2\pi ft}$
2.	Inverse Fourier Transform	$[\tilde{\mathbf{F}}^{-1}\tilde{\mathsf{x}}](t)$	$=$	$\int_f \tilde{\mathsf{x}}(f)e^{i2\pi ft}\,\mathrm{d}f$	$\kappa(f,t)$	$=$	$e^{i2\pi ft}$
3.	Laplace operator	$[\mathbf{L}\mathsf{x}](s)$	$=$	$\int_t \mathsf{x}(t)e^{-st}\,\mathrm{d}t$	$\kappa(t,s)$	$=$	e^{-st}
4.	autocorrelation operator	$[\mathbf{R}\mathsf{x}](t)$	$=$	$\int_s R(t,s)\mathsf{x}(s)\,\mathrm{d}s$	$\kappa(t,s)$	$=$	$R(t,s)$

Theorem D.3. *Let* $\mathbf{K}$ *be a Fredholm operator with kernel* $\kappa(t,s)$ *and adjoint* $\mathbf{K}^*$.

THM

$$[\mathbf{K}\mathsf{f}](t) = \int_A \kappa(t,s)\mathsf{f}(s)\,\mathrm{d}s \iff \left[\mathbf{K}^*\mathsf{f}\right](t) = \int_A \kappa^*(s,t)\mathsf{f}(s)\,\mathrm{d}s$$

[5] Michel and Herget (1993), page 440

[6] Michel and Herget (1993), page 457, Bollobás (1999), page 200, Hilbert (1906), Hilbert (1912), von Neumann (1929), de Witt (1659)

[7] Michel and Herget (1993), page 425

[8] The equation $\int_u \kappa(t,s)\mathsf{f}(s)\,\mathrm{d}s$ is a **Fredholm integral equation of the first kind** and $\kappa(t,u)$ is the **kernel** of the equation. References: Fredholm (1900), Fredholm (1903), page 365, Michel and Herget (1993), page 97, Keener (1988), page 101

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

✎PROOF:

$$
\begin{aligned}
& [\mathbf{K}\mathsf{f}](t) = \int_A \kappa(t,s)\mathsf{f}(s)\,\mathrm{d}s \\
\iff \langle [\mathbf{K}\mathsf{f}](t) \mid \mathsf{g}(t)\rangle &= \left\langle \int_s \kappa(t,s)\mathsf{f}(s)\,\mathrm{d}s \mid \mathsf{g}(t)\right\rangle && \text{by left hypothesis} \\
&= \int_s \mathsf{f}(s)\,\langle \kappa(t,s) \mid \mathsf{g}(t)\rangle\,\mathrm{d}s && \text{by additivity property of } \langle\triangle \mid \triangledown\rangle \text{ (Definition K.1 page 271)} \\
&= \int_s \mathsf{f}(s)\,\langle \mathsf{g}(t) \mid \kappa(t,s)\rangle^*\,\mathrm{d}s && \text{by conjugate symmetry property of } \langle\triangle \mid \triangledown\rangle \text{ (Definition K.1 page 271)} \\
&= \langle \mathsf{f}(s) \mid \langle \mathsf{g}(t) \mid \kappa(t,s)\rangle\rangle && \text{by local definition of } \langle\triangle \mid \triangledown\rangle \\
&= \left\langle \mathsf{f}(s) \mid \underbrace{\int_t \kappa^*(t,s)\mathsf{g}(t)\,\mathrm{d}t}_{[\mathbf{K}^*\mathsf{g}](s)}\right\rangle && \text{by local definition of } \langle\triangle \mid \triangledown\rangle \\
\iff [\mathbf{K}^*\mathsf{g}](s) &= \int_A \kappa^*(t,s)\mathsf{g}(t)\,\mathrm{d}t && \text{by right hypothesis} \\
\iff [\mathbf{K}^*\mathsf{g}](\sigma) &= \int_A \kappa^*(\tau,\sigma)\mathsf{g}(\tau)\,\mathrm{d}\tau && \text{by change of variable: } \tau = t,\ \sigma = s \\
\iff [\mathbf{K}^*\mathsf{f}](t) &= \int_A \kappa^*(s,t)\mathsf{f}(s)\,\mathrm{d}s && \text{by change of variable: } t = \sigma,\ s = \tau,\ \mathsf{f} = \mathsf{g}
\end{aligned}
$$

✏

Corollary D.1. [9] *Let* $\mathbf{K}$ *be an Fredholm operator with kernel* $\kappa(t,s)$ *and adjoint* $\mathbf{K}^*$.

COR

$$\underbrace{\mathbf{K} = \mathbf{K}^*}_{\mathbf{K}\text{ is self-adjoint}} \iff \underbrace{\kappa(t,s) = \kappa^*(s,t)}_{\text{kernel is conjugate symmetric}}$$

✎PROOF:

$$
\begin{aligned}
\mathbf{K} = \mathbf{K}^* &\iff \int_A \kappa(t,s)\mathsf{f}(s)\,\mathrm{d}s = \int_A \kappa^*(s,t)\mathsf{f}(s)\,\mathrm{d}s && \text{by Theorem D.3 page 193} \\
&\iff \kappa(t,s) = \kappa^*(s,t)
\end{aligned}
$$

✏

Theorem D.4 (Mercer's Theorem). [10] *Let* $\mathbf{K}$ *be an Fredholm operator with kernel* $\kappa(t,s)$ *and eigensystem* $(\!(\lambda_n, \phi_n(t))\!)_{n\in\mathbb{Z}}$.

THM

$$
\left\{
\begin{array}{ll}
(A). & \underbrace{\displaystyle\int_a^b\int_a^b \kappa(t,s)\mathsf{f}(t)\mathsf{f}^*(s)\,\mathrm{d}t \geq 0}_{\text{positive}} \quad \textit{and} \\
(B). & \kappa(t,s) \textit{ is } \text{CONTINUOUS } \textit{on} \\
 & [a:b]\times[a:b]
\end{array}
\right\}
\implies
\left\{
\begin{array}{ll}
(1). & \kappa(t,s) = \displaystyle\sum_n \lambda_n\phi_n(t)\phi_n^*(s) \quad \textit{and} \\
(2). & \kappa(t,s) \text{ CONVERGES ABSOLUTELY} \\
 & \textit{and } \text{UNIFORMLY } \textit{on} \\
 & [a:b]\times[a:b]
\end{array}
\right\}
$$

[9] Michel and Herget (1993), page 430

[10] Gohberg et al. (2003), page 198, Courant and Hilbert (1930), pages 138–140, Mercer (1909), page 439

COPYRIGHT © 2019 DANIEL J. GREENHOE

APPENDIX E

MATRIX CALCULUS

Optimization problems often require finding the value of some parameter which results in some measure reaching a minimum or maximum value. Often this optimal parameter value can be found by solving the single equation generated by the partial derivative of the measure with respect to the parameter. When there are several parameters, optimization often requires several simultaneous equations generated by the partial derivatives of the measure with respect to each parameter. The need for several partial derivatives and several simultaneous equations leads to a natual union of two branches of mathematics— partial differential equations and linear algebra. In general, we would like to not only be able to take the partial derivative of a scalar with respect to another scalar, but to be able to take the partial derivative of a vector with respect to another vector. This generalization is the problem addressed in this section. Other references are also available.[1]

E.1 First derivative of a vector with respect to a vector

Definition E.1.

DEF

$\boldsymbol{x}$ is a vector with the following properties:

1. $$\boldsymbol{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \textit{(n element column vector)}$$

2. $$\frac{\partial}{\partial x_k} x_j = \bar{\delta}_{kj} \qquad \textit{((}x_1, x_2, \ldots, x_n\textit{) are mutually independent)}$$

Definition E.2 (Jacobian matrix). [2] *The **gradient of $\boldsymbol{y}$ with respect to $\boldsymbol{x}$**, as well as the **gradient of $\boldsymbol{y}^T$ with respect to $\boldsymbol{x}$**, is defined as*

[1] Graham (1981) ⟨Chapter 4⟩, Haykin (2001) ⟨Appendix B⟩, Moon and Stirling (2000) ⟨Appendix E⟩, Scharf (1991), pages 274–276, Trees (2002) ⟨Section A.7⟩, Felippa (1999)

[2] Graham (1981), page 52, Graham (2018), page 5297804868241 78§"4.2 The Derivatives of Vectors", Scharf (1991), page 274, Trees (2002), page 1398, Anderson (1984) page 13 ⟨§"2.2.5 Transformation of Variables"⟩, Anderson (1958), page 11 ⟨§"2.2.5 Transformation of Variables"⟩

DEF

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} \triangleq \frac{\partial \boldsymbol{y}^T}{\partial \boldsymbol{x}} \triangleq \underbrace{\begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}}_{n \times m \text{ matrix}} \qquad \forall \boldsymbol{y} \in \mathbb{C}^m$$

Remark E.1. Depending on whether $\boldsymbol{x}$ and $\boldsymbol{y}$ are scalars or vectors, $\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}}$ takes on the following forms:[3]

	y scalar	y vector
x scalar	$\frac{\partial y}{\partial x}$	$\frac{\partial \boldsymbol{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} & \frac{\partial y_2}{\partial x} & \cdots & \frac{\partial y_m}{\partial x} \end{bmatrix}$
x vector	$\frac{\partial y}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}$	$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} \triangleq \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$

Lemma E.1. *Let $\boldsymbol{x} \in \mathbb{R}^n$ be a vector. Then*

LEM

$$\frac{\partial}{\partial x_k} x_i x_j = \bar{\delta}_{ik} x_j + \bar{\delta}_{jk} x_i = \begin{cases} 2x_k & \text{for } i = j = k \\ x_j & \text{for } i = k \text{ and } j \neq k \\ x_i & \text{for } i \neq k \text{ and } j = k \\ 0 & \text{otherwise} \end{cases}$$

Lemma E.2.

LEM

$$\left(\boldsymbol{x}^H \mathbf{A} \boldsymbol{x}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i^* x_j \qquad \forall \quad \begin{matrix} \mathbf{A} \in (\mathbb{C}^n \times \mathbb{C}^n) & \text{(}n \times n\text{ array)} & \text{and} \\ \boldsymbol{x} \in \mathbb{C}^n & \text{(}n\text{ element column vector)} & \end{matrix}$$

✎PROOF:

$$\begin{aligned} \boldsymbol{x}^H \mathbf{A} \boldsymbol{x} &\triangleq \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^* \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} && \text{by definitions of } \mathbf{A} \text{ and } \boldsymbol{x} \\ &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^* \sum_{i=1}^{n} x_i \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} \\ &= \sum_{i=1}^{n} x_i \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^* \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} \\ &= \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ji} x_j^* \end{aligned}$$

[3] For the generalization of the partial derivative of a matrix with respect to a matrix, see Graham (1981) ⟨chapter 6⟩. Graham uses *kronecker products* to handle the additional dimensions(?)

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i^* x_j$$

✏

Lemma E.3.

LEM

$$\frac{\partial}{\partial \boldsymbol{x}}[\mathsf{a}(\boldsymbol{x})\,\mathsf{b}(\boldsymbol{x})] = \mathsf{a}(\boldsymbol{x})\left[\frac{\partial}{\partial \boldsymbol{x}}\mathsf{b}(\boldsymbol{x})\right] + \left[\frac{\partial}{\partial \boldsymbol{x}}\mathsf{a}(\boldsymbol{x})\right]\mathsf{b}(\boldsymbol{x}) \qquad \underbrace{\forall \mathsf{a},\mathsf{b} : \mathbb{R}^n \to \mathbb{R}}_{\mathsf{a}(\boldsymbol{x}),\,\mathsf{b}(\boldsymbol{x})\ \textit{are functions from a vector } \boldsymbol{x} \textit{ to a scalar in } \mathbb{R}}$$

✎PROOF:

$$\frac{\partial}{\partial \boldsymbol{x}}[\mathsf{a}(\boldsymbol{x})\,\mathsf{b}(\boldsymbol{x})] = \begin{bmatrix} \frac{\partial}{\partial x_1}[\mathsf{a}(\boldsymbol{x})\,\mathsf{b}(\boldsymbol{x})] \\ \frac{\partial}{\partial x_2}[\mathsf{a}(\boldsymbol{x})\,\mathsf{b}(\boldsymbol{x})] \\ \vdots \\ \frac{\partial}{\partial x_n}[\mathsf{a}(\boldsymbol{x})\,\mathsf{b}(\boldsymbol{x})] \end{bmatrix} \qquad \text{by definition of } \frac{\partial}{\partial \boldsymbol{x}} \qquad \text{(Definition E.2 page 195)}$$

$$= \begin{bmatrix} \mathsf{a}(\boldsymbol{x})\frac{\partial \mathsf{b}(\boldsymbol{x})}{\partial x_1} + \mathsf{b}(\boldsymbol{x})\frac{\partial \mathsf{a}(\boldsymbol{x})}{\partial x_1} \\ \mathsf{a}(\boldsymbol{x})\frac{\partial \mathsf{b}(\boldsymbol{x})}{\partial x_2} + \mathsf{b}(\boldsymbol{x})\frac{\partial \mathsf{a}(\boldsymbol{x})}{\partial x_2} \\ \vdots \\ \mathsf{a}(\boldsymbol{x})\frac{\partial \mathsf{b}(\boldsymbol{x})}{\partial x_n} + \mathsf{b}(\boldsymbol{x})\frac{\partial \mathsf{a}(\boldsymbol{x})}{\partial x_n} \end{bmatrix} \qquad \text{by } \textit{linearity} \text{ of } \frac{\partial}{\partial \boldsymbol{x}}$$

$$= \begin{bmatrix} \mathsf{a}(\boldsymbol{x})\frac{\partial \mathsf{b}(\boldsymbol{x})}{\partial x_1} \\ \mathsf{a}(\boldsymbol{x})\frac{\partial \mathsf{b}(\boldsymbol{x})}{\partial x_2} \\ \\ \mathsf{a}(\boldsymbol{x})\frac{\partial \mathsf{b}(\boldsymbol{x})}{\partial x_n} \end{bmatrix} + \begin{bmatrix} \frac{\partial \mathsf{a}(\boldsymbol{x})}{\partial x_1}\mathsf{b}(\boldsymbol{x}) \\ \frac{\partial \mathsf{a}(\boldsymbol{x})}{\partial x_2}\mathsf{b}(\boldsymbol{x}) \\ \vdots \\ \frac{\partial \mathsf{a}(\boldsymbol{x})}{\partial x_n}\mathsf{b}(\boldsymbol{x}) \end{bmatrix} \qquad \text{by } \textit{linearity} \text{ of } \textit{vector addition}$$

$$= \mathsf{a}(\boldsymbol{x})\left[\frac{\partial \mathsf{b}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right] + \left[\frac{\partial \mathsf{a}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right]\mathsf{b}(\boldsymbol{x})$$

✏

Theorem E.1. [4]

LEM

$$\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x} = \mathbf{I} \qquad \forall \boldsymbol{x} \in \mathbb{R}^n$$

✎PROOF:

$$\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x} = \begin{bmatrix} \frac{\partial x_1}{\partial x_1} & \frac{\partial x_2}{\partial x_1} & \cdots & \frac{\partial x_n}{\partial x_1} \\ \frac{\partial x_1}{\partial x_2} & \frac{\partial x_2}{\partial x_2} & \cdots & \frac{\partial x_n}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial x_n} & \frac{\partial x_2}{\partial x_n} & \cdots & \frac{\partial x_n}{\partial x_n} \end{bmatrix} \qquad \text{by Definition E.2 page 195}$$

$$= \begin{bmatrix} \bar{\delta}_{11} & \bar{\delta}_{21} & \cdots & \bar{\delta}_{n1} \\ \bar{\delta}_{12} & \bar{\delta}_{22} & \cdots & \bar{\delta}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\delta}_{1n} & \bar{\delta}_{2n} & \cdots & \bar{\delta}_{nn} \end{bmatrix} \qquad \text{by Definition E.1 page 195 (mutual independence property)}$$

[4] Scharf (1991), page 274, Trees (2002), page 1398

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \qquad \text{by definition of kronecker delta function } \bar{\delta}$$

$$= \mathbf{I} \qquad \text{by definition of identity operator } \mathbf{I}$$

✏

Theorem E.2.

THM

$$\frac{\partial}{\partial \boldsymbol{x}}(\mathbf{A}\boldsymbol{x}) = \mathbf{A}^T + \sum_{i=1}^{n}\left(\frac{\partial}{\partial \boldsymbol{x}}\begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{mi} \end{bmatrix}\right)x_i \qquad \forall \boldsymbol{x} \in \mathbb{R}^n,\ \mathbf{A} \in \mathbb{C}^m \times \mathbb{C}^n$$

✎PROOF: Let

$$\boldsymbol{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{A} \triangleq \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\frac{\partial}{\partial \boldsymbol{x}}(\mathbf{A}\boldsymbol{x}) = \frac{\partial}{\partial \boldsymbol{x}}\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) \qquad \text{by definition of } A \text{ and } \boldsymbol{x}$$

$$= \frac{\partial}{\partial \boldsymbol{x}}\begin{bmatrix} \sum_{i=1}^{n} a_{1i}x_i \\ \sum_{i=1}^{n} a_{2i}x_i \\ \vdots \\ \sum_{i=1}^{n} a_{mi}x_i \end{bmatrix} \qquad \text{by matrix multiplication}$$

$$= \frac{\partial}{\partial \boldsymbol{x}}\sum_{i=1}^{n}\begin{bmatrix} a_{1i}x_i \\ a_{2i}x_i \\ \vdots \\ a_{mi}x_i \end{bmatrix}$$

$$= \sum_{i=1}^{n}\frac{\partial}{\partial \boldsymbol{x}}\begin{bmatrix} a_{1i}x_i \\ a_{2i}x_i \\ \vdots \\ a_{mi}x_i \end{bmatrix}$$

$$= \sum_{i=1}^{n}\begin{bmatrix} \frac{\partial a_{1i}x_i}{\partial x_1} & \frac{\partial a_{2i}x_i}{\partial x_1} & \cdots & \frac{\partial a_{mi}x_i}{\partial x_1} \\ \frac{\partial a_{1i}x_i}{\partial x_2} & \frac{\partial a_{2i}x_i}{\partial x_2} & \cdots & \frac{\partial a_{mi}x_i}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial a_{1i}x_i}{\partial x_n} & \frac{\partial a_{2i}x_i}{\partial x_n} & \cdots & \frac{\partial a_{mi}x_i}{\partial x_n} \end{bmatrix} \qquad \text{by Definition E.2 page 195}$$

$$= \sum_{i=1}^{n}\begin{bmatrix} a_{1i}\frac{\partial x_i}{\partial x_1} + \frac{\partial a_{1i}}{\partial x_1}x_i & a_{2i}\frac{\partial x_i}{\partial x_1} + \frac{\partial a_{2i}}{\partial x_1}x_i & \cdots & a_{mi}\frac{\partial x_i}{\partial x_1} + \frac{\partial a_{mi}}{\partial x_1}x_i \\ a_{1i}\frac{\partial x_i}{\partial x_2} + \frac{\partial a_{1i}}{\partial x_2}x_i & a_{2i}\frac{\partial x_i}{\partial x_2} + \frac{\partial a_{2i}}{\partial x_2}x_i & \cdots & a_{mi}\frac{\partial x_i}{\partial x_2} + \frac{\partial a_{mi}}{\partial x_2}x_i \\ \vdots & \vdots & \ddots & \vdots \\ a_{1i}\frac{\partial x_i}{\partial x_n} + \frac{\partial a_{1i}}{\partial x_n}x_i & a_{2i}\frac{\partial x_i}{\partial x_n} + \frac{\partial a_{2i}}{\partial x_n}x_i & \cdots & a_{mi}\frac{\partial x_i}{\partial x_n} + \frac{\partial a_{mi}}{\partial x_n}x_i \end{bmatrix} \qquad \text{by Lemma E.3 page 197}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \sum_{i=1}^{n} \begin{bmatrix} a_{1i}\frac{\partial x_i}{\partial x_1} & a_{2i}\frac{\partial x_i}{\partial x_1} & \cdots & a_{mi}\frac{\partial x_i}{\partial x_1} \\ a_{1i}\frac{\partial x_i}{\partial x_2} & a_{2i}\frac{\partial x_i}{\partial x_2} & \cdots & a_{mi}\frac{\partial x_i}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1i}\frac{\partial x_i}{\partial x_n} & a_{2i}\frac{\partial x_i}{\partial x_n} & \cdots & a_{mi}\frac{\partial x_i}{\partial x_n} \end{bmatrix} + \sum_{i=1}^{n} \begin{bmatrix} \frac{\partial a_{1i}}{\partial x_1}x_i & \frac{\partial a_{2i}}{\partial x_1}x_i & \cdots & \frac{\partial a_{mi}}{\partial x_1}x_i \\ \frac{\partial a_{1i}}{\partial x_2}x_i & \frac{\partial a_{2i}}{\partial x_2}x_i & \cdots & \frac{\partial a_{mi}}{\partial x_2}x_i \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial a_{1i}}{\partial x_n}x_i & \frac{\partial a_{2i}}{\partial x_n}x_i & \cdots & \frac{\partial a_{mi}}{\partial x_n}x_i \end{bmatrix}$$

$$= \sum_{i=1}^{n} \begin{bmatrix} a_{1i}\bar{\delta}_{i1} & a_{2i}\bar{\delta}_{i1} & \cdots & a_{mi}\bar{\delta}_{i1} \\ a_{1i}\bar{\delta}_{i2} & a_{2i}\bar{\delta}_{i2} & \cdots & a_{mi}\bar{\delta}_{i2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1i}\bar{\delta}_{in} & a_{2i}\bar{\delta}_{in} & \cdots & a_{mi}\bar{\delta}_{in} \end{bmatrix} + \sum_{i=1}^{n} \left(\frac{\partial}{\partial \boldsymbol{x}} \begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{mi} \end{bmatrix}\right) x_i \qquad \text{by Lemma E.1}$$

$$= \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} + \sum_{i=1}^{n} \left(\frac{\partial}{\partial \boldsymbol{x}} \begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{mi} \end{bmatrix}\right) x_i \qquad \text{by definition of } \bar{\delta}$$

$$= \mathbf{A}^T + \sum_{i=1}^{n} \left(\frac{\partial}{\partial \boldsymbol{x}} \begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{mi} \end{bmatrix}\right) x_i$$

✏

Theorem E.3 (Affine equations). [5]

THM *$\mathbf{A}$ and $\mathbf{B}$ are independent of $\boldsymbol{x}$* $\implies$

$$\begin{cases} \frac{\partial}{\partial \boldsymbol{x}}(\mathbf{A}\boldsymbol{x}) = \mathbf{A}^T & \forall \boldsymbol{x}\in\mathbb{R}^n,\ \mathbf{A}\in\mathbb{C}^m\times\mathbb{C}^n \\ \frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T\mathbf{B}) = \mathbf{B} & \forall \boldsymbol{x}\in\mathbb{R}^n,\ \mathbf{B}\in\mathbb{C}^n\times\mathbb{C}^m \end{cases}$$

✎PROOF: Let $\mathbf{B} \triangleq \mathbf{A}^T$.

1. Proof that $\frac{\partial}{\partial \boldsymbol{x}}(A\boldsymbol{x}) = \mathbf{A}^T$:

$$\frac{\partial}{\partial \boldsymbol{x}}(\mathbf{A}\boldsymbol{x}) = \mathbf{A}^T + \sum_{i=1}^{n} \left(\frac{\partial}{\partial \boldsymbol{x}} \begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{mi} \end{bmatrix}\right) x_i \qquad \text{by Theorem E.2 page 198}$$

$$= \mathbf{A}^T + \sum_{i=1}^{n} \begin{bmatrix} \frac{\partial}{\partial \boldsymbol{x}}a_{1i} & \frac{\partial}{\partial \boldsymbol{x}}a_{2i} & \cdots & \frac{\partial}{\partial \boldsymbol{x}}a_{mi} \end{bmatrix} x_i$$

$$= \mathbf{A}^T + \sum_{i=1}^{n} \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix} x_i \qquad \text{by left hypothesis}$$

$$= \mathbf{A}^T$$

2. Proof that $\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T\mathbf{B}) = \mathbf{B}$:

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T\mathbf{B}) = \frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T\mathbf{A}^T) \qquad \text{by definition of } \mathbf{B}$$

$$= \frac{\partial}{\partial \boldsymbol{x}}\left[(\mathbf{A}\boldsymbol{x})^T\right]$$

$$= \frac{\partial}{\partial \boldsymbol{x}}(\mathbf{A}\boldsymbol{x}) \qquad \text{by Definition E.2 page 195}$$

$$= \mathbf{A}^T \qquad \text{by Theorem E.3 page 199}$$

$$= \mathbf{B} \qquad \text{by definition of } \mathbf{B}$$

✏

[5] Graham (1981), page 54, Graham (2018), page 549780486824178§"4.2 The Derivatives of Vectors"

COPYRIGHT © 2019 DANIEL J. GREENHOE

Theorem E.4 (Product rule). [6] *Let $\boldsymbol{y}$ and $\boldsymbol{z}$ be functions of $\boldsymbol{x}$ and*

THM $$\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{z}^T\boldsymbol{y} = \frac{\partial \boldsymbol{z}}{\partial \boldsymbol{x}}\boldsymbol{y} + \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}}\boldsymbol{z} \qquad \forall \boldsymbol{x}\in\mathbb{R}^n,\ \boldsymbol{y}\in\mathbb{R}^m,\ \boldsymbol{z}\in\mathbb{R}^m$$

✎PROOF:

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{z}^T\boldsymbol{y} &= \frac{\partial}{\partial \boldsymbol{x}}\sum_{k=1}^{m} z_k y_k \\
&= \sum_{k=1}^{m}\frac{\partial}{\partial \boldsymbol{x}} z_k y_k \\
&= \sum_{k=1}^{m}\frac{\partial z_k}{\partial \boldsymbol{x}} y_k + \sum_{k=1}^{m}\frac{\partial y_k}{\partial \boldsymbol{x}} z_k \qquad \text{by Lemma E.3 page 197} \\
&= \begin{bmatrix} \frac{\partial z_1}{\partial x_1}y_1 & + & \frac{\partial z_2}{\partial x_1}y_2 & + & \cdots & + & \frac{\partial z_n}{\partial x_1}y_n \\ \frac{\partial z_1}{\partial x_1}y_1 & + & \frac{\partial z_2}{\partial x_1}y_2 & + & \cdots & + & \frac{\partial z_n}{\partial x_1}y_n \\ \vdots & & & & \ddots & & \vdots \\ \frac{\partial z_1}{\partial x_1}y_1 & + & \frac{\partial z_2}{\partial x_1}y_2 & + & \cdots & + & \frac{\partial z_n}{\partial x_1}y_n \end{bmatrix} + \begin{bmatrix} \frac{\partial y_1}{\partial x_1}z_1 & + & \frac{\partial y_2}{\partial x_1}z_2 & + & \cdots & + & \frac{\partial y_n}{\partial x_1}z_n \\ \frac{\partial y_1}{\partial x_1}z_1 & + & \frac{\partial y_2}{\partial x_1}z_2 & + & \cdots & + & \frac{\partial y_n}{\partial x_1}z_n \\ \vdots & & & & \ddots & & \vdots \\ \frac{\partial y_1}{\partial x_1}z_1 & + & \frac{\partial y_2}{\partial x_1}z_2 & + & \cdots & + & \frac{\partial y_n}{\partial x_1}z_n \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial z_1}{\partial x_1} & \frac{\partial z_2}{\partial x_1} & \cdots & \frac{\partial z_n}{\partial x_1} \\ \frac{\partial z_1}{\partial x_1} & \frac{\partial z_2}{\partial x_1} & \cdots & \frac{\partial z_n}{\partial x_1} \\ \vdots & & \ddots & \vdots \\ \frac{\partial z_1}{\partial x_1} & \frac{\partial z_2}{\partial x_1} & \cdots & \frac{\partial z_n}{\partial x_1} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \vdots & & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \\
&= \frac{\partial \boldsymbol{z}}{\partial \boldsymbol{x}}\boldsymbol{y} + \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}}\boldsymbol{z}
\end{aligned}$$

✏

Theorem E.5.

THM $$\frac{\partial}{\partial \boldsymbol{x}}\left(\boldsymbol{x}^T A\boldsymbol{x}\right) = \mathbf{A}\boldsymbol{x} + \mathbf{A}^T\boldsymbol{x} + \left[\sum_{i=1}^{n}\left(\frac{\partial}{\partial \boldsymbol{x}}\begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{ni}\end{bmatrix}\right)x_i\right]\boldsymbol{x} \qquad \forall \boldsymbol{x}\in\mathbb{R}^n, \mathbf{A}\in\mathbb{R}^n\times\mathbb{R}^n$$

✎PROOF:

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{x}}\left(\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}\right) &= \left[\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x}\right]\mathbf{A}\boldsymbol{x} + \left[\frac{\partial}{\partial \boldsymbol{x}}\mathbf{A}\boldsymbol{x}\right]\boldsymbol{x} && \text{by Theorem E.4 page 200} \\
&= \mathbf{I}\mathbf{A}\boldsymbol{x} + \left[\mathbf{A}^T + \sum_{i=1}^{n}\left(\frac{\partial}{\partial \boldsymbol{x}}\begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{ni}\end{bmatrix}\right)x_i\right]\boldsymbol{x} && \text{by Theorem E.1 and Theorem E.2} \\
&= \mathbf{A}\boldsymbol{x} + \mathbf{A}^T\boldsymbol{x} + \left[\sum_{i=1}^{n}\left(\frac{\partial}{\partial \boldsymbol{x}}\begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{ni}\end{bmatrix}\right)x_i\right]\boldsymbol{x} && \text{by definition of identity operator } \mathbf{I}
\end{aligned}$$

✏

Theorem E.6 (Quadratic form). [7]

THM $$\mathbf{A} \text{ is independent of } \boldsymbol{x} \quad\Longrightarrow\quad \frac{\partial}{\partial \boldsymbol{x}}\left(\boldsymbol{x}^T A\boldsymbol{x}\right) = A\boldsymbol{x} + A^T\boldsymbol{x} \qquad \forall \boldsymbol{x}\in\mathbb{R}^n, \mathbf{A}\in\mathbb{R}^n\times\mathbb{R}^n$$

[6] Scharf (1991), page 274, Trees (2002), page 1398
[7] Graham (1981), page 54

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}) &= \left[\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x}\right]\mathbf{A}\boldsymbol{x} + \left[\frac{\partial}{\partial \boldsymbol{x}}\mathbf{A}\boldsymbol{x}\right]\boldsymbol{x} && \text{by Theorem E.4 page 200}\\ &= \mathbf{I}\mathbf{A}\boldsymbol{x} + \mathbf{A}^T\boldsymbol{x} && \text{by Theorem E.1 page 197 and Theorem E.3 page 199}\end{aligned}$$

✏

Corollary E.1. [8]

COR $$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T\boldsymbol{x}) = 2\boldsymbol{x} \qquad \forall \boldsymbol{x} \in \mathbb{R}^n$$

✎PROOF:

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T\boldsymbol{x}) &= \frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T\mathbf{I}\boldsymbol{x}) && \text{by property of identity operator } I\\ &= \mathbf{I}\boldsymbol{x} + \mathbf{I}^T\boldsymbol{x} && \text{by previous result 3.}\\ &= \boldsymbol{x} + \boldsymbol{x} && \text{by property of identity operator } I\\ &= 2\boldsymbol{x}\end{aligned}$$

✏

Theorem E.7 (Chain rule). [9] *Let z be a function of $\boldsymbol{y}$ and $\boldsymbol{y}$ a function of $\boldsymbol{x}$ and*

$$\boldsymbol{x} \triangleq \begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix} \qquad \boldsymbol{y} \triangleq \begin{bmatrix} y_1\\ y_2\\ \vdots\\ y_m\end{bmatrix} \qquad \boldsymbol{z} \triangleq \begin{bmatrix} z_1\\ z_2\\ \vdots\\ z_k\end{bmatrix}$$

THM $$\frac{\partial}{\partial \boldsymbol{x}}z = \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{y}}$$

✎PROOF:

$$\begin{aligned}\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{x}} &= \begin{bmatrix} \frac{\partial z_1}{\partial x_1} & \frac{\partial z_2}{\partial x_1} & \cdots & \frac{\partial z_k}{\partial x_1}\\ \frac{\partial z_1}{\partial x_2} & \frac{\partial z_2}{\partial x_2} & \cdots & \frac{\partial z_k}{\partial x_2}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial z_1}{\partial x_n} & \frac{\partial z_2}{\partial x_n} & \cdots & \frac{\partial z_k}{\partial x_n}\end{bmatrix}\\ &= \begin{bmatrix} \sum_{j=0}^{m}\frac{\partial z_1}{\partial y_j}\frac{\partial y_j}{\partial x_1} & \sum_{j=0}^{m}\frac{\partial z_2}{\partial y_j}\frac{\partial y_j}{\partial x_1} & \cdots & \sum_{j=0}^{m}\frac{\partial z_k}{\partial y_j}\frac{\partial y_j}{\partial x_1}\\ \sum_{j=0}^{m}\frac{\partial z_1}{\partial y_j}\frac{\partial y_j}{\partial x_2} & \sum_{j=0}^{m}\frac{\partial z_2}{\partial y_j}\frac{\partial y_j}{\partial x_2} & \cdots & \sum_{j=0}^{m}\frac{\partial z_k}{\partial y_j}\frac{\partial y_j}{\partial x_2}\\ \vdots & \vdots & \ddots & \vdots\\ \sum_{j=0}^{m}\frac{\partial z_1}{\partial y_j}\frac{\partial y_j}{\partial x_n} & \sum_{j=0}^{m}\frac{\partial z_2}{\partial y_j}\frac{\partial y_j}{\partial x_n} & \cdots & \sum_{j=0}^{m}\frac{\partial z_k}{\partial y_j}\frac{\partial y_j}{\partial x_n}\end{bmatrix}\\ &= \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1}\\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n}\end{bmatrix}\begin{bmatrix} \frac{\partial z_1}{\partial y_1} & \frac{\partial z_2}{\partial y_1} & \cdots & \frac{\partial z_k}{\partial y_1}\\ \frac{\partial z_1}{\partial y_2} & \frac{\partial z_2}{\partial y_2} & \cdots & \frac{\partial z_k}{\partial y_2}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial z_1}{\partial y_m} & \frac{\partial z_2}{\partial y_m} & \cdots & \frac{\partial z_k}{\partial y_m}\end{bmatrix}\\ &= \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{y}}\end{aligned}$$

[8] Graham (1981), page 54
[9] Graham (1981), pages 54–55

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

E.2 First derivative of a matrix with respect to a scalar

Definition E.3. *Let* $x \in \mathbb{R}$, $\{y_{jk} \in \mathbb{C} | j = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n\}$ *and*

$$Y = \underbrace{\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix}}_{m \times n \text{ matrix}}$$

The ***derivative of*** Y ***with respect to*** x *is*

DEF

$$\frac{\mathrm{d}Y}{\mathrm{d}x} \triangleq \underbrace{\begin{bmatrix} \frac{\mathrm{d}y_{11}}{\mathrm{d}x} & \frac{\mathrm{d}y_{12}}{\mathrm{d}x} & \cdots & \frac{\mathrm{d}y_{1n}}{\mathrm{d}x} \\ \frac{\mathrm{d}y_{21}}{\mathrm{d}x} & \frac{\mathrm{d}y_{22}}{\mathrm{d}x} & \cdots & \frac{\mathrm{d}y_{2n}}{\mathrm{d}x} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathrm{d}y_{m1}}{\mathrm{d}x} & \frac{\mathrm{d}y_{m2}}{\mathrm{d}x} & \cdots & \frac{\mathrm{d}y_{mn}}{\mathrm{d}x} \end{bmatrix}}_{m \times n \text{ matrix}}$$

Theorem E.8. [10] *Let* $x \in \mathbb{R}$, $\{y_{jp} \in \mathbb{C} | j = 1, 2, \ldots, m;\ p = 1, 2, \ldots, n\}$, $\{w_{jp} \in \mathbb{C} | j = 1, 2, \ldots, n;\ p = 1, 2, \ldots, k\}$, *and*

$$Y = \underbrace{\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix}}_{m \times n \text{ matrix}} \qquad W = \underbrace{\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1k} \\ w_{21} & w_{22} & \cdots & w_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ w_{p1} & w_{p2} & \cdots & w_{pk} \end{bmatrix}}_{p \times k \text{ matrix}}$$

THM

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\left(Y + W\right) &= \frac{\mathrm{d}}{\mathrm{d}x}Y + \frac{\mathrm{d}}{\mathrm{d}x}W && \text{(for } p = m,\ k = n\text{)} \\ \frac{\mathrm{d}}{\mathrm{d}x}\left(YW\right) &= \left(\frac{\mathrm{d}}{\mathrm{d}x}Y\right)W + Y\left(\frac{\mathrm{d}}{\mathrm{d}x}W\right) && \text{(for } p = n\text{)} \\ \frac{\mathrm{d}}{\mathrm{d}x}\left(Y^T\right) &= \left(\frac{\mathrm{d}}{\mathrm{d}x}Y\right)^T \\ \frac{\mathrm{d}}{\mathrm{d}x}\left(Y^{-1}\right) &= -Y^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}x}Y\right)Y^{-1} && \text{(for } m = n \text{ and } Y \text{ invertible)} \end{aligned}$$

✎PROOF:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\left(Y + W\right) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} + \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mn} \end{bmatrix}\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} y_{11} + w_{11} & y_{12} + w_{12} & \cdots & y_{1n} + w_{1n} \\ y_{21} + w_{21} & y_{22} + w_{22} & \cdots & y_{2n} + w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} + w_{m1} & y_{m2} + w_{m2} & \cdots & y_{mn} + w_{mn} \end{bmatrix} \end{aligned}$$

[10] Gradshteyn and Ryzhik (1980), pages 1106–1107

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$
\begin{aligned}
&= \begin{bmatrix} (y_{11}+w_{11})' & (y_{12}+w_{12})' & \cdots & (y_{1n}+w_{1n})' \\ (y_{21}+w_{21})' & (y_{22}+w_{22})' & \cdots & (y_{2n}+w_{2n})' \\ \vdots & \vdots & \ddots & \vdots \\ (y_{m1}+w_{m1})' & (y_{m2}+w_{m2})' & \cdots & (y_{mn}+w_{mn})' \end{bmatrix} \\
&= \begin{bmatrix} y'_{11}+w'_{11} & y'_{12}+w'_{12} & \cdots & y'_{1n}+w'_{1n} \\ y'_{21}+w'_{21} & y'_{22}+w'_{22} & \cdots & y'_{2n}+w'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y'_{m1}+w'_{m1} & y'_{m2}+w'_{m2} & \cdots & y'_{mn}+w'_{mn} \end{bmatrix} \\
&= \begin{bmatrix} y'_{11} & y'_{12} & \cdots & y'_{1n} \\ y'_{21} & y'_{22} & \cdots & y'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y'_{m1} & y'_{m2} & \cdots & y'_{mn} \end{bmatrix} + \begin{bmatrix} w'_{11} & w'_{12} & \cdots & w'_{1n} \\ w'_{21} & w'_{22} & \cdots & w'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w'_{m1} & w'_{m2} & \cdots & w'_{mn} \end{bmatrix} \\
&= \frac{\mathrm{d}}{\mathrm{d}x} \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} + \frac{\mathrm{d}}{\mathrm{d}x} \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mn} \end{bmatrix} \\
&= \frac{\mathrm{d}}{\mathrm{d}x} Y + \frac{\mathrm{d}}{\mathrm{d}x} W
\end{aligned}
$$

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\Big(YW\Big) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1k} \\ w_{21} & w_{22} & \cdots & w_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nk} \end{bmatrix} \right) \\
&= \frac{\mathrm{d}}{\mathrm{d}x} \begin{bmatrix} \sum_{j=1}^{n} y_{1j}w_{j1} & \sum_{j=1}^{n} y_{1j}w_{j2} & \cdots & \sum_{j=1}^{n} y_{1j}w_{jk} \\ \sum_{j=1}^{n} y_{2j}w_{j1} & \sum_{j=1}^{n} y_{2j}w_{j2} & \cdots & \sum_{j=1}^{n} y_{2j}w_{jk} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} y_{mj}w_{j1} & \sum_{j=1}^{n} y_{mj}w_{j2} & \cdots & \sum_{j=1}^{n} y_{mj}w_{jk} \end{bmatrix} \\
&= \frac{\mathrm{d}}{\mathrm{d}x} \sum_{j=1}^{n} \begin{bmatrix} y_{1j}w_{j1} & y_{1j}w_{j2} & \cdots & y_{1j}w_{jk} \\ y_{2j}w_{j1} & y_{2j}w_{j2} & \cdots & y_{2j}w_{jk} \\ \vdots & \vdots & \ddots & \vdots \\ y_{mj}w_{j1} & y_{mj}w_{j2} & \cdots & y_{mj}w_{jk} \end{bmatrix} \\
&= \sum_{j=1}^{n} \frac{\mathrm{d}}{\mathrm{d}x} \begin{bmatrix} y_{1j}w_{j1} & y_{1j}w_{j2} & \cdots & y_{1j}w_{jk} \\ y_{2j}w_{j1} & y_{2j}w_{j2} & \cdots & y_{2j}w_{jk} \\ \vdots & \vdots & \ddots & \vdots \\ y_{mj}w_{j1} & y_{mj}w_{j2} & \cdots & y_{mj}w_{jk} \end{bmatrix} \\
&= \sum_{j=1}^{n} \begin{bmatrix} \frac{\mathrm{d}}{\mathrm{d}x}(y_{1j}w_{j1}) & \frac{\mathrm{d}}{\mathrm{d}x}(y_{1j}w_{j2}) & \cdots & \frac{\mathrm{d}}{\mathrm{d}x}(y_{1j}w_{jk}) \\ \frac{\mathrm{d}}{\mathrm{d}x}(y_{2j}w_{j1}) & \frac{\mathrm{d}}{\mathrm{d}x}(y_{2j}w_{j2}) & \cdots & \frac{\mathrm{d}}{\mathrm{d}x}(y_{2j}w_{jk}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathrm{d}}{\mathrm{d}x}(y_{mj}w_{j1}) & \frac{\mathrm{d}}{\mathrm{d}x}(y_{mj}w_{j2}) & \cdots & \frac{\mathrm{d}}{\mathrm{d}x}(y_{mj}w_{jk}) \end{bmatrix} \\
&= \sum_{j=1}^{n} \begin{bmatrix} \vdots & \vdots & \ddots & \vdots \\ y'_{1j}w_{j1}+y_{1j}w'_{j1} & y'_{1j}w_{j2}+y_{1j}w'_{j2} & \cdots & y'_{1j}w_{jk}+y_{1j}w'_{jk} \\ y'_{2j}w_{j1}+y_{2j}w'_{j1} & y'_{2j}w_{j2}+y_{2j}w'_{j2} & \cdots & y'_{2j}w_{jk}+y_{2j}w'_{jk} \\ y'_{mj}w_{j1}+y_{mj}w'_{j1} & y'_{mj}w_{j2}+y_{mj}w'_{j2} & \cdots & y'_{mj}w_{jk}+y_{mj}w'_{jk} \end{bmatrix}
\end{aligned}
$$

Copyright © 2019 Daniel J. Greenhoe

$$
\begin{aligned}
&= \sum_{j=1}^{n}\left(\begin{bmatrix} y'_{1j}w_{j1} & y'_{1j}w_{j2} & \cdots & y'_{1j}w_{jk} \\ y'_{2j}w_{j1} & y'_{2j}w_{j2} & \cdots & y'_{2j}w_{jk} \\ \vdots & \vdots & \ddots & \vdots \\ y'_{mj}w_{j1} & y'_{mj}w_{j2} & \cdots & y'_{mj}w_{jk} \end{bmatrix} + \begin{bmatrix} y_{1j}w'_{j1} & y_{1j}w'_{j2} & \cdots & y_{1j}w'_{jk} \\ y_{2j}w'_{j1} & y_{2j}w'_{j2} & \cdots & y_{2j}w'_{jk} \\ \vdots & \vdots & \ddots & \vdots \\ y_{mj}w'_{j1} & y_{mj}w'_{j2} & \cdots & y_{mj}w'_{jk} \end{bmatrix}\right) \\
&= \left(\frac{\mathrm{d}}{\mathrm{d}x}Y\right)W + Y\left(\frac{\mathrm{d}}{\mathrm{d}x}W\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left(Y^T\right) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix}^T\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} y_{11} & y_{21} & \cdots & y_{n1} \\ y_{12} & y_{22} & \cdots & y_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1n} & y_{2n} & \cdots & y_{nm} \end{bmatrix} \\
&= \begin{bmatrix} y'_{11} & y'_{21} & \cdots & y'_{n1} \\ y'_{12} & y'_{22} & \cdots & y'_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ y'_{1n} & y'_{2n} & \cdots & y'_{nm} \end{bmatrix} \\
&= \begin{bmatrix} y'_{11} & y'_{12} & \cdots & y'_{1n} \\ y'_{21} & y'_{22} & \cdots & y'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y'_{m1} & y'_{m2} & \cdots & y'_{mn} \end{bmatrix}^T \\
&= \left(\frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix}\right)^T
\end{aligned}
$$

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left(Y^{-1}\right) &= \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{adj}Y}{|Y|} \\
&\vdots \\
&\text{no proof at this time} \\
&\vdots \\
&= -Y^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}x}Y\right)Y^{-1}
\end{aligned}
$$

✍

E.3 Second derivative of a scalar with respect to a vector

Definition E.4. [11] *Let*

$$
\boldsymbol{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
$$

[11] Lieb and Loss (2001), page 240, Horn and Johnson (1990), page 167

Copyright © 2019 Daniel J. Greenhoe

*The **Hessian matrix** of a scalar y with respect to the vector $\boldsymbol{x}$ is*

DEF

$$\frac{\partial^2 y}{\partial \boldsymbol{x}^2} = \frac{\partial}{\partial \boldsymbol{x}}\left(\frac{\partial y}{\partial \boldsymbol{x}}\right) = \frac{\partial}{\partial \boldsymbol{x}}\begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\partial}{\partial x_1}\frac{\partial y}{\partial x_1} & \frac{\partial}{\partial x_1}\frac{\partial y}{\partial x_2} & \cdots & \frac{\partial}{\partial x_1}\frac{\partial y}{\partial x_n} \\ \frac{\partial}{\partial x_2}\frac{\partial y}{\partial x_1} & \frac{\partial}{\partial x_2}\frac{\partial y}{\partial x_2} & \cdots & \frac{\partial}{\partial x_2}\frac{\partial y}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_n}\frac{\partial y}{\partial x_1} & \frac{\partial}{\partial x_n}\frac{\partial y}{\partial x_2} & \cdots & \frac{\partial}{\partial x_n}\frac{\partial y}{\partial x_n} \end{bmatrix}}_{n \times n \text{ matrix}}$$

E.4 Multiple derivatives of a vector with respect to a scalar

Definition E.5. *Let*

$$\boldsymbol{y} \triangleq \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

The derivative of a vector $\boldsymbol{y}$ with respect to the scalar x is

DEF

$$\begin{bmatrix} \boldsymbol{y} \\ \frac{\mathrm{d}}{\mathrm{d}x}\boldsymbol{y} \\ \frac{\mathrm{d}^2}{\mathrm{d}x^2}\boldsymbol{y} \\ \vdots \\ \frac{\mathrm{d}^n}{\mathrm{d}x^n}\boldsymbol{y} \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 & y_2 & \cdots & y_m \\ \frac{\mathrm{d}}{\mathrm{d}x}y_1 & \frac{\mathrm{d}}{\mathrm{d}x}y_2 & \cdots & \frac{\mathrm{d}}{\mathrm{d}x}y_m \\ \frac{\mathrm{d}^2}{\mathrm{d}x^2}y_1 & \frac{\mathrm{d}^2}{\mathrm{d}x^2}y_2 & \cdots & \frac{\mathrm{d}^2}{\mathrm{d}x^2}y_m \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathrm{d}^n}{\mathrm{d}x^n}y_1 & \frac{\mathrm{d}^n}{\mathrm{d}x^n}y_2 & \cdots & \frac{\mathrm{d}^n}{\mathrm{d}x^n}y_m \end{bmatrix}}_{(n+1) \times m \text{ matrix}}$$

Copyright © 2019 Daniel J. Greenhoe

Copyright © 2019 Daniel J. Greenhoe

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:

$$\textit{group} \supseteq \textit{ring} \supseteq \textit{commutative ring} \supseteq \textit{integral domain} \supseteq \textit{field}$$

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 *and*
2. $xy = yx \quad \forall x,y \in X$ (COMMUTATIVE).

Definition F.4. [5] *Let* R *be a* COMMUTATIVE RING *(Definition F.3 page 207).*

DEF

A function $|\cdot|$ *in* $\mathbb{R}^{\mathbb{R}}$ *is an* ***absolute value*** *(or* ***modulus****) if*

1. $|x| \geq 0 \quad x \in \mathbb{R}$ (NON-NEGATIVE) *and*
2. $|x| = 0 \iff x = 0 \quad x \in \mathbb{R}$ (NONDEGENERATE) *and*
3. $|xy| = |x| \cdot |y| \quad x,y \in \mathbb{R}$ (HOMOGENEOUS / SUBMULTIPLICATIVE) *and*
4. $|x+y| \leq |x| + |y| \quad x,y \in \mathbb{R}$ (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. $xy = yx \quad \forall x,y \in X$ *(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. $(\boldsymbol{ux})\boldsymbol{y} = \boldsymbol{u}(\boldsymbol{xy}) \quad \forall \boldsymbol{u},\boldsymbol{x},\boldsymbol{y} \in \boldsymbol{V}$ (ASSOCIATIVE) *and*
2. $\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{y}) = (\boldsymbol{ux}) + (\boldsymbol{uy}) \quad \forall \boldsymbol{u},\boldsymbol{x},\boldsymbol{y} \in \boldsymbol{V}$ (LEFT DISTRIBUTIVE) *and*
3. $(\boldsymbol{u}+\boldsymbol{x})\boldsymbol{y} = (\boldsymbol{uy}) + (\boldsymbol{xy}) \quad \forall \boldsymbol{u},\boldsymbol{x},\boldsymbol{y} \in \boldsymbol{V}$ (RIGHT DISTRIBUTIVE) *and*
4. $\alpha(\boldsymbol{xy}) = (\alpha\boldsymbol{x})\boldsymbol{y} = \boldsymbol{x}(\alpha\boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{V}$ *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

COPYRIGHT © 2019 DANIEL J. GREENHOE

APPENDIX G

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]

G.1 Definition and basic results

A *metric space* is a *set* together with nothing else save a *metric* that gives the space a *topology* (Definition **??** page **??**). A *linear space* (next definition) in general has no topology but does have some additional *algebraic* structure (APPENDIX F page 207) 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 **??** page **??**), a *metric* giving a *metric linear space*, an *inner product* giving an *inner product space* (Definition K.1 page 271), or a *norm* giving a *normed linear space* (Definition L.1 page 287).

Definition G.1. [2] *Let* $(\mathbb{F}, +, \cdot, 0, 1)$ *be a* FIELD *(Definition F.5 page 208). Let* X *be a set, let* $+$ *be an* OPERATOR *(Definition O.1 page 319) 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 G.1: Lattice of mathematical spaces

DEF

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

1.	$\exists \mathbb{0} \in X$ such that	$\boldsymbol{x} + \mathbb{0}$	$=$	$\boldsymbol{x}$	$\forall \boldsymbol{x} \in X$	(+ IDENTITY) *
2.	$\exists \boldsymbol{y} \in X$ such that	$\boldsymbol{x} + \boldsymbol{y}$	$=$	$\mathbb{0}$	$\forall \boldsymbol{x} \in X$	(+ INVERSE)
3.		$(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$	$=$	$\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$	$\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$	(+ *is* ASSOCIATIVE)
4.		$\boldsymbol{x} + \boldsymbol{y}$	$=$	$\boldsymbol{y} + \boldsymbol{x}$	$\forall \boldsymbol{x}, \boldsymbol{y} \in X$	(+ *is* COMMUTATIVE)
5.		$1 \cdot \boldsymbol{x}$	$=$	$\boldsymbol{x}$	$\forall \boldsymbol{x} \in X$	($\cdot$ IDENTITY)
6.		$\alpha \cdot (\beta \cdot \boldsymbol{x})$	$=$	$(\alpha \cdot \beta) \cdot \boldsymbol{x}$	$\forall \alpha, \beta \in S$ *and* $\boldsymbol{x} \in X$	($\cdot$ ASSOCIATES *with* $\cdot$)
7.		$\alpha \cdot (\boldsymbol{x} + \boldsymbol{y})$	$=$	$(\alpha \cdot \boldsymbol{x}) + (\alpha \cdot \boldsymbol{y})$	$\forall \alpha \in S$ *and* $\boldsymbol{x}, \boldsymbol{y} \in X$	($\cdot$ DISTRIBUTES *over* $+$)
8.		$(\alpha + \beta) \cdot \boldsymbol{x}$	$=$	$(\alpha \cdot \boldsymbol{x}) + (\beta \cdot \boldsymbol{x})$	$\forall \alpha, \beta \in S$ *and* $\boldsymbol{x} \in X$	($\cdot$ 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 G.2. *Let* $\boldsymbol{L}_1 \triangleq \big(X_1, +, \cdot, (\mathbb{F}_1, \dot{+}, \dot{\times})\big)$ *and* $\boldsymbol{L}_2 \triangleq \big(X_2, +, \cdot, (\mathbb{F}_2, \dot{+}, \dot{\times})\big)$.

DEF

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

1. $\boldsymbol{L}_1$ *is a* LINEAR SPACE $(Definition G.1 page 209)$ *and*
2. $\boldsymbol{L}_2$ *is a* LINEAR SPACE $(Definition G.1 page 209)$ *and*
3. $\mathbb{F}_2 \subseteq \mathbb{F}_1$ *and*
4. $X_2 \subseteq X_1$ *and*

Remark G.1. [3] By the first four conditions $(*])$ listed in Definition G.1, $(X, +)$ is a **commutative group** (or **abelian group**).

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

COPYRIGHT © 2019 DANIEL J. GREENHOE

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 G.1 (tuples in $\mathbb{F}^N$). [5] Let $(\!|x_n|\!)_1^N$ be an N-*tuple* (Definition R.1 page 369) over a *field* (Definition F.5 page 208) $(\mathbb{F}, +, \cdot, 0, 1)$.

EX

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

Example G.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 G.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
$$[\mathsf{f} + \mathsf{g}](x) \triangleq \mathsf{f}(x) + \mathsf{g}(x) \qquad \forall \mathsf{f}, \mathsf{g} \in Y^X \qquad (\textit{pointwise addition}) \quad \text{and}$$
$$[\alpha \cdot \mathsf{f}](x) \triangleq \alpha \cdot [\mathsf{f}(x)] \qquad \forall \mathsf{f} \in Y^X, \alpha \in \mathbb{F}.$$
Then the structure $\big(Y^X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ is a *linear space*.

Example G.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
$$[\mathsf{f} + \mathsf{g}](x) \triangleq \mathsf{f}(x) + \mathsf{g}(x) \qquad \forall \mathsf{f}, \mathsf{g} \in \mathbb{F}^{\boldsymbol{X}} \qquad (\textit{pointwise addition}) \quad \text{and}$$
$$[\alpha \cdot \mathsf{f}](x) \triangleq \alpha \cdot [\mathsf{f}(x)] \qquad \forall \mathsf{f} \in \mathbb{F}^{\boldsymbol{X}}, \alpha \in \mathbb{F}.$$
Then the structure $\big(\mathbb{F}^{\boldsymbol{X}}, +, \cdot, (\mathbb{F}, +, \cdot)\big)$ is a *linear space*.

Theorem G.1 (Additive identity properties). [9] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a linear space,* 0 *the* ADDITIVE IDENTITY ELEMENT *(Definition F.1 page 207) with respect to* $\dot{+}$, *and* $\mathbb{0}$ *the* ADDITIVE IDENTITY ELEMENT *with respect to* $+$.

THM

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

✎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

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

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

$$\begin{aligned}
0\boldsymbol{x} &= 0\boldsymbol{x} + 0\mathbb{0} && \text{by definition of } + \text{ additive identity element}\\
&= 0\boldsymbol{x} + 0\boldsymbol{x} + (-0\boldsymbol{x}) && \text{by definition of } + \text{ additive inverse}\\
&= (0+0)\boldsymbol{x} + (-0\cdot\boldsymbol{x}) && \text{by definition of } + \text{ additive identity element}\\
&= 0\boldsymbol{x} + (-0\boldsymbol{x}) && \text{by Definition G.1 property 4}\\
&= \mathbb{0} && \text{by definition of } + \text{ additive identity element}
\end{aligned}$$

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

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

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 $\boldsymbol{x} + \boldsymbol{x} = \boldsymbol{x} \implies \boldsymbol{x} = \mathbb{0}$:

$$\begin{aligned}
\boldsymbol{x} &= \boldsymbol{x} + \mathbb{0} && \text{by } \textit{additive identity} \text{ property (Definition G.1 page 209)}\\
&= \boldsymbol{x} + [\boldsymbol{x} + (-\boldsymbol{x})] && \text{by } \textit{additive inverse} \text{ property (Definition G.1 page 209)}\\
&= [\boldsymbol{x} + \boldsymbol{x}] + (-\boldsymbol{x}) && \text{by } \textit{associative} \text{ property (Definition G.1 page 209)}\\
&= \boldsymbol{x} + (-\boldsymbol{x}) && \text{by left hypothesis}\\
&= \mathbb{0} && \text{by } \textit{additive inverse} \text{ property (Definition G.1 page 209)}
\end{aligned}$$

✏

Definition G.3. [10] *Let $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ 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 G.2 (Additive inverse properties). [11] *Let $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ be a linear space, $\mathbb{0}$ the* ADDITIVE IDENTITY ELEMENT *with respect to $+$, and $-\boldsymbol{x}$ the* ADDITIVE INVERSE *(Definition F.1 page 207) 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

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

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

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

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

$$\begin{aligned}
\mathbb{0} &= 0\boldsymbol{x} && \text{by Theorem G.1 page 211} \\
&= (\alpha-\alpha)\boldsymbol{x} && \text{by field property of } \mathbb{F} \\
&= [\alpha+(-\alpha)]\boldsymbol{x} && \text{by field property of } \mathbb{F} \\
&= \alpha\boldsymbol{x}+(-\alpha)\boldsymbol{x} && \text{by Definition G.1 page 209} \\
\implies -(\alpha\boldsymbol{x}) &= (-\alpha)\boldsymbol{x} && \text{by item (1) page 213}
\end{aligned}$$

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

$$\begin{aligned}
\mathbb{0} &= \alpha\mathbb{0} && \text{by Theorem G.1 page 211} \\
&= \alpha[\boldsymbol{x}+(-\boldsymbol{x})] && \text{by definition of additive identity element } -\boldsymbol{x} \\
&= \alpha\boldsymbol{x}+\alpha(-\boldsymbol{x}) && \text{by Definition G.1 page 209} \\
&= \alpha\boldsymbol{x}+\alpha(-\boldsymbol{x}) \\
\implies -(\alpha\boldsymbol{x}) &= \alpha(-\boldsymbol{x}) && \text{by item (1) page 213}
\end{aligned}$$

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

$$\begin{aligned}
\alpha(\boldsymbol{x}-\boldsymbol{y}) &= \alpha\big[\boldsymbol{x}+(-\boldsymbol{y})\big] && \text{by Definition G.3 page 212} \\
&= \alpha\boldsymbol{x}+\alpha(-\boldsymbol{y}) && \text{by Definition G.1 page 209} \\
&= \alpha\boldsymbol{x}+(-\alpha\boldsymbol{y}) && \text{by item (3) page 213} \\
&= \alpha\boldsymbol{x}-\alpha\boldsymbol{y} && \text{by Definition G.3 page 212}
\end{aligned}$$

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

$$\begin{aligned}
(\alpha-\beta)\boldsymbol{x} &= [\alpha+(-\beta)]\boldsymbol{x} && \text{by field properties of } \mathbb{F} \\
&= \alpha\boldsymbol{x}+(-\beta)\boldsymbol{x} && \text{by Definition G.1} \\
&= \alpha\boldsymbol{x}+[-(\beta\boldsymbol{x})] && \text{by item (2) page 213} \\
&= \alpha\boldsymbol{x}-(\beta\boldsymbol{x}) && \text{by Definition G.3 page 212}
\end{aligned}$$

✏

Theorem G.3. [12] *Let $\big(X,+,\cdot,(\mathbb{F},\dot{+},\dot{\times})\big)$ be a linear space, $\mathbb{0}$ the additive identity element with respect to $+$, and $-\boldsymbol{x}$ additive inverse of $\boldsymbol{x}$ with respect to $+$.*

THM						
	1.	$\alpha\boldsymbol{x}=\alpha\boldsymbol{y}$ *and* $\alpha\neq 0$	$\implies$	$\boldsymbol{x}$	$=$	$\boldsymbol{y}$	$\forall \boldsymbol{x},\boldsymbol{y}\in X$
	2.	$\alpha\boldsymbol{x}=\beta\boldsymbol{x}$ *and* $\boldsymbol{x}\neq\mathbb{0}$	$\implies$	α	$=$	β	$\forall \boldsymbol{x},\boldsymbol{y}\in X,\ \alpha,\beta\in\mathbb{F}$
	3.	$\boldsymbol{z}+\boldsymbol{x}=\boldsymbol{z}+\boldsymbol{y}$	$\implies$	$\boldsymbol{x}$	$=$	$\boldsymbol{y}$	$\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in X$

[12] Michel and Herget (1993) page 77, Peano (1888a) ⟨Chapter IX⟩, Peano (1888b) pages 119–120, Banach (1922) pages 134–135

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

✎PROOF:

1. Proof that $\alpha\boldsymbol{x} = \alpha\boldsymbol{y}$ and $\alpha \neq 0 \implies \boldsymbol{x} = \boldsymbol{y}$:

$$\begin{aligned}
\mathbb{0} &= \frac{1}{\alpha}(\mathbb{0}) && \text{by left hypothesis } (\alpha \neq 0)\\
&= \frac{1}{\alpha}\Big(\alpha\boldsymbol{x} - \alpha\boldsymbol{y}\Big) && \text{by left hypothesis } (\alpha\boldsymbol{x} = \alpha\boldsymbol{y})\\
&= \frac{1}{\alpha}\alpha\Big(\boldsymbol{x} - \boldsymbol{y}\Big) && \text{by Definition G.1 page 209}\\
&= \boldsymbol{x} - \boldsymbol{y}
\end{aligned}$$

2. Proof that $\alpha\boldsymbol{x} = \beta\boldsymbol{x}$ and $\boldsymbol{x} \neq \mathbb{0} \implies \alpha = \beta$:

$$\begin{aligned}
\mathbb{0} &= \alpha\boldsymbol{x} + (-\alpha\boldsymbol{x}) && \text{by definition of additive inverse}\\
&= \beta\boldsymbol{x} + (-\alpha\boldsymbol{x}) && \text{by left hypothesis}\\
&= \beta\boldsymbol{x} + (-\alpha)\boldsymbol{x} && \text{by Theorem G.2 page 212}\\
&= [\beta + (-\alpha)]\boldsymbol{x} && \text{by Definition G.1 page 209}\\
\implies \beta - \alpha &= 0 && \text{by Theorem G.1 page 211}\\
\implies \alpha &= \beta && \text{by field properties of } \mathbb{F}
\end{aligned}$$

3. Proof that $\boldsymbol{z} + \boldsymbol{x} = \boldsymbol{z} + \boldsymbol{y} \implies \boldsymbol{x} = \boldsymbol{y}$:

$$\begin{aligned}
\mathbb{0} &= (\boldsymbol{z} + \boldsymbol{x}) - (\boldsymbol{z} + \boldsymbol{y})\\
&= (\boldsymbol{x} + \boldsymbol{z}) - (\boldsymbol{z} + \boldsymbol{y}) && \text{by Definition G.1 property 1}\\
&= (\boldsymbol{x} + \boldsymbol{z}) + [(-1)\boldsymbol{z} + (-1)\boldsymbol{y}] && \text{by Definition G.1 property 3}\\
&= (\boldsymbol{x} + \boldsymbol{z}) + (-\boldsymbol{z} - \boldsymbol{y}) && \text{by previous result 2.}\\
&= \boldsymbol{x} + (\boldsymbol{z} - \boldsymbol{z}) - \boldsymbol{y}\\
&= \boldsymbol{x} - \boldsymbol{y}
\end{aligned}$$

✏

G.2 Order on Linear Spaces

Definition G.4. [13] *Let $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times})\big)$ be a real linear space.*

DEF

*The pair $(\boldsymbol{\Omega}, \leq)$ is an **ordered linear space** if*

1. $\boldsymbol{x} \leq \boldsymbol{y} \implies \boldsymbol{x} + \boldsymbol{z} \leq \boldsymbol{y} + \boldsymbol{z} \quad \forall z \in X$ *and*
2. $\boldsymbol{x} \leq \boldsymbol{y} \implies \alpha\boldsymbol{x} \leq \alpha\boldsymbol{y} \quad \forall \alpha \in \mathbb{F}$

*A vector $\boldsymbol{x}$ is **positive** if $\mathbb{0} \leq \boldsymbol{x}$.*

*The **positive cone** $\boldsymbol{X}^+$ of $(\boldsymbol{X}, \leq)$ is the set $\boldsymbol{X}^+ \triangleq \{\boldsymbol{x} \in \boldsymbol{X} | \mathbb{0} \leq \boldsymbol{x}\}$.*

Definition G.5. [14] *Let $(\boldsymbol{X}, \leq)$ be an ordered linear space.*

DEF

*The tuple $\boldsymbol{L} \triangleq (\boldsymbol{X}, \vee, \wedge; \leq)$ is a **Riesz space** if $\boldsymbol{L}$ is a **lattice**.*

A RIESZ SPACE *is also called a **vector lattice**.*

Theorem G.4. [15] *Let $(\boldsymbol{X}, \vee, \wedge; \leq)$ be a Riesz space (Definition G.5 page 214).*

THM

$\boldsymbol{x} \vee \boldsymbol{y}$	$=$	$-\big[(-\boldsymbol{x}) \wedge (-\boldsymbol{y})\big]$	$\boldsymbol{x} \wedge \boldsymbol{y}$	$=$	$-\big[(-\boldsymbol{x}) \vee (-\boldsymbol{y})\big]$	$\forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{X}$
$\boldsymbol{x} + (\boldsymbol{y} \vee \boldsymbol{z})$	$=$	$(\boldsymbol{x} + \boldsymbol{y}) \vee (\boldsymbol{x} + \boldsymbol{z})$	$\boldsymbol{x} + (\boldsymbol{y} \wedge \boldsymbol{z})$	$=$	$(\boldsymbol{x} + \boldsymbol{y}) \wedge (\boldsymbol{x} + \boldsymbol{z})$	$\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \in \boldsymbol{X}$
$\alpha(\boldsymbol{x} \vee \boldsymbol{y})$	$=$	$(\alpha\boldsymbol{x}) \vee (\alpha\boldsymbol{y})$	$\alpha(\boldsymbol{x} \wedge \boldsymbol{y})$	$=$	$(\alpha\boldsymbol{x}) \wedge (\alpha\boldsymbol{y})$	$\forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{X}, \alpha \geq 0$
$\boldsymbol{x} + \boldsymbol{y}$	$=$	$(\boldsymbol{x} \wedge \boldsymbol{y}) + (\boldsymbol{x} \vee \boldsymbol{y})$				$\forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{X}, \alpha \in \mathbb{F}$

[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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

1. Proof that $x \vee y = -\big[(-x) \wedge (-y)\big]$:

$\begin{aligned} (-x) \wedge (-y) &\leq -x \\ x &\leq -\big[(-x) \wedge (-y)\big] \\ x \vee y &\leq -\big[(-x) \wedge (-y)\big] \end{aligned}$	$\begin{aligned} (-x) \wedge (-y) &\leq -y \\ y &\leq -\big[(-x) \wedge (-y)\big] \end{aligned}$
$\begin{aligned} x &\leq x \vee y \\ -(x \vee y) &\leq -x \\ -(x \vee y) &\leq (-x) \wedge (-y) \\ -\big[(-x) \wedge (-y)\big] &\leq x \vee y \end{aligned}$	$\begin{aligned} y &\leq x \vee y \\ -(x \vee y) &\leq -y \end{aligned}$

2. Proof that $x \wedge y = -\big[(-x) \vee (-y)\big]$:

$x \vee y = -\big[(-x) \wedge (-y)\big]$	by item (1)
$(-x) \vee (-y) = -\big[(-(-x)) \wedge (-(-y))\big]$	replace x with $-x$ and y with y
$(-x) \vee (-y) = -\big[x \wedge y\big]$	$-(-x) = x$
$-\big[x \wedge y\big] = (-x) \vee (-y)$	by symmetry of $=$ relation
$x \wedge y = -\big[(-x) \vee (-y)\big]$	multipy both sides by -1

3. Proof that $x + (y \vee z) = (x + y) \vee (x + z)$:

$\begin{aligned} x + y &\leq x + (y \vee z) \\ (x + y) \vee (x + z) &\leq x + (y \vee z) \end{aligned}$	$x + z \leq x + (y \vee z)$
$\begin{aligned} y &= -x + (x + y) \\ &\leq -x + \big[(x + y) \vee (x + z)\big] \\ y \vee z &\leq -x + \big[(x + y) \vee (x + z)\big] \\ x + (y \vee z) &\leq (x + y) \vee (x + z) \end{aligned}$	$\begin{aligned} z &= -x + (x + z) \\ &\leq -x + \big[(x + y) \vee (x + z)\big] \end{aligned}$

4. Proof that $x + (y \wedge z) = (x + y) \wedge (x + z)$:

$\begin{aligned} x + y &\geq x + (y \wedge z) \\ (x + y) \wedge (x + z) &\geq x + (y \wedge z) \end{aligned}$	$x + z \geq x + (y \wedge z)$
$\begin{aligned} y &= -x + (x + y) \\ &\geq -x + \big[(x + y) \wedge (x + z)\big] \\ y \wedge z &\geq -x + \big[(x + y) \wedge (x + z)\big] \\ x + (y \wedge z) &\geq (x + y) \wedge (x + z) \end{aligned}$	$\begin{aligned} z &= -x + (x + z) \\ &\geq -x + \big[(x + y) \wedge (x + z)\big] \end{aligned}$

5. Proof that $\alpha(x \vee y) = (\alpha x) \vee (\alpha y)$ for $\alpha \geq 0$:

$\begin{aligned} x &\leq x \vee y \\ \alpha x &\leq \alpha(x \vee y) \\ (\alpha x) \vee (\alpha y) &\leq \alpha(x \vee y) \end{aligned}$	$\begin{aligned} y &\leq x \vee y \\ \alpha y &\leq \alpha(x \vee y) \end{aligned}$	by Definition G.4 page 214
$\begin{aligned} \alpha x &\leq (\alpha x) \vee (\alpha y) \\ x &\leq \alpha^{-1}(\alpha x) \vee (\alpha y) \\ x \vee y &\leq \alpha^{-1}(\alpha x) \vee (\alpha y) \\ \alpha(x \vee y) &\leq (\alpha x) \vee (\alpha y) \end{aligned}$	$\begin{aligned} \alpha y &\leq (\alpha x) \vee (\alpha y) \\ y &\leq \alpha^{-1}(\alpha x) \vee (\alpha y) \end{aligned}$	

6. Proof that $\alpha(x \wedge y) = (\alpha x) \wedge (\alpha y)$ for $\alpha \geq 0$:

$\begin{aligned} x &\geq x \wedge y \\ \alpha x &\geq \alpha(x \wedge y) \\ (\alpha x) \wedge (\alpha y) &\geq \alpha(x \wedge y) \end{aligned}$	$\begin{aligned} y &\geq x \wedge y \\ \alpha y &\geq \alpha(x \wedge y) \end{aligned}$	by Definition G.4 page 214

COPYRIGHT © 2019 DANIEL J. GREENHOE

$\begin{aligned} \alpha x &\geq (\alpha x)\wedge(\alpha y) \\ x &\geq \alpha^{-1}(\alpha x)\wedge(\alpha y) \\ x\wedge y &\geq \alpha^{-1}(\alpha x)\wedge(\alpha y) \\ \alpha(x\wedge y) &\geq (\alpha x)\wedge(\alpha y) \end{aligned}$	$\begin{aligned} \alpha y &\geq (\alpha x)\wedge(\alpha y) \\ y &\geq \alpha^{-1}(\alpha x)\wedge(\alpha y) \end{aligned}$	

7. Proof that $x+y=(x\wedge y)+(x\vee y)$:

$\begin{aligned} x &\leq x\vee y \\ x+y &\leq (x\vee y)+y \\ x+y-(x\vee y) &\leq y \\ x+y-(x\vee y) &\leq x\wedge y \\ x+y &\leq (x\vee y)+(x\wedge y) \end{aligned}$	$\begin{aligned} y &\leq x\vee y \\ x+vy &\leq x+(x\vee y) \\ x+vy-(x\vee y) &\leq x \end{aligned}$
$\begin{aligned} x\wedge y &\leq x \\ 0 &\leq x-(x\wedge y) \\ y &\leq y+x-(x\wedge y) \\ y &\leq x+y-(x\wedge y) \\ x\vee y &\leq x+y-(x\wedge y) \\ (x\wedge y)+(x\vee y) &\leq x+y \end{aligned}$	$\begin{aligned} x\wedge y &\leq y \\ 0 &\leq y-(x\wedge y) \\ x &\leq x+y-(x\wedge y) \\ x &\leq x+y-(x\wedge y) \end{aligned}$

✏

Definition G.6. [16] *Let* $(\boldsymbol{X},\vee,\wedge;\leq)$ *be a* RIESZ SPACE *(Definition G.5 page 214)*.

DEF

x^+	*is defined as* $x^+\triangleq x\vee\mathbb{0}$	*and is called the* ***positive part*** *of* x.
x^-	*is defined as* $x^-\triangleq(-x)\vee\mathbb{0}$	*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 G.5. [17] *Let* $(\boldsymbol{X},\vee,\wedge;\leq)$ *be a* RIESZ SPACE *(Definition G.5 page 214)*.

THM

$$\left.\begin{aligned} y-z &= x \quad \textit{and} \\ y\wedge z &= \mathbb{0} \end{aligned}\right\} \iff \left\{\begin{aligned} y &= x^+ \quad \textit{and} \\ z &= x^- \end{aligned}\right.$$

✎PROOF:

1. Proof that left hypothesis $\implies$ right hypothesis:

$$\begin{aligned}
x^+ &= x\vee\mathbb{0} && \text{by definition of } x^+ \text{ Definition G.6 page 216} \\
&= (y-z)\vee\mathbb{0} && \text{by left hypothesis} \\
&= (y-z)\vee(z-z) \\
&= (y\vee z)-z && \text{by Theorem G.4 page 214} \\
&= \left[y+z-(y\wedge z)\right]-z && \text{by Theorem G.4 page 214} \\
&= y-(y\wedge z) \\
&= y-\mathbb{0} && \text{by left hypothesis} \\
&= y \\
x^- &= (-x)\vee\mathbb{0} && \text{by definition of } x^- \text{ Definition G.6 page 216} \\
&= (z-y)\vee\mathbb{0} && \text{by left hypothesis} \\
&= (z-y)\vee(y-y) \\
&= (z\vee y)-y && \text{by Theorem G.4 page 214}
\end{aligned}$$

[16] Aliprantis and Burkinshaw (2006) page 4, Istrăţescu (1987) page 129
[17] Aliprantis and Burkinshaw (2006) page 4 ⟨Theorem 1.3⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$
\begin{aligned}
&= \left[z + y - (z \wedge y)\right] - z && \text{by Theorem G.4 page 214}\\
&= z - (z \wedge y) \\
&= z - \mathbb{0} && \text{by left hypothesis}\\
&= z
\end{aligned}
$$

2. Proof that left hypothesis $\Longleftarrow$ right hypothesis:

$$
\begin{aligned}
y - z &= x^+ - x^- && \text{by right hypothesis}\\
&= [x \vee \mathbb{0}] - [(-x) \vee \mathbb{0}] && \text{by Definition G.6 page 216}\\
&= (x \vee \mathbb{0}) + (x \wedge \mathbb{0}) && \text{by Theorem G.4 page 214}\\
&= x && \text{by Theorem G.4 page 214}\\
y \wedge z &= x^+ \wedge x^- && \text{by right hypothesis}\\
&= \left[x^- + \left(x^+ - x^-\right)\right] \wedge [x^- + \mathbb{0}] \\
& x^- + \left[\left(x^+ - x^-\right) \wedge \mathbb{0}\right] && \text{by Theorem G.4 page 214}\\
& x^- + \left[(y - z) \wedge \mathbb{0}\right] && \text{by right hypothesis}\\
& x^- + (x \wedge \mathbb{0}) && \text{by previous result}\\
& x^- - [-x \vee \mathbb{0}] && \text{by Theorem G.4 page 214}\\
& x^- - x^- && \text{by definition of } x^- \text{ (Definition G.6 page 216)}\\
& \mathbb{0}
\end{aligned}
$$

✏

Theorem G.6. [18] *Let* $(\boldsymbol{X}, \vee, \wedge; \leq)$ *be a* RIESZ SPACE *(Definition G.5 page 214). 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 G.6 page 216) 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:

$$
\begin{aligned}
|x| &= x \vee (-x) && \text{by definition of } |x| \text{ (Definition G.6 page 216)}\\
&= (2x - x) \vee (\mathbb{0} - x) \\
&= (-x + 2x) \vee (-x + \mathbb{0}) && \text{by } \textit{commutative} \text{ property (Definition G.1 page 209)}\\
&= (-x) + (2x \vee \mathbb{0}) && \text{by Theorem G.4 page 214}\\
&= (2x \vee 2\mathbb{0}) - x && \text{by the } \textit{commutative} \text{ property (Definition G.1 page 209)}\\
&= 2(x \vee \mathbb{0}) - x && \text{by Theorem G.4 page 214}\\
&= 2x^+ - x && \text{by definition of } x^+ \text{ (Definition G.6 page 216)}\\
&= 2x^+ - \left(x^+ - x^-\right) && \text{by 1}\\
&= x^+ + x^- \\
x &= x + \mathbb{0} && x + y - y\\
&= (x \vee y) + (x \wedge y) - y && \text{by Theorem G.4 page 214}\\
&= \left[(x - y) \vee (y - y)\right] + (x \wedge y) && \text{by Theorem G.4 page 214}\\
&= \left[(x - y) \vee \mathbb{0}\right] + (x \wedge y) \\
&= (x - y)^+ + (x \wedge y) && \text{by definition of } x^+ \text{ (Definition G.6 page 216)}
\end{aligned}
$$

✏

[18] Aliprantis and Burkinshaw (2006) page 4

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Theorem G.7. [19] *Let* $(\boldsymbol{X}, \vee, \wedge; \leq)$ *be a* RIESZ SPACE *(Definition G.5 page 214). 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 G.6 page 216) of* $x \in \boldsymbol{X}$.

THM

1. $x \vee y = \frac{1}{2}(x + y + |x - y|) \quad \forall x,y \in \boldsymbol{X}$
2. $x \wedge y = \frac{1}{2}(x + y - |x - y|) \quad \forall x,y \in \boldsymbol{X}$
3. $|x - y| = (x \vee y) - (x \wedge y) \quad \forall x,y \in \boldsymbol{X}$
4. $|x| \vee |y| = \frac{1}{2}(|x + y| + |x - y|) \quad \forall x,y \in \boldsymbol{X}$
5. $|x| \vee |y| = \frac{1}{2}||x + y| - |x - y|| \quad \forall x,y \in \boldsymbol{X}$

✎PROOF:

$$
\begin{aligned}
(x + y + |x - y|) &= (x + y) + \big[(x - y) \vee (y - x)\big] && \text{by Definition G.6 page 216}\\
&= \big[(x + y) + (x - y)\big] \vee \big[(x + y) + (y - x)\big] && \text{by Theorem G.4 page 214}\\
&= (2x) \vee (2y)\\
&= 2(x \vee y) && \text{by Theorem G.4 page 214}\\
(x + y - |x - y|) &= (x + y) - \big[(x - y) \vee (y - x)\big] && \text{by Definition G.6 page 216}\\
&= (x + y) - \big[(-(y - x)) \vee (-(x - y))\big]\\
&= (x + y) + \big[(y - x) \wedge (x - y)\big] && \text{by Theorem G.4 page 214}\\
&= \big[(x + y) + (y - x)\big] \wedge \big[(x + y) + (x - y)\big] && \text{by Theorem G.4 page 214}\\
&= (2y) \wedge (2x)\\
&= 2(y \wedge x) && \text{by Theorem G.4 page 214}\\
&= 2(x \wedge y)\\
|x - y| &= \frac{1}{2}(x + y + |x - y|) - \frac{1}{2}(x + y + |x - y|)\\
&= (x \vee y) - (x \wedge y) && \text{by 1 and 2}\\
|x + y| + |x - y| &= \frac{1}{2}(\mathbb{0} + |2x + 2y|) + |x - y|\\
&= \frac{1}{2}\big[(x + y) + (-x - y) + |(x + y) - (-x - y)|\big] + |x - y|\\
&= \big[(x + y) \vee (-x - y)\big] + |x - y| && \text{by 1}\\
&= \big[(x + y) + |x - y|\big] \vee \big[(-x - y) + |x - y|\big] && \text{by Theorem G.4 page 214}\\
&= 2(x \vee y) \vee 2\big[(-y) + (-x) + |(-y) - (-x)|\big] && \text{by 1}\\
&= 2(x \vee y) \vee 2\big[(-y) \vee (-x)\big] && \text{by 1}\\
&= 2([x \vee (-x)] \vee (y \vee (-y)))\\
&= 2(|x| \vee |y|) && \text{by Definition G.6 page 216}\\
||x + y| - |x - y|| &= 2(|x + y| \vee |x - y|) - (|x + y| + |x - y|) && \text{by 1}\\
&= (|x + y + x - y| + |x + y - x + y|) - 2(|x| \vee |y|) && \text{by 3}\\
&= 2(|x| + |y|) - 2(|x| \vee |y|)\\
&= 2(|x| \vee |y|) && \text{by Theorem G.4 page 214}
\end{aligned}
$$

✏

Definition G.7. [20] *Let* $(\boldsymbol{X}, \vee, \wedge; \leq)$ *be a* RIESZ SPACE *(Definition G.5 page 214). 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 G.6 page 216) of* $x \in \boldsymbol{X}$.

DEF

*x and y are **disjoint**, denoted by $x \perp y$, if*

$$|x| \wedge |y| = \mathbb{0}.$$

*Two subsets $\boldsymbol{U}$ and $\boldsymbol{V}$ of $\boldsymbol{X}$ are **disjoint**, denoted by $\boldsymbol{U} \perp \boldsymbol{V}$ if*

$$x \perp y \qquad \forall x \in \boldsymbol{U} \text{ and } y \in \boldsymbol{V}$$

[19] Aliprantis and Burkinshaw (2006) page 5 ⟨Theorem 1.4⟩
[20] Aliprantis and Burkinshaw (2006) page 5

COPYRIGHT © 2019 DANIEL J. GREENHOE

Definition G.8. [21] *Let* $(\boldsymbol{X}, \vee, \wedge; \leq)$ *be a* RIESZ SPACE *(Definition G.5 page 214). 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 G.6 page 216) 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 G.9. [22] *Let* $(\boldsymbol{X}, \vee, \wedge; \leq)$ *be a* RIESZ SPACE *(Definition G.5 page 214). 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 G.6 page 216) 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

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

APPENDIX H

LINEAR COMBINATIONS

H.1 Linear combinations in linear spaces

A *linear space* (Definition G.1 page 209) in general is not equipped with a *topology*. Without a topology, it is not possible to determine whether an *infinite sum* of vectors converges. Therefore in this section (dealing with linear spaces), all definitions related to sums of vectors will be valid for *finite* sums(Definition N.1 page 305) only (finite "N").

Definition H.1. [1] *Let* $\left\{ \boldsymbol{x}_n \in X \middle|_{n=1,2,\dots,N} \right\}$ *be a set of vectors in a* LINEAR SPACE $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$.

DEF

A vector $\boldsymbol{x} \in X$ *is a* ***linear combination*** *of the vectors in* $\left\{\boldsymbol{x}_n\right\}$ *if*

there exists $\left\{ \alpha_n \in \mathbb{F} \middle|_{n=1,2,\dots,N} \right\}$ *such that* $$\boldsymbol{x} = \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n .$$

Definition H.2. [2] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *be a linear space and* Y *be a subset of* X.

DEF

The ***linear span*** *of* Y *is defined as* $$\operatorname{span} Y \triangleq \left\{ \sum_{\gamma \in \Gamma} \alpha_\gamma \boldsymbol{y}_\gamma \middle| \alpha_\gamma \in \mathbb{F},\ \boldsymbol{y}_\gamma \in Y \right\}.$$

The set Y ***spans*** *a set* A *if* $$A \subseteq \operatorname{span} Y.$$

Proposition H.1. [3] *Let* $\left\{ \boldsymbol{x}_n \in X \middle|_{n=1,2,\dots,N} \right\}$ *be a set of vectors in a* LINEAR SPACE $\boldsymbol{L} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$.

PRP

1. $\operatorname{span}\left\{\boldsymbol{x}_n\right\}$ *is a* LINEAR SPACE *(Definition G.1 page 209) and*
2. $\operatorname{span}\left\{\boldsymbol{x}_n\right\}$ *is a* LINEAR SUBSPACE *of* $\boldsymbol{L}$ *(Definition G.2 page 210).*

Definition H.3. [4] *Let* $\boldsymbol{L} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *be a* LINEAR SPACE.

DEF

The set $Y \triangleq \left\{ \boldsymbol{x}_n \in X \middle|_{n=1,2,\dots,N} \right\}$ *is* ***linearly independent*** *in* $\boldsymbol{L}$ *if*

$$\left\{ \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n = 0 \right\} \implies \left\{ \alpha_1 = \alpha_2 = \cdots = \alpha_N = 0 \right\}.$$

The set Y *is* ***linearly dependent*** *in* $\boldsymbol{L}$ *if* Y *is not linearly independent in* $\boldsymbol{L}$.

[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 H.1: Lattice of *linear combinations*

Copyright © 2019 Daniel J. Greenhoe

Definition H.4. [5] *Let* $\{\boldsymbol{x}_n \in X \mid_{n=1,2,\ldots,N}\}$ *be a set of vectors in a* LINEAR SPACE $\boldsymbol{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$.

DEF

The set $\{\boldsymbol{x}_n\}$ *is a* ***Hamel basis*** *for* $\boldsymbol{L}$ *if*

1. $\{\boldsymbol{x}_n\}$ SPANS $\boldsymbol{L}$ *(Definition H.2 page 221)* *and*
2. $\{\boldsymbol{x}_n\}$ *is* LINEARLY INDEPENDENT *in* $\boldsymbol{L}$ *(Definition H.1 page 221)* .

A HAMEL BASIS *is also called a* ***linear basis***.

Definition H.5. [6] *Let* $\boldsymbol{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE. *Let* $\boldsymbol{x}$ *be a* VECTOR *in* $\boldsymbol{L}$ *and* $Y \triangleq \{\boldsymbol{x}_n \in X \mid_{n=1,2,\ldots,N}\}$ *be a set of vectors in* $\boldsymbol{L}$.

DEF

The expression $\sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n$ *is the* ***expansion*** *of* $\boldsymbol{x}$ *on* Y *in* $\boldsymbol{L}$ *if* $\quad \boldsymbol{x} = \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n$.

In this case, the sequence $(\!(\alpha_n)\!)_{n=1}^{N}$ *is the* ***coordinates*** *of* $\boldsymbol{x}$ *with respect to* Y *in* $\boldsymbol{L}$.

If $\alpha_N \neq 0$, *then* N *is the* ***dimension*** $\dim \boldsymbol{L}$ *of* $\boldsymbol{L}$.

Theorem H.1. [7] *Let* $\{\boldsymbol{x}_n \mid_{n=1,2,\ldots,N}\}$ *be a* HAMEL BASIS *(Definition H.4 page 223) for a* LINEAR SPACE $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$.

THM

$$\left\{\boldsymbol{x} = \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n = \sum_{n=1}^{N} \beta_n \boldsymbol{x}_n\right\} \quad \Longrightarrow \quad \underbrace{\alpha_n = \beta_n \quad \forall n=1,2,\ldots,N}_{\text{coordinates of } \boldsymbol{x} \text{ are UNIQUE}} \qquad \forall \boldsymbol{x} \in X$$

✎PROOF:

$$\begin{aligned}
\mathbb{0} &= \boldsymbol{x} - \boldsymbol{x} \\
&= \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n - \sum_{n=1}^{N} \beta_n \boldsymbol{x}_n \\
&= \sum_{n=1}^{N} (\alpha_n - \beta_n) \boldsymbol{x}_n \\
&\Longrightarrow \{\boldsymbol{x}_n\} \text{ is } \textit{linearly dependent} \text{ if } (\alpha_n - \beta_n) \neq 0 \qquad \forall n = 1, 2, \ldots, N \\
&\Longrightarrow (\alpha_n - \beta_n) = 0 \qquad \forall n = 1, 2, \ldots, N \qquad \text{(because } \{\boldsymbol{x}_n\} \text{ is a } \textit{basis} \text{ and therefore must be } \textit{linearly independent}\text{)} \\
&\Longrightarrow \alpha_n = \beta_n \text{ for } n = 1, 2, \ldots, N
\end{aligned}$$

✏

Theorem H.2. [8] *Let* $\boldsymbol{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE.

THM

$$\left\{\begin{array}{ll} 1. & \{\boldsymbol{x}_n \in X|_{n=1,2,\ldots,N}\} \textit{ is a } \text{HAMEL BASIS} \textit{ for } \boldsymbol{L} \qquad \textit{and} \\ 2. & \{\boldsymbol{y}_n \in X|_{n=1,2,\ldots,M}\} \textit{ is a set of } \text{LINEARLY INDEPENDENT} \textit{ vectors in } \boldsymbol{L} \end{array}\right\}$$

$$\Longrightarrow \quad \left\{\begin{array}{ll} 1. & M \leq N \qquad \textit{and} \\ 2. & M = N \Longrightarrow \{\boldsymbol{y}_n|_{n=1,2,\ldots,M}\} \textit{ is a } \text{BASIS} \textit{ for } \boldsymbol{L} \qquad \textit{and} \\ 3. & M \neq N \Longrightarrow \{\boldsymbol{y}_n|_{n=1,2,\ldots,M}\} \textit{ is } \text{NOT} \textit{ a basis for } \boldsymbol{L} \end{array}\right\}$$

✎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*.

Copyright © 2019 Daniel J. Greenhoe

(d) This implies that $\{\boldsymbol{y}_n|_{n=1,2,\ldots,N+1}\}$ is *not* a basis.

(e) This implies that $M \not> N$.

5. Proof that $M \neq N \implies \{\boldsymbol{y}_n|_{n=1,2,\ldots,M}\}$ is *not* a basis for $\boldsymbol{L}$:

   (a) Proof that $M > N \implies \{\boldsymbol{y}_n|_{n=1,2,\ldots,M}\}$ is *not* a basis for $\boldsymbol{L}$: same as in item (4).

   (b) Proof that $M < N \implies \{\boldsymbol{y}_n|_{n=1,2,\ldots,M}\}$ is *not* a basis for $\boldsymbol{L}$:

      i. Suppose $m = N - 1$.

      ii. Then $\{\boldsymbol{y}_n|_{n=1,2,\ldots,N-1}\}$ is a *basis* and there exists λ such that
      $$\sum_{n=1}^{N} \lambda_{n\in\mathbb{Z}} \boldsymbol{y}_{n\in\mathbb{Z}} = 0.$$

      iii. This implies that $\{\boldsymbol{y}_n|_{n=1,2,\ldots,N}\}$ 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 H.1. [9] *Let* $\boldsymbol{L} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a linear space.*

COR
$$\underbrace{\left\{\begin{array}{ll} 1. & \{\boldsymbol{x}_n \in X|_{n=1,2,\ldots,N}\} \text{ is a HAMEL BASIS for } \boldsymbol{L} \quad \text{and} \\ 2. & \{\boldsymbol{y}_n \in X|_{n=1,2,\ldots,M}\} \text{ is a HAMEL BASIS for } \boldsymbol{L} \end{array}\right\} \implies \{N = M\}}_{\text{(all Hamel bases for } \boldsymbol{L} \text{ have the same number of vectors)}}$$

✎PROOF: This follows from Theorem H.2 (page 223). ✏

H.2 Bases in topological linear spaces

A linear space supports the concept of the *span* of a set of vectors (Definition H.2 page 221). In a topological linear space $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \boldsymbol{T})$, 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*" .[10] In this text, except for these comments and Definition H.6, "complete" refers to the metric space definition only.

If a set is both *total* and *linearly independent* (Definition H.3 page 221) in $\boldsymbol{\Omega}$, then that set is a *Hamel basis* (Definition H.4 page 223) for $\boldsymbol{\Omega}$.

Definition H.6. [11] *Let* A^- *be the* CLOSURE *of a* A *in a* TOPOLOGICAL LINEAR SPACE $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \boldsymbol{T}$ *Let* $\operatorname{span} A$ *be the* SPAN *(Definition H.2 page 221) of a set* A.

DEF *A set of vectors* A *is* ***total*** *(or* ***complete*** *or* ***fundamental****) in* $\boldsymbol{\Omega}$ *if*
$$(\operatorname{span} A)^- = \boldsymbol{\Omega} \qquad \text{(SPAN of } A \text{ is DENSE in } \boldsymbol{\Omega}\text{)}.$$

[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⟩

BY-NC-ND

H.3 Schauder bases in Banach spaces

Definition H.7. [12] *Let* $\boldsymbol{B} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* BANACH SPACE. *Let* $\overset{\star}{=}$ *represent* STRONG CONVERGENCE *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 $\boldsymbol{x} \overset{\star}{=} \sum_{n=1}^{\infty} \alpha_n \boldsymbol{x}_n$ (STRONG CONVERGENCE *in* $\boldsymbol{B}$) *and*
2. $\left\{ \sum_{n=1}^{\infty} \alpha_n \boldsymbol{x}_n \overset{\star}{=} \sum_{n=1}^{\infty} \beta_n \boldsymbol{x}_n \right\} \implies \{(\!(\alpha_n)\!) = (\!(\beta_n)\!)\}$ (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 H.3 page 226). 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 H.11 page 238).

Theorem H.3. [15] *Let* $\boldsymbol{B} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* BANACH SPACE. *Let* $\mathbb{Q}$ *be the field of rational numbers.*

THM

$$\left\{ \begin{array}{ll} 1. & \boldsymbol{B} \text{ has a SCHAUDER BASIS} \quad \textit{and} \\ 2. & \mathbb{Q} \text{ is DENSE in } \mathbb{F}. \end{array} \right\} \implies \{\, \boldsymbol{B} \text{ is SEPARABLE} \,\}$$

✎PROOF:

1. lemma:

$$\left| \left\{ \boldsymbol{x} \,|\, \exists\, (\!(\alpha_n \mathbb{Q})\!)_{n\in\mathbb{N}} \text{ such that } \lim_{N\to\infty} \left\| \boldsymbol{x} - \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n \right\| = 0 \right\} \right| = |\,\mathbb{Q}\times\mathbb{N}\,|$$
$$= |\,\mathbb{Z}\times\mathbb{Z}\,|$$
$$= |\,\mathbb{Z}\,|$$
$$= \textit{countably infinite}$$

[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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

2. remainder of proof:

$\boldsymbol{B}$ has a *Schauder basis* $(\!(\boldsymbol{x}_n)\!)_{n\in\mathbb{N}}$

$$\Longrightarrow \text{ for every } \boldsymbol{x}\in\boldsymbol{B}\text{, there exists } (\!(\alpha_n\in\mathbb{F})\!)_{n\in\mathbb{N}} \text{ such that } \boldsymbol{x} \triangleq \sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n \qquad \text{by Definition H.7 page 226}$$

$$\Longrightarrow \text{ for every } \boldsymbol{x}\in\boldsymbol{B}\text{, there exists } (\!(\alpha_n\in\mathbb{F})\!)_{n\in\mathbb{N}} \text{ such that } \lim_{N\to\infty}\left\|\boldsymbol{x}-\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right\| = 0$$

$$\Longrightarrow \text{ for every } \boldsymbol{x}\in\boldsymbol{B}\text{, there exists } (\!(\alpha_n\in\mathbb{Q})\!)_{n\in\mathbb{N}} \text{ such that } \lim_{N\to\infty}\left\|\boldsymbol{x}-\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right\| = 0 \quad \text{because } \mathbb{Q}^{-}=\mathbb{F}$$

$$\Longrightarrow \boldsymbol{B} = \left\{\boldsymbol{x}\,|\,\exists\,(\!(\alpha_n\mathbb{Q})\!)_{n\in\mathbb{N}} \text{ such that } \lim_{N\to\infty}\left\|\boldsymbol{x}-\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right\| = 0\right\}$$

$$\Longrightarrow \boldsymbol{B} = \left\{\boldsymbol{x}\,|\,\exists\,(\!(\alpha_n\mathbb{Q})\!)_{n\in\mathbb{N}} \text{ such that } \boldsymbol{x} = \lim_{N\to\infty}\sum_{n=1}^{N}\alpha_n\boldsymbol{x}_n\right\}^{-}$$

$$\Longrightarrow \boldsymbol{B} \text{ is } \textit{separable} \qquad \text{by (1) lemma page 226}$$

✏

Definition H.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 $\big(X,+,\cdot,(\mathbb{F},\dot{+},\dot{\times}),\|\cdot\|\big)$.

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* $\quad \mathbf{R}\boldsymbol{x}_n = \boldsymbol{y}_n \qquad \forall n\in\mathbb{Z}$

Theorem H.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 $\big(X,+,\cdot,(\mathbb{F},\dot{+},\dot{\times}),\|\cdot\|\big)$.

THM

$$\Big\{\{\boldsymbol{x}_n\} \text{ is EQUIVALENT to } \{\boldsymbol{y}_n\}\Big\}$$
$$\Longleftrightarrow \quad \left\{\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n \text{ is CONVERGENT} \Longleftrightarrow \sum_{n=1}^{\infty}\alpha_n\boldsymbol{y}_n \text{ is CONVERGENT}\right\}$$

Lemma H.1. [18] *Let* $\boldsymbol{\Omega} \triangleq \big(X,+,\cdot,(\mathbb{F},\dot{+},\dot{\times}),\boldsymbol{T}\big)$ *be a topological linear space. Let* $\operatorname{span}A$ *be the* SPAN *of a set* A *(Definition H.2 page 221). Let* $\tilde{\mathsf{f}}(\omega)$ *and* $\tilde{\mathsf{g}}(\omega)$ *be the* FOURIER TRANSFORMS *(Definition P.2 page 349) 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 Q.1 page 359) of a sequence* $(\!(a_n)\!)_{n\in\mathbb{Z}}$ *in* $\boldsymbol{\ell}^2_{\mathbb{R}}$ *(Definition R.2 page 369).*

LEM

$$\left\{\begin{array}{ll}(1). & \{\mathbf{T}^n\mathsf{f}|_{n\in\mathbb{Z}}\} \text{ is a SCHAUDER BASIS for } \boldsymbol{\Omega} \quad \text{and}\\ (2). & \{\mathbf{T}^n\mathsf{g}|_{n\in\mathbb{Z}}\} \text{ is a SCHAUDER BASIS for } \boldsymbol{\Omega}\end{array}\right\} \Longrightarrow \left\{\begin{array}{l}\exists\,(\!(a_n)\!)_{n\in\mathbb{Z}} \quad \text{such that}\\ \tilde{\mathsf{f}}(\omega) = \breve{\mathsf{a}}(\omega)\tilde{\mathsf{g}}(\omega)\end{array}\right\}$$

✎PROOF: Let $\boldsymbol{V}_0'$ be the space spanned by $\{\mathbf{T}^n\phi|_{n\in\mathbb{Z}}\}$.

$$\begin{aligned}\tilde{\mathsf{f}}(\omega) &\triangleq \tilde{\mathbf{F}}\mathsf{f} && \text{by definition of } \tilde{\mathbf{F}} \qquad \text{(Definition P.2 page 349)}\\ &= \tilde{\mathbf{F}}\sum_{n\in\mathbb{Z}}a_n\mathbf{T}\mathsf{g} && \text{by (2)}\\ &= \sum_{n\in\mathbb{Z}}a_n\tilde{\mathbf{F}}\mathbf{T}\mathsf{g}\end{aligned}$$

[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

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

$$= \underbrace{\sum_{n\in\mathbb{Z}} a_n e^{-i\omega n}}_{\check{\mathsf{a}}(\omega)} \underbrace{\tilde{\mathbf{F}}\mathsf{g}}_{\tilde{\mathsf{g}}(\omega)} \qquad \text{by Corollary ?? page ??}$$

$$= \check{\mathsf{a}}(\omega)\tilde{\mathsf{g}}(\omega) \qquad \text{by definition of } \check{\mathbf{F}} \text{ and } \tilde{\mathbf{F}} \quad \text{by (Definition Q.1 page 359, Definition P.2 page 349)}$$

$$\begin{aligned}
V_0 &\triangleq \left\{ \mathsf{f}(x) | \mathsf{f}(x) = \sum_{n\in\mathbb{Z}} b_n \mathbf{T}^n \mathsf{g}(x) \right\} \\
&= \left\{ \mathsf{f}(x) | \tilde{\mathbf{F}}\mathsf{f}(x) = \tilde{\mathbf{F}} \sum_{n\in\mathbb{Z}} b_n \mathbf{T}^n \mathsf{g}(x) \right\} \\
&= \left\{ \mathsf{f}(x) | \tilde{\mathsf{f}}(\omega) = \tilde{\mathsf{b}}(\omega)\tilde{\mathsf{g}}(\omega) \right\} \\
&= \left\{ \mathsf{f}(x) | \tilde{\mathsf{f}}(\omega) = \tilde{\mathsf{b}}(\omega)\check{\mathsf{a}}(\omega)\tilde{\mathsf{f}}(\omega) \right\} \\
&= \left\{ \mathsf{f}(x) | \tilde{\mathsf{f}}(\omega) = \tilde{\mathsf{c}}(\omega)\tilde{\mathsf{f}}(\omega) \right\} \qquad \text{where } \tilde{\mathsf{c}}(\omega) \triangleq \tilde{\mathsf{b}}(\omega)\check{\mathsf{a}}(\omega) \\
&= \left\{ \mathsf{f}(x) | \mathsf{f}(x) = \sum_{n\in\mathbb{Z}} \mathsf{c}_n \mathsf{f}(x-n) \right\} \\
&\triangleq V_0'
\end{aligned}$$

✏

H.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 H.5. [19] *Let* $\left\{ \boldsymbol{x}_n \in X \middle|_{n=1,2,\ldots,N} \right\}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition K.1 page 271)* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle | \triangledown \rangle\right)$.

THM $\left\{ \begin{array}{c} \{\boldsymbol{x}_n\} \text{ is ORTHOGONAL} \\ \text{(Definition K.4 page 283)} \end{array} \right\} \implies \left\{ \begin{array}{c} \{\boldsymbol{x}_n\} \text{ is LINEARLY INDEPENDENT} \\ \text{(Definition H.1 page 221)} \end{array} \right\}$

✎PROOF:

1. Proof using *Pythagorean theorem* (Theorem K.10 page 285):
   Let $(\!(\alpha_n)\!)_{n\in\mathbb{N}}$ be a sequence with at least one nonzero element.

$$\begin{aligned}
\left\| \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n \right\|^2 &= \sum_{n=1}^{N} \|\alpha_n \boldsymbol{x}_n\|^2 && \text{by left hypoth. and } \textit{Pythagorean Theorem} \quad \text{(Theorem K.10 page 285)} \\
&= \sum_{n=1}^{N} |\alpha_n|^2 \|\boldsymbol{x}_n\|^2 && \text{by definition of } \|\cdot\| \quad \text{(Definition L.1 page 287)} \\
&> 0 \\
\implies \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n &\neq 0 \\
\implies (\!(\boldsymbol{x}_n)\!)_{n\in\mathbb{N}} \text{ is } \textit{linearly independent} && \text{by definition of } \textit{linear independence} \quad \text{(Definition H.3 page 221)}
\end{aligned}$$

[19] Aliprantis and Burkinshaw (1998) page 283 ⟨Corollary 32.8⟩, Kubrusly (2001) page 352 ⟨Proposition 5.34⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

2. Alternative proof:

$$\begin{aligned}
\sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n = \mathbb{0} &\Longrightarrow \left\langle \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n \,\middle|\, \boldsymbol{x}_m \right\rangle = \langle \mathbb{0} \mid \boldsymbol{x}_m \rangle \\
&\Longrightarrow \sum_{n=1}^{N} \alpha_n \langle \boldsymbol{x}_n \mid \boldsymbol{x}_m \rangle = 0 \\
&\Longrightarrow \sum_{n=1}^{N} \alpha_n \bar{\delta}(k-m) = 0 \\
&\Longrightarrow \alpha_m = 0 \qquad \text{for } m = 1, 2, \ldots, N
\end{aligned}$$

✏

Theorem H.6 (Bessel's Equality).[20] *Let* $\left\{ \boldsymbol{x}_n \in X \,\middle|\, {\scriptstyle n=1,2,\ldots,N} \right\}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition K.1 page 271)* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\big)$ *and with* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$ *(Definition K.2 page 277)*.

THM

$$\left\{ \begin{array}{c} \{\boldsymbol{x}_n\} \text{ is ORTHONORMAL} \\ \text{(Definition K.4 page 283)} \end{array} \right\} \Longrightarrow \left\{ \underbrace{\left\| \boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2}_{\text{approximation error}} = \|\boldsymbol{x}\|^2 - \sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 \quad \forall \boldsymbol{x} \in X \right\}$$

✎PROOF:

$$\begin{aligned}
&\left\| \boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 \\
&= \|\boldsymbol{x}\|^2 + \left\| \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 - 2\Re \left\langle \boldsymbol{x} \,\middle|\, \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\rangle && \text{by } \textit{polar identity} && \text{(Lemma K.1 page 275)} \\
&= \|\boldsymbol{x}\|^2 + \left\| \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 - 2\Re \left[\left(\sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right)^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right] && \text{by property of } \langle \triangle \mid \triangledown \rangle && \text{(Definition K.1 page 271)} \\
&= \|\boldsymbol{x}\|^2 + \sum_{n=1}^{N} \left\| \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 - 2\Re \left[\left(\sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right)^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right] && \text{by } \textit{Pythagorean Theorem} && \text{(Theorem K.10 page 285)} \\
&= \|\boldsymbol{x}\|^2 + \sum_{n=1}^{N} \left\| \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 - 2\Re \left(\sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right) \\
&= \|\boldsymbol{x}\|^2 + \sum_{n=1}^{N} \left| \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right|^2 \underbrace{\|\boldsymbol{x}_n\|^2}_{1} - 2\Re \left(\sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right) && \text{by property of } \|\cdot\| && \text{(Definition L.1 page 287)} \\
&= \|\boldsymbol{x}\|^2 + \sum_{n=1}^{N} \left| \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right|^2 \cdot 1 - 2\Re \left(\sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right) && \text{by def. of } \textit{orthonormality} && \text{(Definition K.4 page 283)} \\
&= \|\boldsymbol{x}\|^2 + \sum_{n=1}^{N} \left| \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right|^2 - 2\Re \sum_{n=1}^{N} \left| \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right|^2 \\
&= \|\boldsymbol{x}\|^2 + \sum_{n=1}^{N} \left| \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right|^2 - 2 \sum_{n=1}^{N} \left| \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right|^2 && \text{because } |\cdot| \text{ is real}
\end{aligned}$$

[20] Bachman et al. (2000) page 103, Pedersen (2000) pages 38–39

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

$$= \|\boldsymbol{x}\|^2 - \sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2$$

✏

Theorem H.7 (Bessel's inequality). [21] *Let* $\left\{ \boldsymbol{x}_n \in X \middle|_{n=1,2,\ldots,N} \right\}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition K.1 page 271)* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *and with* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$ *(Definition K.2 page 277)*.

THM

$$\left\{ \begin{array}{c} \{\boldsymbol{x}_n\} \text{ is ORTHONORMAL} \\ \text{(Definition K.4 page 283)} \end{array} \right\} \implies \left\{ \sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2 \leq \|\boldsymbol{x}\|^2 \quad \forall \boldsymbol{x} \in X \right\}$$

✎PROOF:

$$0 \leq \left\| \boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 \qquad \text{by definition of } \|\cdot\| \qquad \text{(Definition L.1 page 287)}$$

$$= \|\boldsymbol{x}\|^2 - \sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2 \qquad \text{by } \textit{Bessel's Equality} \qquad \text{(Theorem H.6 page 229)}$$

✏

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 H.8 (Best Approximation Theorem). [22] *Let* $\left\{ \boldsymbol{x}_n \in X \middle|_{n=1,2,\ldots,N} \right\}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition K.1 page 271)* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *and with* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$ *(Definition K.2 page 277)*.

THM

$$\left\{ \begin{array}{c} \{\boldsymbol{x}_n\} \text{ is} \\ \text{ORTHONORMAL} \\ \text{(Definition K.4 page 283)} \end{array} \right\} \implies \left\{ \begin{array}{lll} 1. & \arg\min\limits_{(\alpha_n)_{n=1}^N} \left\| \boldsymbol{x} - \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n \right\| = \underbrace{(\!(\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle)\!)_{n=1}^N}_{\text{best } \alpha_n = \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle} & \forall \boldsymbol{x} \in X \quad \textit{and} \\ 2. & \underbrace{\left(\sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right)}_{\textit{approximation}} \perp \underbrace{\left(\boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right)}_{\textit{approximation error}} & \forall \boldsymbol{x} \in X \end{array} \right\}$$

✎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)

COPYRIGHT © 2019 DANIEL J. GREENHOE

1. Proof that $(\!(\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle)\!)$ is the best sequence:

$$
\begin{aligned}
&\left\| \boldsymbol{x} - \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n \right\|^2 \\
&= \|\boldsymbol{x}\|^2 - 2\Re\left\langle \boldsymbol{x} \mid \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n \right\rangle + \left\| \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n \right\|^2 \\
&= \|\boldsymbol{x}\|^2 - 2\Re\left(\sum_{n=1}^{N} \alpha_n^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right) + \sum_{n=1}^{N} \|\alpha_n \boldsymbol{x}_n\|^2 \qquad \text{by } \textit{Pythagorean Theorem} \qquad \text{(Theorem K.10 page 285)} \\
&= \|\boldsymbol{x}\|^2 - 2\Re\left(\sum_{n=1}^{N} \alpha_n^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right) + \sum_{n=1}^{N} |\alpha_n|^2 + \underbrace{\left[\sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2 - \sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2 \right]}_{0} \\
&= \left[\|\boldsymbol{x}\|^2 - \sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 \right] + \sum_{n=1}^{N} \left[\left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2 - 2\mathbf{R}_{\mathrm{e}} \left[\alpha_n^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right] + \left|\alpha_n\right|^2 \right] \\
&= \left[\|\boldsymbol{x}\|^2 - \sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 \right] + \sum_{n=1}^{N} \left[|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 - \alpha_n^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle - \alpha_n \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle^* + |\alpha_n|^2 \right] \\
&= \left\| \boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 + \sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle - \alpha_n|^2 \qquad \text{by } \textit{Bessel's Equality} \qquad \text{(Theorem H.6 page 229)} \\
&\geq \left\| \boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2
\end{aligned}
$$

2. Proof that the approximation and approximation error are orthogonal:

$$
\begin{aligned}
\left\langle \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \,\middle|\, \boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\rangle &= \left\langle \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \,\middle|\, \boldsymbol{x} \right\rangle - \left\langle \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \,\middle|\, \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\rangle \\
&= \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle^* \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle - \sum_{n=1}^{N} \sum_{m=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \langle \boldsymbol{x} \mid \boldsymbol{x}_m \rangle^* \langle \boldsymbol{x}_n \mid \boldsymbol{x}_m \rangle \\
&= \sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 - \sum_{n=1}^{N} \sum_{m=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \langle \boldsymbol{x} \mid \boldsymbol{x}_m \rangle^* \bar{\delta}_{nm} \\
&= \sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 - \sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 \\
&= 0
\end{aligned}
$$

✍

H.5 Orthonormal bases in Hilbert spaces

Definition H.9. *Let* $\left\{ \boldsymbol{x}_n \in X \,\middle|\, {\scriptstyle n=1,2,\ldots,N} \right\}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition K.1 page 271)* $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle \right)$.

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

DEF

The set $\{\boldsymbol{x}_n\}$ is an ***orthogonal basis*** *for $\boldsymbol{\Omega}$ if $\{\boldsymbol{x}_n\}$ is* ORTHOGONAL *and is a* SCHAUDER BASIS *for $\boldsymbol{\Omega}$.*
The set $\{\boldsymbol{x}_n\}$ is an ***orthonormal basis*** *for $\boldsymbol{\Omega}$ if $\{\boldsymbol{x}_n\}$ is* ORTHONORMAL *and is a* SCHAUDER BASIS *for $\boldsymbol{\Omega}$.*

Definition H.10. [23] *Let $\boldsymbol{H} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ be a Hilbert space.*

DEF

Suppose there exists a set $\{\boldsymbol{x}_n \in X \,|\, {}_{n\in\mathbb{N}}\}$ such that $\boldsymbol{x} \overset{*}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \,|\, \boldsymbol{x}_n\rangle\, \boldsymbol{x}_n$.

Then the quantities $\langle \boldsymbol{x} \,|\, \boldsymbol{x}_n\rangle$ are called the ***Fourier coefficients*** *of $\boldsymbol{x}$ and the sum $\sum_{n=1}^{\infty} \langle \boldsymbol{x} \,|\, \boldsymbol{x}_n\rangle\, \boldsymbol{x}_n$ is called the* ***Fourier expansion*** *of $\boldsymbol{x}$ or the* ***Fourier series*** *for $\boldsymbol{x}$.*

Lemma H.2 (Perfect reconstruction)**.** *Let $\{\boldsymbol{x}_n \in X \,|\, {}_{n\in\mathbb{N}}\}$ be a set of vectors in a* HILBERT SPACE *$\boldsymbol{H} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$.*

LEM

$$\left\{\begin{array}{ll} (1). & (\boldsymbol{x}_n) \text{ is a BASIS for } \boldsymbol{H} \quad \text{and} \\ (2). & (\boldsymbol{x}_n) \text{ is ORTHONORMAL} \end{array}\right\} \implies \boldsymbol{x} \overset{*}{=} \overbrace{\sum_{n=1}^{\infty} \underbrace{\langle \boldsymbol{x} \,|\, \boldsymbol{x}_n\rangle}_{\text{Fourier coefficient}}\, \boldsymbol{x}_n}^{\text{Fourier expansion}} \quad \forall \boldsymbol{x}\in X$$

✎PROOF:

$$\begin{aligned}
\langle \boldsymbol{x} \,|\, \boldsymbol{x}_n\rangle &= \left\langle \sum_{m\in\mathbb{Z}} \alpha_m \boldsymbol{x}_m \,|\, \boldsymbol{x}_n \right\rangle && \text{by left hypothesis (1)} \\
&= \sum_{m\in\mathbb{Z}} \alpha_m \langle \boldsymbol{x}_m \,|\, \boldsymbol{x}_n\rangle && \text{by } \textit{homogeneous} \text{ property of } \langle\triangle\,|\,\triangledown\rangle && \text{(Definition K.1 page 271)} \\
&= \sum_{m\in\mathbb{Z}} \alpha_m \bar{\delta}_{n-m} && \text{by left hypothesis (2)} && \text{(Definition K.4 page 283)} \\
&= \alpha_n
\end{aligned}$$

✏

Proposition H.2. [24] *Let $\{\boldsymbol{x}_n \in X \,|\, {}_{n\in\mathbb{N}}\}$ be a set of vectors in a* HILBERT SPACE *$\boldsymbol{H} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$.*

PRP

$$\underbrace{\|\boldsymbol{x}\|^2 \overset{*}{=} \sum_{n=1}^{\infty} \big|\langle \boldsymbol{x} \,|\, \boldsymbol{x}_n\rangle\big|^2}_{\text{PARSEVAL FRAME}} \iff \underbrace{\boldsymbol{x} \overset{*}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \,|\, \boldsymbol{x}_n\rangle\, \boldsymbol{x}_n}_{\text{FOURIER EXPANSION (Definition H.10 page 232)}} \quad \forall \boldsymbol{x}\in X$$

✎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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

1. Proof that *Parseval frame* $\Longleftarrow$ *Fourier expansion*

$$\begin{aligned}
\|\boldsymbol{x}\|^2 &\triangleq \langle \boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by definition of } \|\cdot\| && \text{(Definition L.1 page 287)}\\
&= \left\langle \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\, \boldsymbol{x} \mid \boldsymbol{x}_n \right\rangle && \text{by right hypothesis} && \\
&\overset{\star}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle && \text{by property of } \langle \triangle \mid \triangledown\rangle && \text{(Definition K.1 page 271)}\\
&\overset{\star}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle^* && \text{by property of } \langle \triangle \mid \triangledown\rangle && \text{(Definition K.1 page 271)}\\
&\overset{\star}{=} \sum_{n=1}^{\infty} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 && \text{by property of } \mathbb{C} && \text{(Definition J.7 page 269)}
\end{aligned}$$

2. Proof that *Parseval frame* $\Longrightarrow$ *Fourier expansion*

   (a) Let $(\!(\boldsymbol{e}_n)\!)_{n\in\mathbb{N}}$ be the *standard othornormal basis* such that the nth element of $\boldsymbol{e}_n$ is 1 and all other elements are 0.

   (b) Let $\mathbf{M}$ be an operator in $\boldsymbol{H}$ such that $\mathbf{M}\boldsymbol{x} \triangleq \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\, \boldsymbol{e}_n$.

   (c) lemma: $\mathbf{M}$ is *isometric*. Proof:

$$\begin{aligned}
\|\mathbf{M}\boldsymbol{x}\|^2 &= \left\| \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\, \boldsymbol{e}_n \right\|^2 && \text{by definition of } \mathbf{M} && \text{(item (2b) page 233)}\\
&= \sum_{n=1}^{\infty} \left\| \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\, \boldsymbol{e}_n \right\|^2 && \text{by } \textit{Pythagorean Theorem} && \text{(Theorem K.10 page 285)}\\
&= \sum_{n=1}^{\infty} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 \left\|\boldsymbol{e}_n\right\|^2 && \text{by } \textit{homogeneous} \text{ property of } \|\cdot\| && \text{(Definition L.1 page 287)}\\
&= \sum_{n=1}^{\infty} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 && \text{by definition of } \textit{orthonormal} && \text{(Definition K.4 page 283)}\\
&= \|\boldsymbol{x}\|^2 && \text{by Parseval frame hypothesis} && \\
&\Longrightarrow \quad \mathbf{M} \text{ is } \textit{isometric} && \text{by definition of } \textit{isometric} && \text{(Definition O.10 page 339)}
\end{aligned}$$

   (d) Let $(\!(\boldsymbol{u}_n)\!)_{n\in\mathbb{N}}$ be an *orthornormal basis* for $\boldsymbol{H}$.

   (e) Proof for *Fourier expansion*:

$$\begin{aligned}
\boldsymbol{x} &= \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{u}_n\rangle\, \boldsymbol{u}_n && \text{by } \textit{Fourier expansion} \text{ (Proposition H.3 page 236)}\\
&= \sum_{n=1}^{\infty} \langle \mathbf{M}\boldsymbol{x} \mid \mathbf{M}\boldsymbol{u}_n\rangle\, \boldsymbol{u}_n && \text{by (2c) lemma page 233 and Theorem O.23 page 340}\\
&= \sum_{n=1}^{\infty} \left\langle \sum_{m=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_m\rangle\, \boldsymbol{e}_m \mid \sum_{k=1}^{\infty} \langle \boldsymbol{u}_n \mid \boldsymbol{x}_k\rangle\, \boldsymbol{e}_k \right\rangle \boldsymbol{u}_n && \text{by item (2b) page 233}\\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_m\rangle \sum_{k=1}^{\infty} \langle \boldsymbol{u}_n \mid \boldsymbol{x}_k\rangle^* \langle \boldsymbol{e}_m \mid \boldsymbol{e}_k\rangle\, \boldsymbol{u}_n && \text{by Definition K.1 page 271}\\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_m\rangle \langle \boldsymbol{u}_n \mid \boldsymbol{x}_m\rangle^*\, \boldsymbol{u}_n && \text{by item (2a) page 233 and Definition K.4 page 283}
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

$$= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_m\rangle \langle \boldsymbol{x}_m \mid \boldsymbol{u}_n\rangle \boldsymbol{u}_n \qquad \text{by Definition K.1 page 271}$$

$$= \sum_{m=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_m\rangle \sum_{n=1}^{\infty} \langle \boldsymbol{x}_m \mid \boldsymbol{u}_n\rangle \boldsymbol{u}_n$$

$$= \sum_{m=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_m\rangle \boldsymbol{x}_m \qquad \text{by item (2d) page 233}$$

✏

When is a set of orthonormal vectors in a Hilbert space $\boldsymbol{H}$ *total*? Theorem H.9 (next) offers some help.

Theorem H.9 (The Fourier Series Theorem). [25] *Let* $\left\{\boldsymbol{x}_n \in X \,\middle|\, n\in\mathbb{N}\right\}$ *be a set of vectors in a* HILBERT SPACE $\boldsymbol{H} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle \mid \triangledown\rangle\right)$ *and let* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle}$ *(Definition K.2 page 277)*.

THM

(A) $\left\{\boldsymbol{x}_n\right\}$ *is* ORTHONORMAL *in* $\boldsymbol{H}$ $\implies$

$$\left\{\begin{array}{llrcll}
 & (1). & (\operatorname{span}\{\boldsymbol{x}_n\})^{-} & = & \boldsymbol{H} & & (\{\boldsymbol{x}_n\} \text{ is TOTAL in } \boldsymbol{H}) \\
\iff & (2). & \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle & \stackrel{\star}{=} & \displaystyle\sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \langle \boldsymbol{y} \mid \boldsymbol{x}_n\rangle^* & \forall \boldsymbol{x},\boldsymbol{y}\in \boldsymbol{X} & \text{(GENERALIZED PARSEVAL'S IDENTITY)} \\
\iff & (3). & \|\boldsymbol{x}\|^2 & \stackrel{\star}{=} & \displaystyle\sum_{n=1}^{\infty} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 & \forall \boldsymbol{x}\in X & \text{(PARSEVAL'S IDENTITY)} \\
\iff & (4). & \boldsymbol{x} & \stackrel{\star}{=} & \displaystyle\sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \boldsymbol{x}_n & \forall \boldsymbol{x}\in X & \text{(FOURIER SERIES EXPANSION)}
\end{array}\right\}$$

✎PROOF:

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

$$\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle \stackrel{\star}{=} \left\langle \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \boldsymbol{x}_n \,\middle|\, \sum_{m=1}^{\infty} \langle \boldsymbol{y} \mid \boldsymbol{x}_m\rangle \boldsymbol{x}_m \right\rangle \qquad \text{by (A) and (1)}$$

$$\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \left\langle \boldsymbol{x}_n \,\middle|\, \sum_{m=1}^{\infty} \langle \boldsymbol{y} \mid \boldsymbol{x}_m\rangle \boldsymbol{x}_m \right\rangle \qquad \text{by property of } \langle\cdot\mid\cdot\rangle \quad \text{(Definition K.1 page 271)}$$

$$\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \sum_{m=1}^{\infty} \langle \boldsymbol{y} \mid \boldsymbol{x}_m\rangle^* \langle \boldsymbol{x}_n \mid \boldsymbol{x}_m\rangle \qquad \text{by property of } \langle\cdot\mid\cdot\rangle \quad \text{(Definition K.1 page 271)}$$

$$\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \sum_{m=1}^{\infty} \langle \boldsymbol{y} \mid \boldsymbol{x}_m\rangle^* \bar{\delta}_{mn} \qquad \text{by (A)}$$

$$\stackrel{\star}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \langle \boldsymbol{y} \mid \boldsymbol{x}_n\rangle^* \qquad \text{by definition of } \bar{\delta}_n \quad \text{(Definition K.3 page 283)}$$

[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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

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

$$\begin{aligned}
\|\boldsymbol{x}\|^2 &\triangleq \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle && \text{by definition of } \textit{induced norm} && \text{(Theorem K.4 page 276)}\\
&= \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle^* && \text{by (2)}\\
&= \sum_{n=1}^{\infty} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2
\end{aligned}$$

3. Proof that (3) $\Longleftrightarrow$ (4) *not* using (A): by Proposition H.2 page 232

4. Proof that (3) $\Longrightarrow$ (1) (proof by contradiction):

   (a) Suppose $\{\boldsymbol{x}_n\}$ is *not total*.

   (b) Then there must exist a vector $\boldsymbol{y}$ in $\boldsymbol{H}$ such that the set $B \triangleq \{\boldsymbol{x}_n\} \cup \boldsymbol{y}$ is *orthonormal*.

   (c) Then $1 = \|y\|^2 \neq \sum_{n=1}^{\infty} \left|\langle \boldsymbol{y} \mid \boldsymbol{x}_n \rangle\right|^2 = 0$.

   (d) But this contradicts (3), and so $\{\boldsymbol{x}_n\}$ must be *total* and (3) $\Longrightarrow$ (1).

5. Extraneous proof that (3) $\Longrightarrow$ (4) (this proof is not really necessary here):

$$\begin{aligned}
\left\| \boldsymbol{x} - \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 &= \|\boldsymbol{x}\|^2 - \sum_{n=1}^{\infty} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 && \text{by } \textit{Bessel's Equality} && \text{(Theorem H.6 page 229)}\\
&= 0 && \text{by (3)}\\
\Longrightarrow \quad \boldsymbol{x} &\stackrel{*}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n && \text{by definition of } \stackrel{*}{=}
\end{aligned}$$

6. Extraneous proof that (A) $\Longrightarrow$ (4) (this proof is not really necessary here)

   (a) The sequence $\sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2$ is *monotonically increasing* in n.

   (b) By Bessel's inequality (page 230), the sequence is upper bounded by $\|\boldsymbol{x}\|^2$:

$$\sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2 \leq \|\boldsymbol{x}\|^2$$

   (c) Because this sequence is both monotonically increasing and bounded in n, it must equal its bound in the limit as n approaches infinity:

$$\lim_{N\to\infty} \sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle\right|^2 = \|\boldsymbol{x}\|^2 \tag{H.1}$$

   (d) If we combine this result with *Bessel's Equality* (Theorem H.6 page 229) we have

$$\begin{aligned}
\lim_{N\to\infty} \left\| \boldsymbol{x} - \sum_{n=1}^{N} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 &= \|\boldsymbol{x}\|^2 - \lim_{N\to\infty} \sum_{n=1}^{N} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 && \text{by } \textit{Bessel's Equality} \text{ (Theorem H.6 page 229)}\\
&= \|\boldsymbol{x}\|^2 - \|\boldsymbol{x}\|^2 && \text{by equation (H.1) page 235}\\
&= 0
\end{aligned}$$

Proposition H.3 (Fourier expansion). *Let* $\left\{ \boldsymbol{x}_n \in X \,\middle|\, {}_{n\in\mathbb{N}} \right\}$ *be a set of vectors in a* HILBERT SPACE $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$.

PRP

$$\underbrace{\{\boldsymbol{x}_n\} \text{ is an ORTHONORMAL BASIS for } \boldsymbol{H}}_{(A)} \quad \Longrightarrow \quad \left\{ \underbrace{\boldsymbol{x} \overset{*}{=} \sum_{n=1}^{\infty} \alpha_n \boldsymbol{x}_n}_{(1)} \quad \Longleftrightarrow \quad \underbrace{\alpha_n = \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle}_{(2)} \right\}$$

✎PROOF:

1. Proof that (1) $\Longrightarrow$ (2): by Lemma H.2 page 232
2. Proof that (1) $\Longleftarrow$ (2):

$$\begin{aligned}
\left\| \boldsymbol{x} - \sum_{n=1}^{\infty} \alpha_n \boldsymbol{x}_{n\in\mathbb{Z}} \right\|^2 &= \left\| \boldsymbol{x} - \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_{n\in\mathbb{Z}} \right\|^2 && \text{by right hypothesis} \\
&= \|\boldsymbol{x}\|^2 - \sum_{n=1}^{\infty} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle|^2 && \text{by Bessel's equality} && \text{(Theorem H.6 page 229)} \\
&= 0 && \text{by } \textit{Parseval's Identity} && \text{(Theorem H.9 page 234)} \\
&\overset{\text{def}}{\Longleftrightarrow} \quad \boldsymbol{x} \overset{*}{=} \sum_{n=1}^{\infty} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n && \text{by definition of } \textit{strong convergence}
\end{aligned}$$

✏

Proposition H.4 (Riesz-Fischer Theorem). [26] *Let* $\left\{ \boldsymbol{x}_n \in X \,\middle|\, {}_{n\in\mathbb{N}} \right\}$ *be a set of vectors in a* HILBERT SPACE $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$.

PRP

$$\underbrace{\{\boldsymbol{x}_n\} \text{ is an ORTHONORMAL BASIS for } \boldsymbol{H}}_{(A)} \quad \Longrightarrow \quad \left\{ \underbrace{\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty}_{(1)} \quad \Longleftrightarrow \quad \underbrace{\begin{array}{l} \exists \boldsymbol{x}\in\boldsymbol{H} \text{ such that} \\ \alpha_n = \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \end{array}}_{(2)} \right\}$$

✎PROOF:

1. Proof that (1) $\Longrightarrow$ (2):
   (a) If (1) is true, then let $\boldsymbol{x} \triangleq \sum_{n\in\mathbb{N}} \alpha_n \boldsymbol{x}_n$.
   (b) Then

$$\begin{aligned}
\langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle &= \left\langle \sum_{m\in\mathbb{N}} \alpha_m \boldsymbol{x}_m \,\middle|\, \boldsymbol{x}_n \right\rangle && \text{by definition of } \boldsymbol{x} \\
&= \sum_{m\in\mathbb{N}} \alpha_m \langle \boldsymbol{x}_m \mid \boldsymbol{x}_n \rangle && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \mid \triangledown \rangle && \text{(Definition K.1 page 271)} \\
&= \sum_{m\in\mathbb{N}} \alpha_m \bar{\delta}_{mn} && \text{by (A)} \\
&= \sum_{m\in\mathbb{N}} \alpha_n && \text{by definition of } \bar{\delta} && \text{(Definition K.3 page 283)}
\end{aligned}$$

[26] Young (2001) page 6

COPYRIGHT © 2019 DANIEL J. GREENHOE

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

$$\begin{aligned}
\sum_{n\in\mathbb{N}} |\alpha_n|^2 &= \sum_{n\in\mathbb{N}} |\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle|^2 && \text{by (2)} \\
&\le \|\boldsymbol{x}\|^2 && \text{by } \textit{Bessel's Inequality} \qquad \text{(Theorem H.7 page 230)} \\
&\le \infty
\end{aligned}$$

✏

Theorem H.10. [27]

THM

All SEPARABLE HILBERT SPACES *are* ISOMORPHIC. *That is,*

$$\left\{\begin{array}{l} \boldsymbol{X} \text{ is a separable} \\ \text{Hilbert space} \quad \textit{and} \\ \boldsymbol{Y} \text{ is a separable} \\ \text{Hilbert space} \end{array}\right\} \Longrightarrow \left\{\begin{array}{lllll} \textit{there is a BIJECTIVE operator } \mathbf{M} \in \boldsymbol{Y}^{\boldsymbol{X}} \textit{ such that} & & & & \\ (1). \quad \boldsymbol{y} & = & \mathbf{M}\boldsymbol{x} & \forall \boldsymbol{x}\in\boldsymbol{X}, \boldsymbol{y}\in\boldsymbol{Y} & \textit{and} \\ (2). \quad \|\mathbf{M}\boldsymbol{x}\| & = & \|\boldsymbol{x}\| & \forall \boldsymbol{x}\in\boldsymbol{X} & \textit{and} \\ (3). \quad \langle \mathbf{M}\boldsymbol{x} \mid \mathbf{M}\boldsymbol{y}\rangle & = & \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle & \forall \boldsymbol{x}\in\boldsymbol{X}, \boldsymbol{y}\in\boldsymbol{Y} & \end{array}\right\}$$

✎PROOF:

1. Let $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown\rangle\big)$ be a *separable Hilbert space* with *orthonormal basis* $\{\boldsymbol{x}_n \mid n\in\mathbb{N}\}$.
   Let $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown\rangle\big)$ be a *separable Hilbert space* with *orthonormal basis* $\{\boldsymbol{y}_n \mid n\in\mathbb{N}\}$.

2. Proof that there exists *bijective* operator $\mathbf{M}$ and its inverse $\mathbf{M}^{-1}$ between $\{\boldsymbol{x}_n\}$ and $\{\boldsymbol{y}_n\}$:

   (a) Let $\mathbf{M}$ be defined such that $\boldsymbol{y}_n \triangleq \mathbf{M}\boldsymbol{x}_n$.

   (b) Thus $\mathbf{M}$ is a *bijection* between $\{\boldsymbol{x}_n\}$ and $\{\boldsymbol{y}_n\}$.

   (c) Because $\mathbf{M}$ is a *bijection* between $\{\boldsymbol{x}_n\}$ and $\{\boldsymbol{y}_n\}$, $\mathbf{M}$ has an inverse operator $\mathbf{M}^{-1}$ between $\{\boldsymbol{x}_n\}$ and $\{\boldsymbol{y}_n\}$ such that $\boldsymbol{x}_n = \mathbf{M}^{-1}\boldsymbol{y}_n$.

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* $\boldsymbol{Y}$:

$$\begin{aligned}
\boldsymbol{x} \in \boldsymbol{X} \iff \boldsymbol{x} &\stackrel{*}{=} \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \boldsymbol{x}_n && \text{by } \textit{Fourier expansion} \qquad \text{(Theorem H.9 page 234)} \\
\Longrightarrow \exists \boldsymbol{y} \in \boldsymbol{Y} \quad &\text{such that} \quad \langle \boldsymbol{y} \mid \boldsymbol{y}_n\rangle = \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle && \text{by } \textit{Riesz-Fischer Thm.} \qquad \text{(Proposition H.4 page 236)} \\
\Longrightarrow & && \\
\boldsymbol{y} &= \sum_{n\in\mathbb{N}} \langle \boldsymbol{y} \mid \boldsymbol{y}_n\rangle \boldsymbol{y}_n && \text{by } \textit{Fourier expansion} \qquad \text{(Theorem H.9 page 234)} \\
&= \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \boldsymbol{y}_n && \text{by } \textit{Riesz-Fischer Thm.} \qquad \text{(Proposition H.4 page 236)} \\
&= \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \mathbf{M}\boldsymbol{x}_n && \text{by definition of } \mathbf{M} \qquad \text{(item (2a) page 237)} \\
&= \mathbf{M} \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle \boldsymbol{x}_n && \text{by prop. of linear ops.} \qquad \text{(Theorem O.1 page 320)} \\
&= \mathbf{M}\boldsymbol{x} && \text{by definition of } \boldsymbol{x}
\end{aligned}$$

[27] Young (2001) page 6

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

(b) Proof that $\mathbf{M}^{-1}$ maps $\boldsymbol{Y}$ *into* $\boldsymbol{X}$:

$$\begin{aligned}
\boldsymbol{y} \in \boldsymbol{Y} \iff \boldsymbol{y} \stackrel{*}{=} \sum_{n\in\mathbb{N}} \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle \boldsymbol{y}_n & & \text{by } \textit{Fourier expansion} & \text{(Theorem H.9 page 234)}\\
\implies \exists \boldsymbol{x} \in \boldsymbol{X} \quad \text{such that} \quad \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle = \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle & & \text{by } \textit{Riesz-Fischer Thm.} & \text{(Proposition H.4 page 236)}\\
\implies & & & \\
\boldsymbol{x} = \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n & & \text{by } \textit{Fourier expansion} & \text{(Theorem H.9 page 234)}\\
= \sum_{n\in\mathbb{N}} \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle \boldsymbol{x}_n & & \text{by } \textit{Riesz-Fischer Thm.} & \text{(Proposition H.4 page 236)}\\
= \sum_{n\in\mathbb{N}} \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle \mathbf{M}^{-1} \boldsymbol{y}_n & & \text{by definition of } \mathbf{M}^{-1} & \text{(item (2c) page 237)}\\
= \mathbf{M}^{-1} \sum_{n\in\mathbb{N}} \langle \boldsymbol{y} \mid \boldsymbol{y}_n \rangle \boldsymbol{y}_n & & \text{by prop. of } \textit{linear ops.} & \text{(Theorem O.1 page 320)}\\
= \mathbf{M}^{-1} \boldsymbol{y} & & \text{by definition of } \boldsymbol{y} &
\end{aligned}$$

4. Proof for (2):

$$\begin{aligned}
\|\mathbf{M}\boldsymbol{x}\|^2 &= \left\| \mathbf{M} \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 && \text{by } \textit{Fourier expansion} && \text{(Theorem H.9 page 234)}\\
&= \left\| \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \mathbf{M}\boldsymbol{x}_n \right\|^2 && \text{by property of } \textit{linear operators} && \text{(Theorem O.1 page 320)}\\
&= \left\| \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{y}_n \right\|^2 && \text{by definition of } \mathbf{M} && \text{(item (2a) page 237)}\\
&= \sum_{n\in\mathbb{N}} \left| \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \right|^2 && \text{by } \textit{Parseval's Identity} && \text{(Proposition H.4 page 236)}\\
&= \left\| \sum_{n\in\mathbb{N}} \langle \boldsymbol{x} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \right\|^2 && \text{by } \textit{Parseval's Identity} && \text{(Proposition H.4 page 236)}\\
&= \|\boldsymbol{x}\|^2 && \text{by } \textit{Fourier expansion} && \text{(Theorem H.9 page 234)}
\end{aligned}$$

5. Proof for (3): by (2) and Theorem O.23 page 340

✍

Theorem H.11. [28] *Let* $\boldsymbol{H}$ *be a* HILBERT SPACE.

THM

$\boldsymbol{H}$ *has a* SCHAUDER BASIS $\iff$ $\boldsymbol{H}$ *is* SEPARABLE

Theorem H.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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

H.6 Riesz bases in Hilbert spaces

Definition H.11. [30] *Let* $\{\boldsymbol{x}_n \in X \mid n\in\mathbb{N}\}$ *be a set of vectors in a* SEPARABLE HILBERT SPACE $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}$

DEF $\{\boldsymbol{x}_n\}$ *is a* ***Riesz basis*** *for* $\boldsymbol{H}$ *if* $\{\boldsymbol{x}_n\}$ *is* EQUIVALENT *(Definition H.8 page 227) to some* ORTHONORMAL BASIS *(Definition H.9 page 231) in* $\boldsymbol{H}$.

Definition H.12. [31] *Let* $(\!(\boldsymbol{x}_n \in X)\!)_{n\in\mathbb{N}}$ *be a sequence of vectors in a* SEPARABLE HILBERT SPACE $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\nabla\rangle)$.

DEF *The sequence* $(\!(\boldsymbol{x}_n)\!)$ *is a* ***Riesz sequence*** *for* $\boldsymbol{H}$ *if*

$$\exists A, B \in \mathbb{R}^+ \quad \text{such that} \quad A\sum_{n=1}^{\infty}|\alpha_n|^2 \le \left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 \le B\sum_{n=1}^{\infty}|\alpha_n|^2 \qquad \forall (\!(\alpha_n)\!)\in\ell^2_{\mathbb{F}}.$$

Definition H.13. *Let* $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\nabla\rangle)$ *be an* INNER PRODUCT SPACE *(Definition K.1 page 271)*.

DEF *The sequences* $(\!(\boldsymbol{x}_n \in X)\!)_{n\in\mathbb{Z}}$ *and* $(\!(\boldsymbol{y}_n \in X)\!)_{n\in\mathbb{Z}}$ *are* ***biorthogonal*** *with respect to each other in* $\boldsymbol{X}$ *if* $\langle \boldsymbol{x}_n \,|\, \boldsymbol{y}_m\rangle = \bar{\delta}_{nm}$

Lemma H.3. [32] *Let* $\{\boldsymbol{x}_n \mid n\in\mathbb{N}\}$ *be a sequence in a* HILBERT SPACE $\boldsymbol{X} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\nabla\rangle)$. *Let* $\{\boldsymbol{y}_n \mid n\in\mathbb{N}\}$ *be a sequence in a* HILBERT SPACE $\boldsymbol{Y} \triangleq (Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\nabla\rangle)$. *Let*

LEM

$$\left\{\begin{array}{lll}
\text{(i).} & \{\boldsymbol{x}_n\} \text{ is TOTAL in } \boldsymbol{X} & \text{and} \\
\text{(ii).} & \text{There exists } A>0 \text{ such that } A\sum_{n\in C}|a_n|^2 \le \left\|\sum_{n\in C}a_n\boldsymbol{x}_n\right\|^2 \text{ for finite } C & \text{and} \\
\text{(iii).} & \text{There exists } B>0 \text{ such that } \left\|\sum_{n=1}^{\infty}b_n\boldsymbol{y}_n\right\|^2 \le B\sum_{n=1}^{\infty}|b_n|^2 \quad \forall (\!(b_n)\!)_{n\in\mathbb{N}}\in\ell^2_{\mathbb{F}} &
\end{array}\right\} \implies$$

$$\left\{\begin{array}{lll}
\text{(1).} & \mathbf{R}^\circ \text{ is a linear bounded operator that maps from } \operatorname{span}\{\boldsymbol{x}_n\} \text{ to } \operatorname{span}\{\boldsymbol{y}_n\} & \\
 & \text{where } \mathbf{R}^\circ\sum_{n\in C}c_n\boldsymbol{x}_n \triangleq \sum_{n\in C}c_n\boldsymbol{y}_n, \text{ for some sequence } (\!(c_n)\!) \text{ and finite set } C & \text{and} \\
\text{(2).} & \mathbf{R} \text{ has a unique extension to a bounded operator } \mathbf{R} \text{ that maps from } \boldsymbol{X} \text{ to } \boldsymbol{Y} & \text{and} \\
\text{(3).} & |||\mathbf{R}^\circ||| \le \frac{B}{A} & \text{and} \\
\text{(4).} & |||\mathbf{R}||| \le \frac{B}{A} &
\end{array}\right\}$$

Theorem H.13. [33] *Let* $\{\boldsymbol{x}_n \in X \mid n\in\mathbb{N}\}$ *be a set of vectors in a* SEPARABLE HILBERT SPACE $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\nabla\rangle)$.

THM

$$\left\{\begin{array}{l}\{\boldsymbol{x}_n\} \text{ is a RIESZ BASIS} \\ \text{for } \boldsymbol{H}\end{array}\right\} \iff \left\{\begin{array}{lll}
\text{(1).} & \{\boldsymbol{x}_n\} \text{ is TOTAL in } \boldsymbol{H} & \text{and} \\
\text{(2).} & \exists A,B\in\mathbb{R}^+ \quad \text{such that} \quad \forall (\!(\alpha_n)\!)\in\ell^2_{\mathbb{F}}, & \\
 & A\sum_{n=1}^{\infty}|\alpha_n|^2 \le \left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 \le B\sum_{n=1}^{\infty}|\alpha_n|^2 &
\end{array}\right\}$$

✎PROOF:

[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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

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:

$$\begin{aligned}
\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 &= \left\|\sum_{n=1}^{\infty}\alpha_n\mathbf{R}\boldsymbol{u}_n\right\|^2 && \text{by definition of } \mathbf{R} && \text{(item (1(b)ii))}\\
&= \left\|\mathbf{R}\sum_{n=1}^{\infty}\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by Theorem O.1 page 320} && \\
&\le |||\mathbf{R}|||^2\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by Theorem O.6 page 326} && \\
&= |||\mathbf{R}|||^2\sum_{n=1}^{\infty}\left\|\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by } \textit{Pythagorean Theorem} && \text{(Theorem K.10 page 285)}\\
&= |||\mathbf{R}|||^2\sum_{n=1}^{\infty}|\alpha|^2\left\|\boldsymbol{u}_n\right\|^2 && \text{by } \textit{homogeneous} \text{ property of norms} && \text{(Definition L.1 page 287)}\\
&= \underbrace{|||\mathbf{R}|||^2}_{B}\sum_{n=1}^{\infty}|\alpha|^2 && \text{by definition of } \textit{orthonormality} && \text{(Definition K.4 page 283)}
\end{aligned}$$

      iv. Proof for lower bound A:

$$\begin{aligned}
\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 &= \frac{|||\mathbf{R}^{-1}|||^2}{|||\mathbf{R}^{-1}|||^2}\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 && \text{because } |||\mathbf{R}^{-1}||| > 0 && \text{(Proposition O.1 page 324)}\\
&\ge \frac{1}{|||\mathbf{R}^{-1}|||^2}\left\|\mathbf{R}^{-1}\sum_{n=1}^{\infty}\alpha_n\boldsymbol{x}_n\right\|^2 && \text{by Theorem O.6 page 326} && \\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\left\|\mathbf{R}^{-1}\sum_{n=1}^{\infty}\alpha_n\mathbf{R}\boldsymbol{u}_n\right\|^2 && \text{by definition of } \mathbf{R} && \text{(item (1(b)ii) page 240)}\\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\left\|\mathbf{R}^{-1}\mathbf{R}\sum_{n=1}^{\infty}\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by property of } \textit{linear operators} && \text{(Theorem O.1 page 320)}\\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\left\|\sum_{n=1}^{\infty}\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by definition of inverse op.} && \text{(Definition O.2 page 319)}\\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\sum_{n=1}^{\infty}\left\|\alpha_n\boldsymbol{u}_n\right\|^2 && \text{by } \textit{Pythagorean Theorem} && \text{(Theorem K.10 page 285)}\\
&= \frac{1}{|||\mathbf{R}^{-1}|||^2}\sum_{n=1}^{\infty}\left|\alpha_n\right|^2\left\|\boldsymbol{u}_n\right\|^2 && \text{by } \|\cdot\| \textit{ homogeneous} \text{ prop.} && \text{(Definition L.1 page 287)}\\
&= \underbrace{\frac{1}{|||\mathbf{R}^{-1}|||^2}}_{A}\sum_{n=1}^{\infty}\left|\alpha_n\right|^2 && \text{by def. of } \textit{orthonormality} && \text{(Definition K.4 page 283)}
\end{aligned}$$

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

Copyright © 2019 Daniel J. Greenhoe

(a) Let $\{\boldsymbol{u}_n \mid n\in\mathbb{N}\}$ be an *orthonormal basis* for $\boldsymbol{H}$.

(b) Using (2) and Lemma H.3 (page 239), construct an bounded extension operator $\mathbf{R}$ such that $\mathbf{R}\boldsymbol{u}_n = \boldsymbol{x}_n$ for all $n \in \mathbb{N}$.

(c) Using (2) and Lemma H.3 (page 239), construct an bounded extension operator $\mathbf{S}$ such that $\mathbf{S}\boldsymbol{x}_n = \boldsymbol{u}_n$ for all $n \in \mathbb{N}$.

(d) Then, $\mathbf{RV}\boldsymbol{x} = \mathbf{VR}\boldsymbol{x} \implies \mathbf{V} = \mathbf{R}^{-1}$, and so $\mathbf{R}$ is a bounded invertible operator

(e) and $\{\boldsymbol{x}_n\}$ is a *Riesz sequence*.

✏

Theorem H.14. [34] *Let* $\boldsymbol{H} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle \mid \triangledown\rangle\big)$ *be a* SEPARABLE HILBERT SPACE.

THM

$$\left\{\begin{array}{l}(\!(\boldsymbol{x}_n \in \boldsymbol{H})\!)_{n\in\mathbb{Z}} \text{ is a} \\ \text{RIESZ BASIS for } \boldsymbol{H}\end{array}\right\} \implies \left\{\begin{array}{l}\text{There exists } (\!(\boldsymbol{y}_n \in \boldsymbol{H})\!)_{n\in\mathbb{Z}} \text{ such that} \\ (1).\quad (\!(\boldsymbol{x}_n)\!) \text{ and } (\!(\boldsymbol{y}_n)\!) \text{ are BIORTHOGONAL} \quad \text{and} \\ (2).\quad (\!(\boldsymbol{y}_n)\!) \text{ is also a RIESZ BASIS for } \boldsymbol{H} \quad \text{and} \\ (3).\quad \exists B>A>0 \quad \text{such that} \\ \qquad A\sum_{n=1}^{\infty}|a_n|^2 \le \left\|\sum_{n=1}^{\infty} a_n\boldsymbol{x}_n\right\|^2 = \left\|\sum_{n=1}^{\infty} a_n\boldsymbol{y}_n\right\|^2 \le B\sum_{n=1}^{\infty}|a_n|^2 \\ \qquad \forall (\!(a_n)\!)_{n\in\mathbb{N}} \in \ell^2_{\mathbb{F}}\end{array}\right\}$$

✎PROOF:

1. Proof for (1):

(a) Let $\boldsymbol{e}_n$ be the *unit vector* in $\boldsymbol{H}$ such that the nth element of $\boldsymbol{e}_n$ is 1 and all other elements are 0.

(b) Let $\mathbf{M}$ be an operator on $\boldsymbol{H}$ such that $\mathbf{M}\boldsymbol{e}_n = \boldsymbol{x}_n$.

(c) Note that $\mathbf{M}$ is *isometric*, and as such $\|\mathbf{M}\boldsymbol{x}\| = \|\boldsymbol{x}\| \quad \forall \boldsymbol{x}\in\boldsymbol{H}$.

(d) Let $\boldsymbol{y}_n \triangleq \left(\mathbf{M}^{-1}\right)^*$.

(e) Then,

$$\begin{aligned}\langle \boldsymbol{y}_n \mid \boldsymbol{x}_m\rangle &= \left\langle \left(\mathbf{M}^{-1}\right)^*\boldsymbol{e}_n \mid \mathbf{M}\boldsymbol{e}_m\right\rangle \\ &= \left\langle \boldsymbol{e}_n \mid \mathbf{M}^{-1}\mathbf{M}\boldsymbol{e}_m\right\rangle \\ &= \langle \boldsymbol{e}_n \mid \boldsymbol{e}_m\rangle \\ &= \bar{\delta}_{nm} \\ &\implies \{\boldsymbol{x}_n\} \text{ and } \{\boldsymbol{y}_n\} \text{ are } \textit{biorthogonal} \qquad \text{by Definition K.4 page 283}\end{aligned}$$

[34] Wojtaszczyk (1997) page 20 ⟨Lemma 2.7(a)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

2. Proof for (3):

$$\begin{aligned}
\left\|\sum_{n\in\mathbb{Z}}\alpha_n\boldsymbol{y}_n\right\| &= \left\|\sum_{n\in\mathbb{Z}}\alpha_n\left(\mathbf{M}^{-1}\right)^*\boldsymbol{e}_n\right\| && \text{by definition of } \boldsymbol{y}_n && \text{(Proposition 1d page 241)}\\
&= \left\|\left(\mathbf{M}^{-1}\right)^*\sum_{n\in\mathbb{Z}}\alpha_n\boldsymbol{e}_n\right\| && \text{by property of } \textit{linear ops.} \\
&= \left\|\sum_{n\in\mathbb{Z}}\alpha_n\boldsymbol{e}_n\right\| && \text{because } \left(\mathbf{M}^{-1}\right)^* \text{ is } \textit{isometric} && \text{(Definition O.10 page 339)}\\
&= \left\|\mathbf{M}\sum_{n\in\mathbb{Z}}\alpha_n\boldsymbol{e}_n\right\| && \text{because } \mathbf{M} \text{ is } \textit{isometric} && \text{(Definition O.10 page 339)}\\
&= \left\|\sum_{n\in\mathbb{Z}}\alpha_n\mathbf{M}\boldsymbol{e}_n\right\| && \text{by property of } \textit{linear operators} \\
&= \left\|\sum_{n\in\mathbb{Z}}\alpha_n\boldsymbol{x}_n\right\| && \text{by definition of } \mathbf{M} \\
&\Longrightarrow \{\boldsymbol{y}_n\} \text{ is a } \textit{Riesz basis} && \text{by left hypothesis}
\end{aligned}$$

3. Proof for (2): by (3) and definition of *Riesz basis* (Definition H.11 page 239)

✏

Proposition H.5. [35] *Let* $\{\boldsymbol{x}_n \mid n\in\mathbb{N}\}$ *be a set of vectors in a* HILBERT SPACE $\boldsymbol{H} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\mid\triangledown\rangle\right)$.

PRP

$$\left\{\begin{array}{l}\{\boldsymbol{x}_n\} \textit{ is a } \text{RIESZ BASIS} \textit{ for } \boldsymbol{H} \textit{ with} \\ A\sum_{n=1}^{\infty}|a_n|^2 \le \left\|\sum_{n=1}^{\infty}a_n\boldsymbol{x}_n\right\|^2 \le B\sum_{n=1}^{\infty}|a_n|^2 \\ \forall\{a_n\}\in\ell^2_{\mathbb{F}}\end{array}\right\} \Longrightarrow \left\{\begin{array}{l}\{\boldsymbol{x}_n\} \textit{ is a } \text{FRAME} \textit{ for } \boldsymbol{H} \textit{ with} \\ \underbrace{\frac{1}{B}\|\boldsymbol{x}\|^2 \le \sum_{n=1}^{\infty}|\langle\boldsymbol{x}\mid\boldsymbol{x}_n\rangle|^2 \le \frac{1}{A}\|\boldsymbol{x}\|^2}_{\text{STABILITY CONDITION}} \\ \forall\boldsymbol{x}\in\boldsymbol{H}\end{array}\right\}$$

✎PROOF:

1. Let $\{\boldsymbol{y}_n \mid n\in\mathbb{N}\}$ be a *Riesz basis* that is *biorthogonal* to $\{\boldsymbol{x}_n \mid n\in\mathbb{N}\}$ (Theorem H.14 page 241).

2. Let $\boldsymbol{x} \triangleq \sum_{n=1}^{\infty} a_n\boldsymbol{y}_n$.

3. lemma:

$$\begin{aligned}
\sum_{n=1}^{\infty}|\langle\boldsymbol{x}\mid\boldsymbol{x}_n\rangle|^2 &= \sum_{n=1}^{\infty}\left|\left\langle\sum_{m=1}^{\infty}a_n\boldsymbol{y}_m \mid \boldsymbol{x}_n\right\rangle\right|^2 && \text{by definition of } \boldsymbol{x} && \text{(item (2) page 242)}\\
&= \sum_{n=1}^{\infty}\left|\sum_{m=1}^{\infty}a_n\langle\boldsymbol{y}_m\mid\boldsymbol{x}_n\rangle\right|^2 && \text{by } \textit{homogeneous} \text{ property of } \langle\triangle\mid\triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= \sum_{n=1}^{\infty}\left|\sum_{m=1}^{\infty}a_n\bar{\delta}_{mn}\right|^2 && \text{by definition of } \textit{biorthogonal} && \text{(Definition H.13 page 239)}\\
&= \sum_{n=1}^{\infty}|a_n|^2 && \text{by definition of } \bar{\delta} && \text{(Definition K.3 page 283)}
\end{aligned}$$

[35] Igari (1996) page 220 ⟨Lemma 9.8⟩, Wojtaszczyk (1997) pages 20–21 ⟨Lemma 2.7(a)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

4. Then

$$\begin{aligned}
& A\sum_{n=1}^{\infty}|a_n|^2 && \le && \left\|\sum_{n=1}^{\infty} a_n \boldsymbol{x}_n\right\|^2 && \le && B\sum_{n=1}^{\infty}|a_n|^2 && \text{by definition of } \{\boldsymbol{y}_n\} \text{ (item (1) page 242)} \\
\Longrightarrow\ & A\sum_{n=1}^{\infty}|a_n|^2 && \le && \left\|\sum_{n=1}^{\infty} a_n \boldsymbol{y}_n\right\|^2 && \le && B\sum_{n=1}^{\infty}|a_n|^2 && \text{by definition of } \{\boldsymbol{y}_n\} \text{ (item (1) page 242)} \\
\Longrightarrow\ & A\sum_{n=1}^{\infty}|a_n|^2 && \le && \|\boldsymbol{x}\|^2 && \le && B\sum_{n=1}^{\infty}|a_n|^2 && \text{by definition of } \boldsymbol{x} \text{ (item (2) page 242)} \\
\Longrightarrow\ & A\sum_{n=1}^{\infty}|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle|^2 && \le && \|\boldsymbol{x}\|^2 && \le && B\sum_{n=1}^{\infty}|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle|^2 && \text{by (3) lemma} \\
\Longrightarrow\ & \frac{1}{B}\|\boldsymbol{x}\|^2 && \le && \sum_{n=1}^{\infty}|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle|^2 && \le && \frac{1}{A}\|\boldsymbol{x}\|^2 &&
\end{aligned}$$

✏

Theorem H.15 (Battle-Lemarié orthogonalization). [36] *Let* $\tilde{\mathsf{f}}(\omega)$ *be the* FOURIER TRANSFORM *(Definition P.2 page 349) of a function* $\mathsf{f} \in \boldsymbol{L}^2_{\mathbb{R}}$.

THM

$$\left\{\begin{array}{ll} 1. & \{\mathbf{T}^n\mathsf{g} \mid n\in\mathbb{Z}\} \text{ is a RIESZ BASIS for } \boldsymbol{L}^2_{\mathbb{R}} \text{ and} \\ 2. & \tilde{\mathsf{f}}(\omega) \triangleq \dfrac{\tilde{\mathsf{g}}(\omega)}{\sqrt{2\pi\sum\limits_{n\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi n)|^2}} \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \{\mathbf{T}^n\mathsf{f} \mid n\in\mathbb{Z}\} \\ \text{is an ORTHONORMAL BASIS for } \boldsymbol{L}^2_{\mathbb{R}} \end{array}\right\}$$

✎PROOF:

1. Proof that $\{\mathbf{T}^n\mathsf{f} \mid n\in\mathbb{Z}\}$ is orthonormal:

$$\begin{aligned}
\tilde{\mathsf{S}}_{\phi\phi}(\omega) &= 2\pi\sum_{n\in\mathbb{Z}}\left|\tilde{\mathsf{f}}(\omega+2\pi n)\right|^2 && \text{by Theorem ?? page ??} \\
&= 2\pi\sum_{n\in\mathbb{Z}}\left|\frac{\tilde{\mathsf{g}}(\omega+2\pi n)}{\sqrt{2\pi\sum\limits_{m\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi(n-m))|^2}}\right|^2 && \text{by left hypothesis} \\
&= \sum_{n\in\mathbb{Z}}\left|\frac{\tilde{\mathsf{g}}(\omega+2\pi n)}{\sqrt{\sum\limits_{m\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi m)|^2}}\right|^2 \\
&= \sum_{n\in\mathbb{Z}}\left|\frac{1}{\sqrt{\sum\limits_{m\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi m)|^2}}\right|^2|\tilde{\mathsf{g}}(\omega+2\pi n)|^2 \\
&= \frac{1}{\sum\limits_{m\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi m)|^2}\sum_{n\in\mathbb{Z}}|\tilde{\mathsf{g}}(\omega+2\pi n)|^2 \\
&= 1 \\
\Longrightarrow \{\mathsf{f}_n \mid n\in\mathbb{Z}\} &\text{ is orthonormal} && \text{by Theorem ?? page ??}
\end{aligned}$$

[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)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

2. Proof that $\{\mathbf{T}^n \mathsf{f} \mid n\in\mathbb{Z}\}$ is a basis for V_0: by Lemma H.1 page 227.

✏

H.7 Frames in Hilbert spaces

Definition H.14. [37] *Let* $\{\boldsymbol{x}_n \in X \mid n\in\mathbb{N}\}$ *be a set of vectors in a* HILBERT SPACE $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\mid\triangledown\rangle)$.

DEF

The set $\{\boldsymbol{x}_n\}$ *is a* ***frame*** *for* $\boldsymbol{H}$ *if* (STABILITY CONDITION)

$$\exists A, B \in \mathbb{R}^+ \quad \text{such that} \quad A\,\|\boldsymbol{x}\|^2 \le \sum_{n=1}^{\infty} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 \le B\,\|\boldsymbol{x}\|^2 \qquad \forall \boldsymbol{x}\in X.$$

The quantities A *and* B *are* ***frame bounds***.
The quantity A' *is the* ***optimal lower frame bound*** *if*
$A' = \sup\{A \in \mathbb{R}^+ \mid A \text{ is a lower frame bound}\}$.
The quantity B' *is the* ***optimal upper frame bound*** *if*
$B' = \inf\{B \in \mathbb{R}^+ \mid B \text{ is an upper frame bound}\}$.
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 $\{\boldsymbol{x}_n \mid n\in\mathbb{N}\}$ *is an* ***exact frame*** *if for some* $m \in \mathbb{Z}$, $\{\boldsymbol{x}_n \mid n\in\mathbb{N}\}\setminus\{\boldsymbol{x}_m\}$ *is* NOT *a frame.*

A frame is a *Parseval frame* (Definition H.14) if it satisfies *Parseval's Identity* (Theorem H.9 page 234). All orthonormal bases are Parseval frames (Theorem H.9 page 234); but not all Parseval frames are orthonormal bases.

Definition H.15. *Let* $\{\boldsymbol{x}_n\}$ *be a* ***frame*** *(Definition H.14 page 244) for the* HILBERT SPACE $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\mid\triangledown\rangle)$. *Let* $\mathbf{S}$ *be an* OPERATOR *on* $\boldsymbol{H}$.

DEF

$\mathbf{S}$ *is a* ***frame operator*** *for* $\{\boldsymbol{x}_n\}$ *if* $\quad \mathbf{S}\mathsf{f}(x) = \sum_{n\in\mathbb{Z}} \langle \mathsf{f} \mid \boldsymbol{x}_n\rangle\, \boldsymbol{x}_n \qquad \forall \mathsf{f}\in\boldsymbol{H}.$

Theorem H.16. [38] *Let* $\mathbf{S}$ *be a* FRAME OPERATOR *(Definition H.15 page 244) of a* FRAME $\{\boldsymbol{x}_n\}$ *(Definition H.14 page 244) for the* HILBERT SPACE $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\mid\triangledown\rangle)$.

THM

(1). $\mathbf{S}$ *is* INVERTIBLE. *and*

(2). $\mathsf{f}(x) = \sum_{n\in\mathbb{Z}} \left\langle \mathsf{f} \mid \mathbf{S}^{-1}\boldsymbol{x}_n\right\rangle \boldsymbol{x}_n = \sum_{n\in\mathbb{Z}} \langle \mathsf{f} \mid \boldsymbol{x}_n\rangle\, \mathbf{S}^{-1}\boldsymbol{x}_n \quad \forall \mathsf{f} \in \boldsymbol{H}$

Theorem H.17. [39] *Let* $\{\boldsymbol{x}_n \in X \mid n=1,2,\ldots,N\}$ *be a set of vectors in a* HILBERT SPACE $\boldsymbol{H} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\mid\triangledown\rangle)$.

THM

$\{\boldsymbol{x}_n\}$ *is a* FRAME *for* $\operatorname{span}\{\boldsymbol{x}_n\}$.

✎PROOF:

[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

COPYRIGHT © 2019 DANIEL J. GREENHOE

1. Upper bound: Proof that there exists B such that $\sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 \le B\,\|\boldsymbol{x}\|^2 \quad \forall \boldsymbol{x} \in \boldsymbol{H}$:

$$\begin{aligned}\sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 &\le \sum_{n=1}^{N} \langle \boldsymbol{x}_n \mid \boldsymbol{x}_n\rangle \langle \boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by } \textit{Cauchy-Schwarz inequality} \text{ (Theorem K.2 page 272)}\\ &= \underbrace{\left\{\sum_{n=1}^{N} \|\boldsymbol{x}_n\|^2\right\}}_{B} \|\boldsymbol{x}\|^2\end{aligned}$$

2. Lower bound: Proof that there exists A such that $A\,\|\boldsymbol{x}\|^2 \le \sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 \quad \forall \boldsymbol{x} \in \boldsymbol{H}$:

$$\begin{aligned}\sum_{n=1}^{N} \left|\langle \boldsymbol{x} \mid \boldsymbol{x}_n\rangle\right|^2 &= \sum_{n=1}^{N} \left|\left\langle \boldsymbol{x}_n \mid \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right\rangle\right|^2 \|\boldsymbol{x}\|^2\\ &\ge \underbrace{\left(\inf_{\boldsymbol{y}} \left\{\sum_{n=1}^{N} \left|\langle \boldsymbol{x}_n \mid \boldsymbol{y}\rangle\right|^2 \,\middle|\, \|\boldsymbol{y}\| = 1\right\}\right)}_{A} \|\boldsymbol{x}\|^2\end{aligned}$$

✏

Example H.1. Let $\left(\mathbb{R}^2, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown\rangle\right)$ be an inner product space with $\left\langle \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \,\middle|\, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right\rangle \triangleq x_1x_2 + y_1y_2$. Let $\mathbf{S}$ be the *frame operator* (Definition H.15 page 244) 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, $\mathbf{S} = \frac{Nr^2}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{S}^{-1} = \frac{2}{Nr^2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and $\boldsymbol{v} = \dfrac{2}{Nr^2}\displaystyle\sum_{n=0}^{N-1} \langle \boldsymbol{v} \mid \boldsymbol{x}_n\rangle\, \boldsymbol{x}_n \quad \forall_{\boldsymbol{v}\in\mathbb{R}^2}$.

✎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 $\boldsymbol{v} \triangleq (x, y) \in \mathbb{R}^2$.

$$\begin{aligned}\sum_{n=0}^{N-1} \left|\langle \boldsymbol{v} \mid \boldsymbol{x}_n\rangle\right|^2 &\triangleq \sum_{n=0}^{N-1} \left|\boldsymbol{v}^{\mathbf{H}} r\begin{bmatrix} \cos\left(\theta + \frac{2n\pi}{N}\right) \\ \sin\left(\theta + \frac{2n\pi}{N}\right) \end{bmatrix}\right|^2 && \text{by definitions of } \boldsymbol{v} \text{ of } \langle \boldsymbol{y} \mid \boldsymbol{x}\rangle\\ &\triangleq \sum_{n=0}^{N-1} r^2\left|x\cos\left(\theta + \frac{2n\pi}{N}\right) + y\sin\left(\theta + \frac{2n\pi}{N}\right)\right|^2 && \text{by definition of } \boldsymbol{y}^{\mathbf{H}}\boldsymbol{x} \text{ operation}\\ &= r^2x^2 \sum_{n=0}^{N-1} \cos^2\left(\theta + \frac{2n\pi}{N}\right) + r^2y^2 \sum_{n=0}^{N-1} \sin^2\left(\theta + \frac{2n\pi}{N}\right) + r^2xy \sum_{n=0}^{N-1} \cos\left(\theta + \frac{2n\pi}{N}\right)\sin\left(\theta + \frac{2n\pi}{N}\right)\\ &= r^2x^2\frac{N}{2} + r^2y^2\frac{N}{2} + r^2xy0 && \text{by Corollary ?? page ??}\\ &= (x^2 + y^2)\frac{Nr^2}{2} = \left(\frac{Nr^2}{2}\right)\boldsymbol{v}^{\mathbf{H}}\boldsymbol{v} \triangleq \underbrace{\left(\frac{Nr^2}{2}\right)}_{A} \|\boldsymbol{v}\|^2 && \text{by definition of } \|\boldsymbol{v}\|\end{aligned}$$

2. Proof that $\mathbf{S} = \frac{Nr^2}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$:

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

(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* $\mathbf{S} = \mathbf{Q}\Lambda\mathbf{Q}^{-1}$:

$$\mathbf{S}\boldsymbol{e}_1 = \frac{Nr^2}{2}\boldsymbol{e}_1 \qquad \text{by (2c) lemma}$$

$$\implies \boldsymbol{e}_1 \text{ is an } \textit{eigenvector} \text{ of } \mathbf{S} \text{ with } \textit{eigenvalue } \frac{Nr^2}{2}$$

$$\mathbf{S}\boldsymbol{e}_2 = \frac{Nr^2}{2}\boldsymbol{e}_2 \qquad \text{by (2c) lemma}$$

$$\implies \boldsymbol{e}_2 \text{ is an } \textit{eigenvector} \text{ of } \mathbf{S} \text{ with } \textit{eigenvalue } \frac{Nr^2}{2}$$

$$\mathbf{S} = \overbrace{\underbrace{\begin{bmatrix} | & | \\ \boldsymbol{e}_1 & \boldsymbol{e}_2 \\ | & | \end{bmatrix}}_{\mathbf{Q}} \underbrace{\begin{bmatrix} \frac{Nr^2}{2} & 0 \\ 0 & \frac{Nr^2}{2} \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} | & | \\ \boldsymbol{e}_1 & \boldsymbol{e}_2 \\ | & | \end{bmatrix}^{-1}}_{\mathbf{Q}^{-1}}}^{\textit{Eigendecomposition} \text{ of } \mathbf{S}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{Nr^2}{2} & 0 \\ 0 & \frac{Nr^2}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \frac{Nr^2}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

3. Proof that $\mathbf{S}^{-1} = \frac{2}{Nr^2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$:

$$\mathbf{S}\mathbf{S}^{-1} = \frac{Nr^2}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\frac{2}{Nr^2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \triangleq \mathbf{I}_2 \qquad \text{by item (2)}$$

$$\mathbf{S}^{-1}\mathbf{S} = \frac{2}{Nr^2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\frac{Nr^2}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \triangleq \mathbf{I}_2 \qquad \text{by item (2)}$$

4. Proof that $\boldsymbol{v} = \frac{2}{Nr^2}\sum_{n=0}^{N-1}\left(\boldsymbol{v}^{\mathbf{H}}\boldsymbol{x}_n\right)\boldsymbol{x}_n$:

$$\boldsymbol{v} = \sum_{n=0}^{N-1} \left\langle \boldsymbol{v} \,\middle|\, \mathbf{S}^{-1}\boldsymbol{x}_n \right\rangle \boldsymbol{x}_n = \sum_{n=0}^{N-1} \left\langle \boldsymbol{v} \,\middle|\, \frac{2}{Nr^2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\boldsymbol{x}_n \right\rangle \boldsymbol{x}_n \qquad \text{by item (3)}$$

$$= \frac{2}{Nr^2}\sum_{n=0}^{N-1} \langle \boldsymbol{v} \,|\, \boldsymbol{x}_n \rangle \boldsymbol{x}_n = \frac{2}{Nr^2}\sum_{n=0}^{N-1} \left(\boldsymbol{v}^{\mathbf{H}}\boldsymbol{x}_n\right)\boldsymbol{x}_n \qquad \text{by definition of } \langle \boldsymbol{y} \,|\, \boldsymbol{x} \rangle$$

Copyright © 2019 Daniel J. Greenhoe

✏

Example H.2 (Peace Frame/Mercedes Frame). [40] Let $\left(\mathbb{R}^2, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ 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 H.15 page 244) 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, $\mathbf{S} = \frac{3}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\quad \mathbf{S}^{-1} = \frac{2}{3}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,

and $\quad \boldsymbol{v} = \frac{2}{3}\sum_{n=1}^{3} \langle \boldsymbol{v} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n \triangleq \frac{2}{3}\sum_{n=1}^{3} \left(\boldsymbol{v}^{\mathbf{H}} \boldsymbol{x}_n\right) \boldsymbol{x}_n \qquad \forall \boldsymbol{v} \in \mathbb{R}^2.$

✎PROOF:

1. This frame is simply a special case of the frame presented in Example H.1 (page 245) with $r = 1$, $N = 3$, and $\theta = \pi/2$.

2. Let's give it a try! Let $\boldsymbol{v} \triangleq \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$$\begin{aligned}
\sum_{n=1}^{3} \left\langle \boldsymbol{v} \mid \mathbf{S}^{-1} \boldsymbol{x}_n \right\rangle \boldsymbol{x}_n &= \frac{2}{3}\sum_{n=1}^{3} \left(\boldsymbol{v}^{\mathbf{H}} \boldsymbol{x}_n\right) \boldsymbol{x}_n && \text{by Example H.1 page 245} \\
&= \left(\boldsymbol{v}^{\mathbf{H}} \boldsymbol{x}_1\right)\boldsymbol{x}_1 + \left(\boldsymbol{v}^{\mathbf{H}} \boldsymbol{x}_2\right)\boldsymbol{x}_2 + \left(\boldsymbol{v}^{\mathbf{H}} \boldsymbol{x}_3\right)\boldsymbol{x}_3 \\
&= \frac{2}{3}\left(\left(\boldsymbol{v}^{\mathbf{H}} \frac{1}{2}\begin{bmatrix} 0 \\ 2 \end{bmatrix}\right)\boldsymbol{x}_1 + \left(\boldsymbol{v}^{\mathbf{H}} \frac{1}{2}\begin{bmatrix} -\sqrt{3} \\ -1 \end{bmatrix}\right)\boldsymbol{x}_2 + \left(\boldsymbol{v}^{\mathbf{H}} \frac{1}{2}\begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}\right)\boldsymbol{x}_3\right) \\
&= \frac{2}{3}\cdot\frac{1}{2}\left(\left(\boldsymbol{v}^{\mathbf{H}} \begin{bmatrix} 0 \\ 2 \end{bmatrix}\right)\boldsymbol{x}_1 + \left(\boldsymbol{v}^{\mathbf{H}} \begin{bmatrix} -\sqrt{3} \\ -1 \end{bmatrix}\right)\boldsymbol{x}_2 + \left(\boldsymbol{v}^{\mathbf{H}} \begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}\right)\boldsymbol{x}_3\right) \\
&= \frac{1}{3}\left((2)\frac{1}{2}\begin{bmatrix} 0 \\ 2 \end{bmatrix} + \left(-\sqrt{3}-1\right)\frac{1}{2}\begin{bmatrix} -\sqrt{3} \\ -1 \end{bmatrix} + \left(\sqrt{3}-1\right)\frac{1}{2}\begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}\right) \\
&= \frac{1}{6}\begin{bmatrix} 2(0) & + & (-\sqrt{3}-1)(-\sqrt{3}) & + & (\sqrt{3}-1)(\sqrt{3}) \\ 2(2) & + & (-\sqrt{3}-1)(-1) & + & (\sqrt{3}-1)(-1) \end{bmatrix} \\
&= \frac{1}{6}\begin{bmatrix} 0 & + & (3+\sqrt{3}) & + & (3-\sqrt{3}) \\ 4 & + & (1+\sqrt{3}) & + & (1-\sqrt{3}) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \triangleq \boldsymbol{v}
\end{aligned}$$

✏

In Example H.1 (page 245) and Example H.2 (page 247), 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 H.3 (next).

Example H.3. [41] Let $\left(\mathbb{R}^2, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ 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 H.15 page 244) with *inverse* $\mathbf{S}^{-1}$.

[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⟩

[41] Christensen (2003) pages 7–8 ⟨?⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

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,

$$\mathbf{S}^{-1}\boldsymbol{x}_1 = \tfrac{1}{3}\boldsymbol{p} \quad \text{and}$$
$$\mathbf{S}^{-1}\boldsymbol{x}_2 = \tfrac{1}{3}\boldsymbol{p} + \tfrac{1}{2}\boldsymbol{q} \quad \text{and}$$
$$\mathbf{S}^{-1}\boldsymbol{x}_3 = \tfrac{1}{3}\boldsymbol{p} - \tfrac{1}{2}\boldsymbol{q} \quad .$$

✎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} \,|\, \boldsymbol{x}_n\rangle\right|^2 &\triangleq |\langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle|^2 + |\langle \boldsymbol{v} \,|\, \boldsymbol{p}+\boldsymbol{q}\rangle|^2 + |\langle \boldsymbol{v} \,|\, \boldsymbol{p}-\boldsymbol{q}\rangle|^2 && \text{by definitions of } \boldsymbol{x}_1, \boldsymbol{x}_2, \text{ and } \boldsymbol{x}_3 \\
&= |\langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle|^2 + |\langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle + \langle \boldsymbol{v} \,|\, \boldsymbol{q}\rangle|^2 + |\langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle - \langle \boldsymbol{v} \,|\, \boldsymbol{q}\rangle|^2 && \text{by } \textit{additivity} \text{ of } \langle \triangle \,|\, \triangledown \rangle \text{ (Definition K.1 page 271)} \\
&= |\langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle|^2 + \left(|\langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle|^2 + |\langle \boldsymbol{v} \,|\, \boldsymbol{q}\rangle|^2 + \langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle \langle \boldsymbol{v} \,|\, \boldsymbol{q}\rangle^* + \langle \boldsymbol{v} \,|\, \boldsymbol{q}\rangle \langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle^*\right) \\
&\quad + \left(|\langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle|^2 + |\langle \boldsymbol{v} \,|\, \boldsymbol{q}\rangle|^2 - \langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle \langle \boldsymbol{v} \,|\, \boldsymbol{q}\rangle^* - \langle \boldsymbol{v} \,|\, \boldsymbol{q}\rangle \langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle^*\right) \\
&= 3|\langle \boldsymbol{v} \,|\, \boldsymbol{p}\rangle|^2 + 2|\langle \boldsymbol{v} \,|\, \boldsymbol{q}\rangle|^2 \\
&\boxed{\le} 3\,\|\boldsymbol{v}\|\,\|\boldsymbol{p}\| + 2\,\|\boldsymbol{v}\|\,\|\boldsymbol{q}\| && \text{by } \textit{CS Inequality} \text{ (Theorem K.2 page 272)} \\
&= \|\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:

$$\begin{aligned}
\mathbf{S}\boldsymbol{p} &\triangleq \sum_{n=1}^{3} \langle \boldsymbol{p} \,|\, \boldsymbol{x}_n\rangle\, \boldsymbol{x}_n \\
&= \langle \boldsymbol{p} \,|\, \boldsymbol{p}\rangle\, \boldsymbol{p} + \langle \boldsymbol{p} \,|\, \boldsymbol{p}+\boldsymbol{q}\rangle\,(\boldsymbol{p}+\boldsymbol{q}) + \langle \boldsymbol{p} \,|\, \boldsymbol{p}-\boldsymbol{q}\rangle\,(\boldsymbol{p}-\boldsymbol{q}) \\
&= (1)\boldsymbol{p} + (1+0)(\boldsymbol{p}+\boldsymbol{q}) + (1-0)(\boldsymbol{p}-\boldsymbol{q}) \\
&= 3\boldsymbol{p} \\
\implies \mathbf{S}^{-1}\boldsymbol{p} &= \frac{1}{3}\boldsymbol{p} \\
\mathbf{S}\boldsymbol{q} &\triangleq \sum_{n=1}^{3} \langle \boldsymbol{q} \,|\, \boldsymbol{x}_n\rangle\, \boldsymbol{x}_n \\
&= \langle \boldsymbol{q} \,|\, \boldsymbol{p}\rangle\, \boldsymbol{p} + \langle \boldsymbol{q} \,|\, \boldsymbol{p}+\boldsymbol{q}\rangle\,(\boldsymbol{p}+\boldsymbol{q}) + \langle \boldsymbol{q} \,|\, \boldsymbol{p}-\boldsymbol{q}\rangle\,(\boldsymbol{p}-\boldsymbol{q}) \\
&= (0)\boldsymbol{q} + (0+1)(\boldsymbol{p}+\boldsymbol{q}) + (0-1)(\boldsymbol{p}-\boldsymbol{q}) \\
&= 2\boldsymbol{q} \\
\implies \mathbf{S}^{-1}\boldsymbol{q} &= \frac{1}{2}\boldsymbol{q}
\end{aligned}$$

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 $\mathbf{S}^{-1}\boldsymbol{x}_3 = \frac{1}{3}\boldsymbol{p} - \frac{1}{2}\boldsymbol{q}$:

$$\begin{aligned}
\mathbf{S}^{-1}\boldsymbol{x}_1 &\triangleq \mathbf{S}^{-1}\boldsymbol{p} && \text{by definition of } \boldsymbol{x}_1 \\
&= \frac{1}{3}\boldsymbol{p} && \text{by (2) lemma} \\
\mathbf{S}^{-1}\boldsymbol{x}_2 &\triangleq \mathbf{S}^{-1}(\boldsymbol{p}+\boldsymbol{q}) && \text{by definition of } \boldsymbol{x}_2 \\
&= \frac{1}{3}\boldsymbol{p} + \frac{1}{2}\boldsymbol{q} && \text{by (2) lemma}
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$\begin{aligned}
\mathbf{S}^{-1}\boldsymbol{x}_3 &\triangleq \mathbf{S}^{-1}(\boldsymbol{p}-\boldsymbol{q}) && \text{by definition of } \boldsymbol{x}_2\\
&= \frac{1}{3}\boldsymbol{p} - \frac{1}{2}\boldsymbol{q} && \text{by (2) lemma}
\end{aligned}$$

5. Check that $\boldsymbol{v} = \sum_n \langle \boldsymbol{v} \mid \boldsymbol{x}_n \rangle \boldsymbol{x}_n = \langle \boldsymbol{v} \mid \boldsymbol{p} \rangle \boldsymbol{p} + \langle \boldsymbol{v} \mid \boldsymbol{q} \rangle \boldsymbol{q}$:

$$\begin{aligned}
\boldsymbol{v} &= \sum_{n=1}^{3} \left\langle \boldsymbol{v} \mid \mathbf{S}^{-1}\boldsymbol{x}_n \right\rangle \boldsymbol{x}_n\\
&= \left\langle \boldsymbol{v} \mid \mathbf{S}^{-1}\boldsymbol{p} \right\rangle \boldsymbol{p} + \left\langle \boldsymbol{v} \mid \mathbf{S}^{-1}(\boldsymbol{p}+\boldsymbol{q}) \right\rangle (\boldsymbol{p}+\boldsymbol{q}) + \left\langle \boldsymbol{v} \mid \mathbf{S}^{-1}(\boldsymbol{p}-\boldsymbol{q}) \right\rangle (\boldsymbol{p}-\boldsymbol{q})\\
&= \left\langle \boldsymbol{v} \mid \frac{1}{3}\boldsymbol{p} \right\rangle \boldsymbol{p} + \left\langle \boldsymbol{v} \mid \frac{1}{3}\boldsymbol{p} + \frac{1}{2}\boldsymbol{q} \right\rangle (\boldsymbol{p}+\boldsymbol{q}) + \left\langle \boldsymbol{v} \mid \frac{1}{3}\boldsymbol{p} - \frac{1}{2}\boldsymbol{q}) \right\rangle (\boldsymbol{p}-\boldsymbol{q})\\
&= \left(\frac{1}{3}+\frac{1}{3}+\frac{1}{3}\right)\langle \boldsymbol{v} \mid \boldsymbol{p} \rangle \boldsymbol{p} + \left(\frac{1}{3}-\frac{1}{3}\right)\langle \boldsymbol{v} \mid \boldsymbol{p} \rangle \boldsymbol{q} + \left(\frac{1}{2}-\frac{1}{2}\right)\langle \boldsymbol{v} \mid \boldsymbol{q} \rangle \boldsymbol{p} + \left(\frac{1}{2}+\frac{1}{2}\right)\langle \boldsymbol{v} \mid \boldsymbol{q} \rangle \boldsymbol{q}\\
&= \langle \boldsymbol{v} \mid \boldsymbol{p} \rangle \boldsymbol{p} + \langle \boldsymbol{v} \mid \boldsymbol{q} \rangle \boldsymbol{q}
\end{aligned}$$

✍

Copyright © 2019 Daniel J. Greenhoe

APPENDIX I

LINEAR SUBSPACES

I.1 Subspaces of a linear space

Linear spaces (Definition G.1 page 209) can be decomposed into a collection of *linear subspaces* (Definition I.1 page 252). Often such a collection along with an *order relation* forms a *lattice*.

Figure I.1: lattice of subspaces of $\mathbb{R}^3$ (Example I.1 page 251)

EX

Example I.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 I.1 (page 251).

Definition I.1. [1] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a* LINEAR SPACE *(Definition G.1 page 209).*

DEF

A tupple $\big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *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 G.1 page 209) $\boldsymbol{X}$ has at least two *linear subspaces*—itself and $\mathbf{0}$ (Proposition I.1 page 252), called the *trivial linear space*. The *linear span* (Definition H.2 page 221) of every subset of a linear linear space is a subspace (Proposition I.2 page 253). Every *linear subspace* contains the "zero" vector $\mathbb{0}$, and is *convex* (Definition M.6 page 298, Proposition I.3 page 253).

Proposition I.1. [2] *Let* $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *and* $\mathbf{0} \triangleq \big(\{\mathbb{0}\}, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$.

PRP

$$\left\{\begin{array}{l} \boldsymbol{X} \text{ is a LINEAR SPACE} \\ \text{(Definition G.1 page 209)} \end{array}\right\} \implies \left\{\begin{array}{ll} 1. & \mathbf{0} \text{ is a LINEAR SUBSPACE of } \boldsymbol{X} \quad \text{and} \\ 2. & \boldsymbol{X} \text{ is a LINEAR SUBSPACE of } \boldsymbol{X} \end{array}\right\}$$

✎PROOF: For a structure to be a linear subspace of $\boldsymbol{X}$, it must satisfy the requirements of Definition I.1 (page 252).

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

COPYRIGHT © 2019 DANIEL J. GREENHOE

✏

Proposition I.2. [3] *Let* $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a* LINEAR SPACE *(Definition G.1 page 209). Let* span *be the* LINEAR SPAN *of a set* Y *in* $\boldsymbol{X}$.

PRP
$$\left\{\begin{array}{l} Y \text{ is a SUBSET of the set } X \\ (Y \subseteq X) \end{array}\right\} \Longrightarrow \left\{ \operatorname{span} Y \quad \text{is a LINEAR SUBSPACE of } \boldsymbol{X}. \right\}$$

Proposition I.3. [4] *Let* $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a* LINEAR SPACE *and* $\mathbb{0}$ *the zero vector of* $\boldsymbol{X}$.

PRP
$$\left\{ \boldsymbol{Y} \text{ is a LINEAR SUBSPACE of } \boldsymbol{X} \right\} \Longrightarrow \left\{\begin{array}{ll} 1. & \mathbb{0} \in \boldsymbol{Y} \qquad \text{and} \\ 2. & \boldsymbol{Y} \text{ is CONVEX in } \boldsymbol{X} \end{array}\right\}$$

✎PROOF:

$$\begin{aligned} \boldsymbol{Y} \text{ is a } \textit{subspace} &\Longrightarrow \exists(\alpha \boldsymbol{y}) \in \boldsymbol{Y} \quad \forall \alpha \in \mathbb{F} && \text{by Definition I.1 page 252} \\ &\Longrightarrow \exists \mathbb{0} \in \boldsymbol{Y} && \text{because } \alpha = 0 \in \mathbb{F} \end{aligned}$$

$$\begin{aligned} \boldsymbol{Y} \text{ is a linear subspace} &\Longrightarrow \boldsymbol{x} + \boldsymbol{y} \in Y \; \forall \boldsymbol{x}, \boldsymbol{y} \in Y \\ &\Longrightarrow \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in Y \; \forall \boldsymbol{x}, \boldsymbol{y} \in Y \\ &\Longrightarrow \boldsymbol{Y} \text{ is } \textit{convex} \end{aligned}$$

✏

Definition I.2. [5] *Let* $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *and* $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be* LINEAR SUBSPACES *(Definition I.1 page 252) of a* LINEAR SPACE *(Definition G.1 page 209)* $\boldsymbol{\Omega} \triangleq \big(\Omega, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$.

DEF
$$\begin{array}{lcll} \boldsymbol{X} \hat{+} \boldsymbol{Y} & \triangleq & \big(\{\boldsymbol{x} + \boldsymbol{y} | \boldsymbol{x} \in X \text{ and } \boldsymbol{y} \in Y\}, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big) & \textbf{\textit{(Minkowski addition)}} \\ \boldsymbol{X} \uplus \boldsymbol{Y} & \triangleq & \big(X \cup Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big) & \textbf{\textit{(subspace union)}} \\ \boldsymbol{X} \Cap \boldsymbol{Y} & \triangleq & \big(X \cap Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big) & \textbf{\textit{(subspace intersection)}} \end{array}$$

Example I.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

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Remark I.1.

Notice the similarities between the properties of linear subspaces in a linear space (Proposition I.4 page 254) and the properties of closed sets in a topological space:

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 ($\cup\!\!\!\!\!\!\cup$) 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 ($\cap\!\!\!\!\!\!\cap$). Proposition I.4 (next) shows four useful objects are always subspaces. Some of these in Euclidean space $\mathbb{R}^3$ are illustrated to the right.

Proposition I.4. [7] *Let* $\boldsymbol{X}$ *be a* LINEAR SPACE *(Definition G.1 page 209)*.

PRP

$$\left.\begin{array}{l}\left\{\boldsymbol{X}_n \middle|_{n=1,2,\ldots,N}\right\} \textit{ are} \\ \text{LINEAR SUBSPACES } \textit{of } \boldsymbol{X}\end{array}\right\} \Longrightarrow \left\{\begin{array}{lll} 1. & \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \hat{+} \cdots \hat{+} \boldsymbol{X}_N & \textit{is a } \text{LINEAR SUBSPACE } \textit{of } \boldsymbol{X} \\ & \textit{and} & \\ 2. & \boldsymbol{X}_1 \cap\!\!\!\!\!\!\cap \boldsymbol{X}_2 \cap\!\!\!\!\!\!\cap \cdots \cap\!\!\!\!\!\!\cap \boldsymbol{X}_N & \textit{is a } \text{LINEAR SUBSPACE } \textit{of } \boldsymbol{X}\end{array}\right.$$

✎PROOF: For a structure to be a linear subspace of $\boldsymbol{X}$, it must satisfy the requirements of Definition I.1 (page 252).

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

COPYRIGHT © 2019 DANIEL J. GREENHOE

i. proof that $\boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \neq \varnothing$:

$$\begin{aligned}
\boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 &= \left\{\boldsymbol{v}+\boldsymbol{w} | \boldsymbol{v} \in \boldsymbol{X}_1 \text{ and } \boldsymbol{w} \in Y\right\} && \text{by Definition I.2 page 253}\\
&\supseteq \left\{\boldsymbol{v}+\boldsymbol{w} | \boldsymbol{v} \in \{\mathbb{0}\} \subseteq \boldsymbol{X}_1 \text{ and } \boldsymbol{w} \in \{\mathbb{0}\} \subseteq \boldsymbol{X}_2\right\}\\
&= \{\mathbb{0}+\mathbb{0}\}\\
&= \{\mathbb{0}\}\\
&\neq \varnothing
\end{aligned}$$

ii. proof that $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \quad \Longrightarrow \quad \boldsymbol{x}+\boldsymbol{y} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2$:

$$\begin{aligned}
\boldsymbol{x}+\boldsymbol{y} &= (\boldsymbol{v}_1+\boldsymbol{w}_1)+(\boldsymbol{v}_2+\boldsymbol{w}_2) && \text{by } \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \text{ hypothesis}\\
&= \underbrace{(\boldsymbol{v}_1+\boldsymbol{v}_2)}_{\text{in } \boldsymbol{X}_1} + \underbrace{(\boldsymbol{w}_1+\boldsymbol{w}_2)}_{\text{in } \boldsymbol{X}_2 \text{ because } \boldsymbol{X}_2 \text{ is a linear subspace}} && \text{by Definition G.1 page 209}\\
&\in \left\{\boldsymbol{v}+\boldsymbol{w} | \boldsymbol{v} \in \boldsymbol{X}_1 \text{ and } \boldsymbol{w} \in Y\right\}\\
&= \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 && \text{by Definition I.2 page 253}
\end{aligned}$$

iii. proof that $\boldsymbol{v} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2, \alpha \in F \quad \Longrightarrow \quad \alpha\boldsymbol{v} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2$:

$$\begin{aligned}
\alpha\boldsymbol{x} &= \alpha(\boldsymbol{v}_1+\boldsymbol{w}_1) && \text{by } \boldsymbol{x} \in \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \text{ hypothesis}\\
&= \underbrace{\alpha\boldsymbol{v}_1}_{\text{in } \boldsymbol{X}_1} + \underbrace{\alpha\boldsymbol{w}_1}_{\text{in } \boldsymbol{X}_2 \text{ because } \boldsymbol{X}_2 \text{ is a linear subspace}} && \text{by Definition G.1 page 209}\\
&\in \left\{\boldsymbol{v}+\boldsymbol{w} | \boldsymbol{v} \in \boldsymbol{X}_1 \text{ and } \boldsymbol{w} \in Y\right\}\\
&= \boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 && \text{by Definition I.2 page 253}
\end{aligned}$$

(c) Proof that [N case] $\Longrightarrow$ [$N+1$ case]:

$$\boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \hat{+} \cdots \hat{+} \boldsymbol{X}_{N+1} = \underbrace{\left(\boldsymbol{X}_1 \hat{+} \boldsymbol{X}_2 \hat{+} \cdots \hat{+} \boldsymbol{X}_N\right)}_{\text{linear subspace by } N \text{ case hypothsis}} \hat{+} \boldsymbol{X}_{N+1}$$

$\Longrightarrow$ linear subspace by $N=2$ case (item (1b) page 254)

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 $\boldsymbol{X} \Cap \boldsymbol{Y} \neq \varnothing$:

$$\begin{aligned}
\boldsymbol{X} \Cap \boldsymbol{Y} &= \left\{\boldsymbol{x} \in X | \boldsymbol{x} \in X \text{ and } \boldsymbol{w} \in Y\right\}\\
&\supseteq \left\{\boldsymbol{x} \in X | \boldsymbol{x} \in \{\mathbb{0}\} \subseteq \boldsymbol{X} \text{ and } \boldsymbol{x} \in \{\mathbb{0}\} \subseteq \boldsymbol{Y}\right\}\\
&= \{\mathbb{0}+\mathbb{0}\}\\
&= \{\mathbb{0}\}\\
&\neq \varnothing
\end{aligned}$$

ii. proof that $\boldsymbol{x}, \boldsymbol{y} \in X \Cap \boldsymbol{Y} \quad \Longrightarrow \quad \boldsymbol{x}+\boldsymbol{y} \in X \Cap \boldsymbol{Y}$:

$$\begin{aligned}
\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X} \Cap \boldsymbol{Y} &\Longrightarrow \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X} \text{ and } \boldsymbol{x}, \boldsymbol{y} \in Y\\
&\Longrightarrow \boldsymbol{x}+\boldsymbol{y} \in \boldsymbol{X} \text{ and } \boldsymbol{x}+\boldsymbol{y} \in Y \qquad \text{because } \boldsymbol{X} \text{ and } \boldsymbol{Y} \text{ are linear subspaces}\\
&\Longrightarrow \boldsymbol{x}+\boldsymbol{y} \in \boldsymbol{X} \Cap \boldsymbol{Y}
\end{aligned}$$

iii. proof that $\boldsymbol{v} \in X \Cap \boldsymbol{Y}, \alpha \in F \quad \Longrightarrow \quad \alpha\boldsymbol{v} \in X \Cap \boldsymbol{Y}$:

$$\begin{aligned}
\boldsymbol{x} \in \boldsymbol{X} \Cap \boldsymbol{Y} &\Longrightarrow \boldsymbol{x} \in \boldsymbol{X} \text{ and } \boldsymbol{x} \in Y\\
&\Longrightarrow \alpha\boldsymbol{x} \in \boldsymbol{X} \text{ and } \alpha\boldsymbol{x} \in Y \qquad \text{because } \boldsymbol{X} \text{ and } \boldsymbol{Y} \text{ are linear subspaces}\\
&\Longrightarrow \alpha\boldsymbol{x} \in \boldsymbol{X} \Cap \boldsymbol{Y}
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

(c) Proof that [N case] $\implies$ [$N+1$ case]:

$$\boldsymbol{X}_1 \cap \boldsymbol{X}_2 \cap \cdots \cap \boldsymbol{X}_{N+1} = \underbrace{\left(\boldsymbol{X}_1 \cap \boldsymbol{X}_2 \cap \cdots \cap \boldsymbol{X}_N\right)}_{\text{linear subspace by } N \text{ case hypothsis}} \cap \boldsymbol{X}_{N+1}$$

$\implies$ linear subspace by $N = 2$ case (item (2b) page 255)

✏

Every linear subspace contains the zero vector $\mathbb{0}$ (Proposition I.3 page 253). 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 I.1. [8] *Let* $\boldsymbol{X} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *and* $\boldsymbol{Y} \triangleq \left(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *be* LINEAR SUBSPACES *(Definition I.1 page 252) of a* LINEAR SPACE *(Definition G.1 page 209)* $\boldsymbol{\Omega} \triangleq \left(\Omega, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$.

THM

$$X \cap Y = \{\mathbb{0}\} \iff \left\{\begin{array}{l} \textit{for every } \boldsymbol{u} \in \boldsymbol{X} \hat{+} \boldsymbol{Y} \textit{ there exist } \boldsymbol{x} \in X \textit{ and } \boldsymbol{y} \in Y \textit{ such that} \\ \quad 1.\ \ \boldsymbol{u} = \boldsymbol{x} + \boldsymbol{y} \qquad \textit{and} \\ \quad 2.\ \ \boldsymbol{x} \textit{ and } \boldsymbol{y} \textit{ are } \text{UNIQUE}. \end{array}\right\}$$

✎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 I.2. [9] *Let* $\boldsymbol{\Omega}$ *be a linear subspace and* 2^{Ω} *the set of closed linear subspaces of* $\boldsymbol{\Omega}$.

THM

$\left(2^{\Omega}, \hat{+}, \cap, \mathbf{0}, \boldsymbol{\Omega}; \subseteq\right)$ *is a* LATTICE. *In particular*

$$\begin{array}{rclrcll} \boldsymbol{X} \hat{+} \boldsymbol{X} &=& \boldsymbol{X} & \boldsymbol{X} \cap \boldsymbol{X} &=& \boldsymbol{X} & \forall \boldsymbol{X} \in 2^{\Omega} \\ \boldsymbol{X} \hat{+} \boldsymbol{Y} &=& \boldsymbol{Y} \hat{+} \boldsymbol{X} & \boldsymbol{X} \cap \boldsymbol{Y} &=& \boldsymbol{Y} \cap \boldsymbol{X} & \forall \boldsymbol{X}, \boldsymbol{Y} \in 2^{\Omega} \\ (\boldsymbol{X} \hat{+} \boldsymbol{Y}) \hat{+} \boldsymbol{Z} &=& \boldsymbol{X} \hat{+} (\boldsymbol{Y} \hat{+} \boldsymbol{Z}) & (\boldsymbol{X} \cap \boldsymbol{Y}) \cap \boldsymbol{Z} &=& \boldsymbol{X} \cap (\boldsymbol{Y} \cap \boldsymbol{Z}) & \forall \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \in 2^{\Omega} \\ \boldsymbol{X} \hat{+} (\boldsymbol{X} \cap \boldsymbol{Y}) &=& \boldsymbol{X} & \boldsymbol{X} \cap (\boldsymbol{X} \hat{+} \boldsymbol{Y}) &=& \boldsymbol{X} & \forall \boldsymbol{X}, \boldsymbol{Y} \in 2^{\Omega} \end{array}$$

✎PROOF: These results follow directly from the properties of lattices. ✏

[8] Michel and Herget (1993) page 83 ⟨Theorem 3.2.12⟩, Kubrusly (2001) page 67 ⟨Theorem 2.14⟩
[9] Iturrioz (1985) pages 56–57

COPYRIGHT © 2019 DANIEL J. GREENHOE

I.2 Subspaces of an inner product space

Definition I.3. [10] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \triangledown \rangle\big)$ *be an* INNER PRODUCT SPACE *(Definition K.1 page 271)*.

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* $\big(A^{\perp}\big)^{\perp}$.

Proposition I.5. [11] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \triangledown \rangle\big)$ *be an* INNER PRODUCT SPACE *(Definition K.1 page 271)*.

PRP

$$A \subseteq B \quad \Longrightarrow \quad B^{\perp} \subseteq A^{\perp} \quad \forall A,B \in 2^X \quad \text{(ANTITONE)}$$

✎PROOF:

$$\begin{aligned}
B^{\perp} &\triangleq \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in B\} && \text{by definition of } B^{\perp} \text{ (Definition I.3 page 257)} \\
&\subseteq \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in A\} && \text{by } A \subseteq B \text{ hypothesis} \\
&= A^{\perp} && \text{by definition of } A^{\perp} \text{ (Definition I.3 page 257)}
\end{aligned}$$

✏

Every *linear space* $\boldsymbol{X}$ contains $\mathbf{0}$ and $\boldsymbol{X}$ as *linear subspaces* (Proposition I.1 page 252). If $\boldsymbol{X}$ is also an *inner product space*, then $\mathbf{0}$ and $\boldsymbol{X}$ are *orthogonal complements* of each other (next proposition).

Proposition I.6. [12] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \triangledown \rangle\big)$ *be an* INNER PRODUCT SPACE *(Definition K.1 page 271) and* $\mathbb{0}$ *the* VECTOR ADDITIVE IDENTITY ELEMENT *(Definition G.1 page 209) in* $\boldsymbol{\Omega}$.

PRP

1. $\{\mathbb{0}\}^{\perp} = X$
2. $X^{\perp} = \{\mathbb{0}\}$

✎PROOF:

$$\begin{aligned}
\{\mathbb{0}\}^{\perp} &= \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in \{\mathbb{0}\}\} && \text{by definition of } \perp \text{ (Definition I.3 page 257)} \\
&= \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \mathbb{0} \rangle = 0\} \\
&= X \\
X^{\perp} &= \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y} \rangle = 0 \quad \forall \boldsymbol{y} \in X\} && \text{by definition of } \perp \text{ Definition I.3 page 257} \\
&= \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{x} \rangle = 0\} \\
&= \{\mathbb{0}\}
\end{aligned}$$

✏

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

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Theorem I.3. [13] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be an* INNER PRODUCT SPACE *(Definition K.1 page 271). Let* $\operatorname{span}A$ *be the span of a set* A *(Definition H.2 page 221).*

THM

$$\left\{\begin{array}{l} A \text{ is a } \textbf{subset} \text{ of } X \\ (A \subseteq X) \end{array}\right\} \implies \left\{\begin{array}{lll} 1. & A \cap A^{\perp} = \begin{cases} \{\mathbb{0}\} & \text{if } \mathbb{0} \in A \\ \varnothing & \text{if } \mathbb{0} \notin A \end{cases} & \text{and} \\ 2. & A \subseteq A^{\perp\perp} = \operatorname{span}A & \text{and} \\ 3. & A^{\perp} = A^{\perp\perp\perp} = A^{\perp-} = A^{-\perp} = (\operatorname{span}A)^{\perp} & \text{and} \\ 4. & A^{\perp} \text{ is a } \textbf{subspace} \text{ of } \boldsymbol{\Omega} & \end{array}\right\}$$

✎PROOF:

1. Proof that $A \cap A^{\perp} = \cdots$:

$$\begin{aligned} A \cap A^{\perp} &= \{\boldsymbol{x} \in X \,|\, \boldsymbol{x} \in A\} \cap \{\boldsymbol{x} \in X \,|\, \langle \boldsymbol{x} \,|\, \boldsymbol{y}\rangle \quad \forall \boldsymbol{y} \in A\} && \text{by definition of } A^{\perp} \\ &= \big\{\boldsymbol{x} \in X \,|\, \boldsymbol{x} \in A \quad \text{and} \quad \langle \boldsymbol{x} \,|\, \boldsymbol{y}\rangle \quad \forall \boldsymbol{y} \in A\big\} \\ &= \begin{cases} \{\mathbb{0}\} & \text{if } \mathbb{0} \in A \\ \varnothing & \text{if } \mathbb{0} \notin A \end{cases} \end{aligned}$$

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 $A^{\perp} = \big\{\boldsymbol{x} \in X \,|\, \boldsymbol{x} \in \text{yz plane}\big\}$.

(c) Then $A^{\perp\perp} = \{\boldsymbol{x} \in X \,|\, \boldsymbol{x} \in \text{x axis}\}$.

(d) Therefore, $A \subsetneq A^{\perp\perp}$

3. Proof for $A^{\perp}$ equivalent expressions:

(a) Proof that $A^{\perp} = A^{\perp\perp\perp}$:

$$\begin{aligned} A^{\perp} &\subseteq \big(A^{\perp}\big)^{\perp\perp} && \text{by item (2)} \\ &= \big(A^{\perp\perp}\big)^{\perp} \\ &= A^{\perp\perp\perp} && \text{by Definition I.3 page 257} \\ A^{\perp\perp\perp} &= \big(A^{\perp\perp}\big)^{\perp} && \text{by Definition I.3 page 257} \\ &\subseteq A^{\perp} && \text{by item (2) and Proposition I.5 (page 257)} \end{aligned}$$

(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

COPYRIGHT © 2019 DANIEL J. GREENHOE

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 **??** page **??**), 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

$$A \subseteq A^{-} \implies A^{\perp^-} \subseteq A^{\perp} \qquad \text{by } \textit{antitone} \text{ property (Proposition I.5 page 257)}$$

vi. And so $A^{\perp} = A^{\perp^-}$.

4. Proof that $A^{\perp}$ is a **subspace** of $\boldsymbol{\Omega}$ (must satisfy the conditions of Definition I.1 page 252):

(a) Proof that $A^{\perp} \neq \varnothing$: $A^{\perp}$ has at least one element, the element $\mathbb{0}$…

$$\begin{aligned}\langle \mathbb{0} \mid \boldsymbol{y}\rangle = 0 \quad \forall \boldsymbol{y} \in A &\qquad \text{by definition of } \mathbb{0}\\ \implies \mathbb{0} \in A^{\perp} &\qquad \text{by definition of } A^{\perp} \text{ (Definition I.3 page 257)}\end{aligned}$$

(b) Proof that $A^{\perp} \subseteq X$:

$$\begin{aligned}\boldsymbol{u} \in A^{\perp} &\implies \boldsymbol{u} \in \{\boldsymbol{x} \in X \mid \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle = 0 \quad \forall \boldsymbol{y} \in A\} && \text{by definition of } A^{\perp} \text{ (Definition I.3 page 257)}\\ &\implies \boldsymbol{u} \in X && \text{by definition of sets}\end{aligned}$$

(c) Proof that $\boldsymbol{u}, \boldsymbol{v} \in A^{\perp} \implies (\boldsymbol{u} + \boldsymbol{v}) \in A^{\perp}$:

$$\begin{aligned}\boldsymbol{u}, \boldsymbol{v} \in A^{\perp} &\implies \langle \boldsymbol{u} \mid \boldsymbol{y}\rangle = \langle \boldsymbol{v} \mid \boldsymbol{y}\rangle = 0 \quad \forall \boldsymbol{y} \in A && \text{by definition of } A^{\perp} \text{ (Definition I.3 page 257)}\\ &\implies \langle \boldsymbol{u} \mid \boldsymbol{y}\rangle + \langle \boldsymbol{v} \mid \boldsymbol{y}\rangle = 0 \quad \forall \boldsymbol{y} \in A \\ &\implies \langle \boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{y}\rangle = 0 \quad \forall \boldsymbol{y} \in A && \text{by } \textit{additive} \text{ property of } \langle \triangle \mid \triangledown\rangle \text{ (Definition K.1 page 271)}\\ &\implies \boldsymbol{u} + \boldsymbol{v} \in A^{\perp} && \text{by definition of } A^{\perp} \text{ (Definition I.3 page 257)}\end{aligned}$$

(d) Proof that $\boldsymbol{v} \in \Omega \implies \alpha\boldsymbol{v} \in A^{\perp}$:

$$\begin{aligned}\boldsymbol{v} \in A^{\perp} &\implies \langle \boldsymbol{v} \mid \boldsymbol{y}\rangle = 0 \quad \forall \boldsymbol{y} \in A && \text{by definition of } A^{\perp} \text{ (Definition I.3 page 257)}\\ &\implies \alpha\langle \boldsymbol{v} \mid \boldsymbol{y}\rangle = \alpha \cdot 0 \quad \forall \boldsymbol{y} \in A \\ &\implies \langle \alpha\boldsymbol{v} \mid \boldsymbol{y}\rangle = 0 \quad \forall \boldsymbol{y} \in A && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \mid \triangledown\rangle \text{ (Definition K.1 page 271)}\\ &\implies \alpha\boldsymbol{v} \in A^{\perp} && \text{by definition of } A^{\perp} \text{ (Definition I.3 page 257)}\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe

Corollary I.1. *Let* $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *and* $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be* INNER PRODUCT SPACES. *Let* $\operatorname{span} Y$ *be the span of the set* Y *(Definition H.2 page 221).*

COR

$$\{\ \boldsymbol{Y} \text{ is a } \textbf{linear subspace} \text{ of } \boldsymbol{X}\ \} \quad\Longrightarrow\quad \left\{\begin{array}{lrcll} 1. & Y \cap Y^{\perp} & = & \{\mathbb{0}\} & \text{and} \\ 2. & Y & = & Y^{\perp\perp} = \operatorname{span} Y & \text{and} \\ 3. & Y^{\perp} & = & Y^{\perp\perp\perp} & \text{and} \\ 4. & Y^{\perp} \text{ is a } \textbf{subspace} \text{ of } \boldsymbol{X} & & & \end{array}\right\}$$

✎PROOF:

1. Proof that $\boldsymbol{Y} \cap Y^{\perp} = \{\mathbb{0}\}$: This follows from Theorem I.3 (page 258) and the fact that all subspaces contain the zero vector $\mathbb{0}$ (Proposition I.3 page 253).
2. Proof that $\boldsymbol{Y} = \boldsymbol{Y}^{\perp\perp} = \operatorname{span}\boldsymbol{Y}$: This follows directly from Theorem I.3 (page 258).
3. Proof that $\boldsymbol{Y}^{\perp} = \boldsymbol{Y}^{\perp\perp\perp}$: This follows directly from Theorem I.3 (page 258).
4. Proof that $Y^{\perp}$ is a **subspace** of $\boldsymbol{X}$: This follows directly from Theorem I.3 (page 258).

✏

Theorem I.4. [14] *Let* $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *and* $\boldsymbol{Z} \triangleq \big(Z, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be* LINEAR SUBSPACES *of an* INNER PRODUCT SPACE $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$.

THM

$$\boldsymbol{Y} \perp \boldsymbol{Z} \quad\Longrightarrow\quad Y \cap Z = \{\mathbb{0}\}$$

✎PROOF:

$$\begin{aligned} \boldsymbol{x} \in \boldsymbol{Y} \cap \boldsymbol{Z} &\Longrightarrow \boldsymbol{x} \in \boldsymbol{Y} \text{ and } \boldsymbol{x} \in \boldsymbol{Z} && \text{by definition of } \cap \\ &\Longrightarrow \langle \boldsymbol{x}\,|\,\boldsymbol{x}\rangle = 0 && \text{by hypothesis } \boldsymbol{Y} \perp \boldsymbol{Z} \\ &\Longrightarrow \boldsymbol{x} = \mathbb{0} && \text{by } \textit{non-isotropic} \text{ property of } \langle\triangle\,|\,\triangledown\rangle \text{ (Definition K.1 page 271)} \end{aligned}$$

✏

Theorem I.5. [15] *Let* $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *and* $\boldsymbol{Z} \triangleq \big(Z, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be linear subspaces of an* INNER PRODUCT SPACE $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$.

THM

$$\left\{\begin{array}{ll} 1. & \boldsymbol{Y} \perp \boldsymbol{Z} \text{ and} \\ 2. & \boldsymbol{x} \in \boldsymbol{Y} \hat{+} \boldsymbol{Z} \end{array}\right\} \Longrightarrow \left\{\begin{array}{ll} 1. & \text{There exists } \boldsymbol{y} \in \boldsymbol{Y} \text{ and } \boldsymbol{z} \in \boldsymbol{Z} \text{ such that } \boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z} \text{ and} \\ 2. & \boldsymbol{y} \text{ and } \boldsymbol{z} \text{ are UNIQUE.} \end{array}\right\}$$

✎PROOF:

1. Proof that $\boldsymbol{y}$ and $\boldsymbol{z}$ exist: by definition of Minkowski addition operator $\hat{+}$ (Definition I.2 page 253).
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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

(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 K.9 (page 284), $\boldsymbol{y}_1 - \boldsymbol{y}_2 = \mathbb{0}$ 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*.

✏

I.3 Subspaces of a Hilbert Space

Theorem I.6. [16] *Let* $\boldsymbol{H} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \triangledown \rangle\big)$ *be a* HILBERT SPACE *(Definition ?? page ??).*
Let Y *be a* SUBSET *of* X*, and let* $\mathsf{d}(\boldsymbol{x}, Y) \triangleq \inf\limits_{\boldsymbol{y} \in Y} \|\boldsymbol{x} - \boldsymbol{y}\|$.

THM

$$\left\{\begin{array}{lll} 1. & Y \neq \varnothing & \textit{and}\\ 2. & Y \textit{ is } \text{CLOSED} & \textit{and}\\ 3. & Y \textit{ is } \text{CONVEX} \quad \textit{(Definition M.6 page 298)} & \end{array}\right\} \implies \left\{\begin{array}{ll} \textit{There exists } \boldsymbol{p} \in Y \textit{ such that} & \\ \quad 1. \ \mathsf{d}(\boldsymbol{x}, Y) = \|\boldsymbol{x} - \boldsymbol{p}\| & \textit{and}\\ \quad 2. \ \boldsymbol{p} \textit{ is } \text{UNIQUE}. & \end{array}\right\}$$

✎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*:

$$\begin{aligned}
&\lim_{m,n\to\infty} \|\boldsymbol{y}_n - \boldsymbol{y}_m\|^2\\
&= \lim_{m,n\to\infty} \|(\boldsymbol{y}_n - x) + (x - \boldsymbol{y}_m)\|^2\\
&= \lim_{m,n\to\infty} \left\{-\|(\boldsymbol{y}_n - x) - (x - \boldsymbol{y}_m)\|^2 + 2\|\boldsymbol{y}_n - x\|^2 + 2\|x - \boldsymbol{y}_m\|^2\right\} && \text{by } \textit{parallelogram law} \text{ (page 279)}\\
&= \lim_{m,n\to\infty} \left\{-4\left\|\underbrace{\left(\frac{1}{2}\boldsymbol{y}_n + \frac{1}{2}\boldsymbol{y}_m\right)}_{\text{in } \boldsymbol{Y} \text{ by } \textit{convexity}} - x\right\|^2 + 2\|\boldsymbol{y}_n - x\|^2 + 2\|x - \boldsymbol{y}_m\|^2\right\}\\
&\le \lim_{m,n\to\infty} \left\{-4\delta^2 + 2\|\boldsymbol{y}_n - x\|^2 + 2\|x - \boldsymbol{y}_m\|^2\right\} && \text{by definition of } \delta \text{ (item (1))}\\
&= -4\delta^2 + \lim_{m,n\to\infty} \left\{2\|\boldsymbol{y}_n - x\|^2\right\} + \lim_{m,n\to\infty} \left\{2\|x - \boldsymbol{y}_m\|^2\right\}\\
&= -4\delta^2 + 2\delta^2 + 2\delta^2 && \text{by definition of } \delta \text{ (item (1))}\\
&= 0
\end{aligned}$$

[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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

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*.

✏

Theorem I.7. [17] *Let* $\boldsymbol{H} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be a* HILBERT SPACE *(Definition ?? page ??). Let* $\mathsf{d}(\boldsymbol{x}, Y) \triangleq \inf_{\boldsymbol{y}\in Y} \|\boldsymbol{x} - \boldsymbol{y}\|$. *Let* $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *and* $Y^\perp$ *the* ORTHOGONAL COMPLEMENT *of* Y.

THM

$$\{\ \boldsymbol{Y} \text{ is a SUBSPACE of } \boldsymbol{H}\ \} \quad\Longrightarrow\quad \left\{\begin{array}{ll} \multicolumn{2}{l}{\textit{There exists } \boldsymbol{p} \in Y \textit{ such that}} \\ 1.\ \mathsf{d}(\boldsymbol{x}, Y) = \|\boldsymbol{x} - \boldsymbol{p}\| & \textit{and} \\ 2.\ \boldsymbol{p} \textit{ is UNIQUE} & \textit{and} \\ 3.\ \boldsymbol{x} - \boldsymbol{p} \in Y^\perp. & \end{array}\right\}$$

Theorem I.8 (Projection Theorem). [18] *Let* $\boldsymbol{H} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be a Hilbert space.*

THM

$$\{\ \boldsymbol{Y} \text{ is a SUBSPACE of } \boldsymbol{H}\ \} \quad\Longrightarrow\quad \{\ \boldsymbol{Y} \hat{+} \boldsymbol{Y}^\perp = \boldsymbol{H}\ \}$$

✎PROOF:

$$\begin{aligned} \boldsymbol{Y} \hat{+} \boldsymbol{Y}^\perp &= \big[\boldsymbol{Y} \hat{+} \boldsymbol{Y}^\perp\big]^{\perp\perp} && \text{by Corollary I.1 page 260} \\ &= \big[\boldsymbol{Y}^\perp \Cap \boldsymbol{Y}^{\perp\perp}\big]^\perp && \text{by Proposition I.5 (page 257)} \\ &= \{\mathbb{0}\}^\perp && \text{by Corollary I.1 page 260} \\ &= \boldsymbol{H} && \text{by Proposition I.6 page 257} \end{aligned}$$

✏

The inclusion relation $\subseteq$ is an order relation on the set of subspaces of a linear space $\boldsymbol{\Omega}$.

Proposition I.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* and because

$$\left.\begin{array}{llll} 1.\ \boldsymbol{X} \subseteq \boldsymbol{X} & \forall \boldsymbol{X}\in\boldsymbol{S} & \text{(reflexive)} & \text{and} \\ 2.\ \boldsymbol{X} \subseteq \boldsymbol{Y} \text{ and } \boldsymbol{Y} \subseteq \boldsymbol{Z} \implies \boldsymbol{X} \subseteq \boldsymbol{Z} & \forall \boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}\in\boldsymbol{S} & \text{(transitive)} & \text{and} \end{array}\right] \text{preorder}$$

$$3.\ \boldsymbol{X} \subseteq \boldsymbol{Y} \text{ and } \boldsymbol{Y} \subseteq \boldsymbol{X} \implies \boldsymbol{X} = \boldsymbol{Y} \quad \forall \boldsymbol{X},\boldsymbol{Y}\in\boldsymbol{S} \quad \text{(anti-symmetric)}$$

✏

Theorem I.9. [19] *Let* $\boldsymbol{H}$ *be a Hilbert space and* $2^{\boldsymbol{H}}$ *the set of closed linear subspaces of* $\boldsymbol{H}$.

THM

$\big(2^{\boldsymbol{H}}, \hat{+}, \Cap, \mathbf{0}, \boldsymbol{H}; \subseteq\big)$ *is an* ORTHOMODULAR LATTICE. *In particular*

$$\begin{array}{llll} 1.\ \boldsymbol{X} \hat{+} \boldsymbol{X}^\perp = \boldsymbol{H} & \forall \boldsymbol{X}\in\boldsymbol{H} & \textit{(COMPLEMENTED)} \\ 2.\ \boldsymbol{X} \Cap \boldsymbol{X}^\perp = \mathbf{0} & \forall \boldsymbol{X}\in\boldsymbol{H} & \textit{(COMPLEMENTED)} \\ 3.\ \big(\boldsymbol{X}^\perp\big)^\perp = \boldsymbol{X} & \forall \boldsymbol{X}\in\boldsymbol{H} & \textit{(INVOLUTORY)} \\ 4.\ \boldsymbol{X} \le \boldsymbol{Y} \implies \boldsymbol{Y}^\perp \le \boldsymbol{X}^\perp & \forall \boldsymbol{X},\boldsymbol{Y}\in\boldsymbol{H} & \textit{(ANTITONE)} \\ 5.\ \boldsymbol{X} \le \boldsymbol{Y} \implies \boldsymbol{X} \hat{+} \big(\boldsymbol{X}^\perp \Cap \boldsymbol{Y}\big) = \boldsymbol{Y} & \forall \boldsymbol{X},\boldsymbol{Y}\in X & \textit{(ORTHOMODULAR IDENTITY)} \end{array}$$

[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

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

1. Proof for *complemented* (1) property: by *Projection Theorem* (Theorem I.8 page 262).
2. Proof for *complemented* (2) property: by Corollary I.1 (page 260).
3. Proof for *involutory* property: by Corollary I.1 (page 260).
4. Proof for *antitone* property: by Proposition I.5 (page 257).
5. Proof for *orthomodular identity* property:
6. Proof that lattice is *orthomodular*: by 5 properties and definition of *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

$$\boldsymbol{X} \subseteq \boldsymbol{Y} \quad \Longrightarrow \quad \boldsymbol{Y} = \boldsymbol{X} \hat{+} \left(\boldsymbol{X}^{\perp} \cap Y\right).$$

Corollary I.2. *Let $\boldsymbol{H}$ be a Hilbert space with orthogonality operation $\perp$. Let $\left(2^{\boldsymbol{H}}, \hat{+}, \cap, \mathbf{0}, \boldsymbol{H}; \subseteq\right)$ be the lattice of subspaces of $\boldsymbol{H}$.*

COR	$(\boldsymbol{X} \hat{+} \boldsymbol{Y})^{\perp}$	$=$	$\boldsymbol{X}^{\perp} \cap \boldsymbol{Y}^{\perp}$	$\forall \boldsymbol{X}, \boldsymbol{Y} \in 2^{\boldsymbol{H}}$	(DE MORGAN) *and*
	$(\boldsymbol{X} \cap \boldsymbol{Y})^{\perp}$	$=$	$\boldsymbol{X}^{\perp} \hat{+} \boldsymbol{Y}^{\perp}$	$\forall \boldsymbol{X}, \boldsymbol{Y} \in 2^{\boldsymbol{H}}$	(DE MORGAN)

✎PROOF: By properties of *orthocomplemented lattices* . ✎

I.4 Subspace Metrics

Definition I.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	$\mathsf{d}(\boldsymbol{V}, \boldsymbol{W})$	$\triangleq$	$\|\mathbf{P} - \mathbf{Q}\|$ $\forall \boldsymbol{V}, \boldsymbol{W} \in \boldsymbol{S}$	*(the distance between subspaces $\boldsymbol{V}$ and $\boldsymbol{W}$ is the size of the difference of their projection operators)*
	where $\boldsymbol{V}$	$\triangleq$	$\mathbf{P}\boldsymbol{X}$	*($\mathbf{P}$ is the projection operator that generates the subspace $\boldsymbol{V}$)*
	and $\boldsymbol{W}$	$\triangleq$	$\mathbf{Q}\boldsymbol{X}$	*($\mathbf{Q}$ is the projection operator that generates the subspace $\boldsymbol{W}$).*

Definition I.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	$\mathsf{d}(\boldsymbol{V}, \boldsymbol{W})$	$\triangleq$	$\max\left\{\sup\limits_{\boldsymbol{v} \in V, \|\boldsymbol{v}\|=1} \mathsf{p}(\boldsymbol{v}, \boldsymbol{W}), \sup\limits_{\boldsymbol{w} \in W, \|\boldsymbol{w}\|=1} \mathsf{p}(\boldsymbol{w}, \boldsymbol{V})\right\}$	$\forall \boldsymbol{V}, \boldsymbol{W} \in \boldsymbol{S}$
	where $\mathsf{p}(\boldsymbol{v}, \boldsymbol{W})$	$\triangleq$	$\inf\limits_{\boldsymbol{w} \in W} \|\boldsymbol{v} - \boldsymbol{w}\|$	*(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)

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Definition I.6 (Schäffer's metric). [22]

DEF

$$\mathsf{d}(\boldsymbol{V},\boldsymbol{W}) = \log(1+\max\{\mathsf{r}(\boldsymbol{V},\boldsymbol{W}),\mathsf{r}(\boldsymbol{W},\boldsymbol{V})\}) \quad \textit{where}$$

$$\mathsf{r}(\boldsymbol{V},\boldsymbol{W}) \triangleq \begin{cases} \inf\{\|\mathbf{A}-\mathbf{I}\| \,|\, \mathbf{AV}=\mathbf{W}\} & \textit{if } \mathbf{A} \textit{ and } \mathbf{A}^{-1} \textit{ both exist} \\ 1 & \textit{otherwise} \end{cases}$$

I.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

Copyright © 2019 Daniel J. Greenhoe

APPENDIX J

NORMED ALGEBRAS

J.1 Algebras

All *linear spaces* (Definition G.1 page 209) 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 J.1. [3] *Let* $\boldsymbol{A}$ *be an* ALGEBRA.

DEF *An algebra* $\boldsymbol{A}$ *is* ***unital*** *if* $\exists u \in \boldsymbol{A}$ such that $ux = xu = x \quad \forall x \in \boldsymbol{A}$

Definition J.2. [4] *Let* $\boldsymbol{A}$ *be an* UNITAL ALGEBRA *(Definition J.1 page 265) 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* $\rho_x(\lambda) \triangleq (\lambda e - x)^{-1} \quad \forall \lambda \notin \sigma(x)$.

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

J.2 Star-Algebras

Definition J.3. [5] *Let $\boldsymbol{A}$ be an* ALGEBRA.

DEF

*The pair $(\boldsymbol{A}, *)$ is a $*$-**algebra**, or **star-algebra**, if*

1.	$(x+y)^* = x^* + y^*$	$\forall x,y \in \boldsymbol{A}$	(DISTRIBUTIVE)	*and*
2.	$(\alpha x)^* = \bar{\alpha} x^*$	$\forall x \in \boldsymbol{A}, \alpha \in \mathbb{C}$	(CONJUGATE LINEAR)	*and*
3.	$(xy)^* = y^* x^*$	$\forall x,y \in \boldsymbol{A}$	(ANTIAUTOMORPHIC)	*and*
4.	$x^{**} = x$	$\forall x \in \boldsymbol{A}$	(INVOLUTORY)	

*The operator $*$ is called an **involution** on the algebra $\boldsymbol{A}$.*

Proposition J.1. [6] *Let $(\boldsymbol{A}, *)$ be an* UNITAL $*$-ALGEBRA.

PRP

$$x \text{ is invertible} \quad \Longrightarrow \quad \begin{cases} 1. \quad x^* \text{ is INVERTIBLE} & \forall x \in \boldsymbol{A} \quad \text{and} \\ 2. \quad \left(x^*\right)^{-1} = \left(x^{-1}\right)^* & \forall x \in \boldsymbol{A} \end{cases}$$

✎PROOF: Let e be the unit element of $(\boldsymbol{A}, *)$.

1. Proof that $e^* = e$:

$$\begin{aligned}
x\,e^* &= \left(x\,e^*\right)^{**} && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition J.3 page 266)} \\
&= \left(x^*\,e^{**}\right)^* && \text{by } \textit{antiautomorphic} \text{ property of } * && \text{(Definition J.3 page 266)} \\
&= \left(x^*\,e\right)^* && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition J.3 page 266)} \\
&= \left(x^*\right)^* && \text{by definition of } e \\
&= x && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition J.3 page 266)} \\
e^*\,x &= \left(e^*\,x\right)^{**} && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition J.3 page 266)} \\
&= \left(e^{**}\,x^*\right)^* && \text{by } \textit{antiautomorphic} \text{ property of } * && \text{(Definition J.3 page 266)} \\
&= \left(e\,x^*\right)^* && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition J.3 page 266)} \\
&= \left(x^*\right)^* && \text{by definition of } e \\
&= x && \text{by } \textit{involutory} \text{ property of } * && \text{(Definition J.3 page 266)}
\end{aligned}$$

2. Proof that $\left(x^*\right)^{-1} = \left(x^{-1}\right)^*$:

$$\begin{aligned}
\left(x^{-1}\right)^* \left(x^*\right) &= \left[x\left(x^{-1}\right)\right]^* && \text{by } \textit{antiautomorphic} \text{ and } \textit{involution} \text{ properties of } * && \text{(Definition J.3 page 266)} \\
&= e^* \\
&= e && \text{by item (1) page 266} \\
\left(x^*\right)\left(x^{-1}\right)^* &= \left[x^{-1}\,x\right]^* && \text{by } \textit{antiautomorphic} \text{ and } \textit{involution} \text{ properties of } * && \text{(Definition J.3 page 266)} \\
&= e^* \\
&= e && \text{by item (1) page 266}
\end{aligned}$$

✏

Definition J.4. [7] *Let $(\boldsymbol{A}, \|\cdot\|)$ be a* $*$-ALGEBRA *(Definition J.3 page 266).*

DEF

- *An element $x \in \boldsymbol{A}$ is **hermitian** or **self-adjoint** if $x^* = x$.*
- *An element $x \in \boldsymbol{A}$ is **normal** if $xx^* = x^*x$.*
- *An element $x \in \boldsymbol{A}$ 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

COPYRIGHT © 2019 DANIEL J. GREENHOE

Theorem J.1. [8] *Let* $(\boldsymbol{A}, \|\cdot\|)$ *be a* $*$-ALGEBRA *(Definition J.3 page 266)*.

THM

$$\underbrace{x = x^* \text{ and } y = y^*}_{x \text{ and } y \text{ are HERMITIAN}} \implies \begin{cases} x + y = (x+y)^* & (x+y \text{ is self adjoint}) \\ x^* = (x^*)^* & (x^* \text{ is self adjoint}) \\ \underbrace{xy = (xy)^*}_{(xy) \text{ is HERMITIAN}} \iff \underbrace{xy = yx}_{\text{commutative}} \end{cases}$$

✎PROOF:

$(x+y)^* = x^* + y^*$	by *distributive* property of $*$	(Definition J.3 page 266)
$= x + y$	by left hypothesis	
$(x^*)^* = x$	by *involutory* property of $*$	(Definition J.3 page 266)

Proof that $xy = (xy)^* \implies xy = yx$

$xy = (xy)^*$	by left hypothesis	
$= y^* x^*$	by *antiautomorphic* property of $*$	(Definition J.3 page 266)
$= yx$	by left hypothesis	

Proof that $xy = (xy)^* \impliedby xy = yx$

$(xy)^* = (yx)^*$	by left hypothesis	
$= x^* y^*$	by *antiautomorphic* property of $*$	(Definition J.3 page 266)
$= xy$	by left hypothesis	

✏

Definition J.5 (Hermitian components). [9] *Let* $(\boldsymbol{A}, \|\cdot\|)$ *be a* $*$-ALGEBRA *(Definition J.3 page 266)*.

DEF

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 J.2. [10] *Let* $(\boldsymbol{A}, *)$ *be a* $*$-ALGEBRA *(Definition J.3 page 266)*.

THM

$$\begin{aligned} \mathbf{R}_{\mathrm{e}} x &= \left(\mathbf{R}_{\mathrm{e}} x\right)^* && \forall x \in \boldsymbol{A} && (\mathbf{R}_{\mathrm{e}} x \text{ is HERMITIAN}) \\ \mathbf{I}_{\mathrm{m}} x &= \left(\mathbf{I}_{\mathrm{m}} x\right)^* && \forall x \in \boldsymbol{A} && (\mathbf{I}_{\mathrm{m}} x \text{ is HERMITIAN}) \end{aligned}$$

✎PROOF:

$\left(\mathbf{R}_{\mathrm{e}} x\right)^* = \left(\frac{1}{2}\left(x + x^*\right)\right)^*$	by definition of $\Re$	(Definition J.5 page 267)
$= \frac{1}{2}\left(x^* + x^{**}\right)$	by *distributive* property of $*$	(Definition J.3 page 266)
$= \frac{1}{2}\left(x^* + x\right)$	by *involutory* property of $*$	(Definition J.3 page 266)
$= \mathbf{R}_{\mathrm{e}} x$	by definition of $\Re$	(Definition J.5 page 267)
$\left(\mathbf{I}_{\mathrm{m}} x\right)^* = \left(\frac{1}{2i}\left(x - x^*\right)\right)^*$	by definition of $\Im$	(Definition J.5 page 267)

[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

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

$$= \frac{1}{2i}\left(x^* - x^{**}\right) \quad \text{by } \textit{distributive} \text{ property of } * \quad \text{(Definition J.3 page 266)}$$
$$= \frac{1}{2i}\left(x^* - x\right) \quad \text{by } \textit{involutory} \text{ property of } * \quad \text{(Definition J.3 page 266)}$$
$$= \mathbf{I}_{\mathsf{m}} x \quad \text{by definition of } \mathfrak{I} \quad \text{(Definition J.5 page 267)}$$

✏

Theorem J.3 (Hermitian representation). [11] *Let* $(\boldsymbol{A}, *)$ *be a* $*$-ALGEBRA *(Definition J.3 page 266)*.

THM $a = x + iy \iff x = \mathbf{R}_{\mathsf{e}} a$ *and* $y = \mathbf{I}_{\mathsf{m}} a$

✎PROOF:

Proof that $a = x + iy \implies x = \mathbf{R}_{\mathsf{e}} a$ and $y = \mathbf{I}_{\mathsf{m}} a$:

$$a = x + iy \quad \text{by left hypothesis}$$
$$\implies a^* = (x+iy)^* \quad \text{by definition of } \textit{adjoint} \quad \text{(Definition J.4 page 266)}$$
$$= x^* - iy^* \quad \text{by } \textit{distributive} \text{ property of } * \quad \text{(Definition J.3 page 266)}$$
$$= x - iy \quad \text{by Theorem J.2 page 267}$$
$$\implies x = a - iy \quad \text{by solving for } x \text{ in } a = x + iy \text{ equation}$$
$$x = a^* + iy \quad \text{by solving for } x \text{ in } a^* = x - iy \text{ equation}$$
$$\implies x + x = a + a^* \quad \text{by adding previous 2 equations}$$
$$\implies 2x = a + a^* \quad \text{by solving for } x \text{ in previous equation}$$
$$\implies x = \frac{1}{2}\left(a + a^*\right)$$
$$= \mathbf{R}_{\mathsf{e}} a \quad \text{by definition of } \mathfrak{R} \quad \text{(Definition J.5 page 267)}$$

$$iy = a - x \quad \text{by solving for } iy \text{ in } a = x + iy \text{ equation}$$
$$iy = -a^* + x \quad \text{by solving for } iy \text{ in } a = x + iy \text{ equation}$$
$$\implies iy + iy = a - a^* \quad \text{by adding previous 2 equations}$$
$$\implies y = \frac{1}{2i}\left(a - a^*\right) \quad \text{by solving for } iy \text{ in previous equations}$$
$$= \mathbf{I}_{\mathsf{m}} a \quad \text{by definition of } \mathfrak{I} \quad \text{(Definition J.5 page 267)}$$

Proof that $a = x + iy \impliedby x = \mathbf{R}_{\mathsf{e}} a$ and $y = \mathbf{I}_{\mathsf{m}} a$:

$$x + iy = \mathbf{R}_{\mathsf{e}} a + i\,\mathbf{I}_{\mathsf{m}} a \quad \text{by right hypothesis}$$
$$= \underbrace{\frac{1}{2}\left(a + a^*\right)}_{\mathbf{R}_{\mathsf{e}} a} + i\underbrace{\frac{1}{2i}\left(a - a^*\right)}_{\mathbf{I}_{\mathsf{m}} a} \quad \text{by definition of } \mathfrak{R} \text{ and } \mathfrak{I} \quad \text{(Definition J.5 page 267)}$$
$$= \left(\frac{1}{2}a + \frac{1}{2}a\right) + \cancelto{0}{\left(\frac{1}{2}a^* - \frac{1}{2}a^*\right)}$$
$$= a$$

✏

[11] Michel and Herget (1993) page 430, Rickart (1960), page 179, Gelfand and Neumark (1943b), page 7

COPYRIGHT © 2019 DANIEL J. GREENHOE

J.3 Normed Algebras

Definition J.6. [12] *Let $\boldsymbol{A}$ be an algebra.*

DEF

The pair $(\boldsymbol{A}, \|\cdot\|)$ is a ***normed algebra*** *if*

$$\|xy\| \le \|x\|\,\|y\| \qquad \forall x, y \in \boldsymbol{A} \qquad \textit{(multiplicative condition)}$$

A normed algebra $(\boldsymbol{A}, \|\cdot\|)$ is a ***Banach algebra*** *if $(\boldsymbol{A}, \|\cdot\|)$ is also a Banach space.*

Proposition J.2.

PRP

$$(\boldsymbol{A}, \|\cdot\|) \text{ is a normed algebra} \quad \Longrightarrow \quad \text{multiplication is } \textbf{continuous} \text{ in } (\boldsymbol{A}, \|\cdot\|)$$

✎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 $(\boldsymbol{A}, \|\cdot\|)$:

$$\begin{aligned}
\|\mathsf{f}(x) - \mathsf{f}(y)\| &\triangleq \|zx - zy\| && \text{by definition of } \mathsf{f} && \text{(item (1) page 269)}\\
&= \|z(x - y)\| \\
&\le \|z\|\,\|x - y\| && \text{by definition of } \textit{normed algebra} && \text{(Definition J.6 page 269)}\\
&\triangleq \|z\|\,\delta && \text{by definition of } \delta && \text{(item (2) page 269)}\\
&\le \epsilon && \text{for some value of } \delta > 0
\end{aligned}$$

✏

Theorem J.4 (Gelfand-Mazur Theorem)**.** [13] *Let $\mathbb{C}$ be the field of complex numbers.*

THM

$$\left.\begin{array}{l}(\boldsymbol{A}, \|\cdot\|) \text{ is a Banach algebra}\\ \text{every nonzero } x \in \boldsymbol{A} \text{ is invertible}\end{array}\right\} \quad \Longrightarrow \quad \boldsymbol{A} \equiv \mathbb{C} \quad (\boldsymbol{A} \text{ is isomorphic to } \mathbb{C})$$

J.4 C* Algebras

Definition J.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. $\left\|x^* x\right\| = \|x\|^2 \quad \forall x \in \boldsymbol{A}$.

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)

COPYRIGHT © 2019 DANIEL J. GREENHOE

Theorem J.5. [15] *Let $\boldsymbol{A}$ be an algebra.*

THM $(\boldsymbol{A}, \|\cdot\|, *)$ *is a* C^* ***algebra*** $\Longrightarrow$ $\|x^*\| = \|x\|$

✎PROOF:

$$\begin{aligned}
\|x\| &= \frac{1}{\|x\|}\ \|x\|^2 \\
&= \frac{1}{\|x\|}\ \|x^*x\| && \text{by definition of } C^*\text{-}\textit{algebra} && \text{(Definition J.7 page 269)} \\
&\le \frac{1}{\|x\|}\ \|x^*\|\,\|x\| && \text{by definition of } \textit{normed algebra} && \text{(Definition J.6 page 269)} \\
&= \|x^*\| \\
\|x^*\| &\le \|x^{**}\| && \text{by previous result} \\
&= \|x\| && \text{by } \textit{involution} \text{ property of } * && \text{(Definition J.3 page 266)}
\end{aligned}$$

✏

[15] Folland (1995) page 1, Gelfand and Neumark (1943b), page 4, Gelfand and Neumark (1943a)

COPYRIGHT © 2019 DANIEL J. GREENHOE

APPENDIX K

INNER PRODUCT SPACES

K.1 Definition and basic results

Definition K.1. [1] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a* LINEAR SPACE *(Definition G.1 page 209).*

DEF

A FUNCTIONAL $\langle \triangle \mid \triangledown \rangle \in \mathbb{F}^{X \times X}$ *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 $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\big)$ *is called an* ***inner product space****.*

Theorem K.1. [2] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\big)$ *be a* LINEAR SPACE *(Definition G.1 page 209).*

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:

$\langle \boldsymbol{x} \mid \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{y} + \boldsymbol{z} \mid \boldsymbol{x} \rangle^*$	by *conjugate symmetric* property of $\langle \triangle \mid \triangledown \rangle$	(Definition K.1 page 271)
$= (\langle \boldsymbol{y} \mid \boldsymbol{x} \rangle + \langle \boldsymbol{z} \mid \boldsymbol{x} \rangle)^*$	by *additive* property of $\langle \triangle \mid \triangledown \rangle$	(Definition K.1 page 271)
$= \langle \boldsymbol{y} \mid \boldsymbol{x} \rangle^* + \langle \boldsymbol{z} \mid \boldsymbol{x} \rangle^*$	by *distributive* property of $*$	(Definition J.3 page 266)
$= \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + \langle \boldsymbol{x} \mid \boldsymbol{z} \rangle$	by *conjugate symmetric* property of $\langle \triangle \mid \triangledown \rangle$	(Definition K.1 page 271)
$\langle \boldsymbol{x} \mid \alpha \boldsymbol{y} \rangle = \langle \alpha \boldsymbol{y} \mid \boldsymbol{x} \rangle^*$	by *conjugate symmetric* property of $\langle \triangle \mid \triangledown \rangle$	(Definition K.1 page 271)

[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} \mid \boldsymbol{x}\rangle)^* && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \mid \triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= \alpha^*\,\langle \boldsymbol{y} \mid \boldsymbol{x}\rangle^* && \text{by } \textit{antiautomorphic} \text{ property of } * && \text{(Definition J.3 page 266)}\\
&= \alpha^*\,\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by } \textit{conjugate symmetric} \text{ property of } \langle \triangle \mid \triangledown\rangle && \text{(Definition K.1 page 271)}\\
\langle \boldsymbol{x} \mid \mathbb{0}\rangle &= \langle \mathbb{0} \mid \boldsymbol{x}\rangle^* && \text{by } \textit{conjugate symmetric} \text{ property of } \langle \triangle \mid \triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= \langle 0\cdot \boldsymbol{y} \mid \boldsymbol{x}\rangle^* && &&\\
&= (0\cdot\langle \boldsymbol{y} \mid \boldsymbol{x}\rangle)^* && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \mid \triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= 0 && &&\\
\langle \mathbb{0} \mid \boldsymbol{x}\rangle &= \langle 0\cdot \boldsymbol{y} \mid \boldsymbol{x}\rangle && &&\\
&= (0\cdot\langle \boldsymbol{y} \mid \boldsymbol{x}\rangle) && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \mid \triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= 0 && &&\\
\langle \boldsymbol{x}-\boldsymbol{y} \mid \boldsymbol{z}\rangle &= \langle \boldsymbol{x}+(-\boldsymbol{y}) \mid \boldsymbol{z}\rangle && \text{by definition of } + && \\
&= \langle \boldsymbol{x} \mid \boldsymbol{z}\rangle + \langle -\boldsymbol{y} \mid \boldsymbol{z}\rangle && \text{by } \textit{additive} \text{ property of } \langle \triangle \mid \triangledown\rangle && \text{(Definition K.1 page 271)}\\
&= \langle \boldsymbol{x} \mid \boldsymbol{z}\rangle - \langle \boldsymbol{y} \mid \boldsymbol{z}\rangle && \text{by } \textit{homogeneous} \text{ property of } \langle \triangle \mid \triangledown\rangle && \text{(Definition K.1 page 271)}\\
\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 K.1 page 271)}\\
&= (\langle \boldsymbol{y} \mid \boldsymbol{x}\rangle - \langle \boldsymbol{z} \mid \boldsymbol{x}\rangle)^* && \text{by 4.} && \\
&= \langle \boldsymbol{y} \mid \boldsymbol{x}\rangle^* - \langle \boldsymbol{z} \mid \boldsymbol{x}\rangle^* && \text{by } \textit{distributive} \text{ property of } * && \text{(Definition J.3 page 266)}\\
&= \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 K.1 page 271)}
\end{aligned}
$$

$$
\begin{aligned}
& \langle \boldsymbol{x} \mid \boldsymbol{z}\rangle = \langle \boldsymbol{y} \mid \boldsymbol{z}\rangle && \forall \boldsymbol{z} && \\
\iff\quad & \langle \boldsymbol{x} \mid \boldsymbol{z}\rangle - \langle \boldsymbol{y} \mid \boldsymbol{z}\rangle = 0 && \forall \boldsymbol{z} && \text{by property of complex numbers}\\
\iff\quad & \langle \boldsymbol{x}-\boldsymbol{y} \mid \boldsymbol{z}\rangle = 0 && \forall \boldsymbol{z} && \text{by 4.}\\
\iff\quad & \boldsymbol{x}-\boldsymbol{y} = \mathbb{0} && \forall \boldsymbol{z} && \text{by } \textit{non-isotropic} \text{ property of } \langle \triangle \mid \triangledown\rangle \text{ (Definition K.1 page 271)}
\end{aligned}
$$

Proof that $\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle = 0 \implies \boldsymbol{y} = \mathbb{0}$:

1. Suppose $\boldsymbol{y} \neq \mathbb{0}$;
2. Then $\langle \boldsymbol{y} \mid \boldsymbol{y}\rangle \neq \mathbb{0}$ by the *non-isotropic* property of $\langle \triangle \mid \triangledown\rangle$ (Definition K.1 page 271)
3. But because $\boldsymbol{y} \in X$, the left hypothesis implies that $\langle \boldsymbol{y} \mid \boldsymbol{y}\rangle = 0$.
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 $\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle = 0 \impliedby \boldsymbol{y} = \mathbb{0}$:

$$
\begin{aligned}
\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle &= \langle \boldsymbol{x} \mid \mathbb{0}\rangle && \text{by right hypothesis}\\
&= 0 && \text{by Theorem K.1 page 271}
\end{aligned}
$$

✏

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 K.2 (Cauchy-Schwarz Inequality). [3] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown\rangle\big)$ *be an* INNER PRODUCT SPACE *and* $|\cdot| \in \mathbb{R}^{\mathbb{C}}$ *an* ABSOLUTE VALUE *function* *(Definition F.4 page 208)*. *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)

COPYRIGHT © 2019 DANIEL J. GREENHOE

that $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$.[4]

THM

$$
\begin{array}{rcl}
|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 & \leq & \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \ \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle \qquad \forall \boldsymbol{x},\boldsymbol{y} \in X \\
|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 & = & \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \ \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle \iff \exists \alpha \in \mathbb{F} \text{ such that } \boldsymbol{y} = \alpha \boldsymbol{x} \qquad \forall \boldsymbol{x},\boldsymbol{y} \in X \\
|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle| & \leq & \|\boldsymbol{x}\| \ \|\boldsymbol{y}\| \qquad \forall \boldsymbol{x},\boldsymbol{y} \in X \\
|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle| & = & \|\boldsymbol{x}\| \ \|\boldsymbol{y}\| \iff \exists \alpha \in \mathbb{F} \text{ such that } \boldsymbol{y} = \alpha \boldsymbol{x} \qquad \forall \boldsymbol{x},\boldsymbol{y} \in X
\end{array}
$$

✎PROOF:

1. Proof that $|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\| \ \|\boldsymbol{y}\|$: [5]

   (a) $\boldsymbol{y} = \mathbb{0}$ case:

$$
\begin{aligned}
|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 &= |\langle \boldsymbol{x} \mid \mathbb{0} \rangle|^2 && \text{by } \boldsymbol{y} = \mathbb{0} \text{ hypothesis} \\
&= |\langle \mathbb{0} \mid \boldsymbol{x} \rangle|^2 && \text{by Definition K.1 page 271} \\
&= |\langle 0\mathbb{0} \mid \boldsymbol{x} \rangle|^2 && \text{by Definition G.1 page 209} \\
&= |0 \langle \mathbb{0} \mid \boldsymbol{x} \rangle|^2 && \text{by Definition K.1 page 271} \\
&= 0 \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \langle \mathbb{0} \mid \mathbb{0} \rangle \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle && \text{by } \boldsymbol{y} = \mathbb{0} \text{ hypothesis}
\end{aligned}
$$

   (b) $\boldsymbol{y} \neq \mathbb{0}$ case: Let $\lambda \triangleq \frac{\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle}{\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle}$.

$$
\begin{aligned}
0 &\leq \langle \boldsymbol{x} - \lambda \boldsymbol{y} \mid \boldsymbol{x} - \lambda \boldsymbol{y} \rangle && \text{by Definition K.1} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} - \lambda \boldsymbol{y} \rangle + \langle -\lambda \boldsymbol{y} \mid \boldsymbol{x} - \lambda \boldsymbol{y} \rangle && \text{by Definition K.1} \\
&= \langle \boldsymbol{x} - \lambda \boldsymbol{y} \mid \boldsymbol{x} \rangle^* - \lambda \langle \boldsymbol{y} \mid \boldsymbol{x} - \lambda \boldsymbol{y} \rangle && \text{by Definition K.1} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle^* + \langle -\lambda \boldsymbol{y} \mid \boldsymbol{x} \rangle^* - \lambda \langle \boldsymbol{x} - \lambda \boldsymbol{y} \mid \boldsymbol{y} \rangle^* && \text{by Definition K.1} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle^* - \lambda^* \langle \boldsymbol{y} \mid \boldsymbol{x} \rangle^* - \lambda \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle^* - \lambda \langle -\lambda \boldsymbol{y} \mid \boldsymbol{y} \rangle^* && \text{by Definition K.1} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle - \lambda^* \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle - \lambda \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle^* + \lambda \lambda^* \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle^* && \text{by Definition K.1} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle + \cancelto{0}{\left[\frac{\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle}{\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle} \lambda^* \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle - \lambda^* \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle \right]} - \frac{\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle}{\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle} \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle^* && \text{by definition of } \lambda \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle - \frac{1}{\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle} |\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 \\
\implies |\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 &\leq \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \ \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle
\end{aligned}
$$

2. Proof that $|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 = \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \ \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle \iff \boldsymbol{y} = a\boldsymbol{x}$:
   Let $\frac{1}{a} \triangleq \lambda \triangleq \frac{\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle}{\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle}$. Then…

$$
\begin{aligned}
& \boldsymbol{y} = a\boldsymbol{x} \\
\iff \quad & \boldsymbol{x} = \lambda \boldsymbol{y} \\
\iff \quad & \boldsymbol{x} - \lambda \boldsymbol{y} = \mathbb{0} \\
\iff \quad & 0 = \langle \boldsymbol{x} - \lambda \boldsymbol{y} \mid \boldsymbol{x} - \lambda \boldsymbol{y} \rangle && \text{by Definition K.1 page 271} \\
& = \langle \boldsymbol{x} \mid \boldsymbol{x} - \lambda \boldsymbol{y} \rangle + \langle -\lambda \boldsymbol{y} \mid \boldsymbol{x} - \lambda \boldsymbol{y} \rangle && \text{by Definition K.1 page 271} \\
& = \langle \boldsymbol{x} - \lambda \boldsymbol{y} \mid \boldsymbol{x} \rangle^* - \lambda \langle \boldsymbol{y} \mid \boldsymbol{x} - \lambda \boldsymbol{y} \rangle && \text{by Definition K.1 page 271} \\
& \vdots && \text{(same steps as in 1(b)}
\end{aligned}
$$

[4] The function $\|\cdot\|$ is a *norm* (Theorem K.4 page 276) and is called the *norm induced by the inner product* $\langle \triangle \mid \triangledown \rangle$ (Definition K.2 page 277).

[5] Haaser and Sullivan (1991), page 278

$$= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle - \frac{1}{\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle} \; |\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2$$

$$\iff \qquad |\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|^2 = \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \; \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle$$

3. Alternate proof for $|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|$: (Note: This is essentially the same proof as used by Schwarz). [6]

   (a) Proof that $\left\{a\lambda^2 + b\lambda + c \geq 0 \qquad \forall \lambda \in \mathbb{R}\right\} \quad \Longrightarrow \quad \left\{b^2 \leq 4ac\right\}$ (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

$$\begin{aligned}
0 &\leq a\lambda^2 + b\lambda + c && \text{by left hypothesis} \\
&= k(\lambda - r_1)(\lambda - r_2) && \text{by definition of } r_1 \text{ and } r_2 \\
&= k(\lambda^2 - r_1\lambda - r_2\lambda + r_1 r_2) \\
\Longrightarrow \qquad \lambda^2 - r_1\lambda - r_2\lambda + r_1 r_2 &\geq 0 \\
\Longrightarrow \qquad r_1 &= r_2^* && \text{because } r_1 r_2 \geq 0 \text{ for } \lambda = 0
\end{aligned}$$

   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 $b^2 \leq 4ac$ (the **discriminate** is non-positive).

   (b) Proof that $\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle \lambda^2 + 2|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|\lambda + \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \geq 0 \quad \forall \lambda \in \mathbb{R}$:

$$\begin{aligned}
0 &\leq \langle \boldsymbol{x} + \alpha \boldsymbol{y} \mid \boldsymbol{x} + \alpha \boldsymbol{y} \rangle && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} + \alpha \boldsymbol{y} \rangle + \langle \alpha \boldsymbol{y} \mid \boldsymbol{x} + \alpha \boldsymbol{y} \rangle && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} + \alpha \boldsymbol{y} \rangle + \alpha \langle \boldsymbol{y} \mid \boldsymbol{x} + \alpha \boldsymbol{y} \rangle && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x} + \alpha \boldsymbol{y} \mid \boldsymbol{x} \rangle^* + \alpha \langle \boldsymbol{x} + \alpha \boldsymbol{y} \mid \boldsymbol{y} \rangle^* && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle^* + \langle \alpha \boldsymbol{y} \mid \boldsymbol{x} \rangle^* + \alpha \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle^* + \alpha \langle \alpha \boldsymbol{y} \mid \boldsymbol{y} \rangle^* && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle^* + \alpha^* \langle \boldsymbol{y} \mid \boldsymbol{x} \rangle^* + \alpha \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle^* + \alpha\alpha^* \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle^* && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle + \alpha^* \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + \left(\alpha^* \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle\right)^* + |\alpha|^2 \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle + 2\Re\left(\alpha^* \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle\right) + |\alpha|^2 \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle \\
&\leq \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle + 2\left|\alpha^* \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle\right| + |\alpha|^2 \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle + 2|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle||\alpha| + \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle |\alpha|^2 \\
&= \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle |\alpha|^2 + 2\,|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|\,|\alpha| + \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \\
&= \underbrace{\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle}_{a} \lambda^2 + \underbrace{2|\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle|}_{b} \lambda + \underbrace{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}_{c} && \text{because } \lambda \triangleq |\alpha| \in \mathbb{R}
\end{aligned}$$

[6] Aliprantis and Burkinshaw (1998), page 278, Steele (2004), page 11

Copyright © 2019 Daniel J. Greenhoe

(c) The above equation is in the quadratic form used in the lemma of part (a).

$$\left(\underbrace{2|\langle x \mid y\rangle|}_{b}\right)^2 \le 4\,\underbrace{\langle y \mid y\rangle}_{a}\,\underbrace{\langle x \mid x\rangle}_{c} \qquad \text{by the results of parts (a) and (b)}$$

$$\Longrightarrow \quad |\langle x \mid y\rangle|^2 \le \langle x \mid x\rangle\,\langle y \mid y\rangle$$

4. Proof that $|\langle x \mid y\rangle| \le \|x\|\,\|y\|$:
   This follows directly from the definition $\|x\| \triangleq \sqrt{\langle x \mid x\rangle}$.

5. Proof that $|\langle x \mid y\rangle| = \|x\|\,\|y\| \quad \Longleftrightarrow \quad \exists \alpha \in \mathbb{C}$ such that $y = \alpha x$:
   This follows directly from the definition $\|x\| \triangleq \sqrt{\langle x \mid x\rangle}$.

✏

Corollary K.1. [7] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla\rangle\big)$ *be an* INNER PRODUCT SPACE.

COR $\langle x \mid y\rangle$ *is* CONTINUOUS *in both* x *and* y.

✎PROOF: Let $\|x\| \triangleq \sqrt{\langle x \mid x\rangle}$.

$$\begin{aligned}
|\langle x+\epsilon \mid y\rangle - \langle x \mid y\rangle|^2 &= |\langle x+\epsilon - x \mid y\rangle|^2 && \text{by } \textit{additivity} \text{ of } \langle \triangle \mid \nabla\rangle && \text{(Definition K.1 page 271)}\\
&= |\langle \epsilon \mid y\rangle|^2 \\
&\le \|\epsilon\|^2\,\|y\| && \text{by } \textit{Cauchy-Schwarz Inequality} && \text{(Theorem K.2 page 272)}
\end{aligned}$$

✏

K.2 Relationship between norms and inner products

K.2.1 Norms induced by inner products

Lemma K.1 (Polar Identity). [8] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \nabla\rangle\big)$ *be an* INNER PRODUCT SPACE *(Definition K.1 page 271). Let* $\Re z$ *represent the real part of* $z \in \mathbb{C}$. *Let* $\|\cdot\|$ *be a function in* $\mathbb{R}^{\mathbb{F}}$ *such that* $\|x\| \triangleq \sqrt{\langle x \mid x\rangle}$.[9]

LEM $\|x+y\|^2 = \|x\|^2 + 2\mathbf{R}_{\mathrm{e}}\left[\langle x \mid y\rangle\right] + \|y\|^2 \qquad \forall x,y \in X$

[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 K.4 page 276) and is called the *norm induced by the inner product* $\langle \triangle \mid \nabla\rangle$ (Definition K.2 page 277).

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

✎PROOF:

$$
\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|^2 &= \langle \boldsymbol{x}+\boldsymbol{y} \mid \boldsymbol{x}+\boldsymbol{y}\rangle && \text{by definition of } \textit{induced norm} && \text{(Theorem K.4 page 276)}\\
&= \langle \boldsymbol{x} \mid \boldsymbol{x}+\boldsymbol{y}\rangle + \langle \boldsymbol{y} \mid \boldsymbol{x}+\boldsymbol{y}\rangle && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x}+\boldsymbol{y} \mid \boldsymbol{x}\rangle^* + \langle \boldsymbol{x}+\boldsymbol{y} \mid \boldsymbol{y}\rangle^* && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x}\rangle^* + \langle \boldsymbol{y} \mid \boldsymbol{x}\rangle^* + \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle^* + \langle \boldsymbol{y} \mid \boldsymbol{y}\rangle^* && \text{by Definition K.1 page 271} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x}\rangle + \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle + \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle^* + \langle \boldsymbol{y} \mid \boldsymbol{y}\rangle && \text{by definition of } \textit{inner product} && \text{(Definition K.1 page 271)}\\
&= \|\boldsymbol{x}\|^2 + 2\Re\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle + \|\boldsymbol{y}\|^2 && \text{by definition of } \textit{induced norm} && \text{(Theorem K.4 page 276)}
\end{aligned}
$$

✏

Theorem K.3 (Minkowski's Inequality). [10] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown\rangle\big)$ *be an* INNER PRODUCT SPACE. *Let* $\|\cdot\|$ *be a function in* $\mathbb{R}^{\mathbb{F}}$ *such that* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle}$.[11]

THM $\|\boldsymbol{x}+\boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \quad \forall \boldsymbol{x},\boldsymbol{y}\in X$

✎PROOF:

$$
\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|^2 &= \|\boldsymbol{x}\|^2 + 2\Re\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle + \|\boldsymbol{y}\|^2 && \text{by } \textit{Polar Identity} && \text{(Lemma K.1 page 275)}\\
&\le \|\boldsymbol{x}\|^2 + 2|\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle| + \|\boldsymbol{y}\|^2 \\
&\le \|\boldsymbol{x}\|^2 + 2\sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle}\sqrt{\langle \boldsymbol{y} \mid \boldsymbol{y}\rangle} + \|\boldsymbol{y}\|^2 && \text{by } \textit{Cauchy-Schwarz Inequality} && \text{(Theorem K.2 page 272)}\\
&= (\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2
\end{aligned}
$$

✏

Theorem K.4 (induced norm). [12] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown\rangle\big)$ *be an* INNER PRODUCT SPACE *(Definition K.1 page 271)*.

THM $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle} \quad \Longrightarrow \quad \|\cdot\|$ *is a* NORM

✎PROOF: For a function to be a norm, it must satisfy the four properties listed in Definition L.1 (page 287).

1. Proof that $\|\cdot\|$ is a norm:
   (a) Proof that $\|\boldsymbol{x}\| > 0$ for $\boldsymbol{x} \neq 0$ (non-negative):
   By Definition K.1 page 271, all inner products have this property.
   (b) Proof that $\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = 0$ (non-isometric):
   By Definition K.1, all inner products have this property.
   (c) Prove $\|a\boldsymbol{x}\| = |a|\,\|\boldsymbol{x}\|$ (homogeneous):

$$\|a\boldsymbol{x}\| \triangleq \sqrt{\langle a\boldsymbol{x} \mid a\boldsymbol{x}\rangle} = \sqrt{aa^*\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle} = \sqrt{|a|^2\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle} = |a|\,\|\boldsymbol{x}\|$$

   (d) Proof that $\|\boldsymbol{x}+\boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ (subadditive): This is true by Minkowski's inequality page **??**

[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 K.4 page 276) and is called the *normed induced by the inner product* $\langle \triangle \mid \triangledown\rangle$ (Definition K.2 page 277).

[12] Aliprantis and Burkinshaw (1998), pages 278–279, Haaser and Sullivan (1991) page 278

COPYRIGHT © 2019 DANIEL J. GREENHOE

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 K.4 (previous theorem) demonstrates that in any inner product space $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$, the function $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x}\,|\,\boldsymbol{x}\rangle}$ is a norm. That is, $\|\boldsymbol{x}\|$ is the *norm induced by the inner product.* This norm is formally defined next.

Definition K.2. *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be an* INNER PRODUCT SPACE *(Definition K.1 page 271).*

DEF *The* ***norm induced by the inner product*** $\langle\triangle\,|\,\triangledown\rangle$ *is defined as*

$$\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x}\,|\,\boldsymbol{x}\rangle}$$

K.2.2 Inner products induced by norms

Theorem K.4 (page 276) demonstrates that if a *linear space* (Definition G.1 page 209) has an *inner product* (Definition K.1 page 271), then that inner product always induces a *norm* (Definition L.1 page 287), and the relationship between the two is simply $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}\,|\,\boldsymbol{x}\rangle}$ (Definition K.2 page 277). 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 K.5 page 277).

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 K.6 page 278).

3. Which norms can induce an inner product and which ones cannot? The answer is that norms that satisfy the *parallelogram law* (Theorem K.7 page 279) **can** induce an inner product; and the ones that don't, cannot (Theorem K.7 page 279).

Theorem K.5. [13] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* NORMED LINEAR SPACE *(Definition L.1 page 287).*

THM
$$\underbrace{\left.\begin{array}{l}\exists\ \langle\triangle\,|\,\triangledown\rangle \text{ and } (\cdot\,|\,\cdot) \quad \text{such that}\\ \|\boldsymbol{x}\|^2 = \langle \boldsymbol{x}\,|\,\boldsymbol{x}\rangle = (\boldsymbol{x}\,|\,\boldsymbol{x}) \quad \forall x\in X\end{array}\right\}}_{\textit{If a norm is induced by two inner products...}} \implies \underbrace{\langle \boldsymbol{x}\,|\,\boldsymbol{y}\rangle = (\boldsymbol{x}\,|\,\boldsymbol{y}) \quad \forall x,y\in X}_{\textit{...then those two inner products are equivalent.}}$$

[13] Aliprantis and Burkinshaw (1998), page 280, Bollobás (1999), page 132, Jordan and von Neumann (1935) page 721

COPYRIGHT © 2019 DANIEL J. GREENHOE

BY-NC-ND

✎PROOF:

$$
\begin{aligned}
2\langle x \mid y\rangle &= \big[\langle x \mid y\rangle + \langle y \mid x\rangle\big] + \big[\langle x \mid y\rangle - \langle y \mid x\rangle\big] \\
&= \big[\langle x \mid y\rangle + \langle y \mid x\rangle\big] + i\big[-i\langle x \mid y\rangle + i\langle y \mid x\rangle\big] \\
&= \big[\langle x \mid y\rangle + \langle y \mid x\rangle\big] + i\big[-\langle ix \mid y\rangle - \langle y \mid ix\rangle\big] \\
&= \left(\underbrace{\big[\langle x \mid y\rangle + \langle y \mid x\rangle + \langle x \mid x\rangle + \langle y \mid y\rangle\big]}_{\langle x+y \mid x+y\rangle} - \underbrace{\big[\langle x \mid x\rangle + \langle y \mid y\rangle\big]}_{\text{residue}}\right) \\
&\quad - i\left(\underbrace{\big[\langle ix \mid y\rangle + \langle y \mid ix\rangle + \langle ix \mid ix\rangle + \langle y \mid y\rangle\big]}_{\langle ix+y \mid ix+y\rangle} - \underbrace{\big[\langle ix \mid ix\rangle + \langle y \mid y\rangle\big]}_{\text{residue}}\right) \\
&= \big(\big[\langle x+y \mid x+y\rangle\big] - \big[\langle x \mid x\rangle + \langle y \mid y\rangle\big]\big) - i\big(\big[\langle ix+y \mid ix+y\rangle\big] - \big[\langle ix \mid ix\rangle + \langle y \mid y\rangle\big]\big) \\
&= \big(\big[(x+y \mid x+y)\big] - \big[(x \mid x) + (y \mid y)\big]\big) - i\big(\big[(ix+y \mid ix+y)\big] - \big[(ix \mid ix) + (y \mid y)\big]\big) \\
&= \left(\underbrace{\big[(x \mid y) + (y \mid x) + (x \mid x) + (y \mid y)\big]}_{(x+y \mid x+y)} - \underbrace{\big[(x \mid x) + (y \mid y)\big]}_{\text{residue}}\right) \\
&\quad - i\left(\underbrace{\big[(ix \mid y) + (y \mid ix) + (ix \mid ix) + (y \mid y)\big]}_{(ix+y \mid ix+y)} - \underbrace{\big[(ix \mid ix) + (y \mid y)\big]}_{\text{residue}}\right) \\
&= \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 K.6 (Polarization Identities). [14] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a linear space,* $\langle \triangle \mid \triangledown\rangle \in \mathbb{F}^{X\times X}$ *a function, and* $\|x\| \triangleq \sqrt{\langle x \mid x\rangle}$.

THM

$$
\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown\rangle\big) \text{ is an inner product space} \implies
$$

$$
\underbrace{4\langle x \mid y\rangle = \begin{cases} \|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2 & \text{for } \mathbb{F} = \mathbb{C} \quad \forall x,y\in X \\ \|x+y\|^2 - \|x-y\|^2 & \text{for } \mathbb{F} = \mathbb{R} \quad \forall x,y\in X \end{cases}}_{\text{inner product induced by norm}}
$$

✎PROOF:

1. These follow directly from properties of *bilinear functionals* (Theorem ?? page ??).

[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

COPYRIGHT © 2019 DANIEL J. GREENHOE

2. Alternative proof for $\mathbb{F} = \mathbb{C}$ case:

$$\begin{aligned}
&\|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2 \\
&= \underbrace{\|x\|^2 + \|y\|^2 + 2\Re\langle x \mid y\rangle}_{\langle x+y \mid x+y\rangle} - \underbrace{\left(\|x\|^2 + \|-y\|^2 + 2\Re\langle x \mid -y\rangle\right)}_{\langle x-y \mid x-y\rangle} \\
&+ \underbrace{i\left(\|x\|^2 + \|iy\|^2 + 2\Re\langle x \mid iy\rangle\right)}_{i\langle x+iy \mid x+iy\rangle} - \underbrace{i\left(\|x\|^2 + \|-iy\|^2 + 2\Re\langle x \mid -iy\rangle\right)}_{i\langle x-iy \mid x-iy\rangle} && \text{by Lemma K.1 page 275} \\
&= \underbrace{\|x\|^2 + \|y\|^2 + 2\Re\langle x \mid y\rangle}_{\langle x+y \mid x+y\rangle} - \underbrace{\left(\|x\|^2 + \|y\|^2 - 2\Re\langle x \mid y\rangle\right)}_{\langle x-y \mid x-y\rangle} \\
&+ \underbrace{i\left(\|x\|^2 + \|y\|^2 + 2\Re\langle x \mid iy\rangle\right)}_{i\langle x+iy \mid x+iy\rangle} - \underbrace{i\left(\|x\|^2 + \|y\|^2 - 2\Re\langle x \mid iy\rangle\right)}_{i\langle x-iy \mid x-iy\rangle} && \text{by Definition K.1 page 271} \\
&= 4\Re\langle x \mid y\rangle + 4i\Re\langle x \mid iy\rangle \\
&= \underbrace{2\left(\langle x \mid y\rangle + \langle x \mid y\rangle^*\right)}_{4\Re\langle x \mid y\rangle} + \underbrace{2i\left(\langle x \mid iy\rangle + \langle x \mid iy\rangle^*\right)}_{4i\Re\langle x \mid iy\rangle} \\
&= 2\left(\langle x \mid y\rangle + \langle x \mid y\rangle^*\right) + 2i\left(i^*\langle x \mid y\rangle + (i^{**})\langle x \mid y\rangle^*\right) \\
&= 2\left(\langle x \mid y\rangle + \langle x \mid y\rangle^*\right) + 2i\left(-i\langle x \mid y\rangle + i\langle x \mid y\rangle^*\right) && \text{by Definition K.1 page 271} \\
&= 2\langle x \mid y\rangle + 2\langle x \mid y\rangle^* + 2\langle x \mid y\rangle - 2\langle x \mid y\rangle^* \\
&= 4\langle x \mid y\rangle
\end{aligned}$$

✏

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.

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 K.7 (Parallelogram law). [15] *Let* $\Omega \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle \mid \triangledown\rangle\right)$ *and* $\|x\| \triangleq \sqrt{\langle x \mid x\rangle}$.

THM *Ω is an inner product space* $\iff$ $\underbrace{2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2}_{\text{PARALLELOGRAM LAW / VON NEUMANN-JORDAN CONDITION}}$ $\forall x,y \in \Omega$

✎PROOF:

1. Proof that $[\exists\langle x \mid y\rangle$ such that $\|x\|^2 = \langle x \mid x\rangle] \implies$ [parallelogram law is true]:

$$\begin{aligned}
\|x+y\|^2 + \|x-y\|^2 &= \left[\|x\|^2 + \|y\|^2 + 2\mathbf{R}_e\left[2\langle x \mid y\rangle\right]\right] + \left[\|x\|^2 + \|-y\|^2 + 2\mathbf{R}_e\left[2\langle x \mid -y\rangle\right]\right] \\
&\qquad \text{by Lemma K.1 page 275} \\
&= \left[\|x\|^2 + \|y\|^2 + 2\mathbf{R}_e\left[2\langle x \mid y\rangle\right]\right] + \left[\|x\|^2 + \|y\|^2 - 2\mathbf{R}_e\left[2\langle x \mid y\rangle\right]\right] \\
&= 2\|x\|^2 + 2\|y\|^2
\end{aligned}$$

[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

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

2. Proof that $[\exists \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle \quad \text{such that} \quad \|\boldsymbol{x}\|^2 = \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle\,] \Longleftarrow$ [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 $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\big)$. And if it is an inner product space, then by Theorem K.6 page 278 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 K.1 page 271.

(a) Proof that $\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle \geq 0$ (non-negative):

$$\begin{aligned}
4\,\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle &\triangleq \|\boldsymbol{x} + \boldsymbol{x}\|^2 - \cancelto{0}{\|\boldsymbol{x} - \boldsymbol{x}\|^2} + i\,\|\boldsymbol{x} + i\boldsymbol{x}\|^2 - i\,\|\boldsymbol{x} - i\boldsymbol{x}\|^2 && \text{by } \textit{Polarization Identity} \\
&= \|2\boldsymbol{x}\|^2 - \mathbb{0} + i\,\big(\,\|\boldsymbol{x} + i\boldsymbol{x}\|^2 - \|\boldsymbol{x} - i\boldsymbol{x}\|^2\,\big) && \text{by Definition L.1 page 287} \\
&= |2|^2\,\|\boldsymbol{x}\|^2 + i\,\big(\,\|\boldsymbol{x} + i\boldsymbol{x}\|^2 - |i|\,\|\boldsymbol{x} - i\boldsymbol{x}\|^2\,\big) \\
&= 4\,\|\boldsymbol{x}\|^2 + i\cancelto{0}{\big(\,\|\boldsymbol{x} + i\boldsymbol{x}\|^2 - \|i\boldsymbol{x} + \boldsymbol{x}\|^2\,\big)} && \text{by Definition L.1 page 287} \\
&= 4\,\|\boldsymbol{x}\|^2 \\
&\geq 0 && \text{by Definition L.1 page 287}
\end{aligned}$$

(b) Proof that $\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle = 0 \iff \boldsymbol{x} = \mathbb{0}$ (non-isotropic):

$$\begin{aligned}
4\,\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle &= 4\,\|\boldsymbol{x}\|^2 && \text{by result of part (a)} \\
&= 0 \qquad \Longleftrightarrow \qquad \boldsymbol{x} = \mathbb{0} && \text{by Definition L.1 page 287}
\end{aligned}$$

(c) Proof that $\langle \boldsymbol{x} + \boldsymbol{u} \mid \boldsymbol{y} \rangle = \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + \langle \boldsymbol{u} \mid \boldsymbol{y} \rangle$ (additive):[16]

$$\begin{aligned}
4\,\langle \boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{z} \rangle &= 8\left\langle \frac{\boldsymbol{x} + \boldsymbol{y}}{2} \mid \boldsymbol{z} \right\rangle && \text{by Definition K.1 page 271} \\
&= 2\left\|\frac{\boldsymbol{x} + \boldsymbol{y}}{2} + \boldsymbol{z}\right\|^2 - 2\left\|\frac{\boldsymbol{x} + \boldsymbol{y}}{2} - \boldsymbol{z}\right\|^2 \\
&\quad + 2i\left\|\frac{\boldsymbol{x} + \boldsymbol{y}}{2} + \boldsymbol{z}\right\|^2 - 2i\left\|\frac{\boldsymbol{x} + \boldsymbol{y}}{2} - i\boldsymbol{z}\right\|^2 && \text{by } \textit{Polarization Identity} \\
&= \left(2\left\|\frac{\boldsymbol{x} + \boldsymbol{y}}{2} + \boldsymbol{z}\right\|^2 + 2\left\|\frac{\boldsymbol{x} - \boldsymbol{y}}{2}\right\|^2\right) \\
&\quad - \left(2\left\|\frac{\boldsymbol{x} + \boldsymbol{y}}{2} - \boldsymbol{z}\right\|^2 + 2\left\|\frac{\boldsymbol{x} - \boldsymbol{y}}{2}\right\|^2\right) \\
&\quad + \left(2i\left\|\frac{\boldsymbol{x} + \boldsymbol{y}}{2} + \boldsymbol{z}\right\|^2 + 2i\left\|\frac{\boldsymbol{x} - \boldsymbol{y}}{2}\right\|^2\right) \\
&\quad - \left(2i\left\|\frac{\boldsymbol{x} + \boldsymbol{y}}{2} - i\boldsymbol{z}\right\|^2 + 2i\left\|\frac{\boldsymbol{x} - \boldsymbol{y}}{2}\right\|^2\right) \\
&= \big(\|\boldsymbol{x} + \boldsymbol{z}\|^2 + \|\boldsymbol{y} + \boldsymbol{z}\|^2\big) - \big(\|\boldsymbol{x} - \boldsymbol{z}\|^2 + \|\boldsymbol{y} - \boldsymbol{z}\|^2\big) \\
&\quad + \big(i\,\|\boldsymbol{x} + \boldsymbol{z}\|^2 + i\,\|\boldsymbol{y} + \boldsymbol{z}\|^2\big) - \big(i\,\|\boldsymbol{x} - i\boldsymbol{z}\|^2 + i\,\|\boldsymbol{y} - i\boldsymbol{z}\|^2\big) && \text{by parallelogram law} \\
&= \big(\|\boldsymbol{x} + \boldsymbol{z}\|^2 - \|\boldsymbol{x} - \boldsymbol{z}\|^2 + i\,\|\boldsymbol{x} + \boldsymbol{z}\|^2 - i\,\|\boldsymbol{x} - i\boldsymbol{z}\|^2\big) \\
&\quad + \big(\|\boldsymbol{y} + \boldsymbol{z}\|^2 - \|\boldsymbol{y} - \boldsymbol{z}\|^2 + i\,\|\boldsymbol{y} + \boldsymbol{z}\|^2 - i\,\|\boldsymbol{y} - i\boldsymbol{z}\|^2\big) \\
&= 4\,\langle \boldsymbol{x} \mid \boldsymbol{z} \rangle + 4\,\langle \boldsymbol{y} \mid \boldsymbol{z} \rangle && \text{by Polarization Identity}
\end{aligned}$$

[16] Aliprantis and Burkinshaw (1998), page 281

Copyright © 2019 Daniel J. Greenhoe

(d) Proof that $\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle^*$ (*conjugate symmetric*):

$$\begin{aligned}
4\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle &\triangleq \|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + i\,\|\boldsymbol{x}+i\boldsymbol{y}\|^2 - i\,\|\boldsymbol{x}-i\boldsymbol{y}\|^2 && \text{by Polarization Identity}\\
&= \|\boldsymbol{y}+\boldsymbol{x}\|^2 - \|\boldsymbol{y}-\boldsymbol{x}\|^2 + i\,\|i(\boldsymbol{y}-i\boldsymbol{x})\|^2 - i\,\|-i(\boldsymbol{y}+i\boldsymbol{x})\|^2 && \text{by Definition G.1 page 209}\\
&= \|\boldsymbol{y}+\boldsymbol{x}\|^2 - \|\boldsymbol{y}-\boldsymbol{x}\|^2 + i\,\|\boldsymbol{y}-i\boldsymbol{x}\|^2 - i\,\|\boldsymbol{y}+i\boldsymbol{x}\|^2 && \text{by Definition L.1 page 287}\\
&= \left(\|\boldsymbol{y}+\boldsymbol{x}\|^2 - \|\boldsymbol{y}-\boldsymbol{x}\|^2 - i\,\|\boldsymbol{y}-i\boldsymbol{x}\|^2 + i\,\|\boldsymbol{y}+i\boldsymbol{x}\|^2\right)^*\\
&= \left(\|\boldsymbol{y}+\boldsymbol{x}\|^2 - \|\boldsymbol{y}-\boldsymbol{x}\|^2 + i\,\|\boldsymbol{y}+i\boldsymbol{x}\|^2 - i\,\|\boldsymbol{y}-i\boldsymbol{x}\|^2\right)^*\\
&\triangleq 4\langle \boldsymbol{y} \mid \boldsymbol{x} \rangle^* && \text{by } \textit{Polarization Identity}
\end{aligned}$$

(e) Proof that $\langle \alpha\boldsymbol{x} \mid \boldsymbol{y} \rangle = \alpha\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$ (*homogeneous*): [17]

i. Proof that $\langle \alpha\boldsymbol{x} \mid \boldsymbol{y} \rangle$ is linear in α:

$$\begin{aligned}
0 &\le \big|\|\alpha\boldsymbol{x}+\boldsymbol{y}\| - \|\beta\boldsymbol{x}+\boldsymbol{y}\|\big| && \text{by Definition F.4 page 208}\\
&\le \big|\|(\alpha\boldsymbol{x}+\boldsymbol{y}) - (\beta\boldsymbol{x}+\boldsymbol{y})\|\big| && \text{by Theorem L.2 page 288}\\
&\le \big|\|(\alpha-\beta)\boldsymbol{x}\|\big|
\end{aligned}$$

This implies that as $a \to \beta$, $\|\alpha\boldsymbol{x}+\boldsymbol{y}\| \to \|\beta\boldsymbol{x}+\boldsymbol{y}\|$, which by definition implies that $\|\alpha\boldsymbol{x}+\boldsymbol{y}\|$ linear in α. And by the parallelogram law, $\langle \alpha\boldsymbol{x} \mid \boldsymbol{y} \rangle$ is also linear in α.

ii. Proof that $\langle n\boldsymbol{x} \mid \boldsymbol{y} \rangle = n\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle$ for $n \in \mathbb{Z}$ (integer case):

A. Proof for $n = \pm 1$:

$$\begin{aligned}
\langle n\boldsymbol{x} \mid \boldsymbol{y} \rangle &= \langle \pm 1\boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by } n = \pm 1 \text{ hypothesis}\\
&= \pm 1\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by definition of } \textit{inner product} \quad \text{(Definition K.1 page 271)}\\
&= n\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by } n = \pm 1 \text{ hypothesis}
\end{aligned}$$

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 $n = \pm 2$:

$$\begin{aligned}
\langle n\boldsymbol{x} \mid \boldsymbol{y} \rangle &= \langle \pm 2\boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by } n = \pm 1 \text{ hypothesis}\\
&= \langle \pm(\boldsymbol{x}+\boldsymbol{x}) \mid \boldsymbol{y} \rangle\\
&= \pm\langle \boldsymbol{x}+\boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by definition of } \textit{inner product} \quad \text{(Definition K.1 page 271)}\\
&= \pm(\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle) && \text{by additive property}\\
&= \pm 2\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle\\
&= n\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by } n = \pm 2 \text{ hypothesis}
\end{aligned}$$

D. Proof that [n case] $\implies$ [$n \pm 1$ case]:

$$\begin{aligned}
\langle (n\pm 1)\boldsymbol{x} \mid \boldsymbol{y} \rangle &= \langle n\boldsymbol{x} \pm 1\boldsymbol{x} \mid \boldsymbol{y} \rangle\\
&= \langle n\boldsymbol{x} \mid \boldsymbol{y} \rangle + \langle \pm 1\boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by additive property}\\
&= n\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle \pm 1\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{by left hypothesis}\\
&= (n\pm 1)\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle
\end{aligned}$$

[17] Aliprantis and Burkinshaw (1998), page 138

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):

$$\begin{aligned}
\frac{n}{m}\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle &= \frac{n}{m}\left\langle \frac{m}{m}\boldsymbol{x} \mid \boldsymbol{y}\right\rangle && \text{where } n, m \in \mathbb{Z} \text{ and } m \neq 0\\
&= \frac{nm}{m}\left\langle \frac{1}{m}\boldsymbol{x} \mid \boldsymbol{y}\right\rangle && \text{by previous result}\\
&= \frac{m}{m}\left\langle \frac{n}{m}\boldsymbol{x} \mid \boldsymbol{y}\right\rangle && \text{by previous result}\\
&= \left\langle \frac{n}{m}\boldsymbol{x} \mid \boldsymbol{y}\right\rangle
\end{aligned}$$

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 K.1. [18] The inner product has already been defined in Definition K.1 (page 271) 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 K.7 page 279) together with the *Polarization Identity* (Theorem K.6 page 278). 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.

Of the uncountably infinite number of $\ell^{\mathsf{p}}_{\mathbb{F}}$ norms, only the norm for $p = 2$ induces an inner product (Proposition K.1, next).

Proposition K.1. [19] *Let $\left\|(\!(x_n)\!)_{n\in\mathbb{Z}}\right\|_p$ be the $\ell^{\mathsf{p}}_{\mathbb{F}}$ norm of the sequence $(\!(x_n)\!)$ in the space $\ell^{\mathsf{p}}_{\mathbb{F}}$.*

PRP	$\left\|(\!(x_n)\!)\right\|_p$ *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 279):

$$\begin{aligned}
&\|\boldsymbol{x} + \boldsymbol{y}\|_p^2 + \|\boldsymbol{x} - \boldsymbol{y}\|_p^2\\
&= \|\boldsymbol{x} + \boldsymbol{y}\|_2^2 + \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 && \text{by right hypothesis}\\
&= \left(\sum_{n\in\mathbb{Z}} |x_n + y_n|^2\right)^{\frac{2}{2}} + \left(\sum_{n\in\mathbb{Z}} |x_n - y_n|^2\right)^{\frac{2}{2}} && \text{by definition of } \|\cdot\|_p\\
&= \sum_{n\in\mathbb{Z}} (x_n + y_n)(x_n + y_n)^* + \sum_{n\in\mathbb{Z}} (x_n - y_n)(x_n - y_n)^*\\
&= \sum_{n\in\mathbb{Z}} \left(|x_n|^2 + |y_n|^2 + 2\Re(x_n y_n)\right) + \sum_{n\in\mathbb{Z}} \left(|x_n|^2 + |y_n|^2 - 2\Re(x_n y_n)\right)\\
&= 2\sum_{n\in\mathbb{Z}} |x_n|^2 + 2\sum_{n\in\mathbb{Z}} |y_n|^2\\
&= 2\,\|\boldsymbol{x}\|_2^2 + 2\,\|\boldsymbol{y}\|_2^2 && \text{by definition of } \|\cdot\|_p\\
&= 2\,\|\boldsymbol{x}\|_p^2 + 2\,\|\boldsymbol{y}\|_p^2 && \text{by right hypothesis}\\
&\Longrightarrow \|\cdot\|_2 \text{ induces an inner product} && \text{by Theorem K.7 page 279}
\end{aligned}$$

[18] Loomis (1953), pages 23–24, Kubrusly (2001) page 317
[19] Kubrusly (2001) pages 318–319 ⟨Example 5B⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

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]

$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|_p^2 + \|\boldsymbol{x}-\boldsymbol{y}\|_p^2 &= \left(\sum_{n\in\mathbb{Z}} |x_n+y_n|^p\right)^{\frac{2}{p}} + \left(\sum_{n\in\mathbb{Z}} |x_n-y_n|^p\right)^{\frac{2}{p}} && \text{by definition of } \|\cdot\|_p \\
&= \left(|1+0|^p+|0+1|^p\right)^{\frac{2}{p}} + \left(|1-0|^p+|0-1|^p\right)^{\frac{2}{p}} && \text{by definitions of } \boldsymbol{x} \text{ and } \boldsymbol{y} \\
&= 2^{\frac{2}{p}} + 2^{\frac{2}{p}} \\
&= 2\cdot 2^{\frac{2}{p}} \\
2\,\|\boldsymbol{x}\|_p^2 + 2\,\|\boldsymbol{y}\|_p^2 &= 2\left(\sum_{n\in\mathbb{Z}} |x_n|^p\right)^{\frac{2}{p}} + 2\left(\sum_{n\in\mathbb{Z}} |y_n|^p\right)^{\frac{2}{p}} && \text{by definition of } \|\cdot\|_p \\
&= 2\left(|1|^p+|0|^p\right)^{\frac{2}{p}} + 2\left(|1|^p+|0|^p\right)^{\frac{2}{p}} && \text{by definitions of } \boldsymbol{x} \text{ and } \boldsymbol{y} \\
&= 2+2 \\
&= 4 \\
2\cdot 2^{\frac{2}{p}} = 4 &= \iff 2^{\frac{2}{p}} = 2 \\
&= \implies p=2
\end{aligned}$$

   (b) Proof that $2^{2/p}$ is monotonic decreasing in p (and so $p=2$ is the only solution):

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}p} 2^{\frac{2}{p}} &= \frac{\mathrm{d}}{\mathrm{d}p} e^{\ln 2^{\frac{2}{p}}} \\
&= \left(e^{\ln 2^{\frac{2}{p}}}\right) \frac{\mathrm{d}}{\mathrm{d}p} \ln 2^{\frac{2}{p}} \\
&= \left(2^{\frac{2}{p}}\right) \frac{\mathrm{d}}{\mathrm{d}p} (2\ln 2)\frac{1}{p} \\
&= \left(2^{\frac{2}{p}}\right) 2\ln 2\left(-\frac{1}{p^2}\right) \\
&< 0 \qquad \forall p \in (0,\infty)
\end{aligned}$$

✍

K.3 Orthogonality

Definition K.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 K.4. *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\right)$ *be an* INNER PRODUCT SPACE *(Definition K.1 page 271)*.

[20] http://groups.google.com/group/sci.math/msg/531b1173f08871e9

DEF

Two vectors $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *in* X *are* ***orthogonal*** *if*

$$\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = \begin{cases} 0 & \text{for } \boldsymbol{x} \neq \boldsymbol{y} \\ c \in \mathbb{F}\backslash 0 & \text{for } \boldsymbol{x} = \boldsymbol{y} \end{cases}$$

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* $c \in \mathbb{R}\backslash 0$.

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 K.8. [21] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\big)$ *be an* INNER PRODUCT SPACE.

THM

1.	$\boldsymbol{x} \perp \boldsymbol{x} \iff \boldsymbol{x} = \mathbb{0}$	$\forall_{\boldsymbol{x}\in X}$	
2.	$\boldsymbol{x} \perp \boldsymbol{y} \implies \alpha\boldsymbol{x} \perp \boldsymbol{y}$	$\forall_{\alpha\in\mathbb{R},\boldsymbol{x},\boldsymbol{y}\in X}$	(HOMOGENEOUS)
3.	$\boldsymbol{x} \perp \boldsymbol{y} \iff \boldsymbol{y} \perp \boldsymbol{x}$	$\forall_{\boldsymbol{x},\boldsymbol{y}\in X}$	(SYMMETRY)
4.	$\boldsymbol{x} \perp \boldsymbol{y}$ *and* $\boldsymbol{y} \perp \boldsymbol{z} \implies \boldsymbol{x} \perp (\boldsymbol{y}+\boldsymbol{z})$	$\forall_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in X}$	(ADDITIVE)
5.	$\exists\beta \in \mathbb{R}$ such that $\boldsymbol{x} \perp (\beta\boldsymbol{x}+\boldsymbol{y})$	$\forall_{\boldsymbol{x}\in X\backslash\mathbb{0},\, \boldsymbol{y}\in X}$	

Theorem K.9. *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\big)$ *be an* INNER PRODUCT SPACE.

THM

$$\left.\begin{array}{lll} 1. & \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = 0 & \textit{and} \\ 2. & \boldsymbol{x}+\boldsymbol{y} = \mathbb{0} & \end{array}\right\} \iff \left\{\begin{array}{lll} 1. & \boldsymbol{x} = \mathbb{0} & \textit{and} \\ 2. & \boldsymbol{y} = \mathbb{0} & \end{array}\right. \quad \forall_{\boldsymbol{x},\boldsymbol{y}\in X}$$

✎PROOF:

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

$$\begin{aligned}
0 &= \langle \mathbb{0} \mid \mathbb{0} \rangle && \text{by } \textit{non-isotropic} \text{ property of } \langle \triangle \mid \triangledown \rangle \text{ (Definition K.1 page 271)} \\
&= \langle \boldsymbol{x}+\boldsymbol{y} \mid \boldsymbol{x}+\boldsymbol{y} \rangle && \text{by left hypothesis 2} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x}+\boldsymbol{y} \rangle + \langle \boldsymbol{y} \mid \boldsymbol{x}+\boldsymbol{y} \rangle && \text{by } \textit{additive} \text{ property of } \langle \triangle \mid \triangledown \rangle \text{ (Definition K.1 page 271)} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{x} \rangle + \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle^* + \langle \boldsymbol{y} \mid \boldsymbol{y} \rangle && \text{by } \textit{conjugate symmetric} \text{ and } \textit{additive} \text{ properties of } \langle \triangle \mid \triangledown \rangle \\
&= \underbrace{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}_{\geq 0} + 0 + 0 + \underbrace{\langle \boldsymbol{y} \mid \boldsymbol{y} \rangle}_{\geq 0} && \text{by left hypothesis 1} \\
&\implies \boldsymbol{x} = \mathbb{0} \text{ and } \boldsymbol{y} = \mathbb{0} && \text{by } \textit{non-negative} \text{ and } \textit{non-isotropic} \text{ properties of } \langle \triangle \mid \triangledown \rangle
\end{aligned}$$

2. Proof that $\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = 0$:

$$\begin{aligned}
\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle &= \langle \mathbb{0} \mid \mathbb{0} \rangle && \text{by right hypotheses} \\
&= 0 && \text{by } \textit{non-isotropic} \text{ property of } \langle \triangle \mid \triangledown \rangle \text{ (Definition K.1 page 271)}
\end{aligned}$$

3. Proof that $\boldsymbol{x} + \boldsymbol{y} = \mathbb{0}$:

$$\begin{aligned}
\boldsymbol{x}+\boldsymbol{y} &= \mathbb{0}+\mathbb{0} && \text{by right hypotheses} \\
&= \mathbb{0}
\end{aligned}$$

✏

The *triangle inequality for vectors* in a *normed linear space* (Theorem L.1 page 287) demonstrates that $\left\| \sum_{n=1}^{N} \boldsymbol{x}_n \right\| \leq \sum_{n=1}^{N} \|\boldsymbol{x}_n\|$. The *Pythagorean Theorem* (next) demonstrates that this *inequality* becomes *equality* when the set $\{\boldsymbol{x}_n\}$ is *orthogonal*.

[21] James (1945), page 292, Drljević (1989) page 232

COPYRIGHT © 2019 DANIEL J. GREENHOE

Theorem K.10 (Pythagorean Theorem). [22] *Let* $\left\{ \boldsymbol{x}_n \in X \,\middle|\, {\scriptstyle n=1,2,\ldots,N} \right\}$ *be a set of vectors in an* INNER PRODUCT SPACE *(Definition K.1 page 271)* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \mid \triangledown \rangle\right)$ *and let* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x} \rangle}$ *(Definition K.2 page 277)*.

THM

$$\left\{ \boldsymbol{x}_n \right\} \text{ is ORTHOGONAL} \quad \Longleftrightarrow \quad \left\| \sum_{n=1}^{N} \boldsymbol{x}_n \right\|^2 = \sum_{n=1}^{N} \left\| \boldsymbol{x}_n \right\|^2 \qquad \forall N \in \mathbb{N}$$

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

$$\begin{aligned}
\left\| \sum_{n=1}^{N} \boldsymbol{x}_n \right\|^2 &= \left\langle \sum_{n=1}^{N} \boldsymbol{x}_n \,\middle|\, \sum_{m=1}^{N} \boldsymbol{x}_m \right\rangle && \text{by def. of } \|\cdot\| && \text{(Definition L.1 page 287)} \\
&= \sum_{n=1}^{N} \sum_{m=1}^{N} \langle \boldsymbol{x}_n \mid \boldsymbol{x}_m \rangle && \text{by def. of } \langle \triangle \mid \triangledown \rangle && \text{(Definition K.1 page 271)} \\
&= \sum_{n=1}^{N} \sum_{m=1}^{N} \langle \boldsymbol{x}_n \mid \boldsymbol{x}_m \rangle \, \bar{\delta}_{n-m} && \text{by left hypothesis} \\
&= \sum_{n=1}^{N} \langle \boldsymbol{x}_n \mid \boldsymbol{x}_n \rangle && \text{by def. of } \bar{\delta} && \text{(Definition K.3 page 283)} \\
&= \sum_{n=1}^{N} \left\| \boldsymbol{x}_n \right\|^2 && \text{by def. of } \|\cdot\| && \text{(Definition L.1 page 287)}
\end{aligned}$$

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

$$\begin{aligned}
&4 \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle \\
&= \|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + i\,\|\boldsymbol{x}+i\boldsymbol{y}\|^2 - i\,\|\boldsymbol{x}-i\boldsymbol{y}\|^2 && \text{by } \textit{polarization identity} \text{ (Theorem K.6 page 278)} \\
&= \left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) - \left(\|\boldsymbol{x}\|^2 + \|-\boldsymbol{y}\|^2\right) + i\left(\|\boldsymbol{x}\|^2 + \|i\boldsymbol{y}\|^2\right) - i\left(\|\boldsymbol{x}\|^2 + \|-i\boldsymbol{y}\|^2\right) && \text{by right hypothesis} \\
&= \left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) - \left(\|\boldsymbol{x}\|^2 + |-1|^2\,\|\boldsymbol{y}\|^2\right) + i\left(\|\boldsymbol{x}\|^2 + |i|^2\,\|\boldsymbol{y}\|^2\right) - i\left(\|\boldsymbol{x}\|^2 + |-i|^2\,\|\boldsymbol{y}\|^2\right) && \text{by definition of } \|\cdot\| \\
&= \left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) - \left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) + i\left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) - i\left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right) && \text{by def. of } |\cdot| \text{ (Definition F.4 page 208)} \\
&= 0
\end{aligned}$$

✏

[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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

APPENDIX L

NORMED LINEAR SPACES

L.1 Definition and basic results

Definition L.1. [1] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a* LINEAR SPACE *(Definition G.1 page 209) and* $|\cdot| \in \mathbb{R}^{\mathbb{F}}$ *the* ABSOLUTE VALUE *function (Definition F.4 page 208).*

DEF

A functional $\|\cdot\|$ *in* $\mathbb{R}^{X}$ *is a* ***norm*** *if*

1.	$\|\boldsymbol{x}\| \geq 0$	$\forall \boldsymbol{x}\in X$	(STRICTLY POSITIVE)	*and*
2.	$\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = \mathbb{0}$	$\forall \boldsymbol{x}\in X$	(NONDEGENERATE)	*and*
3.	$\|\alpha\boldsymbol{x}\| = \lvert\alpha\rvert\,\|\boldsymbol{x}\|$	$\forall \boldsymbol{x}\in X,\ \alpha\in\mathbb{C}$	(HOMOGENEOUS)	*and*
4.	$\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$	$\forall \boldsymbol{x},\boldsymbol{y}\in X$	(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 L.1 page 287) requires that any two vectors in a norm space be *sub-additive* (they satisfy the *triangle inequality* property) such that $\|\boldsymbol{x}_1 + \boldsymbol{x}_2\| \leq \|\boldsymbol{x}_1\| + \|\boldsymbol{x}_2\|$. Acutally, in **any** normed linear space, this property holds true for **any** finite number of vectors—not just two—such that $\|\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_N\| \leq \|\boldsymbol{x}_1\| + \|\boldsymbol{x}_2\| + \cdots + \|\boldsymbol{x}_N\|$ (next theorem).

Theorem L.1 (triangle inequality)**.** [2] *Let* $(\!|\boldsymbol{x}_n \in X|\!)_1^N$ *be an* N-TUPLE *(Definition R.1 page 369) of vectors in a* NORMED LINEAR SPACE $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$.

THM

$$\left\|\sum_{n=1}^{N} \boldsymbol{x}_n\right\| \leq \sum_{n=1}^{N} \|\boldsymbol{x}_n\| \qquad \forall N\in\mathbb{N},\ \boldsymbol{x}_n\in V$$

✎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:

$$\begin{aligned}\left\|\sum_{n=1}^{2} \boldsymbol{x}_n\right\| &= \left\|\sum_{n=1}^{2} \boldsymbol{x}_n\right\| \\ &= \|\boldsymbol{x}_1 + \boldsymbol{x}_2\| \\ &\le \|\boldsymbol{x}_1\| + \|\boldsymbol{x}_2\| && \text{by Definition L.1 page 287 (triangle inequality)} \\ &= \sum_{n=1}^{2} \|\boldsymbol{x}_n\|\end{aligned}$$

3. Proof that [N case] $\implies$ [$N+1$ case]:

$$\begin{aligned}\left\|\sum_{n=1}^{N+1} \boldsymbol{x}_n\right\| &= \left\|\sum_{n=1}^{N} \boldsymbol{x}_n + \boldsymbol{x}_{n+1}\right\| \\ &\le \left\|\sum_{n=1}^{N} \boldsymbol{x}_n\right\| + \|\boldsymbol{x}_{n+1}\| && \text{by Definition L.1 page 287 (triangle inequality)} \\ &\le \sum_{n=1}^{N} \|\boldsymbol{x}_n\| + \|\boldsymbol{x}_{n+1}\| && \text{by left hypothesis} \\ &= \sum_{n=1}^{N+1} \|\boldsymbol{x}_n\|\end{aligned}$$

✏

Theorem L.2 (Reverse Triangle Inequality). [3] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* NORMED LINEAR SPACE *(Definition L.1 page 287)*.

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:

$$\begin{aligned}\big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| &= \big|\|(\boldsymbol{x} - \boldsymbol{y}) + \boldsymbol{y}\| - \|\boldsymbol{y}\|\big| \\ &\le \big|\|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y}\| - \|\boldsymbol{y}\|\big| && \text{by Definition L.1 page 287} \\ &= \big|\|\boldsymbol{x} - \boldsymbol{y}\|\big| \\ &= \|\boldsymbol{x} - \boldsymbol{y}\| && \text{by Definition L.1 page 287}\end{aligned}$$

$$\begin{aligned}\|\boldsymbol{x} - \boldsymbol{y}\| &\le \|\boldsymbol{x} - \mathbb{0}\| + \|\mathbb{0} - \boldsymbol{y}\| && \text{by previous result with } \boldsymbol{u} = 0 \\ &= \|\boldsymbol{x}\| + |-1|\,\|\boldsymbol{y}\| && \text{by Definition L.1 page 287} \\ &= \|\boldsymbol{x}\| + \|\boldsymbol{y}\|\end{aligned}$$

✏

[3] Aliprantis and Burkinshaw (1998), page 218, Giles (2000) page 2, Banach (1922) page 136

COPYRIGHT © 2019 DANIEL J. GREENHOE

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 L.1. [4] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* NORMED LINEAR SPACE *(Definition L.1 page 287)*.

PRP

$$\|\boldsymbol{x}-\boldsymbol{y}\| \quad \leq \quad \|\boldsymbol{x}-\boldsymbol{u}\| + \|\boldsymbol{u}-\boldsymbol{y}\| \qquad \forall_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}\in X}$$

✎PROOF:

$$\begin{aligned}\|\boldsymbol{x}-\boldsymbol{y}\| &= \|(\boldsymbol{x}-\boldsymbol{u})+(\boldsymbol{u}-\boldsymbol{y})\| \\ &\leq \|\boldsymbol{x}-\boldsymbol{u}\| + \|\boldsymbol{u}-\boldsymbol{y}\| \qquad \text{by Definition L.1 page 287}\end{aligned}$$

✏

Example L.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 208) $|\cdot| \in \mathbb{R}^{\mathbb{R}}$ is a *norm*.

Example L.2 (l_p norms). Let $(\!(x_n)\!)_{n\in\mathbb{Z}}$ be a *sequence* (Definition R.1 page 369) of real numbers. An uncountably infinite number of norms is provided by the $\ell^{\mathsf{p}}_{\mathbb{F}}$ norms $\|(\!(x_n)\!)\|_p$:

EX $$\|(\!(x_n)\!)\|_p \triangleq \left(\sum_{n\in\mathbb{Z}} |x_n|^p\right)^{\frac{1}{p}} \qquad \text{is a norm for } p\in[1\,:\,\infty]$$

L.2 Relationship between metrics and norms

L.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 **??** page **??**). 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,

$$\text{normed linear space} \quad\Longrightarrow\quad \text{metric space} \quad\Longrightarrow\quad \text{topological space.}$$

[4] Aliprantis and Burkinshaw (1998), page 218
[5] Giles (1987), page 3

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Theorem L.3. [6] *Let* $\mathsf{d} \in \mathbb{R}^{X \times X}$ *be a function on a* REAL *normed linear space* $\big(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \|\cdot\|\big)$. *Let* $\mathsf{B}(\boldsymbol{x}, r) \triangleq \{\boldsymbol{y} \in \boldsymbol{X} \,|\, \|\boldsymbol{y} - \boldsymbol{x}\| < r\}$ *be the* OPEN BALL *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 (not included in this text) the definition of *norm* (Definition L.1 page 287). ✏

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 L.2. [7] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* NORMED LINEAR SPACE *(Definition L.1 page 287)*.

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 L.1 page 290). 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 L.2 page 288) shows this to be true when $\mathsf{f}(\cdot) \triangleq \|\cdot\|$.

Corollary L.1. [8] *Let* $\boldsymbol{\Omega} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* NORMED LINEAR SPACE *(Definition L.1 page 287)*.

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 L.3 page 290).
2. *Continuity* in a metric space.
3. The *Reverse Triangle Inequality* (Theorem L.2 page 288).

✏

Theorem L.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.

Theorem L.4. [9] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d}\big)$ *be a* METRIC LINEAR SPACE *(Definition* **??** *page* **??***). Let* B *be the* OPEN BALL $\mathsf{B}(\boldsymbol{p}, r) \triangleq \{\boldsymbol{x} \in \boldsymbol{X} \,|\, \mathsf{d}(\boldsymbol{p}, \boldsymbol{x}) < r\}$ *(open ball with respect to metric* d *centered at point* $\boldsymbol{p}$ *and with radius* r*).*

THM

$$\left.\begin{array}{l} \exists \|\cdot\| \in \mathbb{R}^{X} \quad \text{such that} \\ \underbrace{\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{x}\|}_{\mathsf{d}\ \textit{is generated by a norm}} \end{array}\right\} \Longrightarrow \left\{\begin{array}{ll} 1. & \mathsf{B}(\boldsymbol{x}, r) = \boldsymbol{x} + B(0, r) \\ 2. & \mathsf{B}(\mathbb{0}, r) = r\,\mathsf{B}(\mathbb{0}, 1) \\ 3. & \mathsf{B}(\boldsymbol{x}, r) \ \textit{is}\ \text{CONVEX} \\ 4. & \boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) \iff -\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) \quad \text{(SYMMETRIC)} \end{array}\right.$$

[6] Michel and Herget (1993), page 344, Banach (1932a) page 53
[7] Giles (2000) page 1 ⟨1.1 Definition⟩
[8] Giles (2000) page 2
[9] Giles (2000) page 2 ⟨1.2 Remarks⟩, Giles (1987) pages 22–26 ⟨2.4 Theorem, 2.11 Theorem⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

1. Proof that $\mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{v}\boldsymbol{z}) = \mathsf{d}(\boldsymbol{x}, \boldsymbol{y})$ (invariant):

$$\begin{aligned}
\mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{v}\boldsymbol{z}) &= \|(\boldsymbol{y}+\boldsymbol{v}\boldsymbol{z}) - (\boldsymbol{x}+\boldsymbol{z})\| && \text{by left hypothesis}\\
&= \|\boldsymbol{y}-\boldsymbol{x}\| \\
&= \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) && \text{by left hypothesis}
\end{aligned}$$

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

$$\begin{aligned}
\mathsf{B}(\mathbb{0}, r) &= \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{y}) < r\} && \text{by definition of open ball } B\\
&= \left\{\boldsymbol{y} \in \boldsymbol{X} | \frac{1}{r}\mathsf{d}(\mathbb{0}, \boldsymbol{y}) < 1\right\} \\
&= \left\{\boldsymbol{y} \in \boldsymbol{X} | \frac{1}{r}\,\|\boldsymbol{y}-\mathbb{0}\| < 1\right\} && \text{by left hypothesis}\\
&= \left\{\boldsymbol{y} \in \boldsymbol{X} | \left\|\frac{1}{r}\boldsymbol{y} - \frac{1}{r}\mathbb{0}\right\| < 1\right\} && \text{by homogeneous property of } \|\cdot\| \text{ page 287}\\
&= \left\{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}\left(\frac{1}{r}\mathbb{0}, \frac{1}{r}\boldsymbol{y}\right) < 1\right\} && \text{by left hypothesis}\\
&= \{r\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < 1\} && \text{let } \boldsymbol{u} \triangleq \frac{1}{r}\boldsymbol{y}\\
&= r\ \{\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < 1\} \\
&= r\,B(\mathbb{0},\ 1) && \text{by definition of open ball } B
\end{aligned}$$

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 than r.

$$\begin{aligned}
\mathsf{d}(\boldsymbol{p}, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \|\boldsymbol{p} - \lambda\boldsymbol{x} - (1-\lambda)\boldsymbol{y}\| && \text{by left hypothesis}\\
&= \left\|\underbrace{\lambda\boldsymbol{p} + (1-\lambda)\boldsymbol{p}}_{\boldsymbol{p}} - \lambda\boldsymbol{x} - (1-\lambda)\boldsymbol{y}\right\| \\
&= \|\lambda\boldsymbol{p} - \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{p} - (1-\lambda)\boldsymbol{y}\| \\
&\le \|\lambda\boldsymbol{p} - \lambda\boldsymbol{x}\| + \|(1-\lambda)\boldsymbol{p} - (1-\lambda)\boldsymbol{y}\| && \text{by subadditivity property of } \|\cdot\| \text{ page 287}\\
&= |\lambda|\ \|\boldsymbol{p}-\boldsymbol{x}\| + |1-\lambda|\ \|\boldsymbol{p}-\boldsymbol{y}\| && \text{by homogeneous property of } \|\cdot\| \text{ page 287}\\
&= \lambda\ \|\boldsymbol{p}-\boldsymbol{x}\| + (1-\lambda)\ \|\boldsymbol{p}-\boldsymbol{y}\| && \text{because } 0 \le \lambda \le 1\\
&\le \lambda r + (1-\lambda)r && \text{because } \boldsymbol{x}, \boldsymbol{y} \text{ are in the ball } \mathsf{B}(\boldsymbol{p}, r)\\
&= r
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

5. Proof that $\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) \iff -\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r)$ (symmetric):

$$
\begin{aligned}
\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) &\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{y}) < r\} && \text{by definition of open ball } B \\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \, \|\boldsymbol{y} - \mathbb{0}\| < r\} && \text{by left hypothesis} \\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \, \|\boldsymbol{y}\| < r\} \\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \, \|(-1)(-\boldsymbol{y})\| < r\} \\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \, |-1| \, \|-\boldsymbol{y}\| < r\} && \text{by homogeneous property of } \|\cdot\| \text{ page 287} \\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \, \|-\boldsymbol{y} - \mathbb{0}\| < r\} \\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, -\boldsymbol{y}) < r\} && \text{by left hypothesis} \\
&\iff \boldsymbol{x} \in \{-\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\} && \text{let } \boldsymbol{u} \triangleq -\boldsymbol{y} \\
&\iff \boldsymbol{x} \in (-\{\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\}) \\
&\iff \boldsymbol{x} \in (-\mathsf{B}(\mathbb{0}, r)) \\
&\iff -\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r)
\end{aligned}
$$

✏

Theorem L.4 (page 290) demonstrates that if a metric d in a metric space $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d}\big)$ 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.

	taxi-cab metric	Euclidean Metric	sup metric	parabolic metric
$\mathbb{R}^0$				
$\mathbb{R}^1$				
$\mathbb{R}^2$				

Figure L.1: Open balls in $(\mathbb{R}^0, \mathsf{d}_n)$, $(\mathbb{R}, \mathsf{d}_n)$, $(\mathbb{R}^2, \mathsf{d}_n)$, and $(\mathbb{R}^3, \mathsf{d}_n)$.

L.2.2 Norms generated by metrics

Every normed linear space is also a metric linear space (Theorem L.3 page 290). 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 L.5 page 293.

Lemma L.1. [10] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d}\big)$ *be a* METRIC LINEAR SPACE. *Let* $\|\boldsymbol{x}\| \triangleq \mathsf{d}(\boldsymbol{x}, \mathbb{0}) \; \forall \boldsymbol{x} \in X$.

[10] Oikhberg and Rosenthal (2007), page 599

Copyright © 2019 Daniel J. Greenhoe

LEM

$$\underbrace{\mathsf{d}(\boldsymbol{x}+\boldsymbol{z},\boldsymbol{y}+\boldsymbol{z})=\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in X}_{\text{TRANSLATION INVARIANT}} \implies \begin{cases} 1. & \|\boldsymbol{x}\| = \|-\boldsymbol{x}\| & \forall \boldsymbol{x}\in X & \textit{and} \\ 2. & \|\boldsymbol{x}\| = 0 \iff \boldsymbol{x}=0 & \forall \boldsymbol{x}\in X & \textit{and} \\ 3. & \|\boldsymbol{x}+\boldsymbol{y}\| \le \|\boldsymbol{x}\|+\|\boldsymbol{y}\| & \forall \boldsymbol{x},\boldsymbol{y}\in X & \end{cases}$$

✎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 $\|\boldsymbol{x}\| = 0 \implies \boldsymbol{x} = 0$:

$$\begin{aligned} 0 &= \|\boldsymbol{x}\| && \text{by left hypothesis} \\ &= \mathsf{d}(\boldsymbol{x},\mathbb{0}) && \text{by definition of } \|\cdot\| \\ &= \mathsf{d}(\boldsymbol{x},\mathbb{0}) && \text{by definition of } \|\cdot\| \\ &\implies \boldsymbol{x} = \mathbb{0} && \text{by property of metrics} \end{aligned}$$

2b. Proof that $\|\boldsymbol{x}\| = 0 \impliedby \boldsymbol{x} = 0$:

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

3. Proof that $\|\boldsymbol{x}+\boldsymbol{y}\| \le \|\boldsymbol{x}\|+\|\boldsymbol{y}\|$:

$$\begin{aligned} \|\boldsymbol{x}+\boldsymbol{y}\| &= \mathsf{d}(\boldsymbol{x}+\boldsymbol{y},\mathbb{0}) && \text{by definition of } \|\cdot\| \\ &= \mathsf{d}(\boldsymbol{x}+\boldsymbol{y}-\boldsymbol{y},\mathbb{0}-\boldsymbol{y}) && \text{by translation invariance hypothesis} \\ &= \mathsf{d}(\boldsymbol{x},-\boldsymbol{y}) \\ &\le \mathsf{d}(\boldsymbol{x},\mathbb{0})+\mathsf{d}(\mathbb{0},\boldsymbol{y}) && \text{by property of metrics} \\ &= \mathsf{d}(\boldsymbol{x},\mathbb{0})+\mathsf{d}(\boldsymbol{y},\mathbb{0}) && \text{by property of metrics} \\ &= \|\boldsymbol{x}\|+\|\boldsymbol{y}\| && \text{by definition of } \|\cdot\| \end{aligned}$$

✏

Theorem L.5. [11] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a* LINEAR SPACE. *Let* $\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) \triangleq \|\boldsymbol{x}-\boldsymbol{y}\| \ \forall \boldsymbol{x},\boldsymbol{y}\in X$.

THM

$$\left.\begin{array}{lll} 1. & \mathsf{d}(\boldsymbol{x}+\boldsymbol{z},\boldsymbol{y}+\boldsymbol{z})=\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in X & \text{(TRANSLATION INVARIANT)} \quad \textit{and} \\ 2. & \mathsf{d}(\alpha\boldsymbol{x},\alpha\boldsymbol{y})=|\alpha|\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y}\in X, \alpha\in\mathbb{F} & \text{(HOMOGENEOUS)} \end{array}\right\} \iff \|\cdot\| \textit{ is a } \text{NORM}$$

✎PROOF:

1. Proof of $\implies$ 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 L.1 (page 292).
   (c) Proof that $\|\cdot\|$ is *homogeneous*: This follows from the second left hypothesis.

[11] Bollobás (1999), page 21

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

(d) Proof that $\|\cdot\|$ satisfies the *triangle-inequality*: This follows directly from Lemma L.1 (page 292).

2. Proof of $\Longleftarrow$ assertion:

$$\begin{aligned}
\mathsf{d}(\boldsymbol{x}+\boldsymbol{z},\boldsymbol{y}+\boldsymbol{z}) &= \|(\boldsymbol{x}+\boldsymbol{z})-(\boldsymbol{y}+\boldsymbol{z})\| && \text{by definition of } \mathsf{d}\\
&= \|\boldsymbol{x}-\boldsymbol{y}\| \\
&= \mathsf{d}(\boldsymbol{x},\boldsymbol{y}) && \text{by definition of } \mathsf{d}\\
\mathsf{d}(\alpha\boldsymbol{x},\alpha\boldsymbol{y}) &= \|(\alpha\boldsymbol{x})-(\alpha\boldsymbol{y})\| && \text{by definition of } \mathsf{d}\\
&= \|\alpha(\boldsymbol{x}-\boldsymbol{y})\| \\
&= |\alpha|\,\|\boldsymbol{x}-\boldsymbol{y}\| && \text{by definition of } \|\cdot\| \text{ page } 287\\
&= |\alpha|\mathsf{d}(\boldsymbol{x},\boldsymbol{y}) && \text{by definition of } \mathsf{d}
\end{aligned}$$

✍

L.3 Orthogonality on normed linear spaces

Traditionally, *orthogonality* (Definition K.4 page 283) is a property defined in *inner product spaces* (Definition K.1 page 271) However, the concept of orthogonality can be extended to *normed linear spaces* (Definition L.1 page 287). Here are some examples:

① *Isosceles orthogonality*: Definition L.3 page 294
② *Pythagorean orthogonality*: Definition L.4 page 296
③ *Birkhoff orthogonality*: Definition L.5 page 296

Isosceles orthogonality (Definition L.3 page 294) 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 L.3. [12] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* NORMED LINEAR SPACE *(Definition L.1 page 287)*.

DEF

Two vectors $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are* ***orthogonal in the sense of James*** *if*

$$\|\boldsymbol{x}+\boldsymbol{y}\| = \|\boldsymbol{x}-\boldsymbol{y}\|\,.$$

This property is also called ***isosceles orthogonality*** *or* ***James orthogonality***.

Theorem L.6. *Let* $\big(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be an* INNER-PRODUCT SPACE *(Definition K.1 page 271) with induced norm* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle\boldsymbol{x}\,|\,\boldsymbol{x}\rangle}$, ISOSCELES ORTHOGONALITY *(Definition L.3 page 294) relation* $\textcircled{\perp}$, *and inner-product relation* ORTHOGONALITY *(Definition K.4 page 283) relation* $\perp$.

THM

$$\underbrace{\boldsymbol{x}\,\textcircled{\perp}\,\boldsymbol{y}}_{\text{orthogonal in the sense of James}} \quad\Longleftrightarrow\quad \underbrace{\boldsymbol{x}\perp\boldsymbol{y}}_{\text{orthogonal in the sense of inner-product space}}$$

[12] James (1945) page 292 ⟨DEFINITION 2.1⟩, Amir (1986) page 24, Dunford and Schwartz (1957), page 93

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

1. Proof that $x \textcircled{\perp} y \quad \Longrightarrow \quad x \perp y$:

$$
\begin{aligned}
&4\langle x \mid y\rangle \\
&= \underbrace{\|x+y\|^2 - \|x-y\|^2}_{0 \text{ by } x \textcircled{\perp} y \text{ hypothesis}} + i\,\|x+iy\|^2 - i\,\|x-iy\|^2 && \text{by } \textit{polarization identity} \text{ (Theorem K.6 page 278)} \\
&= 0 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2 && \text{by } x \textcircled{\perp} y \text{ hypothesis} \\
&= i\left[\|x\|^2 + \|iy\|^2 + 2\Re\langle x \mid iy\rangle\right] \\
&\qquad - i\left[\|x\|^2 + \|-iy\|^2 + 2\Re\langle x \mid -iy\rangle\right] && \text{by } \textit{Polar Identity} \text{ (Lemma K.1 page 275)} \\
&= i\left[\|x\|^2 + \|y\|^2 + 2\Re\langle x \mid iy\rangle\right] \\
&\qquad - i\left[\|x\|^2 + \|y\|^2 + 2\Re\langle x \mid -iy\rangle\right] && \text{by Definition L.1 page 287 and Definition K.1 page 271} \\
&= 4i\Re\langle x \mid iy\rangle \\
&= 4i\Re\left[i^*\langle x \mid y\rangle\right] \\
&= 0 && \text{because inner-product space is } \textit{real} \ (\mathbb{F} = \mathbb{R})
\end{aligned}
$$

2. Proof that $x \textcircled{\perp} y \quad \Longleftarrow \quad x \perp y$:

$$
\begin{aligned}
\|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\Re\langle x \mid y\rangle && \text{by } \textit{Polar Identity} \text{ (Lemma K.1 page 275)} \\
&= \|x\|^2 + \|y\|^2 + 0 && \text{by } x \perp y \text{ hypothesis} \\
&= \|x\|^2 + \|y\|^2 - 2\Re\cancelto{0 \text{ when } x \perp y}{\langle x \mid y\rangle} && \text{by } x \perp y \text{ hypothesis} \\
&= \|x\|^2 + \|-y\|^2 + 2\Re\langle x \mid -y\rangle \\
&= \|x-y\|^2 && \text{by } \textit{Polar Identity} \text{ (Lemma K.1 page 275)}
\end{aligned}
$$

✏

Theorem L.7. [13] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\right)$ *be a normed linear space and with* ISOSCELES ORTHOGONALITY *(Definition L.3 page 294) relation* $\textcircled{\perp}$.

THM $\quad x \textcircled{\perp} y \quad \Longleftrightarrow \quad y \textcircled{\perp} x \quad \Longleftrightarrow \quad \alpha x \textcircled{\perp} \alpha y \quad \forall \alpha \in \mathbb{F}$

✎PROOF:

$$
\begin{aligned}
x \textcircled{\perp} y &\Longrightarrow \|x+y\| &&= \|x-y\| && \text{by Definition L.3 page 294} \\
&\Longrightarrow \|x+y\| &&= |-1|\ \|x-y\| \\
&\Longrightarrow \|x+y\| &&= \|-(x-y)\| && \text{by Definition L.1 page 287} \\
&\Longrightarrow \|y+x\| &&= \|y-x\| && \text{by Definition G.1 page 209} \\
&\Longrightarrow y \textcircled{\perp} x && && \text{by Definition L.3 page 294}
\end{aligned}
$$

$$
\begin{aligned}
y \textcircled{\perp} x &\Longrightarrow \|y+x\| &&= \|y-x\| && \text{by Definition L.3 page 294} \\
&\Longrightarrow |\alpha|\ \|y+x\| &&= |\alpha|\ \|y-x\| \\
&\Longrightarrow \|\alpha(y+x)\| &&= \|\alpha(y-x)\| && \text{by Definition L.1 page 287} \\
&\Longrightarrow \|\alpha y + \alpha x\| &&= \|\alpha y - \alpha x\| \\
&\Longrightarrow \|\alpha x + \alpha y\| &&= \|-(\alpha x - \alpha y)\| && \text{by Definition G.1 page 209}
\end{aligned}
$$

[13] Amir (1986) page 24

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

$$\begin{aligned}
&\Longrightarrow \|\alpha\boldsymbol{x}+\alpha\boldsymbol{y}\| &&= |-1|\ \|\alpha\boldsymbol{x}-\alpha\boldsymbol{y}\| && \text{by Definition L.1 page 287}\\
&\Longrightarrow \|\alpha\boldsymbol{x}+\alpha\boldsymbol{y}\| &&= \|\alpha\boldsymbol{x}-\alpha\boldsymbol{y}\| && \text{by Definition F.4 page 208}\\
&\Longrightarrow \alpha\boldsymbol{x} \textcircled{\perp} \alpha\boldsymbol{y} && && \text{by Definition L.3 page 294}
\end{aligned}$$

$$\begin{aligned}
\alpha\boldsymbol{x} \textcircled{\perp} \alpha\boldsymbol{y} &\Longrightarrow \|\alpha\boldsymbol{x}+\alpha\boldsymbol{y}\| &&= \|\alpha\boldsymbol{x}-\alpha\boldsymbol{y}\| && \text{by Definition L.3 page 294}\\
&\Longrightarrow \|\alpha(\boldsymbol{x}+\boldsymbol{y})\| &&= \|\alpha(\boldsymbol{x}-\boldsymbol{y})\| && \text{by Definition G.1 page 209}\\
&\Longrightarrow |\alpha|\ \|\boldsymbol{x}+\boldsymbol{y}\| &&= |\alpha|\ \|\boldsymbol{x}-\boldsymbol{y}\| && \text{by Definition L.1 page 287}\\
&\Longrightarrow \|\boldsymbol{x}+\boldsymbol{y}\| &&= \|\boldsymbol{x}-\boldsymbol{y}\| && \text{by Definition L.1 page 287}\\
&\Longrightarrow \boldsymbol{x} \textcircled{\perp} \boldsymbol{y} && && \text{by Definition L.3 page 294}
\end{aligned}$$

✏

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 K.10 page 285), $a^2+b^2=c^2$. 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 L.4. [14] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* NORMED LINEAR SPACE *(Definition L.1 page 287).*

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 L.8. [15] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be an* INNER-PRODUCT SPACE *(Definition K.1 page 271) with induced norm* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle\boldsymbol{x}\,|\,\boldsymbol{x}\rangle}$, PYTHAGOREAN ORTHOGONALITY *(Definition L.4 page 296) relation* $\textcircled{\perp}$, *and inner-product relation* ORTHOGONALITY *(Definition K.4 page 283) relation* $\perp$.

THM

$$\underbrace{\boldsymbol{x} \textcircled{\perp} \boldsymbol{y}}_{\text{orthogonal in the Pythagorean sense}} \quad\Longleftrightarrow\quad \underbrace{\boldsymbol{x} \perp \boldsymbol{y}}_{\text{orthogonal in the sense of inner-product space}}$$

Besides *isosceles orthogonality* (Definition L.3 page 294), orthogonality in normed linear spaces can be defined using *Birkhoff orthogonality*, as defined in Definition L.5 (next) and illustrated to the left.

Definition L.5. [16] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be a* NORMED LINEAR SPACE *(Definition L.1 page 287).*

DEF

Two vectors $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are* ***orthogonal in the sense of Birkhoff*** *if*

$$\|\boldsymbol{x}\| \le \|\boldsymbol{x}+\alpha\boldsymbol{y}\| \quad \forall\alpha\in\mathbb{F}.$$

This relationship is also called ***Birkhoff orthogonality****.*

Theorem L.9. *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle\triangle\,|\,\triangledown\rangle\big)$ *be an* INNER-PRODUCT SPACE *(Definition K.1 page 271) with induced norm* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle\boldsymbol{x}\,|\,\boldsymbol{x}\rangle}$, BIRKHOFF ORTHOGONALITY *relation* $\textcircled{\perp}$ *(Definition L.5 page 296), and inner-product relation* ORTHOGONALITY *relation* $\perp$ *(Definition K.4 page 283).*

THM

$$\underbrace{\boldsymbol{x} \textcircled{\perp} \boldsymbol{y}}_{\text{orthogonal in the sense of Birkhoff}} \quad\Longleftrightarrow\quad \underbrace{\boldsymbol{x} \perp \boldsymbol{y}}_{\text{orthogonal in the sense of inner-product space}}$$

[14] James (1945) page 292 ⟨DEFINITION 2.2⟩, Amir (1986) page 57, Drljević (1989) page 232
[15] Amir (1986) page 57
[16] Amir (1986) page 33, Dunford and Schwartz (1957), page 93, James (1947) page 265

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

APPENDIX M

INTERVALS AND CONVEXITY

M.1 Intervals

In the real number system, for $a \leq b$, the *interval* $[a : b]$ is the set a and b and all the numbers inbetween, as in $[a : b] \triangleq \{x \in \mathbb{R} | a \leq x \leq b\}$. This concept can be easily generalized:

- In an **ordered set**, if two elements x and y are *comparable* and $x \leq y$, then we say that x and y and all the elements inbetween, as determined by the ordering relation $\leq$, are the interval $[a : b]$.
- In a **lattice**, the concept of the *interval* can be generalized even further. In an arbitrary ordered set, the interval $[x : y]$ of item (M.1) is restricted to the case in which x and y are *comparable*. This restriction can be lifted (Definition M.2 page 297) with the additional structure of upper and lower bounds provided by lattices.
- A **metric space** in general has no *order relation* $\leq$. But intervals can still be defined (Definition M.4 page 298) in a metric space in terms of the *triangle inequality*.
- A **linear space** (Definition G.1 page 209) over a real or complex field in general has no *order relation* that compares *vectors* in the space, but the standard order relation $\leq$ for real numbers $\mathbb{R}$ can still be used (Definition M.5 page 298) to define an interval in a linear space.

Definition M.1 (intervals on ordered sets). [1] *Let* $(X, \leq)$ *be an* ORDERED SET.

DEF					
	The set	$[x : y]$	$\triangleq \{z \in X \mid x \leq z \leq y\}$	*is called a* ***closed interval***	*and*
	The set	$(x : y]$	$\triangleq \{z \in X \mid x < z \leq y\}$	*is called a* ***half-open interval***	*and*
	The set	$[x : y)$	$\triangleq \{z \in X \mid x \leq z < y\}$	*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 M.2 (intervals on lattices). [2] *Let* $(X, \vee, \wedge; \leq)$ *be a* LATTICE.

DEF				
	The set $[x : y]$	$\triangleq$	$\{z \in X \mid x \wedge y \leq z \leq x \vee y\}$	*is called a* ***closed interval.***
	The set $(x : y]$	$\triangleq$	$\{z \in X \mid x \wedge y < z \leq x \vee y\}$	*is called a* ***half-open interval.***
	The set $[x : y)$	$\triangleq$	$\{z \in X \mid x \wedge y \leq z < x \vee y\}$	*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.***

When x and y *are* comparable and $x \leq y$, then Definition M.2 (previous) simplifies to item (M.1)

[1] Apostol (1975) page 4, Ore (1935) page 409

[2] Duthie (1942) page 2, Ore (1935) page 425 ⟨*quotient structures*⟩

(page 297).

Definition M.3. [3] *Let* $\boldsymbol{L} \triangleq (X, \vee, \wedge; \leq)$ *be a* LATTICE *with dual* $\boldsymbol{L}^*$. *Let* $[x : y]$ *be a* CLOSED INTERVAL *(Definition M.2 page 297) on set* X. *The sublattices* $\boldsymbol{L}[x : y]$ *and* $\boldsymbol{L}^*[x : y]$ *are defined as follows:*

DEF
$$\boldsymbol{L}[x : y] \triangleq \{z \in \boldsymbol{L} | z \in [x : y]\} \qquad \forall x,y \in X$$
$$\boldsymbol{L}^*[x : y] \triangleq \left\{z \in \boldsymbol{L}^* | z \in [x : y]\right\} \qquad \forall x,y \in X$$

Definition M.4. [4]

DEF
In a METRIC SPACE (X, d),
the set $[a : b]$ *is the* ***closed interval*** *from* x *to* y *and is defined as*
$$[x : y] \triangleq \{z \in X | \mathsf{d}(x, z) + \mathsf{d}(z, y) = \mathsf{d}(x, y)\}.$$
An element $z \in X$ *is* ***geodesically between*** x *and* y *if* $z \in [x : y]$.

Definition M.5. [5]

DEF
In a LINEAR SPACE $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *(Definition G.1 page 209),*
$$\begin{aligned}
[\boldsymbol{x} : \boldsymbol{y}] &\triangleq \{\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} = \boldsymbol{z} | 0 \leq \lambda \leq 1\} && \textit{is called a } \textbf{\textit{closed interval}} && \textit{and}\\
(\boldsymbol{x} : \boldsymbol{y}] &\triangleq \{\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} = \boldsymbol{z} | 0 < \lambda \leq 1\} && \textit{is called a } \textbf{\textit{half-open interval}} && \textit{and}\\
[\boldsymbol{x} : \boldsymbol{y}) &\triangleq \{\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} = \boldsymbol{z} | 0 \leq \lambda < 1\} && \textit{is called a } \textbf{\textit{half-open interval}} && \textit{and}\\
(\boldsymbol{x} : \boldsymbol{y}) &\triangleq \{\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} = \boldsymbol{z} | 0 < \lambda < 1\} && \textit{is called an } \textbf{\textit{open interval}}.
\end{aligned}$$

M.2 Convex sets

Using the concept of the *interval* (previous section), we can define the *convex set* (next definition).

Definition M.6. [6] *Let* X *be a* SET *in an* ORDERED SET $(X, \leq)$, *a* LATTICE $(X, \vee, \wedge; \leq)$, *a* METRIC SPACE (X, d), *or a* LINEAR SPACE $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$.

DEF
A subset $D \subseteq X$ *is a* ***convex set*** *in* X *if*
$$x, y \in D \qquad \Longrightarrow \qquad [x : y] \subseteq D.$$
A set that is ***not*** *convex is* ***concave***.

Example M.1. Consider the *Euclidean space* $\mathbb{R}^2$ (a special case of a *linear space*).

$\Leftarrow \left\{\begin{array}{l}\text{The figure to the left is a}\\ \textit{convex} \text{ set in } \mathbb{R}^2.\end{array}\right.$

$\left.\begin{array}{l}\text{The figure to the right is a}\\ \textit{concave} \text{ set in } \mathbb{R}^2.\end{array}\right\} \Rightarrow$

Example M.2. In a *metric space*, examples of *convex sets* are convex *balls*. Examples include those balls generated by the following metrics:

- Taxi-cab metric
- Euclidean metric
- Sup metric
- Tangential metric

[3] Maeda and Maeda (1970), page 1
[4] van de Vel (1993) page 8
[5] Barvinok (2002) page 2
[6] Barvinok (2002) page 5

COPYRIGHT © 2019 DANIEL J. GREENHOE

Examples of metrics generating balls which are *not* convex include the following:

- Parabolic metric
- Exponential metric

M.3 Convex functions

Figure M.1: Convex and concave functions

Definition M.7. [7] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *be a* LINEAR SPACE *(Definition G.1 page 209) and* D *a* CONVEX SET *(Definition M.6 page 298) in* X.

DEF

A function $\mathsf{f} \in F^D$ *is* ***convex*** *if*

$$\mathsf{f}\left(\lambda \boldsymbol{x} + [1-\lambda]\boldsymbol{y}\right) \quad \leq \quad \lambda \mathsf{f}(\boldsymbol{x}) + (1-\lambda)\,\mathsf{f}(\boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in D \text{ and } \forall \lambda \in (0,1)$$

A function $\mathsf{g} \in F^D$ *is* ***strictly convex*** *if*

$$\mathsf{g}\left(\lambda \boldsymbol{x} + [1-\lambda]\boldsymbol{y}\right) \quad = \quad \lambda \mathsf{g}(\boldsymbol{x}) + (1-\lambda)\,\mathsf{g}(\boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in D,\ \boldsymbol{x} \neq \boldsymbol{y}, \text{ and } \forall \lambda \in (0,1)$$

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 M.3. The function $\mathsf{f}(x) = 2^x$ is a *convex function* (Definition M.7 page 299), as illustrated to the right.

Definition M.8. [8] *Let* $\left(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times})\right)$ *be a* LINEAR SPACE *(Definition G.1 page 209).*

DEF

The ***epigraph*** $\mathsf{epi}(\mathsf{f})$ *and* ***hypograph*** $\mathsf{hyp}(\mathsf{f})$ *of a functional* $\mathsf{f} \in \mathbb{R}^X$ *are defined as*

$$\mathsf{epi}(\mathsf{f}) \triangleq \{(x,y) \in X \times \mathbb{R} \,|\, y \geq \mathsf{f}(x)\}$$
$$\mathsf{hyp}(\mathsf{f}) \triangleq \{(x,y) \in X \times \mathbb{R} \,|\, y \leq \mathsf{f}(x)\}$$

Example M.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

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

EX

Proposition M.1. [9] *Let* $\big(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times})\big)$ *be a* LINEAR SPACE *(Definition G.1 page 209). Let* f *be a* FUNCTIONAL *in* $\mathbb{R}^X$.

PRP

$$\left\{\begin{array}{l}\mathsf{f} \text{ is a}\\ \text{CONVEX FUNCTION}\end{array}\right\} \iff \left\{\begin{array}{l}\text{epi}(\mathsf{f}) \text{ is a}\\ \text{CONVEX SET}\end{array}\right\}$$

Often a function can be proven to be *convex* or *concave*. *Convex* and *concave* functions are defined in Definition M.9 (page 300) (next) and illustrated in Figure M.1 (page 299).

Definition M.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*	$\mathsf{f}(x)$	$\leq$	$\mathsf{y}(x)$	*for* $x \in (a : b)$
(2).	***concave***	*in* $(a : b)$ *if*	$\mathsf{f}(x)$	$\geq$	$\mathsf{y}(x)$	*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 M.1 (Jensen's Inequality). [10] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *be a* LINEAR SPACE *(Definition G.1 page 209), D a subset of X, and* f *a functional in* $\mathbb{F}^D$. *Let* $\sum$ *be the* SUMMATION OPERATOR *(Definition N.1 page 305).*

THM

$$\left\{\begin{array}{lll} 1. & D \text{ is CONVEX} & \text{and}\\ 2. & \mathsf{f} \text{ is CONVEX} & \text{and}\\ 3. & \displaystyle\sum_{n=1}^{N} \lambda_n = 1 & \text{(WEIGHTS)}\end{array}\right\} \implies \mathsf{f}\left(\sum_{n=1}^{N} \lambda_n \boldsymbol{x}_n\right) \leq \sum_{n=1}^{N} \lambda_n \mathsf{f}(\boldsymbol{x}_n) \qquad \forall \boldsymbol{x}_n \in D,\ N \in \mathbb{N}$$

✎PROOF: Proof is by induction:

1. Proof that statement is true for $N = 1$:

$$\begin{aligned}\mathsf{f}\left(\sum_{n=1}^{N=1} \lambda_n \boldsymbol{x}_n\right) &= \mathsf{f}(\lambda_1 \boldsymbol{x}_1)\\ &\leq \mathsf{f}(\lambda_1 \boldsymbol{x}_1)\\ &= \sum_{n=1}^{N=1} \lambda_n \mathsf{f}(\boldsymbol{x}_n)\end{aligned}$$

[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

COPYRIGHT © 2019 DANIEL J. GREENHOE

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 M.2. *Let* $\mathsf{f} : \mathbb{R} \to \mathbb{R}$. *For every* $x_1, x_2 \in (a,b)$ *and* $\lambda \in [0,1]$

THM

$$\mathsf{f} \text{ is convex in } (a,b) \iff \mathsf{f}(\lambda x_1 + (1-\lambda)x_2) \le \lambda \mathsf{f}(x_1) + (1-\lambda)\mathsf{f}(x_2)$$

✎PROOF:

1. prove f is convex $\implies \mathsf{f}(\lambda x_1 + (1-\lambda)x_2) \le \lambda \mathsf{f}(x_1) + (1-\lambda)\mathsf{f}(x_2)$:

$$\begin{aligned}
\mathsf{f}(\lambda x_1 + (1-\lambda)x_2) &\le \frac{\mathsf{f}(b)-\mathsf{f}(a)}{b-a}[\lambda x_1 + (1-\lambda)x_2 - a] + \mathsf{f}(a) \\
&= \frac{\mathsf{f}(x_2)-\mathsf{f}(x_1)}{x_2-x_1}[\lambda x_1 + (1-\lambda)x_2 - x_1] + \mathsf{f}(x_1) \\
&= \frac{\mathsf{f}(x_2)-\mathsf{f}(x_1)}{x_2-x_1}[(x_2-x_1)(1-\lambda)] + \mathsf{f}(x_1) \\
&= (1-\lambda)\mathsf{f}(x_2) - (1-\lambda)\mathsf{f}(x_1) + \mathsf{f}(x_1) \\
&= \lambda \mathsf{f}(x_1) + (1-\lambda)\mathsf{f}(x_2)
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

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 M.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 M.2 (page 301), $\mathsf{f}(x)$ is convex. ✏

M.4 Literature

Literature survey:

[11] Cover and Thomas (1991), pages 24–25

COPYRIGHT © 2019 DANIEL J. GREENHOE

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

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

APPENDIX N

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]

N.1 Summation

Definition N.1. [2] *Let $+$ be an addition operator on a tuple $(\!| x_n |\!)_m^N$.*

DEF

*The **summation** of $(\!| x_n |\!)$ 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 N.1 (Generalized associative property). [3] *Let $+$ be an addition operator on a tuple $(\!| x_n |\!)_m^N$.*

THM

$+$ *is* ASSOCIATIVE $\Longrightarrow$

$$\underbrace{\sum_{n=m}^{L} x_n + \left(\sum_{n=L+1}^{M} x_n + \sum_{n=M+1}^{N} x_n\right) = \left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \sum_{n=M+1}^{N} x_n}_{\sum_{n=m}^{N} \text{ is ASSOCIATIVE}} \qquad \text{for } m < L < M \leq N$$

✎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:

$$\begin{aligned}
\sum_{n=m}^{m+2} x_n &= \left(\sum_{n=m}^{m+1} x_n\right) + x_{m+2} && \text{by Definition N.1 page 305}\\
&= \left(x_m + x_{m+1}\right) + x_{m+2} && \text{by item (3)}\\
&= x_m + \left(x_{m+1} + x_{m+2}\right) && \text{by left hypothesis}
\end{aligned}$$

5. Proof that N case $\implies$ $N+1$ case:

$$\begin{aligned}
\sum_{n=m}^{N+1} x_n &= \underbrace{\left(\sum_{n=m}^{N} x_n\right)}_{associative} + x_{N+1} && \text{by Definition N.1 page 305}\\
&= \left(\sum_{n=m}^{L} x_n + \left(\sum_{n=L+1}^{M} x_n + \sum_{n=M+1}^{N} x_n\right)\right) + x_{N+1} && = \left(\left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \sum_{n=M+1}^{N} x_n\right) + x_{N+1}\\
&= \left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \left(\sum_{n=M+1}^{N} x_n + x_{N+1}\right) && = \left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \left(\sum_{n=M+1}^{N+1} x_n\right)
\end{aligned}$$

✍

N.2 Means

N.2.1 Weighted ϕ-means

Definition N.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 $\sum_{n=1}^{N} \lambda_n = 1$.

[4] Bollobás (1999) page 5

COPYRIGHT © 2019 DANIEL J. GREENHOE

Lemma N.1. [5] *Let* $\mathsf{M}_\phi(\langle x_n \rangle)$ *be the* $\langle \lambda_n \rangle_1^N$ *weighted* ϕ*-mean of a tuple* $\langle x_n \rangle_1^N$*. Let the property* CONVEX *be defined as in Definition M.7 (page 299).*

LEM	$\phi\psi^{-1}$ *is* CONVEX	*and*	ϕ *is* INCREASING	$\Longrightarrow$	$\mathsf{M}_\phi(\langle x_n \rangle) \geq \mathsf{M}_\psi(\langle x_n \rangle)$
	$\phi\psi^{-1}$ *is* CONVEX	*and*	ϕ *is* DECREASING	$\Longrightarrow$	$\mathsf{M}_\phi(\langle x_n \rangle) \leq \mathsf{M}_\psi(\langle x_n \rangle)$
	$\phi\psi^{-1}$ *is* CONCAVE	*and*	ϕ *is* INCREASING	$\Longrightarrow$	$\mathsf{M}_\phi(\langle x_n \rangle) \leq \mathsf{M}_\psi(\langle x_n \rangle)$
	$\phi\psi^{-1}$ *is* CONCAVE	*and*	ϕ *is* DECREASING	$\Longrightarrow$	$\mathsf{M}_\phi(\langle x_n \rangle) \geq \mathsf{M}_\psi(\langle x_n \rangle)$

✎PROOF:

1. Case where $\phi\psi^{-1}$ is *convex* and ϕ is *increasing*:

$$\begin{aligned}
\mathsf{M}_\phi(\langle x_n \rangle) &\triangleq \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi(x_n)\right) && \text{by definition of } \mathsf{M}_\phi && \text{(Definition N.2 page 306)}\\
&= \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi\psi^{-1}\psi(x_n)\right) && \text{by definition of } \psi^{-1} && \\
&\geq \phi^{-1}\left(\phi\psi^{-1}\sum_{n=1}^N \lambda_n \psi(x_n)\right) && \text{by } \textit{Jensen's Inequality} && \text{(Theorem M.1 page 300)}\\
&= \left(\psi^{-1}\sum_{n=1}^N \lambda_n \psi(x_n)\right) && \text{by definition of } \psi^{-1} && \\
&\triangleq \mathsf{M}_\psi(\langle x_n \rangle) && \text{by definition of } \mathsf{M}_\psi && \text{(Definition N.2 page 306)}
\end{aligned}$$

2. Case where $\phi\psi^{-1}$ is *convex* and ϕ is *decreasing*:

$$\begin{aligned}
\mathsf{M}_\phi(\langle x_n \rangle) &\triangleq \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi(x_n)\right) && \text{by definition of } \mathsf{M}_\phi && \text{(Definition N.2 page 306)}\\
&= \phi^{-1}\left(\sum_{n=1}^N \lambda_n \phi\psi^{-1}\psi(x_n)\right) && \text{by definition of } \psi^{-1} && \\
&\leq \phi^{-1}\left(\phi\psi^{-1}\sum_{n=1}^N \lambda_n \psi(x_n)\right) && \text{by } \textit{Jensen's Inequality} \text{ and because } \phi^{-1} \text{ is } \textit{decreasing} && \text{(Theorem M.1 page 300) (by hypothesis)}\\
&= \left(\psi^{-1}\sum_{n=1}^N \lambda_n \psi(x_n)\right) && \text{by definition of } \psi^{-1} && \\
&\triangleq \mathsf{M}_\psi(\langle x_n \rangle) && \text{by definition of } \mathsf{M}_\psi && \text{(Definition N.2 page 306)}
\end{aligned}$$

✏

One of the most well known inequalities in mathematics is *Minkowski's Inequality* (1910, Theorem N.5 page 313). 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 N.2, next).

Theorem N.2. [7]

[5] Pečarić et al. (1992) page 107, Bollobás (1999) page 5, Hardy et al. (1952) page 75
[6] Mulholland (1950)
[7] Milovanović and Milovanović (1979), Bullen (2003) page 306 ⟨Theorem 9⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

THM

$$\left\{\begin{array}{lll|lll} (1). & \phi \text{ is CONVEX} & \text{and} & (3). & \phi(0)=0 & \text{and} \\ (2). & \phi \text{ is STRICTLY MONOTONIC} & \text{and} & (4). & \log\circ\phi\circ\exp \text{ is CONVEX} & \end{array}\right\} \implies$$

$$\left\{\phi^{-1}\left(\sum_{n=1}^{N}\lambda_n\phi(x_n+y_n)\right) \le \phi^{-1}\left(\sum_{n=1}^{N}\lambda_n\phi(x_n)\right)+\phi^{-1}\left(\sum_{n=1}^{N}\lambda_n\phi(y_n)\right)\right\}$$

N.2.2 Power means

Definition N.3. [8] *Let* $\mathsf{M}_{\phi(x;r)}(\langle x_n\rangle)$ *be the* $\langle\lambda_n\rangle_1^N$ *weighted* ϕ*-mean of a* NON-NEGATIVE *tuple* $\langle x_n\rangle_1^N$ *(Definition N.2 page 306).*

DEF

A mean $\mathsf{M}_{\phi(x;r)}(\langle x_n\rangle)$ *is a* ***power mean*** *with parameter* r *if* $\phi(x)\triangleq x^r$*. That is,*

$$\mathsf{M}_{\phi(x;r)}(\langle x_n\rangle) = \left(\sum_{n=1}^{N}\lambda_n(x_n)^r\right)^{\frac{1}{r}}$$

Theorem N.3. [9] *Let* $\mathsf{M}_{\phi(x;r)}(\langle x_n\rangle)$ *be* POWER MEAN *with parameter* r *of an* N*-tuple* $\langle x_n\rangle_1^N$*. Let* $\mathbb{R}^*$ *be the set of extended real numbers* $\left(\mathbb{R}^*\triangleq\mathbb{R}\cup\{-\infty,\infty\}\right)$.[10]

THM

$$\mathsf{M}_{\phi(x;r)}(\langle x_n\rangle) \triangleq \left(\sum_{n=1}^{N}\lambda_n(x_n)^r\right)^{\frac{1}{r}} \text{ is CONTINUOUS and STRICTLY INCREASING in } \mathbb{R}^*.$$

$$\mathsf{M}_{\phi(x;r)}(\langle x_n\rangle) = \begin{cases} \min\limits_{n=1,2,\ldots,N}\langle x_n\rangle & \text{for } r=-\infty \\ \prod\limits_{n=1}^{N} x_n^{\lambda_n} & \text{for } r=0 \\ \max\limits_{n=1,2,\ldots,N}\langle x_n\rangle & \text{for } r=+\infty \end{cases}$$

✎PROOF:

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 N.1 (page 307),
   $$-\infty<r<s<\infty \implies \mathsf{M}_{\phi(x;r)}(\langle x_n\rangle) < \mathsf{M}_{\phi(x;s)}(\langle x_n\rangle).$$
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.

[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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

3. Lemma: $\mathsf{M}_{\phi(x;-r)}\big(\llparenthesis x_n \rrparenthesis\big) = \big\{\mathsf{M}_{\phi(x;r)}\big(\llparenthesis x_n^{-1} \rrparenthesis\big)\big\}^{-1}$. Proof:

$$\begin{aligned}
\big\{\mathsf{M}_{\phi(x;r)}\big(\llparenthesis x_n^{-1} \rrparenthesis\big)\big\}^{-1} &= \left\{\left(\sum_{n=1}^{N} \lambda_n \big(x_n^{-1}\big)^r\right)^{\frac{1}{r}}\right\}^{-1} && \text{by definition of } \mathsf{M}_\phi \\
&= \left(\sum_{n=1}^{N} \lambda_n \big(x_n\big)^{-r}\right)^{\frac{1}{-r}} \\
&= \mathsf{M}_{\phi(x;-r)}\big(\llparenthesis x_n \rrparenthesis\big) && \text{by definition of } \mathsf{M}_\phi
\end{aligned}$$

4. Proof that $\lim\limits_{r\to\infty} \mathsf{M}_\phi\big(\llparenthesis x_n \rrparenthesis\big) = \max\limits_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis$:

   (a) Let $x_m \triangleq \max\limits_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis$

   (b) Note that $\lim\limits_{r\to\infty} \mathsf{M}_\phi \le \max\limits_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis$ because

$$\begin{aligned}
\lim_{r\to\infty} \mathsf{M}_\phi\big(\llparenthesis x_n \rrparenthesis\big) &= \lim_{r\to\infty} \left(\sum_{n=1}^{N} \lambda_n x_n^r\right)^{\frac{1}{r}} && \text{by definition of } \mathsf{M}_\phi \\
&\le \lim_{r\to\infty} \left(\sum_{n=1}^{N} \lambda_n x_m^r\right)^{\frac{1}{r}} && \text{by definition of } x_m \text{ in item (4a) and because } \phi(x) \triangleq x^r \text{ and } \phi^{-1} \text{ are both increasing or both decreasing} \\
&= \lim_{r\to\infty} \left(x_m^r \underbrace{\sum_{n=1}^{N} \lambda_n}_{1}\right)^{\frac{1}{r}} && \text{because } x_m \text{ is a constant} \\
&= \lim_{r\to\infty} \big(x_m^r \cdot 1\big)^{\frac{1}{r}} \\
&= x_m \\
&= \max_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis && \text{by definition of } x_m \text{ in item (4a)}
\end{aligned}$$

   (c) But also note that $\lim\limits_{r\to\infty} \mathsf{M}_\phi \ge \max\limits_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis$ because

$$\begin{aligned}
\lim_{r\to\infty} \mathsf{M}_\phi\big(\llparenthesis x_n \rrparenthesis\big) &= \lim_{r\to\infty} \left(\sum_{n=1}^{N} \lambda_n x_n^r\right)^{\frac{1}{r}} && \text{by definition of } \mathsf{M}_\phi \\
&\ge \lim_{r\to\infty} \big(w_m x_m^r\big)^{\frac{1}{r}} && \text{by definition of } x_m \text{ in item (4a) and because } \phi(x) \triangleq x^r \text{ and } \phi^{-1} \text{ are both increasing or both decreasing} \\
&= \lim_{r\to\infty} w_m^{\frac{1}{r}} x_m^{\frac{r}{r}} \\
&= x_m \\
&= \max_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis && \text{by definition of } x_m \text{ in item (4a)}
\end{aligned}$$

   (d) Combining items (b) and (c) we have $\lim\limits_{r\to\infty} \mathsf{M}_\phi = \max\limits_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis$.

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

5. Proof that $\lim_{r\to-\infty} \mathsf{M}_\phi\left(\llparenthesis x_n \rrparenthesis\right) = \min_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis$:

$$\begin{aligned}
\lim_{r\to-\infty} \mathsf{M}_{\phi(x;r)}\left(\llparenthesis x_n \rrparenthesis\right) &= \lim_{r\to\infty} \mathsf{M}_{\phi(x;-r)}\left(\llparenthesis x_n \rrparenthesis\right) && \text{by change of variable } r\\
&= \lim_{r\to\infty} \left\{\mathsf{M}_{\phi(x;r)}\left(\llparenthesis x_n^{-1} \rrparenthesis\right)\right\}^{-1} && \text{by Lemma in item (3) page 309}\\
&= \lim_{r\to\infty} \frac{1}{\mathsf{M}_{\phi(x;r)}\left(\llparenthesis x_n^{-1} \rrparenthesis\right)}\\
&= \frac{\lim_{r\to\infty} 1}{\lim_{r\to\infty} \mathsf{M}_{\phi(x;r)}\left(\llparenthesis x_n^{-1} \rrparenthesis\right)} && \text{by property of lim [11]}\\
&= \frac{1}{\max_{n\in\mathbb{Z}} \llparenthesis x_n^{-1} \rrparenthesis} && \text{by item (4)}\\
&= \frac{1}{\left(\min_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis\right)^{-1}}\\
&= \min_{n\in\mathbb{Z}} \llparenthesis x_n \rrparenthesis
\end{aligned}$$

6. Proof that $\lim_{r\to 0} \mathsf{M}_\phi\left(\llparenthesis x_n \rrparenthesis\right) = \prod_{n=1}^{N} x_n^{\lambda_n}$:

$$\begin{aligned}
\lim_{r\to 0} \mathsf{M}_\phi\left(\llparenthesis x_n \rrparenthesis\right) &= \lim_{r\to 0} \exp\left\{\ln\left\{\mathsf{M}_\phi\left(\llparenthesis x_n \rrparenthesis\right)\right\}\right\}\\
&= \lim_{r\to 0} \exp\left\{\ln\left\{\left(\sum_{n=1}^{N} \lambda_n \left(x_n^r\right)\right)^{\frac{1}{r}}\right\}\right\} && \text{by definition of } \mathsf{M}_\phi\\
&= \exp\left\{\frac{\frac{\partial}{\partial r}\ln\left(\sum_{n=1}^{N} \lambda_n \left(x_n^r\right)\right)}{\frac{\partial}{\partial r} r}\right\}_{r=0} && \text{by l'Hôpital's rule[12]}\\
&= \exp\left\{\frac{\sum_{n=1}^{N} \lambda_n \frac{\partial}{\partial r}\left(x_n^r\right)}{\sum_{n=1}^{N} \lambda_n \left(x_n^r\right)}\right\}_{r=0} && = \exp\left\{\frac{\sum_{n=1}^{N} \lambda_n \frac{\partial}{\partial r}\exp\left(\ln\left(x_n^r\right)\right)}{\sum_{n=1}^{N} \lambda_n}\right\}_{r=0}\\
&= \exp\left\{\frac{\sum_{n=1}^{N} \lambda_n \frac{\partial}{\partial r}\exp\left(r\ln\left(x_n\right)\right)}{1}\right\}_{r=0} && = \exp\left\{\sum_{n=1}^{N} \lambda_n \frac{\partial}{\partial r}\exp\left(r\ln\left(x_n\right)\right)\right\}_{r=0}\\
&= \exp\left\{\sum_{n=1}^{N} \lambda_n \exp\left\{r\ln x_n\right\}\ln\left(x_n\right)\right\}_{r=0} && = \exp\left\{\sum_{n=1}^{N} \lambda_n \ln\left(x_n\right)\right\}\\
&= \exp\left\{\sum_{n=1}^{N} \ln\left(x_n^{\lambda_n}\right)\right\} && = \exp\left\{\ln\prod_{n=1}^{N} x_n^{\lambda_n}\right\} = \prod_{n=1}^{N} x_n^{\lambda_n}
\end{aligned}$$

[11] Rudin (1976) page 85 ⟨4.4 Theorem⟩
[12] Rudin (1976) page 109 ⟨5.13 Theorem⟩

Copyright © 2019 Daniel J. Greenhoe

✏

Definition N.4. *Let $(\!| x_n |\!)_1^N$ be a tuple. Let $(\!| \lambda_n |\!)_1^N$ be a tuple of weighting values.*

DEF

*The **harmonic mean** of $(\!| x_n |\!)$ is defined as* $\mu_h \triangleq \left(\sum_{n=1}^N \lambda_n \frac{1}{x_n}\right)^{-1}$ *where* $\sum_{n=1}^N \lambda_n = 1$

*The **geometric mean** of $(\!| x_n |\!)$ is defined as* $\mu_g \triangleq \sum_{n=1}^N x_n^{\lambda_n}$ *where* $\sum_{n=1}^N \lambda_n = 1$

*The **arithmetic mean** of $(\!| x_n |\!)$ is defined as* $\mu_a \triangleq \sum_{n=1}^N \lambda_n x_n$ *where* $\sum_{n=1}^N \lambda_n = 1$

*The **average** of $(\!| x_n |\!)$ is defined as* $\mu_a \triangleq \frac{1}{N}\sum_{n=1}^N x_n$

N.3 Inequalities on power means

Corollary N.1. [13] *Let $(\!| x_n |\!)_1^N$ be a tuple. Let $(\!| \lambda_n |\!)_1^N$ be a tuple of weighting values.*

COR

$$\min(\!| x_n |\!) \le \underbrace{\left(\sum_{n=1}^N \lambda_n \frac{1}{x_n}\right)^{-1}}_{\text{harmonic mean}} \le \underbrace{\prod_{n=1}^N x_n^{\lambda_n}}_{\text{geometric mean}} \le \underbrace{\sum_{n=1}^N \lambda_n x_n}_{\text{arithmetic mean}} \le \max(\!| x_n |\!) \quad \text{where} \quad \sum_{n=1}^N \lambda_n = 1$$

✎PROOF:

1. These five means are all special cases of the *power mean* $\mathsf{M}_{\phi(x:r)}$ (Definition N.3 page 308):
   $r = \infty$: $\max(\!| x_n |\!)$
   $r = 1$: arithmetic mean
   $r = 0$: geometric mean
   $r = -1$: harmonic mean
   $r = -\infty$: $\min(\!| x_n |\!)$

2. The inequalities follow directly from Theorem N.3 (page 308).

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*:

$$\sum_{n=1}^N \lambda_n x_n = b^{\log_b\left(\sum_{n=1}^N \lambda_n x_n\right)} \ge b^{\left(\sum_{n=1}^N \lambda_n \log_b x_n\right)} \quad \text{by } \textit{Jensen's Inequality} \text{ (Theorem M.1 page 300)}$$

$$= \prod_{n=1}^N b^{(\lambda_n \log_b x_n)} = \prod_{n=1}^N b^{(\log_b x_n)\lambda_n} = \prod_{n=1}^N x_n^{\lambda_n}$$

✏

[13] Bullen (2003) page 71, Bollobás (1999) page 5, Cauchy (1821), pages 457–459 ⟨Note II, theorem 17⟩, Jensen (1906) page 183

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Lemma N.2 (Young's Inequality). [14]

LEM

$$xy < \frac{x^p}{p} + \frac{y^q}{q} \quad \text{with } \frac{1}{p} + \frac{1}{q} = 1 \quad \forall 1 < p < \infty,\ x, y \geq 0, \quad \text{but } y \neq x^{p-1}$$

$$xy = \frac{x^p}{p} + \frac{y^q}{q} \quad \text{with } \frac{1}{p} + \frac{1}{q} = 1 \quad \forall 1 < p < \infty,\ x, y \geq 0, \quad \text{and } y = x^{p-1}$$

✎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 propemonotonically increasing in u and $u = v^{q-1}$ is propemonotonically 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 $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$:

$$\begin{aligned}
xy &\leq \int_0^x u^{p-1}\,\mathrm{d}u + \int_0^y v^{q-1}\,\mathrm{d}v \\
&= \left.\frac{u^p}{p}\right|_0^x + \left.\frac{v^q}{q}\right|_0^y \\
&= \frac{x^p}{p} + \frac{y^q}{q}
\end{aligned}$$

[14] Carothers (2000), page 43, Tolsted (1964), page 5, Maligranda (1995), page 257, Hardy et al. (1952) ⟨Theorem 24⟩, Young (1912) page 226

COPYRIGHT © 2019 DANIEL J. GREENHOE

✏

Theorem N.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 $\|x_n\|_p \triangleq \left(\sum_{n=1}^{N} |x_n|^p\right)^{\frac{1}{p}}$.

$$\begin{aligned}
\sum_{n=1}^{N} |x_n y_n| &= \|(\!| x_n |\!)\|_p \, \|(\!| y_n |\!)\|_q \sum_{n=1}^{N} \frac{|x_n y_n|}{\|(\!| x_n |\!)\|_p \, \|(\!| y_n |\!)\|_q} \\
&= \|(\!| x_n |\!)\|_p \, \|(\!| y_n |\!)\|_q \sum_{n=1}^{N} \frac{|x_n|}{\|(\!| x_n |\!)\|_p} \frac{|y_n|}{\|(\!| y_n |\!)\|_q} \\
&\le \|(\!| x_n |\!)\|_p \, \|(\!| y_n |\!)\|_q \sum_{n=1}^{N} \left(\frac{1}{p}\frac{|x_n|^p}{\|(\!| x_n |\!)\|_p^p} + \frac{1}{q}\frac{|y_n|^q}{\|(\!| y_n |\!)\|_q^q}\right) && \text{by } \textit{Young's Inequality} \quad \text{(Lemma N.2 page 312)} \\
&= \|(\!| x_n |\!)\|_p \, \|(\!| y_n |\!)\|_q \left(\frac{1}{p}\cdot\frac{\sum |x_n|^p}{\|(\!| x_n |\!)\|_p^p} + \frac{1}{q}\cdot\frac{\sum |y_n|^q}{\|(\!| y_n |\!)\|_q^q}\right) \\
&= \|(\!| x_n |\!)\|_p \, \|(\!| y_n |\!)\|_q \left(\frac{1}{p}\frac{\|(\!| x_n |\!)\|_p^p}{\|(\!| x_n |\!)\|_p^p} + \frac{1}{q}\frac{\|(\!| y_n |\!)\|_q^q}{\|(\!| y_n |\!)\|_q^q}\right) && \text{by definition of } \|\cdot\| \\
&= \|(\!| x_n |\!)\|_p \, \|(\!| y_n |\!)\|_q \underbrace{\left(\frac{1}{p}+\frac{1}{q}\right)}_{1} \\
&= \|(\!| x_n |\!)\|_p \, \|(\!| y_n |\!)\|_q && \text{by } \frac{1}{p}+\frac{1}{q}=1 \text{ constraint}
\end{aligned}$$

✏

Theorem N.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.*

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 $q \triangleq \dfrac{p}{p-1}$

[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

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

2. Define $\|\boldsymbol{x}\|_p \triangleq \left(\sum_{n=1}^{N} |x_n|^p\right)^{\frac{1}{p}}$.

3. Proof that $\|x_n + y_n\|_p \le \|x_n\|_p + \|y_n\|_p$:

$$\boxed{\|x_n + y_n\|_p^p}$$

$$= \sum_{n=1}^{N} |x_n + y_n|^p \qquad \text{by definition of } \|\cdot\|_p \qquad \text{(definition 2 page 314)}$$

$$= \sum_{n=1}^{N} |x_n + y_n||x_n + y_n|^{p-1} \qquad \text{by } \textit{homogeneous} \text{ property of } |\cdot|$$

$$\boxed{\le} \sum_{n=1}^{N} |x_n||x_n + y_n|^{p-1} + \sum_{n=1}^{N} |y_n||x_n + y_n|^{p-1} \qquad \text{by } \textit{subadditive} \text{ property of } |\cdot|$$

$$= \sum_{n=1}^{N} |x_n(x_n + y_n)^{p-1}| + \sum_{n=1}^{N} |y_n(x_n + y_n)^{p-1}| \qquad \text{by } \textit{homogeneous} \text{ property of } |\cdot|$$

$$\boxed{\le} \|x_n\|_p \|(x_n + y_n)^{p-1}\|_q + \|y_n\|_p \|(x_n + y_n)^{p-1}\|_q \qquad \text{by } \textit{Hölder's Inequality} \qquad \text{(Theorem N.4 page 313)}$$

$$= \left(\|x_n\|_p + \|y_n\|_p\right) \|(x_n + y_n)^{p-1}\|_q$$

$$= \left(\|x_n\|_p + \|y_n\|_p\right)\left(\sum_{n=1}^{N} |(x_n + y_n)^{p-1}|^q\right)^{\frac{1}{q}} \qquad \text{by definition of } \|\cdot\|_p \qquad \text{(definition 2 page 314)}$$

$$= \left(\|x_n\|_p + \|y_n\|_p\right)\left(\sum_{n=1}^{N} |(x_n + y_n)^{p-1}|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \qquad \text{by definition 1}$$

$$= \left(\|x_n\|_p + \|y_n\|_p\right)\left(\sum_{n=1}^{N} |(x_n + y_n)|^p\right)^{\frac{p-1}{p}}$$

$$= \left(\|x_n\|_p + \|y_n\|_p\right) \|x_n + y_n\|^{p-1} \qquad \text{by definition of } \|\cdot\|_p \qquad \text{(definition 2 page 314)}$$

$$\Longrightarrow \quad \boxed{\|x_n + y_n\|_p \le \|x_n\|_p + \|y_n\|_p}$$

✍

❝*Cauchy is the only one occupied with pure mathematics: Poisson, Fourier, Ampere, etc., busy themselves exclusively with magnetism and other physical subjects. Mr. Laplace writes nothing now, I believe.*❞

Niels Henrik Abel in an 1826 letter [17]

Theorem N.6 (Cauchy-Schwarz Inequality for sequences). [18] *Let* $(x_n \in \mathbb{C})_1^N$ *and* $(y_n \in \mathbb{C})_1^N$ *be complex* N*-tuples.*

[17] quote: Bell (1986) page 318 ⟨Chapter 17. "Genius and Poverty" "Abel (1802–1829)"⟩, Boyer and Merzbach (2011) page 462 ⟨without "Mr. Laplace…" portion⟩. 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 ⟨???⟩

Copyright © 2019 Daniel J. Greenhoe

THM

$$\left|\sum_{n=1}^{N} x_n y_n^*\right|^2 \leq \left(\sum_{n=1}^{N} |x_n|^2\right)\left(\sum_{n=1}^{N} |y_n|^2\right) \qquad \forall \boldsymbol{x},\boldsymbol{y}\in\boldsymbol{X}$$

$$\left|\sum_{n=1}^{N} x_n y_n^*\right|^2 = \left(\sum_{n=1}^{N} |x_n|^2\right)\left(\sum_{n=1}^{N} |y_n|^2\right) \iff \exists a \in \mathbb{C} \quad \text{such that} \quad \boldsymbol{y} = a\boldsymbol{x} \qquad \forall \boldsymbol{x},\boldsymbol{y}\in\boldsymbol{X}$$

✎PROOF:

1. The *Cauchy-Schwarz Inequality for sequences* is a special case of the *Hölder inequality* (Theorem N.4 page 313) for $p = q = 2$.

2. Alternatively, the *Cauchy-Schwarz inequality for sequences* is a special case of the *Cauchy-Schwarz inequality for 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 $\left(\llparenthesis x_n \rrparenthesis, \langle \triangle \mid \triangledown\rangle\right)$ is an inner-product space.

   (b) By the more general *Cauchy-Schwarz Inequality for inner-product spaces*,

$$\begin{aligned}
\left(\sum_{n=1}^{N} a_n\lambda_n\right)^2 &\triangleq \langle a_n \mid \lambda_n\rangle^2 && \text{by definition of } \langle x_n \mid y_n\rangle \\
&\boxed{\leq} \|x_n\|^2 \|y_n\|^2 && \text{by } \textit{Cauchy-Schwarz Inequality for inner-product spaces} \text{ (Theorem K.2 page 272)} \\
&\triangleq \left(\sum_{n=1}^{N} x_n^2\right)\left(\sum_{n=1}^{N} y_n^2\right) && \text{by definition of } \|\cdot\|
\end{aligned}$$

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 N.1. [20]

PRP

$$(x+y)^p \leq 2^p(x^p + y^p) \qquad \forall x, y \geq 0,\ 1 < p < \infty$$

✎PROOF:

$$\begin{aligned}
(x+y)^p &\leq (2\max\{x,y\})^p \\
&= 2^p(\max\{x,y\})^p \\
&= 2^p(\max\{x^p, y^p\}) \\
&\leq 2^p(x^p + y^p)
\end{aligned}$$

✏

[19] Bullen (2003) pages 183–185 ⟨Theorem 5⟩
[20] Carothers (2000), page 43

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

N.4 Power Sums

Theorem N.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:

$$\boxed{\sum_{k=0}^{n-1} r^k} = \left(\frac{1}{1-r}\right)\left[(1-r)\sum_{k=0}^{n-1} r^k\right] = \left(\frac{1}{1-r}\right)\left[\sum_{k=0}^{n-1} r^k - r\sum_{k=0}^{n-1} r^k\right] = \left(\frac{1}{1-r}\right)\left[\sum_{k=0}^{n-1} r^k - \left(\sum_{k=0}^{n-1} r^k - 1 + r^n\right)\right]$$
$$= \left(\frac{1}{1-r}\right)\left[1-r^n\right] = \boxed{\frac{1-r^n}{1-r}}$$

✏

Lemma N.3. *Let* $\mathsf{f}(x)$ *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:

$$\mathsf{S}(x+\tau) \triangleq \sum_{n\in\mathbb{Z}} \mathsf{f}(x+\tau+n\tau) = \sum_{n\in\mathbb{Z}} \mathsf{f}(x+(n+1)\tau) = \sum_{m\in\mathbb{Z}} \mathsf{f}(x+m\tau) \quad (\text{where } m \triangleq n+1)$$
$$\triangleq \mathsf{S}(x)$$

✏

Proposition N.2 (Power Sums). [22]

PRP
$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \quad \forall n\in\mathbb{N} \qquad \sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4} \quad \forall n\in\mathbb{N}$$
$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \quad \forall n\in\mathbb{N} \qquad \sum_{k=1}^{n} k^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30} \quad \forall n\in\mathbb{N}$$

✎PROOF:

1. Proof that $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$: (proof by induction)

$$\sum_{k=1}^{n=1} k = 1 = \frac{1(1+1)}{2} = \left.\frac{n(n+1)}{2}\right|_{n=1}$$
$$\sum_{k=1}^{n+1} k = \left(\sum_{k=1}^{n} k\right) + \left(n+1\right) = \underbrace{\left(\frac{n(n+1)}{2}\right) + \left(n+1\right)}_{\text{by left hypothesis}} = \left(n+1\right)\left(\frac{n}{2}+1\right)$$
$$= \left(n+1\right)\left(\frac{n+2}{2}\right) = \frac{(n+1)(n+2)}{2}$$

[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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

2. Proof that $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$: (proof by induction)

$$\sum_{k=1}^{n=1} k^2 = 1 = \frac{1(1+1)(2+1)}{6} = \left.\frac{n(n+1)(2n+1)}{6}\right|_{n=1}$$

$$\sum_{k=1}^{n+1} k^2 = \left(\sum_{k=1}^{n} k^2\right) + \left(n+1\right)^2 = \underbrace{\left(\frac{n(n+1)(2n+1)}{6}\right) + \left(n+1\right)^2}_{\text{by left hypothesis}} = \left(n+1\right)\left(\frac{n(2n+1)+6(n+1)}{6}\right)$$

$$= \left(n+1\right)\left(\frac{2n^2+7n+6}{6}\right) = \left(n+1\right)\left(\frac{(n+2)(2n+3)}{6}\right) = \frac{(n+1)[(n+1)+1][2(n+1)+1]}{6}$$

✍

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

APPENDIX O

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]

O.1 Operators on linear spaces

O.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 O.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 G.1 page 209)*.

Two operators $\mathbf{A}$ and $\mathbf{B}$ in $\boldsymbol{Y}^{\boldsymbol{X}}$ are **equal** if $\mathbf{A}\boldsymbol{x} = \mathbf{B}\boldsymbol{x}$ for all $\boldsymbol{x} \in \boldsymbol{X}$. 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 $\mathbf{I}\boldsymbol{x} = \mathbf{I}$ for all $\boldsymbol{x} \in \boldsymbol{X}$.

Definition O.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 O.1: Some operator types

DEF $\mathbf{I}$ *is the* ***identity operator*** *in* $\boldsymbol{X}^{\boldsymbol{X}}$ *if* $\quad \mathbf{I} = \mathbb{1}(\boldsymbol{X})$.

O.1.2 Linear operators

Definition O.3. [4] *Let* $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *and* $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\big)$ *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. $\mathbf{L}(\alpha\boldsymbol{x}) = \alpha\mathbf{L}\boldsymbol{x} \quad \forall \boldsymbol{x} \in \boldsymbol{X}, \quad \forall \alpha \in \mathbb{F}$ (HOMOGENEOUS).

The set of all linear operators from $\boldsymbol{X}$ *to* $\boldsymbol{Y}$ *is denoted* $\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$ *such that*

$$\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}) \triangleq \big\{\mathbf{L} \in \boldsymbol{Y}^{\boldsymbol{X}} \,|\, \mathbf{L} \text{ is linear}\big\} \quad .$$

Theorem O.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

$$\{\mathbf{L} \text{ is LINEAR}\} \quad \Longrightarrow \quad \left\{\begin{array}{llll} 1. & \mathbf{L}\mathbb{0} = \mathbb{0} & & \text{and} \\ 2. & \mathbf{L}(-\boldsymbol{x}) = -(\mathbf{L}\boldsymbol{x}) & \forall \boldsymbol{x} \in \boldsymbol{X} & \text{and} \\ 3. & \mathbf{L}(\boldsymbol{x}-\boldsymbol{y}) = \mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y} & \forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{X} & \text{and} \\ 4. & \mathbf{L}\left(\sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n\right) = \sum_{n=1}^{N} \alpha_n \big(\mathbf{L}\boldsymbol{x}_n\big) & \boldsymbol{x}_n \in \boldsymbol{X}, \alpha_n \in \mathbb{F} & \end{array}\right\}$$

✎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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

1. Proof that $\mathbf{L}\mathbb{0} = \mathbb{0}$:

$$\begin{aligned}
\mathbf{L}\mathbb{0} &= \mathbf{L}(0 \cdot \mathbb{0}) && \text{by additive identity property} && \text{(Theorem G.1 page 211)}\\
&= 0 \cdot (\mathbf{L}\mathbb{0}) && \text{by \textit{homogeneous} property of } \mathbf{L} && \text{(Definition O.3 page 320)}\\
&= \mathbb{0} && \text{by \textit{additive identity} property} && \text{(Theorem G.1 page 211)}
\end{aligned}$$

2. Proof that $\mathbf{L}(-\boldsymbol{x}) = -(\mathbf{L}\boldsymbol{x})$:

$$\begin{aligned}
\mathbf{L}(-\boldsymbol{x}) &= \mathbf{L}(-1 \cdot \boldsymbol{x}) && \text{by \textit{additive inverse} property} && \text{(Theorem G.2 page 212)}\\
&= -1 \cdot (\mathbf{L}\boldsymbol{x}) && \text{by \textit{homogeneous} property of } \mathbf{L} && \text{(Definition O.3 page 320)}\\
&= -(\mathbf{L}\boldsymbol{x}) && \text{by \textit{additive inverse} property} && \text{(Theorem G.2 page 212)}
\end{aligned}$$

3. Proof that $\mathbf{L}(\boldsymbol{x} - \boldsymbol{y}) = \mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y}$:

$$\begin{aligned}
\mathbf{L}(\boldsymbol{x} - \boldsymbol{y}) &= \mathbf{L}(\boldsymbol{x} + (-\boldsymbol{y})) && \text{by \textit{additive inverse} property} && \text{(Theorem G.2 page 212)}\\
&= \mathbf{L}(\boldsymbol{x}) + \mathbf{L}(-\boldsymbol{y}) && \text{by \textit{linearity} property of } \mathbf{L} && \text{(Definition O.3 page 320)}\\
&= \mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y} && \text{by item (2)}
\end{aligned}$$

4. Proof that $\mathbf{L}\left(\sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n\right) = \sum_{n=1}^{N} \alpha_n (\mathbf{L}\boldsymbol{x}_n)$:

   (a) Proof for $N = 1$:

$$\begin{aligned}
\mathbf{L}\left(\sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n\right) &= \mathbf{L}(\alpha_1 \boldsymbol{x}_1) && \text{by } N = 1 \text{ hypothesis}\\
&= \alpha_1 (\mathbf{L}\boldsymbol{x}_1) && \text{by \textit{homogeneous} property of } \mathbf{L} && \text{(Definition O.3 page 320)}
\end{aligned}$$

   (b) Proof that N case $\implies$ $N+1$ case:

$$\begin{aligned}
\mathbf{L}\left(\sum_{n=1}^{N+1} \alpha_n \boldsymbol{x}_n\right) &= \mathbf{L}\left(\alpha_{N+1}\boldsymbol{x}_{N+1} + \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n\right)\\
&= \mathbf{L}(\alpha_{N+1}\boldsymbol{x}_{N+1}) + \mathbf{L}\left(\sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n\right) && \text{by \textit{linearity} property of } \mathbf{L} && \text{(Definition O.3 page 320)}\\
&= \alpha_{N+1}\mathbf{L}(\boldsymbol{x}_{N+1}) + \sum_{n=1}^{N} \mathbf{L}(\alpha_n \boldsymbol{x}_n) && \text{by left } N+1 \text{ hypothesis}\\
&= \sum_{n=1}^{N+1} \mathbf{L}(\alpha_n \boldsymbol{x}_n)
\end{aligned}$$

✏

Theorem O.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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

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) $\mathcal{I}(\mathbf{L}) \triangleq \{\boldsymbol{y} \in \boldsymbol{Y} | \exists \boldsymbol{x} \in \boldsymbol{X} \quad \text{such that} \quad \boldsymbol{y} = \mathbf{L}\boldsymbol{x}\} \subseteq \boldsymbol{Y}$
   (c) $\boldsymbol{x} + \boldsymbol{y} \in \mathcal{I}(\mathbf{L}) \implies \exists \boldsymbol{v} \in \boldsymbol{X} \quad \text{such that} \quad \mathbf{L}\boldsymbol{v} = \boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x} \implies \boldsymbol{y} + \boldsymbol{x} \in \mathcal{I}(\mathbf{L})$
   (d) $\alpha \in \mathbb{F},\ \boldsymbol{x} \in \mathcal{I}(\mathbf{L}) \implies \exists \boldsymbol{x} \in \boldsymbol{X} \quad \text{such that} \quad \boldsymbol{y} = \mathbf{L}\boldsymbol{x} \implies \alpha\boldsymbol{y} = \alpha\mathbf{L}\boldsymbol{x} = \mathbf{L}(\alpha\boldsymbol{x}) \implies \alpha\boldsymbol{x} \in \mathcal{I}(\mathbf{L})$

✏

Example O.1. [7] Let $\mathcal{C}([a : b], \mathbb{R})$ be the set of all *continuous* functions from the closed real interval $[a : b]$ to $\mathbb{R}$.

EX $\mathcal{C}([a : b], \mathbb{R})$ is a linear space.

Theorem O.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

$$\begin{aligned} \mathbf{L}\boldsymbol{x} = \mathbf{L}\boldsymbol{y} \quad &\iff \quad \boldsymbol{x} - \boldsymbol{y} \in \mathcal{N}(\mathbf{L}) \\ \mathbf{L} \textit{ is } \text{INJECTIVE} \quad &\iff \quad \mathcal{N}(\mathbf{L}) = \{\mathbb{0}\} \end{aligned}$$

✎PROOF:

1. Proof that $\mathbf{L}\boldsymbol{x} = \mathbf{L}\boldsymbol{y} \implies \boldsymbol{x} - \boldsymbol{y} \in \mathcal{N}(\mathbf{L})$:

$$\begin{aligned} \mathbf{L}(\boldsymbol{x} - \boldsymbol{y}) &= \mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y} && \text{by Theorem O.1 page 320} \\ &= \mathbb{0} && \text{by left hypothesis} \\ &\implies \quad \boldsymbol{x} - \boldsymbol{y} \in \mathcal{N}(\mathbf{L}) && \text{by definition of } \textit{null space} \end{aligned}$$

2. Proof that $\mathbf{L}\boldsymbol{x} = \mathbf{L}\boldsymbol{y} \impliedby \boldsymbol{x} - \boldsymbol{y} \in \mathcal{N}(\mathbf{L})$:

$$\begin{aligned} \mathbf{L}\boldsymbol{y} &= \mathbf{L}\boldsymbol{y} + \mathbb{0} && \text{by definition of linear space (Definition G.1 page 209)} \\ &= \mathbf{L}\boldsymbol{y} + \mathbf{L}(\boldsymbol{x} - \boldsymbol{y}) && \text{by right hypothesis} \\ &= \mathbf{L}\boldsymbol{y} + (\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y}) && \text{by Theorem O.1 page 320} \\ &= (\mathbf{L}\boldsymbol{y} - \mathbf{L}\boldsymbol{y}) + \mathbf{L}\boldsymbol{x} && \text{by } \textit{associative} \text{ and } \textit{commutative} \text{ properties (Definition G.1 page 209)} \\ &= \mathbf{L}\boldsymbol{x} \end{aligned}$$

3. Proof that $\mathbf{L}$ is *injective* $\iff \mathcal{N}(\mathbf{L}) = \{\mathbb{0}\}$:

$$\begin{aligned} \mathbf{L} \text{ is } \textit{injective} &\iff \{(\mathbf{L}\boldsymbol{x} = \mathbf{L}\boldsymbol{y} \quad \iff \quad \boldsymbol{x} = \boldsymbol{y}) \quad \forall \boldsymbol{x}, \boldsymbol{y} \in X\} \\ &\iff \left\{\left[\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y} = \mathbb{0} \quad \iff \quad (\boldsymbol{x} - \boldsymbol{y}) = \mathbb{0}\right] \quad \forall \boldsymbol{x}, \boldsymbol{y} \in X\right\} \\ &\iff \left\{\left[\mathbf{L}(\boldsymbol{x} - \boldsymbol{y}) = \mathbb{0} \quad \iff \quad (\boldsymbol{x} - \boldsymbol{y}) = \mathbb{0}\right] \quad \forall \boldsymbol{x}, \boldsymbol{y} \in X\right\} \\ &\iff \mathcal{N}(\mathbf{L}) = \{\mathbb{0}\} \end{aligned}$$

[7] Eidelman et al. (2004) page 3
[8] Berberian (1961) page 88 ⟨Theorem IV.1.4⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

Theorem O.4. [9] *Let $\boldsymbol{W}$, $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ be linear spaces over a field $\mathbb{F}$.*

THM					
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.	$\mathbf{L}\big(\mathbf{M}\mathbin{\dot{+}}\mathbf{N}\big)$	$=$	$(\mathbf{LM})\mathbin{\dot{+}}(\mathbf{LN})$	$\forall \mathbf{L}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{M}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), \mathbf{N}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(LEFT DISTRIBUTIVE)
3.	$\big(\mathbf{L}\mathbin{\dot{+}}\mathbf{M}\big)\mathbf{N}$	$=$	$(\mathbf{LN})\mathbin{\dot{+}}(\mathbf{MN})$	$\forall \mathbf{L}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{M}\in\mathcal{L}(\boldsymbol{Y},\boldsymbol{Z}), \mathbf{N}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(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 $\mathbf{L}\big(\mathbf{M}\mathbin{\dot{+}}\mathbf{N}\big) = (\mathbf{LM})\mathbin{\dot{+}}(\mathbf{LN})$:

$$\begin{aligned}\big[\mathbf{L}\big(\mathbf{M}\mathbin{\dot{+}}\mathbf{N}\big)\big]\boldsymbol{x} &= \mathbf{L}\big[\big(\mathbf{M}\mathbin{\dot{+}}\mathbf{N}\big)\boldsymbol{x}\big] \\ &= \mathbf{L}\big[(\mathbf{M}\boldsymbol{x})\mathbin{\dot{+}}(\mathbf{N}\boldsymbol{x})\big] \\ &= [\mathbf{L}(\mathbf{M}\boldsymbol{x})]\mathbin{\dot{+}}[\mathbf{L}(\mathbf{N}\boldsymbol{x})] && \text{by } \textit{additive} \text{ property Definition O.3 page 320} \\ &= [(\mathbf{LM})\boldsymbol{x}]\mathbin{\dot{+}}[(\mathbf{LN})\boldsymbol{x}]\end{aligned}$$

3. Proof that $\big(\mathbf{L}\mathbin{\dot{+}}\mathbf{M}\big)\mathbf{N} = (\mathbf{LN})\mathbin{\dot{+}}(\mathbf{MN})$: 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 $\alpha(\mathbf{LM}) = \mathbf{L}(\alpha\mathbf{M})$:

$$\begin{aligned}[\alpha(\mathbf{LM})]\boldsymbol{x} &= \alpha[(\mathbf{LM})\boldsymbol{x}] \\ &= \mathbf{L}[\alpha(\mathbf{M}\boldsymbol{x})] && \text{by } \textit{homogeneous} \text{ property Definition O.3 page 320} \\ &= \mathbf{L}[(\alpha\mathbf{M})\boldsymbol{x}] \\ &= [\mathbf{L}(\alpha\mathbf{M})]\boldsymbol{x}\end{aligned}$$

Theorem O.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 ??).*

THM	
$\dim\mathcal{I}(\mathbf{L}) + \dim\mathcal{N}(\mathbf{L}) = \dim\boldsymbol{X}$	$\forall \mathbf{L}\in\boldsymbol{Y}^{\boldsymbol{X}}$

✎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⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

$\mathcal{N}(\mathbf{L})$.

$$\begin{aligned}
&\text{Let } p \triangleq \dim \boldsymbol{X}.\\
&\text{Let } n \triangleq \dim \mathcal{N}(\mathbf{L}).
\end{aligned}$$

$$\begin{aligned}
\dim \mathcal{I}(\mathbf{L}) &= \dim \{\boldsymbol{y} \in \boldsymbol{Y} | \exists \boldsymbol{x} \in \boldsymbol{X} \quad \text{such that} \quad \boldsymbol{y} = \mathbf{L}\boldsymbol{x}\}\\
&= \dim \left\{ \boldsymbol{y} \in \boldsymbol{Y} | \exists(\alpha_1, \alpha_2, \dots, \alpha_p) \quad \text{such that} \quad \boldsymbol{y} = \mathbf{L}\sum_{k=1}^{p} \alpha_k \psi_k \right\}\\
&= \dim \left\{ \boldsymbol{y} \in \boldsymbol{Y} | \exists(\alpha_1, \alpha_2, \dots, \alpha_p) \quad \text{such that} \quad \boldsymbol{y} = \sum_{k=1}^{p} \alpha_k \mathbf{L}\psi_k \right\}\\
&= \dim \left\{ \boldsymbol{y} \in \boldsymbol{Y} | \exists(\alpha_1, \alpha_2, \dots, \alpha_p) \quad \text{such that} \quad \boldsymbol{y} = \sum_{k=1}^{p-n} \alpha_k \mathbf{L}\psi_k + \sum_{k=1}^{n} \alpha_k \mathbf{L}\psi_k \right\}\\
&= \dim \left\{ \boldsymbol{y} \in \boldsymbol{Y} | \exists(\alpha_1, \alpha_2, \dots, \alpha_p) \quad \text{such that} \quad \boldsymbol{y} = \sum_{k=1}^{p-n} \alpha_k \mathbf{L}\psi_k + \mathbb{0} \right\}\\
&= p - n\\
&= \dim \boldsymbol{X} - \dim \mathcal{N}(\mathbf{L})
\end{aligned}$$

Note: This "proof" may be missing some necessary detail. ✏

O.2 Operators on Normed linear spaces

O.2.1 Operator norm

Definition O.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*

$$|||\mathbf{A}||| \triangleq \sup_{\boldsymbol{x}\in\boldsymbol{X}} \{\|\mathbf{A}\boldsymbol{x}\| \mid \|\boldsymbol{x}\| \le 1\} \qquad \forall \mathbf{A} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$$

*The pair $(\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}), |||\cdot|||)$ is the **normed space of linear operators** on $(\boldsymbol{X}, \boldsymbol{Y})$.*

Proposition O.1 (next) shows that the functional defined in Definition O.4 (previous) is a *norm* (Definition L.1 page 287).

Proposition O.1. [12] *Let $(\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}), |||\cdot|||)$ be the normed space of linear operators over the normed linear spaces $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ and $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$.*

PRP

*The functional $|||\cdot|||$ is a **norm** on $\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$. In particular,*

1.	$\|\|\|\mathbf{A}\|\|\| \ge 0$		$\forall \mathbf{A}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(NON-NEGATIVE)	*and*
2.	$\|\|\|\mathbf{A}\|\|\| = 0$	$\iff \mathbf{A} \mathrel{\hat{=}} \mathbb{0}$	$\forall \mathbf{A}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(NONDEGENERATE)	*and*
3.	$\|\|\|\alpha\mathbf{A}\|\|\| = \lvert\alpha\rvert \, \|\|\|\mathbf{A}\|\|\|$		$\forall \mathbf{A}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), \alpha\in\mathbb{F}$	(HOMOGENEOUS)	*and*
4.	$\|\|\|\mathbf{A} \mathbin{\dot{+}} \mathbf{B}\|\|\| \le \|\|\|\mathbf{A}\|\|\| + \|\|\|\mathbf{B}\|\|\|$		$\forall \mathbf{A}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y})$	(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] Rudin (1991) page 93

Copyright © 2019 Daniel J. Greenhoe

✎PROOF:

1. Proof that $|||\mathbf{A}||| > 0$ for $\mathbf{A} \neq \mathbb{0}$:

$$\begin{aligned}
|||\mathbf{A}||| &\triangleq \sup_{x\in X}\{\|\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition O.4 page 324)}\\
&> 0
\end{aligned}$$

2. Proof that $|||\mathbf{A}||| = 0$ for $\mathbf{A} \mathrel{\hat{=}} \mathbb{0}$:

$$\begin{aligned}
|||\mathbf{A}||| &\triangleq \sup_{x\in X}\{\|\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition O.4 page 324)}\\
&= \sup_{x\in X}\{\|\mathbb{0}x\| \mid \|x\| \le 1\}\\
&= 0
\end{aligned}$$

3. Proof that $|||\alpha\mathbf{A}||| = |\alpha|\,|||\mathbf{A}|||$:

$$\begin{aligned}
|||\alpha\mathbf{A}||| &\triangleq \sup_{x\in X}\{\|\alpha\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition O.4 page 324)}\\
&= \sup_{x\in X}\{|\alpha|\,\|\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition O.4 page 324)}\\
&= |\alpha| \sup_{x\in X}\{\|\mathbf{A}x\| \mid \|x\| \le 1\} && \text{by definition of sup}\\
&= |\alpha|\;\; |||\mathbf{A}||| && \text{by definition of } |||\cdot||| \text{ (Definition O.4 page 324)}
\end{aligned}$$

4. Proof that $\left|\left|\left|\mathbf{A} \mathbin{\dot{+}} \mathbf{B}\right|\right|\right| \le |||\mathbf{A}||| + |||\mathbf{B}|||$:

$$\begin{aligned}
\left|\left|\left|\mathbf{A} \mathbin{\dot{+}} \mathbf{B}\right|\right|\right| &\triangleq \sup_{x\in X}\left\{\left\|(\mathbf{A} \mathbin{\dot{+}} \mathbf{B})x\right\| \mid \|x\| \le 1\right\} && \text{by definition of } |||\cdot||| \text{ (Definition O.4 page 324)}\\
&= \sup_{x\in X}\{\|\mathbf{A}x + \mathbf{B}x\| \mid \|x\| \le 1\}\\
&\le \sup_{x\in X}\{\|\mathbf{A}x\| + \|\mathbf{B}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition O.4 page 324)}\\
&\le \sup_{x\in X}\{\|\mathbf{A}x\| \mid \|x\| \le 1\} + \sup_{x\in X}\{\|\mathbf{B}x\| \mid \|x\| \le 1\}\\
&\triangleq |||\mathbf{A}||| + |||\mathbf{B}||| && \text{by definition of } |||\cdot||| \text{ (Definition O.4 page 324)}
\end{aligned}$$

✏

Lemma O.1. *Let* $(\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}), |||\cdot|||)$ *be the normed space of linear operators over normed linear spaces* $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *and* $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$.

LEM

$$|||\mathbf{L}||| = \sup_{x}\{\|\mathbf{L}x\| \mid \|x\| = 1\} \qquad \forall x \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$$

✎PROOF: [13]

1. Proof that $\sup_x\{\|\mathbf{L}x\| \mid \|x\| \le 1\} \ge \sup_x\{\|\mathbf{L}x\| \mid \|x\| = 1\}$:

$$\sup_{x}\{\|\mathbf{L}x\| \mid \|x\| \le 1\} \ge \sup_{x}\{\|\mathbf{L}x\| \mid \|x\| = 1\} \qquad \text{because } A \subseteq B \implies \sup A \le \sup B$$

[13] 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.)

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

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 $\sup_x \{\|\mathbf{L}x\| \mid \|x\| \le 1\} \le \sup_x \{\|\mathbf{L}x\| \mid \|x\| = 1\}$:

$$\begin{aligned}
\sup_{x \in X} \{\|\mathbf{L}x\| \mid \|x\| \le 1\} &= \|\mathbf{L}y\| && \text{by definition of set } Y \\
&= \frac{\|y\|}{\|y\|} \|\mathbf{L}y\| \\
&= \|y\| \left\| \frac{1}{\|y\|} \mathbf{L}y \right\| && \text{by homogeneous property (page 287)} \\
&= \|y\| \left\| \mathbf{L} \frac{y}{\|y\|} \right\| && \text{by homogeneous property (page 320)} \\
&\le \|y\| \sup_{y \in Y} \left\{ \left\| \mathbf{L} \frac{y}{\|y\|} \right\| \right\} && \text{by definition of supremum} \\
&= \|y\| \sup_{y \in Y} \left\{ \left\| \mathbf{L} \frac{y}{\|y\|} \right\| \,\middle|\, \left\| \frac{y}{\|y\|} \right\| = 1 \right\} && \text{because } \left\| \frac{y}{\|y\|} \right\| = 1 \text{ for all } y \in Y \\
&\le \sup_{y \in Y} \left\{ \left\| \mathbf{L} \frac{y}{\|y\|} \right\| \,\middle|\, \left\| \frac{y}{\|y\|} \right\| = 1 \right\} && \text{because } 0 < \|y\| \le 1 \\
&\le \sup_{x \in X} \{\|\mathbf{L}x\| \mid \|x\| = 1\} && \text{because } \frac{y}{\|y\|} \in X \quad \forall y \in Y
\end{aligned}$$

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 O.2. [14] *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:

$$\begin{aligned}
|||\mathbf{I}||| &\triangleq \sup \{\|\mathbf{I}x\| \mid \|x\| \le 1\} && \text{by definition of } |||\cdot||| \text{ (Definition O.4 page 324)} \\
&= \sup \{\|x\| \mid \|x\| \le 1\} && \text{by definition of } \mathbf{I} \text{ (Definition O.2 page 319)} \\
&= 1
\end{aligned}$$

✏

Theorem O.6. [15] *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
$$\begin{aligned}
\|\mathbf{L}x\| &\le |||\mathbf{L}|||\ \|x\| && \forall \mathbf{L} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y}),\ x \in \boldsymbol{X} \\
|||\mathbf{KL}||| &\le |||\mathbf{K}|||\ |||\mathbf{L}||| && \forall \mathbf{K}, \mathbf{L} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})
\end{aligned}$$

[14] Michel and Herget (1993) page 410
[15] Rudin (1991) page 103, Aliprantis and Burkinshaw (1998) page 225

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

1. Proof that $\|\mathbf{L}\boldsymbol{x}\| \le |||\mathbf{L}|||\ \|\boldsymbol{x}\|$:

$$\begin{aligned}
\|\mathbf{L}\boldsymbol{x}\| &= \frac{\|\boldsymbol{x}\|}{\|\boldsymbol{x}\|}\|\mathbf{L}\boldsymbol{x}\| \\
&= \|\boldsymbol{x}\|\left\|\frac{1}{\|\boldsymbol{x}\|}\mathbf{L}\boldsymbol{x}\right\| && \text{by property of norms} \\
&= \|\boldsymbol{x}\|\left\|\mathbf{L}\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right\| && \text{by property of linear operators} \\
&\triangleq \|\boldsymbol{x}\|\ \|\mathbf{L}\boldsymbol{y}\| && \text{where } y \triangleq \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \\
&\le \|\boldsymbol{x}\| \sup_{y} \|\mathbf{L}\boldsymbol{y}\| && \text{by definition of supremum} \\
&= \|\boldsymbol{x}\| \sup_{y} \{\|\mathbf{L}\boldsymbol{y}\| \mid \|\boldsymbol{y}\| = 1\} && \text{because } \|\boldsymbol{y}\| = \left\|\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right\| = \frac{\|\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = 1 \\
&\triangleq \|\boldsymbol{x}\|\ \|\mathbf{L}\| && \text{by definition of operator norm}
\end{aligned}$$

2. Proof that $|||\mathbf{KL}||| \le |||\mathbf{K}|||\ |||\mathbf{L}|||$:

$$\begin{aligned}
|||\mathbf{KL}||| &\triangleq \sup_{x\in\boldsymbol{X}} \{\|(\mathbf{KL})\boldsymbol{x}\| \mid \|\boldsymbol{x}\| \le 1\} && \text{by Definition O.4 page 324 } (|||\cdot|||) \\
&= \sup_{x\in\boldsymbol{X}} \{\|\mathbf{K}(\mathbf{L}\boldsymbol{x})\| \mid \|\boldsymbol{x}\| \le 1\} \\
&\le \sup_{x\in\boldsymbol{X}} \{|||\mathbf{K}|||\ \ \|\mathbf{L}\boldsymbol{x}\| \mid \|\boldsymbol{x}\| \le 1\} && \text{by 1.} \\
&\le \sup_{x\in\boldsymbol{X}} \{|||\mathbf{K}|||\ \ |||\mathbf{L}|||\ \ \|\boldsymbol{x}\| \mid \|\boldsymbol{x}\| \le 1\} && \text{by 1.} \\
&= \sup_{x\in\boldsymbol{X}} \{|||\mathbf{K}|||\ \ |||\mathbf{L}|||\ \ 1 \mid \|\boldsymbol{x}\| \le 1\} && \text{by definition of sup} \\
&= |||\mathbf{K}|||\ \ |||\mathbf{L}||| && \text{by definition of sup}
\end{aligned}$$

✏

O.2.2 Bounded linear operators

Definition O.5. [16] *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 O.7. [17] *Let $(\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), |||\cdot|||)$ be the set of linear operators over normed linear spaces $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ and $\boldsymbol{Y} \triangleq \big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$.*

THM

The following conditions are all EQUIVALENT*:*

1. $\mathbf{L}$ *is continuous at* A SINGLE POINT $\boldsymbol{x}_0 \in X$ $\quad \forall \mathbf{L}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}) \quad \Longleftrightarrow$
2. $\mathbf{L}$ *is* CONTINUOUS *(at every point $\boldsymbol{x} \in X$)* $\quad \forall \mathbf{L}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}) \quad \Longleftrightarrow$
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 $\|\mathbf{L}\boldsymbol{x}\| \le M\ \|\boldsymbol{x}\|$ $\quad \forall \mathbf{L}\in\mathcal{L}(\boldsymbol{X},\boldsymbol{Y}), \boldsymbol{x}\in X$

[16] Rudin (1991) pages 92–93
[17] Aliprantis and Burkinshaw (1998) page 227

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✎PROOF:

1. Proof that 1 $\Longrightarrow$ 2:

$$\begin{aligned}
\epsilon &> \left\|\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{x}_0\right\| && \text{by hypothesis 1}\\
&= \left\|\mathbf{L}(\boldsymbol{x} - \boldsymbol{x}_0)\right\| && \text{by linearity (Definition O.3 page 320)}\\
&= \left\|\mathbf{L}(\boldsymbol{x} + \boldsymbol{y} - \boldsymbol{x}_0 - \boldsymbol{y})\right\| && \\
&= \left\|\mathbf{L}(\boldsymbol{x} + \boldsymbol{y}) - \mathbf{L}(\boldsymbol{x}_0 + \boldsymbol{y})\right\| && \text{by linearity (Definition O.3 page 320)}\\
&\Longrightarrow \mathbf{L} \text{ is continuous at point } \boldsymbol{x} + \boldsymbol{y} && \\
&\Longrightarrow \mathbf{L} \text{ is continuous at every point in } X && \text{(hypothesis 2)}
\end{aligned}$$

2. Proof that 2 $\Longrightarrow$ 1: obvious:

3. Proof that 4 $\Longrightarrow$ 2:[18]

$$\begin{aligned}
|||\mathbf{L}\boldsymbol{x}||| \le M\,\|\boldsymbol{x}\| &\Longrightarrow |||\mathbf{L}(\boldsymbol{x} - \boldsymbol{y})||| \le M\,\|\boldsymbol{x} - \boldsymbol{y}\| && \text{by hypothesis 4}\\
&\Longrightarrow |||\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y}||| \le M\,\|\boldsymbol{x} - \boldsymbol{y}\| && \text{by linearity of } \mathbf{L} \text{ (Definition O.3 page 320)}\\
&\Longrightarrow |||\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y}||| \le \epsilon \text{ whenever } M\,\|\boldsymbol{x} - \boldsymbol{y}\| < \epsilon && \\
&\Longrightarrow |||\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y}||| \le \epsilon \text{ whenever } \|\boldsymbol{x} - \boldsymbol{y}\| < \frac{\epsilon}{M} && \text{(hypothesis 2)}
\end{aligned}$$

4. Proof that 3 $\Longrightarrow$ 4:

$$\begin{aligned}
\|\mathbf{L}\boldsymbol{x}\| &\le \underbrace{|||\mathbf{L}|||}_{M}\ \|\boldsymbol{x}\| && \text{by Theorem O.6 page 326}\\
&= M\ \|\boldsymbol{x}\| && \text{where } M \triangleq |||\mathbf{L}||| < \infty \text{ (by hypothesis 1)}
\end{aligned}$$

5. Proof that 1 $\Longrightarrow$ 3:[19]

$$\begin{aligned}
|||\mathbf{L}||| = \infty &\Longrightarrow \{\|\mathbf{L}\boldsymbol{x}\| \mid \|\boldsymbol{x}\| \le 1\} = \infty\\
&\Longrightarrow \exists\,(\boldsymbol{x}_n) \quad \text{such that} \quad \left\|\boldsymbol{x}_n\right\| = 1 \text{ and } |||\mathbf{L}||| = \left\{\left\|\mathbf{L}\boldsymbol{x}_n\right\| \mid \left\|\boldsymbol{x}_n\right\| \le 1\right\} = \infty\\
&\Longrightarrow \left\|\boldsymbol{x}_n\right\| = 1 \text{ and } \infty = |||\mathbf{L}||| = \left\|\mathbf{L}\boldsymbol{x}_n\right\|\\
&\Longrightarrow \left\|\boldsymbol{x}_n\right\| = 1 \text{ and } \left\|\mathbf{L}\boldsymbol{x}_n\right\| \ge n\\
&\Longrightarrow \frac{1}{n}\left\|\boldsymbol{x}_n\right\| = \frac{1}{n} \text{ and } \frac{1}{n}\left\|\mathbf{L}\boldsymbol{x}_n\right\| \ge 1\\
&\Longrightarrow \left\|\frac{\boldsymbol{x}_n}{n}\right\| = \frac{1}{n} \text{ and } \left\|\mathbf{L}\frac{\boldsymbol{x}_n}{n}\right\| \ge 1\\
&\Longrightarrow \lim_{n\to\infty}\left\|\frac{\boldsymbol{x}_n}{n}\right\| = 0 \text{ and } \lim_{n\to\infty}\left\|\mathbf{L}\frac{\boldsymbol{x}_n}{n}\right\| \ge 1\\
&\Longrightarrow \mathbf{L} \text{ is not continuous at } 0
\end{aligned}$$

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*).

✏

[18] Bollobás (1999), page 29
[19] Aliprantis and Burkinshaw (1998), page 227

COPYRIGHT © 2019 DANIEL J. GREENHOE

O.2.3 Adjoints on normed linear spaces

Definition O.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 *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 O.8. [20] *Let $\mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$ be the space of* BOUNDED LINEAR OPERATORS *on* NORMED LINEAR SPACES *$\boldsymbol{X}$ and $\boldsymbol{Y}$.*

THM

$$\begin{aligned}
(\mathbf{A} \mathbin{\mathring{+}} \mathbf{B})^* &= \mathbf{A}^* \mathbin{\mathring{+}} \mathbf{B}^* && \forall \mathbf{A},\mathbf{B} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{Y}) \\
(\lambda\mathbf{A})^* &= \lambda\mathbf{A}^* && \forall \mathbf{A},\mathbf{B} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{Y}) \\
(\mathbf{A}\mathbf{B})^* &= \mathbf{B}^*\mathbf{A}^* && \forall \mathbf{A},\mathbf{B} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{Y})
\end{aligned}$$

✎PROOF:

$$\begin{aligned}
\left[\mathbf{A} \mathbin{\mathring{+}} \mathbf{B}\right]^*\mathsf{f}(\boldsymbol{x}) &= \mathsf{f}\left(\left[\mathbf{A} \mathbin{\mathring{+}} \mathbf{B}\right]\boldsymbol{x}\right) && \text{by definition of adjoint} && \text{(Definition O.6 page 329)} \\
&= \mathsf{f}(\mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{x}) && \text{by definition of linear operators} && \text{(Definition O.3 page 320)} \\
&= \mathsf{f}(\mathbf{A}\boldsymbol{x}) + \mathsf{f}(\mathbf{B}\boldsymbol{x}) && \text{by definition of } \textit{linear functional} \\
&= \mathbf{A}^*\mathsf{f}(\boldsymbol{x}) + \mathbf{B}^*\mathsf{f}(\boldsymbol{x}) && \text{by definition of } \textit{adjoint} && \text{(Definition O.6 page 329)} \\
&= \left[\mathbf{A}^* + \mathbf{B}^*\right]\mathsf{f}(\boldsymbol{x}) && \text{by definition of } \textit{linear functional}
\end{aligned}$$

$$\begin{aligned}
[\lambda\mathbf{A}]^*\mathsf{f}(\boldsymbol{x}) &= \mathsf{f}([\lambda\mathbf{A}]\boldsymbol{x}) && \text{by definition of } \textit{adjoint} && \text{(Definition O.6 page 329)} \\
&= \lambda\mathsf{f}(\mathbf{A}\boldsymbol{x}) && \text{by definition of } \textit{linear functional} \\
&= \left[\lambda\mathbf{A}^*\right]\mathsf{f}(\boldsymbol{x}) && \text{by definition of } \textit{adjoint} && \text{(Definition O.6 page 329)}
\end{aligned}$$

$$\begin{aligned}
[\mathbf{A}\mathbf{B}]^*\mathsf{f}(\boldsymbol{x}) &= \mathsf{f}([\mathbf{A}\mathbf{B}]\boldsymbol{x}) && \text{by definition of } \textit{adjoint} && \text{(Definition O.6 page 329)} \\
&= \mathsf{f}(\mathbf{A}[\mathbf{B}\boldsymbol{x}]) && \text{by definition of } \textit{linear operators} && \text{(Definition O.3 page 320)} \\
&= \left[\mathbf{A}^*\mathsf{f}\right](\mathbf{B}\boldsymbol{x}) && \text{by definition of } \textit{adjoint} && \text{(Definition O.6 page 329)} \\
&= \mathbf{B}^*\left[\mathbf{A}^*\mathsf{f}\right](\boldsymbol{x}) && \text{by definition of } \textit{adjoint} && \text{(Definition O.6 page 329)} \\
&= \left[\mathbf{B}^*\mathbf{A}^*\right]\mathsf{f}(\boldsymbol{x}) && \text{by definition of } \textit{adjoint} && \text{(Definition O.6 page 329)}
\end{aligned}$$

✏

Theorem O.9. [21] *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

$$|||\mathbf{B}||| = \left|\left|\left|\mathbf{B}^*\right|\right|\right| \qquad \forall \mathbf{B} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{Y})$$

✎PROOF:

$$\begin{aligned}
|||\mathbf{B}||| &\triangleq \sup\left\{\|\mathbf{B}\boldsymbol{x}\| \mid \|\boldsymbol{x}\| \le 1\right\} && \text{by Definition O.4 page 324} \\
&\overset{?}{=} \sup\left\{|\mathsf{g}(\mathbf{B}\boldsymbol{x};\boldsymbol{y}^*)| \mid \|\boldsymbol{x}\| \le 1,\ \left\|\boldsymbol{y}^*\right\| \le 1\right\} \\
&= \sup\left\{|\mathsf{f}(\boldsymbol{x};\mathbf{B}^*\boldsymbol{y}^*)| \mid \|\boldsymbol{x}\| \le 1,\ \left\|\boldsymbol{y}^*\right\| \le 1\right\} \\
&\triangleq \sup\left\{\left\|\mathbf{B}^*\boldsymbol{y}^*\right\| \mid \|\boldsymbol{x}\| \le 1,\ \left\|\boldsymbol{y}^*\right\| \le 1\right\} \\
&= \sup\left\{\left\|\mathbf{B}^*\boldsymbol{y}^*\right\| \mid \left\|\boldsymbol{y}^*\right\| \le 1\right\} \\
&\triangleq \left|\left|\left|\mathbf{B}^*\right|\right|\right| && \text{by Definition O.4 page 324}
\end{aligned}$$

✏

[20] Bollobás (1999), page 156
[21] Rudin (1991) page 98

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

O.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 [22]

Theorem O.10 (Mazur-Ulam theorem)**.** [23] *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

$$\left.\begin{array}{ll} 1. & \underbrace{\phi^{-1}\phi = \phi\phi^{-1} = \mathbf{I}}_{\textit{bijective}} \quad \textit{and} \\ 2. & \underbrace{\|\phi\boldsymbol{x} - \phi\boldsymbol{y}\|_{\boldsymbol{Y}} = \|\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{X}} \quad \forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{X}}_{\textit{isometric}} \end{array}\right\} \Longrightarrow \underbrace{\phi([1-\lambda]\boldsymbol{x} + \lambda\boldsymbol{y}) = [1-\lambda]\phi\boldsymbol{x} + \lambda\phi\boldsymbol{y} \forall \lambda \in \mathbb{R}}_{\textit{affine}}$$

✎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 $\hat{\boldsymbol{y}} \triangleq \mathrm{g}^{-1}\boldsymbol{y}$.

$$\begin{aligned} \left\|\mathrm{g}^{-1}\boldsymbol{x} - \mathrm{g}^{-1}\boldsymbol{y}\right\| &= \|\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}\| && \text{by definition of } \hat{\boldsymbol{x}} \text{ and } \hat{\boldsymbol{y}} \\ &= \|\mathrm{g}\hat{\boldsymbol{x}} - \mathrm{g}\hat{\boldsymbol{y}}\| && \text{by left hypothesis} \\ &= \left\|\mathrm{g}\mathrm{g}^{-1}\boldsymbol{x} - \mathrm{g}\mathrm{g}^{-1}\boldsymbol{y}\right\| && \text{by definition of } \hat{\boldsymbol{x}} \text{ and } \hat{\boldsymbol{y}} \\ &= \|\boldsymbol{x} - \boldsymbol{y}\| && \text{by definition of } \mathrm{g}^{-1} \end{aligned}$$

[22] quote: Ulam (1991), page 33
image: http://www-history.mcs.st-andrews.ac.uk/Biographies/Ulam.html

[23] Oikhberg and Rosenthal (2007), page 598, Väisälä (2003), page 634, Giles (2000), page 11, Dunford and Schwartz (1957), page 91, Mazur and Ulam (1932)

COPYRIGHT © 2019 DANIEL J. GREENHOE

4. Proof that $\mathrm{g}z = z$:

$$\begin{aligned}
2\lambda &= 2\sup\{\|\mathrm{g}z - z\|\} && \text{by definition of } \lambda \text{ item (2)}\\
&\le 2\|\mathrm{g}z - z\| && \text{by definition of sup}\\
&= \|2z - 2\mathrm{g}z\| \\
&= \|\psi\mathrm{g}z - \mathrm{g}z\| && \text{by definition of } \psi \text{ item (1)}\\
&= \left\|\mathrm{g}^{-1}\psi\mathrm{g}z - \mathrm{g}^{-1}\mathrm{g}z\right\| && \text{by item (3)}\\
&= \left\|\mathrm{g}^{-1}\psi\mathrm{g}z - z\right\| && \text{by definition of } \mathrm{g}^{-1}\\
&= \left\|\psi\mathrm{g}^{-1}\psi\mathrm{g}z - z\right\| \\
&= \left\|\mathrm{g}^{*}z - z\right\| \\
&\le \lambda && \text{by definition of } \lambda \text{ item (2)}\\
&\implies 2\lambda \le \lambda \\
&\implies \lambda = 0 \\
&\implies \mathrm{g}z = z
\end{aligned}$$

5. Proof that $\phi\left(\frac{1}{2}\boldsymbol{x} + \frac{1}{2}\boldsymbol{y}\right) = \frac{1}{2}\phi\boldsymbol{x} + \frac{1}{2}\phi\boldsymbol{y}$:

$$\begin{aligned}
\phi\left(\frac{1}{2}\boldsymbol{x} + \frac{1}{2}\boldsymbol{y}\right) &= \\
&= \frac{1}{2}\phi\boldsymbol{x} + \frac{1}{2}\phi\boldsymbol{y}
\end{aligned}$$

6. Proof that $\phi([1-\lambda]\boldsymbol{x} + \lambda\boldsymbol{y}) = [1-\lambda]\phi\boldsymbol{x} + \lambda\phi\boldsymbol{y}$:

$$\begin{aligned}
\phi([1-\lambda]\boldsymbol{x} + \lambda\boldsymbol{y}) &= \\
&= [1-\lambda]\phi\boldsymbol{x} + \lambda\phi\boldsymbol{y}
\end{aligned}$$

✍

Theorem O.11 (Neumann Expansion Theorem)**.** [24] *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

$$\left.\begin{array}{ll} 1. & \mathbf{A} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{X}) \quad (\mathbf{A} \text{ is bounded}) \\ 2. & |||\mathbf{A}||| < 1 \end{array}\right\} \implies \left\{\begin{array}{lll} 1. & (\mathbf{I} - \mathbf{A})^{-1} & \text{exists} \\ 2. & \left|\left|\left|(\mathbf{I} - \mathbf{A})^{-1}\right|\right|\right| \le & \dfrac{1}{1 - |||\mathbf{A}|||} \\ 3. & (\mathbf{I} - \mathbf{A})^{-1} = & \displaystyle\sum_{n=0}^{\infty} \mathbf{A}^n \\ & \text{with uniform convergence} & \end{array}\right.$$

O.3 Operators on Inner product spaces

O.3.1 General Results

Theorem O.12. [25] *Let* $\mathbf{A}, \mathbf{B} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{X})$ *be* BOUNDED LINEAR OPERATORS *on an inner product space* $\boldsymbol{X} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \langle \triangle \,|\, \nabla \rangle\right)$.

THM

$$\begin{array}{lllllllll} \langle \mathbf{B}\boldsymbol{x} \,|\, \boldsymbol{x} \rangle & = & 0 & \forall x \in X & \iff & \mathbf{B}\boldsymbol{x} & = & \mathbb{0} & \forall x \in X \\ \langle \mathbf{A}\boldsymbol{x} \,|\, \boldsymbol{x} \rangle & = & \langle \mathbf{B}\boldsymbol{x} \,|\, \boldsymbol{x} \rangle & \forall x \in X & \iff & \mathbf{A} & = & \mathbf{B} & \end{array}$$

[24] Michel and Herget (1993) page 415
[25] Rudin (1991) page 310 ⟨Theorem 12.7, Corollary⟩

CC BY-NC-ND

✎PROOF:

1. Proof that $\langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{x}\rangle = 0 \implies \mathbf{B}\boldsymbol{x} = \mathbb{0}$:

$$\begin{aligned}
0 &= \langle \mathbf{B}(\boldsymbol{x}+\mathbf{B}\boldsymbol{x}) \mid (\boldsymbol{x}+\mathbf{B}\boldsymbol{x})\rangle + i\,\langle \mathbf{B}(\boldsymbol{x}+i\mathbf{B}\boldsymbol{x}) \mid (\boldsymbol{x}+i\mathbf{B}\boldsymbol{x})\rangle && \text{by left hypothesis}\\
&= \left\{\left\langle \mathbf{B}\boldsymbol{x}+\mathbf{B}^2\boldsymbol{x}) \mid \boldsymbol{x}+\mathbf{B}\boldsymbol{x}\right\rangle\right\} + i\left\{\left\langle \mathbf{B}\boldsymbol{x}+i\mathbf{B}^2\boldsymbol{x}) \mid \boldsymbol{x}+i\mathbf{B}\boldsymbol{x}\right\rangle\right\} && \text{by Definition O.3 page 320}\\
&= \left\{\langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{x}\rangle + \langle \mathbf{B}\boldsymbol{x} \mid \mathbf{B}\boldsymbol{x}\rangle + \left\langle \mathbf{B}^2\boldsymbol{x} \mid \boldsymbol{x}\right\rangle + \left\langle \mathbf{B}^2\boldsymbol{x} \mid \mathbf{B}\boldsymbol{x}\right\rangle\right\} && \text{by Definition K.1 page 271}\\
&\quad + i\left\{\langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{x}\rangle - i\,\langle \mathbf{B}\boldsymbol{x} \mid \mathbf{B}\boldsymbol{x}\rangle + i\left\langle \mathbf{B}^2\boldsymbol{x} \mid \boldsymbol{x}\right\rangle - i^2\left\langle \mathbf{B}^2\boldsymbol{x} \mid \mathbf{B}\boldsymbol{x}\right\rangle\right\}\\
&= \left\{0 + \langle \mathbf{B}\boldsymbol{x} \mid \mathbf{B}\boldsymbol{x}\rangle + \left\langle \mathbf{B}^2\boldsymbol{x} \mid \boldsymbol{x}\right\rangle + 0\right\} + i\left\{0 - i\,\langle \mathbf{B}\boldsymbol{x} \mid \mathbf{B}\boldsymbol{x}\rangle + i\left\langle \mathbf{B}^2\boldsymbol{x} \mid \boldsymbol{x}\right\rangle - i^2 0\right\} && \text{by left hypothesis}\\
&= \left\{\langle \mathbf{B}\boldsymbol{x} \mid \mathbf{B}\boldsymbol{x}\rangle + \left\langle \mathbf{B}^2\boldsymbol{x} \mid \boldsymbol{x}\right\rangle\right\} + \left\{\langle \mathbf{B}\boldsymbol{x} \mid \mathbf{B}\boldsymbol{x}\rangle - \left\langle \mathbf{B}^2\boldsymbol{x} \mid \boldsymbol{x}\right\rangle\right\}\\
&= 2\,\langle \mathbf{B}\boldsymbol{x} \mid \mathbf{B}\boldsymbol{x}\rangle\\
&= 2\,\|\mathbf{B}\boldsymbol{x}\|^2\\
&\implies \mathbf{B}\boldsymbol{x} = \mathbb{0} && \text{by Definition L.1 page 287}
\end{aligned}$$

2. Proof that $\langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{x}\rangle = 0 \impliedby \mathbf{B}\boldsymbol{x} = \mathbb{0}$: by property of inner products (Theorem K.1 page 271).

3. Proof that $\langle \mathbf{A}\boldsymbol{x} \mid \boldsymbol{x}\rangle = \langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{x}\rangle \implies \mathbf{A} \mathrel{\hat{=}} \mathbf{B}$:

$$\begin{aligned}
0 &= \langle \mathbf{A}\boldsymbol{x} \mid \boldsymbol{x}\rangle - \langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by left hypothesis}\\
&= \langle \mathbf{A}\boldsymbol{x} - \mathbf{B}\boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by \textit{additivity} property of } \langle \triangle \mid \triangledown\rangle \text{ (Definition K.1 page 271)}\\
&= \langle (\mathbf{A}-\mathbf{B})\boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by definition of operator addition}\\
\implies (\mathbf{A}-\mathbf{B})\boldsymbol{x} &= \mathbb{0} && \text{by item 1}\\
\implies \mathbf{A} &= \mathbf{B} && \text{by definition of operator subtraction}
\end{aligned}$$

4. Proof that $\langle \mathbf{A}\boldsymbol{x} \mid \boldsymbol{x}\rangle = \langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{x}\rangle \impliedby \mathbf{A} \mathrel{\hat{=}} \mathbf{B}$:

$$\langle \mathbf{A}\boldsymbol{x} \mid \boldsymbol{x}\rangle = \langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{x}\rangle \qquad \text{by } \mathbf{A} \mathrel{\hat{=}} \mathbf{B} \text{ hypothesis}$$

✏

O.3.2 Operator adjoint

A fundamental concept of operators on inner product spaces is the *operator adjoint* (Proposition O.3 page 332). The adjoint of an operator is a kind of generalization of the conjugate of a complex number in that

- Both are *star-algebras* (Theorem O.13 page 333).
- Both support decomposition into "real" and "imaginary" parts (Theorem J.3 page 268).

Structurally, the operator adjoint provides a convenient symmetric relationship between the *range space* and *null space* of an operator (Theorem O.14 page 334).

Proposition O.3. [26] *Let $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$ be the space of* BOUNDED LINEAR OPERATORS *(Definition O.5 page 327) on a* HILBERT SPACE $\boldsymbol{H}$.

PRP *An operator $\mathbf{B}^*$ is the **adjoint** of $\mathbf{B} \in \mathcal{B}(\boldsymbol{H},\boldsymbol{H})$ if*
$$\langle \mathbf{B}\boldsymbol{x} \mid \boldsymbol{y}\rangle = \langle \boldsymbol{x} \mid \mathbf{B}^*\boldsymbol{y}\rangle \qquad \forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{H}.$$

[26] Michel and Herget (1993) page 220, Rudin (1991) page 311, Giles (2000), page 182, von Neumann (1929) page 49, Stone (1932) page 41

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

1. For fixed y, $f(x) \triangleq \langle x \mid y\rangle$ is a *functional* in $\mathbb{F}^X$.
2. $\mathbf{B}^*$ is the *adjoint* of $\mathbf{B}$ because

$$\begin{aligned}\langle \mathbf{B}x \mid y\rangle &\triangleq f(\mathbf{B}x)\\ &\triangleq \mathbf{B}^* f(x) && \text{by definition of } \textit{operator adjoint} && \text{(Definition O.6 page 329)}\\ &= \langle x \mid \mathbf{B}^* y\rangle\end{aligned}$$

✏

Example O.2.

EX

In matrix algebra ("linear algebra")

- The inner product operation $\langle x \mid y\rangle$ is represented by $y^H x$.
- The linear operator is represented as a matrix A.
- The operation of A on a vector x is represented as Ax.
- The adjoint of matrix A is the Hermitian matrix A^H.

✎PROOF:

$$\langle Ax \mid y\rangle \triangleq y^H Ax = [(Ax)^H y]^H = [x^H A^H y]^H = (A^H y)^H x \triangleq \left\langle x \mid A^H y\right\rangle$$

✏

Structures that satisfy the four conditions of the next theorem are known as **-algebras* ("*star-algebras*" (Definition J.3 page 266). Other structures which are *-algebras include the *field of complex numbers* $\mathbb{C}$ and any *ring of complex square* $n \times n$ *matrices*.[27]

Theorem O.13 (operator star-algebra). [28] *Let* H *be a* HILBERT SPACE *with operators* $\mathbf{A}, \mathbf{B} \in \mathcal{B}(H, H)$ *and with adjoints* $\mathbf{A}^*, \mathbf{B}^* \in \mathcal{B}(H, H)$. *Let* $\bar{\alpha}$ *be the complex conjugate of some* $\alpha \in \mathbb{C}$.

THM

The pair $(H, *)$ *is a* $*$-ALGEBRA (STAR-ALGEBRA). *In particular,*

1.	$(\mathbf{A} \mathbin{\mathring{+}} \mathbf{B})^*$	$=$	$\mathbf{A}^* + \mathbf{B}^*$	$\forall \mathbf{A},\mathbf{B} \in H$	(DISTRIBUTIVE)	*and*
2.	$(\alpha\mathbf{A})^*$	$=$	$\bar{\alpha}\mathbf{A}^*$	$\forall \mathbf{A},\mathbf{B} \in H$	(CONJUGATE LINEAR)	*and*
3.	$(\mathbf{AB})^*$	$=$	$\mathbf{B}^*\mathbf{A}^*$	$\forall \mathbf{A},\mathbf{B} \in H$	(ANTIAUTOMORPHIC)	*and*
4.	$\mathbf{A}^{**}$	$=$	$\mathbf{A}$	$\forall \mathbf{A},\mathbf{B} \in H$	(INVOLUTARY)	

✎PROOF:

$$\begin{aligned}\left\langle x \mid (\mathbf{A} \mathbin{\mathring{+}} \mathbf{B})^* y\right\rangle &= \left\langle (\mathbf{A} \mathbin{\mathring{+}} \mathbf{B})x \mid y\right\rangle && \text{by definition of adjoint} && \text{(Proposition O.3 page 332)}\\ &= \langle \mathbf{A}x \mid y\rangle + \langle \mathbf{B}x \mid y\rangle && \text{by definition of inner product} && \text{(Definition K.1 page 271)}\\ &= \langle x \mid \mathbf{A}^* y\rangle + \langle x \mid \mathbf{B}^* y\rangle && \text{by definition of operator addition}\\ &= \langle x \mid \mathbf{A}^* y + \mathbf{B}^* y\rangle && \text{by definition of inner product} && \text{(Definition K.1 page 271)}\\ &= \langle x \mid (\mathbf{A}^* + \mathbf{B}^*) y\rangle && \text{by definition of operator addition}\end{aligned}$$

$$\begin{aligned}\langle x \mid (\alpha\mathbf{A})^* y\rangle &= \langle (\alpha\mathbf{A})x \mid y\rangle && \text{by definition of adjoint} && \text{(Proposition O.3 page 332)}\\ &= \langle \alpha(\mathbf{A}x) \mid y\rangle && \text{by definition of scalar multiplication}\\ &= \alpha \langle \mathbf{A}x \mid y\rangle && \text{by definition of inner product} && \text{(Definition K.1 page 271)}\end{aligned}$$

[27] Sakai (1998) page 1
[28] Halmos (1998a), pages 39–40, Rudin (1991) page 311

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

$$
\begin{aligned}
&= \alpha\,\langle \boldsymbol{x} \mid \mathbf{A}^*\boldsymbol{y}\rangle && \text{by definition of adjoint} && \text{(Proposition O.3 page 332)}\\
&= \langle \boldsymbol{x} \mid \alpha^*\mathbf{A}^*\boldsymbol{y}\rangle && \text{by definition of inner product} && \text{(Definition K.1 page 271)}
\end{aligned}
$$

$$
\begin{aligned}
\langle \boldsymbol{x} \mid (\mathbf{AB})^*\boldsymbol{y}\rangle &= \langle (\mathbf{AB})\boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by definition of adjoint} && \text{(Proposition O.3 page 332)}\\
&= \langle \mathbf{A}(\mathbf{B}\boldsymbol{x}) \mid \boldsymbol{y}\rangle && \text{by definition of operator multiplication} && \\
&= \langle (\mathbf{B}\boldsymbol{x}) \mid \mathbf{A}^*\boldsymbol{y}\rangle && \text{by definition of adjoint} && \text{(Proposition O.3 page 332)}\\
&= \langle \boldsymbol{x} \mid \mathbf{B}^*\mathbf{A}^*\boldsymbol{y}\rangle && \text{by definition of adjoint} && \text{(Proposition O.3 page 332)}
\end{aligned}
$$

$$
\begin{aligned}
\langle \boldsymbol{x} \mid \mathbf{A}^{**}\boldsymbol{y}\rangle &= \langle \mathbf{A}^*\boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by definition of adjoint} && \text{(Proposition O.3 page 332)}\\
&= \langle \boldsymbol{y} \mid \mathbf{A}^*\boldsymbol{x}\rangle^* && \text{by definition of inner product} && \text{(Definition K.1 page 271)}\\
&= \langle \mathbf{A}\boldsymbol{y} \mid \boldsymbol{x}\rangle^* && \text{by definition of adjoint} && \text{(Proposition O.3 page 332)}\\
&= \langle \boldsymbol{x} \mid \mathbf{A}\boldsymbol{y}\rangle && \text{by definition of inner product} && \text{(Definition K.1 page 271)}
\end{aligned}
$$

✍

Theorem O.14. [29] *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:

$$
\begin{aligned}
\mathcal{I}(\mathbf{A}^*)^\perp &= \left\{y \in \boldsymbol{H} \mid \langle \boldsymbol{y} \mid \boldsymbol{u}\rangle = 0 \quad \forall u \in \mathcal{I}(\mathbf{A}^*)\right\}\\
&= \left\{y \in \boldsymbol{H} \mid \langle \boldsymbol{y} \mid \mathbf{A}^*\boldsymbol{x}\rangle = 0 \quad \forall \boldsymbol{x} \in \boldsymbol{H}\right\}\\
&= \{y \in \boldsymbol{H} \mid \langle \mathbf{A}\boldsymbol{y} \mid \boldsymbol{x}\rangle = 0 \quad \forall \boldsymbol{x} \in \boldsymbol{H}\} && \text{by definition of } \mathbf{A}^* && \text{(Proposition O.3 page 332)}\\
&= \{y \in \boldsymbol{H} \mid \mathbf{A}\boldsymbol{y} = 0\}\\
&= \mathcal{N}(\mathbf{A}) && \text{by definition of } \mathcal{N}(\mathbf{A})
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{I}(\mathbf{A})^\perp &= \{y \in \boldsymbol{H} \mid \langle \boldsymbol{y} \mid \boldsymbol{u}\rangle = 0 \quad \forall u \in \mathcal{I}(\mathbf{A})\}\\
&= \{y \in \boldsymbol{H} \mid \langle \boldsymbol{y} \mid \mathbf{A}\boldsymbol{x}\rangle = 0 \quad \forall \boldsymbol{x} \in \boldsymbol{H}\} && \text{by definition of } \mathcal{I}\\
&= \left\{y \in \boldsymbol{H} \mid \langle \mathbf{A}^*\boldsymbol{y} \mid \boldsymbol{x}\rangle = 0 \quad \forall \boldsymbol{x} \in \boldsymbol{H}\right\} && \text{by definition of } \mathbf{A}^* && \text{(Proposition O.3 page 332)}\\
&= \left\{y \in \boldsymbol{H} \mid \mathbf{A}^*\boldsymbol{y} = 0\right\}\\
&= \mathcal{N}(\mathbf{A}^*) && \text{by definition of } \mathcal{N}(\mathbf{A})
\end{aligned}
$$

✍

O.4 Special Classes of Operators

O.4.1 Projection operators

Definition O.7. [30] *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 $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$.*

[29] Rudin (1991) page 312

[30] Rudin (1991) page 133 ⟨5.15 Projections⟩, Kubrusly (2001) page 70, Bachman and Narici (1966) page 6, Halmos (1958) page 73 ⟨§41. Projections⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

DEF $\mathbf{P}$ *is a* ***projection*** *operator if* $\mathbf{P}^2 = \mathbf{P}$.

Theorem O.15. [31] *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

$$\left.\begin{array}{llll} 1. & \mathbf{P}^2 = \mathbf{P} & (\mathbf{P} \text{ is a projection operator}) & \text{and} \\ 2. & \boldsymbol{\Omega} = \boldsymbol{X} \hat{+} \boldsymbol{Y} & (\boldsymbol{Y} \text{ compliments } \boldsymbol{X} \text{ in } \boldsymbol{\Omega}) & \text{and} \\ 3. & \mathbf{P}\boldsymbol{\Omega} = \boldsymbol{X} & (\mathbf{P} \text{ projects onto } \boldsymbol{X}) & \end{array}\right\} \Longrightarrow \left\{\begin{array}{llll} 1. & \mathcal{I}(\mathbf{P}) = \boldsymbol{X} & & \text{and} \\ 2. & \mathcal{N}(\mathbf{P}) = \boldsymbol{Y} & & \text{and} \\ 3. & \boldsymbol{\Omega} = \mathcal{I}(\mathbf{P}) \hat{+} \mathcal{N}(\mathbf{P}) & & \end{array}\right.$$

✎PROOF:

$$\begin{aligned} \mathcal{I}(\mathbf{P}) &= \mathbf{P}\boldsymbol{\Omega} \\ &= \mathbf{P}(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2) \\ &= \mathbf{P}\boldsymbol{\Omega}_1 + \mathbf{P}\boldsymbol{\Omega}_2 \\ &= \boldsymbol{\Omega}_1 + \{\mathbb{0}\} \\ &= \boldsymbol{\Omega}_1 \end{aligned}$$

$$\begin{aligned} \mathcal{N}(\mathbf{P}) &= \{\boldsymbol{x} \in \boldsymbol{\Omega} | \mathbf{P}\boldsymbol{x} = \mathbb{0}\} \\ &= \left\{\boldsymbol{x} \in (\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2) | \mathbf{P}\boldsymbol{x} = \mathbb{0}\right\} \\ &= \left\{\boldsymbol{x} \in \boldsymbol{\Omega}_1 | \mathbf{P}\boldsymbol{x} = \mathbb{0}\right\} + \left\{\boldsymbol{x} \in \boldsymbol{\Omega}_2 | \mathbf{P}\boldsymbol{x} = \mathbb{0}\right\} \\ &= \{\mathbb{0}\} + \boldsymbol{\Omega}_2 \\ &= \boldsymbol{\Omega}_2 \end{aligned}$$

✏

Theorem O.16. [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* $\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$.

THM

$$\underbrace{\mathbf{P}^2 = \mathbf{P}}_{\mathbf{P} \text{ is a projection operator}} \quad \Longleftrightarrow \quad \underbrace{(\mathbf{I} - \mathbf{P})^2 = (\mathbf{I} - \mathbf{P})}_{(\mathbf{I} - \mathbf{P}) \text{ is a projection operator}}$$

✎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 $\mathbf{P}^2 = \mathbf{P} \impliedby (\mathbf{I} - \mathbf{P})^2 = (\mathbf{I} - \mathbf{P})$:

$$\begin{aligned} \mathbf{P}^2 &= \underbrace{\mathbf{I} - \mathbf{P} - \mathbf{P} + \mathbf{P}^2}_{(\mathbf{I} - \mathbf{P})^2} - (\mathbf{I} - \mathbf{P} - \mathbf{P}) \\ &= (\mathbf{I} - \mathbf{P})^2 - (\mathbf{I} - \mathbf{P} - \mathbf{P}) \\ &= (\mathbf{I} - \mathbf{P}) - (\mathbf{I} - \mathbf{P} - \mathbf{P}) && \text{by right hypothesis} \\ &= \mathbf{P} \end{aligned}$$

✏

[31] Michel and Herget (1993) pages 120–121
[32] Michel and Herget (1993) page 121

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

Theorem O.17. [33] *Let $\boldsymbol{H}$ be a* HILBERT SPACE *and* $\mathbf{P}$ *an operator in* $\boldsymbol{H}^{\boldsymbol{H}}$ *with adjoint* $\mathbf{P}^*$, NULL SPACE $\mathcal{N}(\mathbf{P})$, *and* IMAGE SET $\mathcal{I}(\mathbf{P})$.

THM

If $\mathbf{P}$ *is a* PROJECTION OPERATOR, *then the following are equivalent:*

1.	$\mathbf{P}^* = \mathbf{P}$		(**P** *is* SELF-ADJOINT)	$\Longleftrightarrow$
2.	$\mathbf{P}^*\mathbf{P} = \mathbf{P}\mathbf{P}^*$		(**P** *is* NORMAL)	$\Longleftrightarrow$
3.	$\mathcal{I}(\mathbf{P}) = \mathcal{N}(\mathbf{P})^\perp$			$\Longleftrightarrow$
4.	$\langle \mathbf{P}\boldsymbol{x} \mid \boldsymbol{x}\rangle = \|\mathbf{P}\boldsymbol{x}\|^2$	$\forall \boldsymbol{x}\in X$		

✎PROOF: This proof is incomplete at this time.

Proof that (1) $\Longrightarrow$ (2):

$$\begin{aligned}\mathbf{P}^*\mathbf{P} &= \mathbf{P}^{**}\mathbf{P}^* && \text{by (1)}\\ &= \mathbf{P}\mathbf{P}^* && \text{by Theorem O.13 page 333}\end{aligned}$$

Proof that (1) $\Longrightarrow$ (3):

$$\begin{aligned}\mathcal{I}(\mathbf{P}) &= \mathcal{N}(\mathbf{P}^*)^\perp && \text{by Theorem O.14 page 334}\\ &= \mathcal{N}(\mathbf{P})^\perp && \text{by (1)}\end{aligned}$$

Proof that (3) $\Longrightarrow$ (4):

Proof that (4) $\Longrightarrow$ (1):

✏

O.4.2 Self Adjoint Operators

Definition O.8. [34] *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* $\mathbf{B} \triangleq \mathbf{B}^*$.

Example O.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 auto-correlation operator $\mathbf{R}$ is *self-adjoint*)

Theorem O.18. [35] *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* $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle}$.

THM

$$\left\{\begin{array}{c}\underbrace{\mathbf{S} = \mathbf{S}^*}\\ \mathbf{S} \text{ is } \textbf{\textit{self adjoint}}\end{array}\right\} \Longrightarrow \left\{\begin{array}{lll} 1. & \langle \mathbf{S}\boldsymbol{x} \mid \boldsymbol{x}\rangle \in \mathbb{R} & \text{(the hermitian quadratic form of } \mathbf{S} \text{ is REAL-VALUED)}\\ 2. & \lambda_n \in \mathbb{R} & \text{(eigenvalues of } \mathbf{S} \text{ are REAL-VALUED)}\\ 3. & \lambda_n \neq \lambda_m \Longrightarrow \langle \psi_n \mid \psi_m\rangle = 0 & \text{(eigenvectors are ORTHOGONAL)}\end{array}\right\}$$

[33] Rudin (1991) page 314

[34] Historical works regarding self-adjoint operators: von Neumann (1929), page 49, "linearer Operator R selbstadjungiert oder Hermitesch", Stone (1932), page 50 ⟨"self-adjoint transformations"⟩

[35] Lax (2002), pages 315–316, Keener (1988), pages 114–119, Bachman and Narici (1966) page 24 ⟨Theorem 2.1⟩,
Bertero and Boccacci (1998) page 225 ⟨§"9.2 SVD of a matrix ...If all eigenvectors are normalized..."⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

1. Proof that $\mathbf{S} = \mathbf{S}^* \implies \langle \mathbf{S}\boldsymbol{x} \mid \boldsymbol{x}\rangle \in \mathbb{R}$:

$$\begin{aligned}
\langle \boldsymbol{x} \mid \mathbf{S}\boldsymbol{x}\rangle &= \langle \mathbf{S}\boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by left hypothesis}\\
&= \langle \boldsymbol{x} \mid \mathbf{S}\boldsymbol{x}\rangle^* && \text{by definition of } \langle \triangle \mid \triangledown\rangle \text{ Definition K.1 page 271}
\end{aligned}$$

2. Proof that $\mathbf{S} = \mathbf{S}^* \implies \lambda_n \in \mathbb{R}$:

$$\begin{aligned}
\lambda_n \left\|\psi_n\right\|^2 &= \lambda_n \langle \psi_n \mid \psi_n\rangle && \text{by definition}\\
&= \langle \lambda_n\psi_n \mid \psi_n\rangle && \text{by definition of } \langle \triangle \mid \triangledown\rangle \text{ Definition K.1 page 271}\\
&= \langle \mathbf{S}\psi_n \mid \psi_n\rangle && \text{by definition of eigenpairs}\\
&= \langle \psi_n \mid \mathbf{S}\psi_n\rangle && \text{by left hypothesis}\\
&= \langle \psi_n \mid \lambda_n\psi_n\rangle && \text{by definition of eigenpairs}\\
&= \lambda_n^* \langle \psi_n \mid \psi_n\rangle && \text{by definition of } \langle \triangle \mid \triangledown\rangle \text{ Definition K.1 page 271}\\
&= \lambda_n^* \left\|\psi_n\right\|^2 && \text{by definition}
\end{aligned}$$

3. Proof that $\mathbf{S} = \mathbf{S}^* \implies [\lambda_n \neq \lambda_m \implies \langle \psi_n \mid \psi_m\rangle = 0]$:

$$\begin{aligned}
\lambda_n \langle \psi_n \mid \psi_m\rangle &= \langle \lambda_n\psi_n \mid \psi_m\rangle && \text{by definition of } \langle \triangle \mid \triangledown\rangle \text{ Definition K.1 page 271}\\
&= \langle \mathbf{S}\psi_n \mid \psi_m\rangle && \text{by definition of eigenpairs}\\
&= \langle \psi_n \mid \mathbf{S}\psi_m\rangle && \text{by left hypothesis}\\
&= \langle \psi_n \mid \lambda_m\psi_m\rangle && \text{by definition of eigenpairs}\\
&= \lambda_m^* \langle \psi_n \mid \psi_m\rangle && \text{by definition of } \langle \triangle \mid \triangledown\rangle \text{ Definition K.1 page 271}\\
&= \lambda_m \langle \psi_n \mid \psi_m\rangle && \text{because } \lambda_m \text{ is real}
\end{aligned}$$

This implies for $\lambda_n \neq \lambda_m \neq 0$, $\langle \psi_n \mid \psi_m\rangle = 0$.

✏

O.4.3 Normal Operators

Definition O.9. [36] *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 O.19. [37] *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 $\underbrace{\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^*}_{\mathbf{N}\text{ is normal}} \iff \left\|\mathbf{N}^*\boldsymbol{x}\right\| = \|\mathbf{N}\boldsymbol{x}\| \quad \forall \boldsymbol{x} \in \boldsymbol{H}$

[36] Rudin (1991) page 312, Michel and Herget (1993) page 431, Dieudonné (1969), page 167, Frobenius (1878), Frobenius (1968), page 391

[37] Rudin (1991) pages 312–313

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

✎PROOF:

1. Proof that $\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^* \implies \|\mathbf{N}^*\boldsymbol{x}\| = \|\mathbf{N}\boldsymbol{x}\|$:

$$\begin{aligned}
\|\mathbf{N}\boldsymbol{x}\|^2 &= \langle \mathbf{N}\boldsymbol{x} \mid \mathbf{N}\boldsymbol{x}\rangle && \text{by definition}\\
&= \langle \boldsymbol{x} \mid \mathbf{N}^*\mathbf{N}\boldsymbol{x}\rangle && \text{by Proposition O.3 page 332 (definition of } \mathbf{N}^*)\\
&= \langle \boldsymbol{x} \mid \mathbf{N}\mathbf{N}^*\boldsymbol{x}\rangle && \text{by left hypothesis } (\mathbf{N} \text{ is normal})\\
&= \langle \mathbf{N}\boldsymbol{x} \mid \mathbf{N}^*\boldsymbol{x}\rangle && \text{by Proposition O.3 page 332 (definition of } \mathbf{N}^*)\\
&= \|\mathbf{N}^*\boldsymbol{x}\|^2 && \text{by definition}
\end{aligned}$$

2. Proof that $\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^* \impliedby \|\mathbf{N}^*\boldsymbol{x}\| = \|\mathbf{N}\boldsymbol{x}\|$:

$$\begin{aligned}
\langle \mathbf{N}^*\mathbf{N}\boldsymbol{x} \mid \boldsymbol{x}\rangle &= \langle \mathbf{N}\boldsymbol{x} \mid \mathbf{N}^{**}\boldsymbol{x}\rangle && \text{by Proposition O.3 page 332 (definition of } \mathbf{N}^*)\\
&= \langle \mathbf{N}\boldsymbol{x} \mid \mathbf{N}\boldsymbol{x}\rangle && \text{by Theorem O.13 page 333 (property of adjoint)}\\
&= \|\mathbf{N}\boldsymbol{x}\|^2 && \text{by definition}\\
&= \|\mathbf{N}^*\boldsymbol{x}\|^2 && \text{by right hypothesis } (\|\mathbf{N}^*\boldsymbol{x}\| = \|\mathbf{N}\boldsymbol{x}\|)\\
&= \langle \mathbf{N}^*\boldsymbol{x} \mid \mathbf{N}^*\boldsymbol{x}\rangle && \text{by definition}\\
&= \langle \mathbf{N}\mathbf{N}^*\boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by Proposition O.3 page 332 (definition of } \mathbf{N}^*)
\end{aligned}$$

✏

Theorem O.20. [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
$$\underbrace{\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^*}_{\mathbf{N}\text{ is normal}} \implies \underbrace{\mathcal{N}(\mathbf{N}^*) = \mathcal{N}(\mathbf{N})}_{\mathbf{N}\text{ and }\mathbf{N}^*\text{ have the same null space}}$$

✎PROOF:

$$\begin{aligned}
\mathcal{N}(\mathbf{N}^*) &= \{\boldsymbol{x} | \mathbf{N}^*\boldsymbol{x} = 0 \quad \forall \boldsymbol{x} \in \boldsymbol{X}\} && \text{(definition of } \mathcal{N})\\
&= \{\boldsymbol{x} | \|\mathbf{N}^*\boldsymbol{x}\| = 0 \quad \forall \boldsymbol{x} \in \boldsymbol{X}\} && \text{by definition of } \|\cdot\| \text{ (Definition L.1 page 287)}\\
&= \{\boldsymbol{x} | \|\mathbf{N}\boldsymbol{x}\| = 0 \quad \forall \boldsymbol{x} \in \boldsymbol{X}\} && \\
&= \{\boldsymbol{x} | \mathbf{N}\boldsymbol{x} = 0 \quad \forall \boldsymbol{x} \in \boldsymbol{X}\} && \text{by definition of } \|\cdot\| \text{ (Definition L.1 page 287)}\\
&= \mathcal{N}(\mathbf{N}) && \text{(definition of } \mathcal{N})
\end{aligned}$$

✏

Theorem O.21. [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
$$\left\{\underbrace{\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^*}_{\mathbf{N}\text{ is normal}}\right\} \implies \left\{\underbrace{\lambda_n \neq \lambda_m \implies \langle \psi_n \mid \psi_m\rangle = 0}_{\text{eigenfunctions associated with distinct eigenvalues are orthogonal}}\right\}$$

✎PROOF: The proof in (1) is flawed. This implies that (2) is also flawed. [Rudin] claims both to be true.(Rudin, 1991)313

[38] Rudin (1991) pages 312–313
[39] Rudin (1991) pages 312–313

COPYRIGHT © 2019 DANIEL J. GREENHOE

1. Proof that $\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^* \implies \mathbf{N}^*\psi = \lambda^*\psi$:

$$\begin{aligned}
&\mathbf{N}\psi = \lambda\psi \\
&\Longleftarrow \\
&\quad 0 = \mathcal{N}(\mathbf{N} - \lambda\mathbf{I}) \\
&\qquad = \mathcal{N}([\mathbf{N} - \lambda\mathbf{I}]^*) && \text{by } \mathcal{N}(\mathbf{N}) = \mathcal{N}(\mathbf{N}^*) \\
&\qquad = \mathcal{N}(\mathbf{N}^* - [\lambda\mathbf{I}]^*) && \text{by Theorem O.13 page 333} \\
&\qquad = \mathcal{N}(\mathbf{N}^* - \lambda^*\mathbf{I}^*) && \text{by Theorem O.13 page 333} \\
&\qquad = \mathcal{N}(\mathbf{N}^* - \lambda^*\mathbf{I}) \\
&\Longrightarrow \\
&(\mathbf{N}^* - \lambda^*\mathbf{I})\psi = 0 \\
&\Longleftrightarrow \mathbf{N}^*\psi = \lambda^*\psi
\end{aligned}$$

2. Proof that $\mathbf{N}^*\mathbf{N} = \mathbf{N}\mathbf{N}^* \implies [\lambda_n \neq \lambda_m \implies \langle\psi_n \,|\, \psi_m\rangle = 0]$:

$$\begin{aligned}
\lambda_n \langle\psi_n \,|\, \psi_m\rangle &= \langle\lambda_n\psi_n \,|\, \psi_m\rangle && \text{by definition of } \langle\triangle \,|\, \triangledown\rangle \text{ Definition K.1 page 271} \\
&= \langle\mathbf{N}\psi_n \,|\, \psi_m\rangle && \text{by definition of eigenpairs} \\
&= \langle\psi_n \,|\, \mathbf{N}^*\psi_m\rangle && \text{by Proposition O.3 page 332 (definition of adjoint)} \\
&= \langle\psi_n \,|\, \lambda_m^*\psi_m\rangle && \text{by (4.)} \\
&= \lambda_m \langle\psi_n \,|\, \psi_m\rangle && \text{by definition of } \langle\triangle \,|\, \triangledown\rangle \text{ Definition K.1 page 271}
\end{aligned}$$

This implies for $\lambda_n \neq \lambda_m \neq 0$, $\langle\psi_n \,|\, \psi_m\rangle = 0$.

✏

O.4.4 Isometric operators

An operator on a pair of normed linear spaces is *isometric* (next definition) if it is an *isometry*.

Definition O.10. *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *and* $\big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be* NORMED LINEAR SPACES *(Definition L.1 page 287)*.

DEF *An operator* $\mathbf{M} \in \mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$ *is* ***isometric*** *if*

$$\|\mathbf{M}\boldsymbol{x}\| = \|\boldsymbol{x}\| \qquad \forall_{\boldsymbol{x}\in X}.$$

Theorem O.22. [40] *Let* $\big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *and* $\big(Y, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$ *be* NORMED LINEAR SPACES. *Let* $\mathbf{M}$ *be a linear operator in* $\mathcal{L}(\boldsymbol{X}, \boldsymbol{Y})$.

THM

$$\underbrace{\|\mathbf{M}\boldsymbol{x}\| = \|\boldsymbol{x}\| \quad \forall_{\boldsymbol{x}\in X}}_{\textit{isometric in length}} \quad \Longleftrightarrow \quad \underbrace{\|\mathbf{M}\boldsymbol{x} - \mathbf{M}\boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{y}\| \quad \forall_{\boldsymbol{x},\boldsymbol{y}\in X}}_{\textit{isometric in distance}}$$

✎PROOF:

1. Proof that $\|\mathbf{M}\boldsymbol{x}\| = \|\boldsymbol{x}\| \implies \|\mathbf{M}\boldsymbol{x} - \mathbf{M}\boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{y}\|$:

$$\begin{aligned}
\|\mathbf{M}\boldsymbol{x} - \mathbf{M}\boldsymbol{y}\| &= \|\mathbf{M}(\boldsymbol{x} - \boldsymbol{y})\| && \text{by definition of linear operators (Definition O.3 page 320)} \\
&= \|\mathbf{M}\boldsymbol{u}\| && \text{let } \boldsymbol{u} \triangleq \boldsymbol{x} - \boldsymbol{y} \\
&= \|\boldsymbol{x} - \boldsymbol{y}\| && \text{by left hypothesis}
\end{aligned}$$

[40] Kubrusly (2001) page 239 ⟨Proposition 4.37⟩, Berberian (1961) page 27 ⟨Theorem IV.7.5⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

2. Proof that $\|\mathbf{M}\boldsymbol{x}\| = \|\boldsymbol{x}\| \Longleftarrow \|\mathbf{M}\boldsymbol{x} - \mathbf{M}\boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{y}\|$:

$$\begin{aligned}
\|\mathbf{M}\boldsymbol{x}\| &= \|\mathbf{M}(\boldsymbol{x} - \mathbb{0})\| \\
&= \|\mathbf{M}\boldsymbol{x} - \mathbf{M}\mathbb{0})\| && \text{by definition of linear operators (Definition O.3 page 320)} \\
&= \|\boldsymbol{x} - \mathbb{0}\| && \text{by right hypothesis} \\
&= \|\boldsymbol{x}\|
\end{aligned}$$

✏

Isometric operators have already been defined (Definition O.10 page 339) in the more general normed linear spaces, while Theorem O.22 (page 339) demonstrated that in a normed linear space $\boldsymbol{X}$, $\|\mathbf{M}\boldsymbol{x}\| = \|\boldsymbol{x}\| \iff \|\mathbf{M}\boldsymbol{x} - \mathbf{M}\boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$. Here in the more specialized inner product spaces, Theorem O.23 (next) demonstrates two additional equivalent properties.

Theorem O.23. [41] *Let $\mathcal{B}(\boldsymbol{X}, \boldsymbol{X})$ be the space of* BOUNDED LINEAR OPERATORS *on a normed linear space $\boldsymbol{X} \triangleq \big(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|\big)$. 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 $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle}$.*

THM

*The following conditions are all **equivalent**:*

1.	$\mathbf{M}^*\mathbf{M}$	$=$	$\mathbf{I}$			$\iff$
2.	$\langle \mathbf{M}\boldsymbol{x} \mid \mathbf{M}\boldsymbol{y}\rangle$	$=$	$\langle \boldsymbol{x} \mid \boldsymbol{y}\rangle$	$\forall \boldsymbol{x},\boldsymbol{y} \in X$	*($\mathbf{M}$ is surjective)*	$\iff$
3.	$\|\mathbf{M}\boldsymbol{x} - \mathbf{M}\boldsymbol{y}\|$	$=$	$\|x - y\|$	$\forall \boldsymbol{x},\boldsymbol{y} \in X$	*(isometric in distance)*	$\iff$
4.	$\|\mathbf{M}\boldsymbol{x}\|$	$=$	$\|\boldsymbol{x}\|$	$\forall \boldsymbol{x} \in X$	*(isometric in length)*	

✎PROOF:

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

$$\begin{aligned}
\langle \mathbf{M}\boldsymbol{x} \mid \mathbf{M}\boldsymbol{y}\rangle &= \langle \boldsymbol{x} \mid \mathbf{M}^*\mathbf{M}\boldsymbol{y}\rangle && \text{by Proposition O.3 page 332 (definition of adjoint)} \\
&= \langle \boldsymbol{x} \mid \mathbf{I}\boldsymbol{y}\rangle && \text{by (1)} \\
&= \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle && \text{by Definition O.2 page 319 (definition of } \mathbf{I}\text{)}
\end{aligned}$$

2. Proof that (2) $\Longrightarrow$ (4):

$$\begin{aligned}
\|\mathbf{M}\boldsymbol{x}\| &= \sqrt{\langle \mathbf{M}\boldsymbol{x} \mid \mathbf{M}\boldsymbol{x}\rangle} && \text{by definition of } \|\cdot\| \\
&= \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle} && \text{by right hypothesis} \\
&= \|\boldsymbol{x}\| && \text{by definition of } \|\cdot\|
\end{aligned}$$

3. Proof that (2) $\Longleftarrow$ (4):

$$\begin{aligned}
4\,\langle \mathbf{M}\boldsymbol{x} \mid \mathbf{M}\boldsymbol{y}\rangle &= \|\mathbf{M}\boldsymbol{x} + \mathbf{M}\boldsymbol{y}\|^2 - \|\mathbf{M}\boldsymbol{x} - \mathbf{M}\boldsymbol{y}\|^2 + i\,\|\mathbf{M}\boldsymbol{x} + i\mathbf{M}\boldsymbol{y}\|^2 - i\,\|\mathbf{M}\boldsymbol{x} - i\mathbf{M}\boldsymbol{y}\|^2 && \text{by polarization id.} \\
&= \|\mathbf{M}(\boldsymbol{x} + \boldsymbol{y})\|^2 - \|\mathbf{M}(\boldsymbol{x} - \boldsymbol{y})\|^2 + i\,\|\mathbf{M}(\boldsymbol{x} + i\boldsymbol{y})\|^2 - i\,\|\mathbf{M}(\boldsymbol{x} - i\boldsymbol{y})\|^2 && \text{by Definition O.3} \\
&= \|\boldsymbol{x} + \boldsymbol{y}\|^2 - \|\boldsymbol{x} - \boldsymbol{y}\|^2 + i\,\|\boldsymbol{x} + i\boldsymbol{y}\|^2 - i\,\|\boldsymbol{x} - i\boldsymbol{y}\|^2 && \text{by left hypothesis}
\end{aligned}$$

4. Proof that (3) $\iff$ (4): by Theorem O.22 page 339

[41] Michel and Herget (1993) page 432 ⟨Theorem 7.5.8⟩, Kubrusly (2001) page 391 ⟨Proposition 5.72⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

5. Proof that (4) $\Longrightarrow$ (1):

$$\begin{aligned}
\langle \mathbf{M}^*\mathbf{M}\boldsymbol{x} \mid \boldsymbol{x}\rangle &= \langle \mathbf{M}\boldsymbol{x} \mid \mathbf{M}^{**}\boldsymbol{x}\rangle && \text{by Proposition O.3 page 332 (definition of adjoint)}\\
&= \langle \mathbf{M}\boldsymbol{x} \mid \mathbf{M}\boldsymbol{x}\rangle && \text{by Theorem O.13 page 333 (property of adjoint)}\\
&= \|\mathbf{M}\boldsymbol{x}\|^2 && \text{by definition}\\
&= \|\boldsymbol{x}\|^2 && \text{by left hypothesis with } y = 0\\
&= \langle \boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by definition}\\
&= \langle \mathbf{I}\boldsymbol{x} \mid \boldsymbol{x}\rangle && \text{by Definition O.2 page 319 (definition of } \mathbf{I})\\
\Longrightarrow \qquad \mathbf{M}^*\mathbf{M} &= \mathbf{I} && \forall \boldsymbol{x} \in X
\end{aligned}$$

✏

Theorem O.24. [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 $\|\boldsymbol{x}\| \triangleq \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{x}\rangle}$.*

THM

$$\underbrace{\mathbf{M}^*\mathbf{M} = \mathbf{I}}_{\mathbf{M}\text{ is isometric}} \quad \Longrightarrow \quad \begin{cases} |||\mathbf{M}||| = 1 & \text{(UNIT LENGTH) and}\\ |\lambda| = 1 & \forall \lambda \in \Lambda \end{cases}$$

✎PROOF:

1. Proof that $\mathbf{M}^*\mathbf{M} = \mathbf{I} \Longrightarrow |||\mathbf{M}||| = 1$:

$$\begin{aligned}
|||\mathbf{M}||| &= \sup_{\boldsymbol{x}\in X}\{\|\mathbf{M}\boldsymbol{x}\| \mid \|\boldsymbol{x}\| = 1\} && \text{by Definition O.4 page 324}\\
&= \sup_{\boldsymbol{x}\in X}\{\|\boldsymbol{x}\| \mid \|\boldsymbol{x}\| = 1\} && \text{by Theorem O.23 page 340}\\
&= \sup_{\boldsymbol{x}\in X}\{1\}\\
&= 1
\end{aligned}$$

2. Proof that $|\lambda| = 1$: Let $(\boldsymbol{x}, \lambda)$ be an eigenvector-eigenvalue pair.

$$\begin{aligned}
1 &= \frac{1}{\|\boldsymbol{x}\|}\ \|\boldsymbol{x}\|\\
&= \frac{1}{\|\boldsymbol{x}\|}\ \|\mathbf{M}\boldsymbol{x}\| && \text{by Theorem O.23 page 340}\\
&= \frac{1}{\|\boldsymbol{x}\|}\ \|\lambda\boldsymbol{x}\| && \text{by definition of } \lambda\\
&= \frac{1}{\|\boldsymbol{x}\|}\ |\lambda|\ \|\boldsymbol{x}\| && \text{by homogeneous property of } \|\cdot\|\\
&= |\lambda|
\end{aligned}$$

✏

Example O.4 (One sided shift operator). [43] Let $\boldsymbol{X}$ be the set of all sequences with range $\mathbb{W}$ $(0, 1, 2, \ldots)$ and shift operators defined as

$$\begin{aligned}
&1. \quad \mathbf{S}_r\,(\!(x_0, x_1, x_2, \ldots)\!) \triangleq (\!(0, x_0, x_1, x_2, \ldots)\!) && \text{(right shift operator)}\\
&2. \quad \mathbf{S}_l\,(\!(x_0, x_1, x_2, \ldots)\!) \triangleq (\!(x_1, x_2, x_3, \ldots)\!) && \text{(left shift operator)}
\end{aligned}$$

EX

1. $\mathbf{S}_r$ is an isometric operator.
2. $\mathbf{S}_r^* = \mathbf{S}_l$

[42] Michel and Herget (1993) page 432
[43] Michel and Herget (1993) page 441

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

✎PROOF:

1. Proof that $\mathbf{S}_r^* = \mathbf{S}_l$:

$$\begin{aligned}
\langle \mathbf{S}_r (\!(x_0, x_1, x_2, \ldots)\!) \mid (\!(y_0, \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots)\!) \rangle &= \langle (\!(0, x_0, x_1, x_2, \ldots)\!) \mid (\!(y_0, \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots)\!) \rangle \\
&= \sum_{n=1}^{\infty} \boldsymbol{x}_{n-1}\ \boldsymbol{y}_n^* \\
&= \sum_{n=0}^{\infty} \boldsymbol{x}_n\ \boldsymbol{y}_{n+1}^* \\
&= \sum_{n=0}^{\infty} \boldsymbol{x}_n\ \boldsymbol{y}_{n+1}^* \\
&= \langle (\!(x_0, x_1, x_2, \ldots)\!) \mid (\!(y_1, \boldsymbol{y}_2, \boldsymbol{y}_3, \ldots)\!) \rangle \\
&= \left\langle (\!(x_0, x_1, x_2, \ldots)\!) \mid \underbrace{\mathbf{S}_l}_{\mathbf{S}_r^*} (\!(y_0, \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots)\!) \right\rangle
\end{aligned}$$

2. Proof that $\mathbf{S}_r$ is isometric ($\mathbf{S}_r^*\mathbf{S}_r = \mathbf{I}$):

$$\begin{aligned}
\mathbf{S}_r^*\mathbf{S}_r &= \mathbf{S}_l\mathbf{S}_r && \text{by 1.} \\
&= \mathbf{I}
\end{aligned}$$

✏

O.4.5 Unitary operators

Definition O.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 O.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 $\left.\begin{array}{l}\mathbf{U} \text{ is UNITARY and} \\ \mathbf{V} \text{ is UNITARY}\end{array}\right\} \implies (\mathbf{UV}) \text{ is UNITARY.}$

✎PROOF:

$$\begin{aligned}
(\mathbf{UV})(\mathbf{UV})^* &= (\mathbf{UV})\left(\mathbf{V}^*\mathbf{U}^*\right) && \text{by Theorem O.8 page 329} \\
&= \mathbf{U}\left(\mathbf{V}\mathbf{V}^*\right)\mathbf{U}^* && \text{by associative property} \\
&= \mathbf{U}\mathbf{I}\mathbf{U}^* && \text{by definition of unitary operators—Definition O.11 page 342} \\
&= \mathbf{I} && \text{by definition of unitary operators—Definition O.11 page 342}
\end{aligned}$$

$$\begin{aligned}
(\mathbf{UV})^*(\mathbf{UV}) &= \left(\mathbf{V}^*\mathbf{U}^*\right)(\mathbf{UV}) && \text{by Theorem O.8 page 329} \\
&= \mathbf{V}^*\left(\mathbf{U}^*\mathbf{U}\right)\mathbf{V} && \text{by associative property} \\
&= \mathbf{V}^*\mathbf{I}\mathbf{V} && \text{by definition of unitary operators—Definition O.11 page 342} \\
&= \mathbf{I} && \text{by definition of unitary operators—Definition O.11 page 342}
\end{aligned}$$

[44] Rudin (1991) page 312, Michel and Herget (1993) page 431, Autonne (1901) page 209, Autonne (1902), Schur (1909), Steen (1973)

COPYRIGHT © 2019 DANIEL J. GREENHOE

✏

Theorem O.25. [45] *Let $\mathcal{B}(\boldsymbol{H},\boldsymbol{H})$ be the space of* BOUNDED LINEAR OPERATORS *on a* HILBERT SPACE $\boldsymbol{H}$. *Let $\mathcal{I}(\mathbf{U})$ be the* IMAGE SET *of* $\mathbf{U}$.

THM

If $\mathbf{U}$ *is a* ***bounded linear operator*** *($\mathbf{U} \in \mathcal{B}(\boldsymbol{H},\boldsymbol{H})$), then the following conditions are* ***equivalent****:*

1. $\mathbf{U}\mathbf{U}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}$ (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 O.23 (page 340).
   (b) Proof that $\mathcal{I}(\mathbf{U}) = X$:

$$\begin{aligned} X &\supseteq \mathcal{I}(\mathbf{U}) && \text{because } \mathbf{U} \in X^X \\ &\supseteq \mathcal{I}(\mathbf{U}\mathbf{U}^*) \\ &= \mathcal{I}(\mathbf{I}) && \text{by left hypothesis } (\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I}) \\ &= X && \text{by Definition O.2 page 319 (definition of } \mathbf{I}) \end{aligned}$$

2. Proof that (2) $\iff$ (3) $\iff$ (4): by Theorem O.23 page 340.
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 O.23 page 340
   (b) Proof that $\left\|\mathbf{U}^*\boldsymbol{x} - \mathbf{U}^*\boldsymbol{y}\right\| = \|\boldsymbol{x} - \boldsymbol{y}\| \implies \mathbf{U}\mathbf{U}^* = \mathbf{I}$:

$$\begin{aligned} \left\|\mathbf{U}^*\boldsymbol{x} - \mathbf{U}^*\boldsymbol{y}\right\| = \|\boldsymbol{x} - \boldsymbol{y}\| \implies \quad \mathbf{U}^{**}\mathbf{U}^* &= \mathbf{I} && \text{by Theorem O.23 page 340} \\ \mathbf{U}\mathbf{U}^* &= \mathbf{I} && \text{by Theorem O.13 page 333} \end{aligned}$$

✏

Theorem O.26. *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

$$\underbrace{\mathbf{U}\mathbf{U}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}}_{\mathbf{U}\text{ is }\textbf{unitary}} \implies \left\{ \begin{array}{rclcll} \mathbf{U}^{-1} &=& \mathbf{U}^* & & & \textit{and} \\ \mathcal{I}(\mathbf{U}) &=& \mathcal{I}(\mathbf{U}^*) &=& X & \textit{and} \\ \mathcal{N}(\mathbf{U}) &=& \mathcal{N}(\mathbf{U}^*) &=& \{\mathbb{0}\} & \textit{and} \\ |||\mathbf{U}||| &=& \left|||\mathbf{U}^*\right||| &=& 1 & \text{(UNIT LENGTH)} \end{array} \right\}$$

✎PROOF:

1. Note that $\mathbf{U}$, $\mathbf{U}^*$, and $\mathbf{U}^{-1}$ are all both *isometric* and *normal*:

$$\begin{array}{rclcrcl} \mathbf{U}^*\mathbf{U} & & & = & \mathbf{I} & \implies & \mathbf{U} \text{ is } \textit{isometric} \\ \mathbf{U}\mathbf{U}^* & = & \mathbf{U}^*\mathbf{U} & = & \mathbf{I} & \implies & \mathbf{U}^* \text{ is } \textit{isometric} \\ \mathbf{U}^{-1} & = & \mathbf{U}^* & & & \implies & \mathbf{U}^{-1} \text{ is } \textit{isometric} \\ & & & & & & \\ \mathbf{U}^*\mathbf{U} & = & \mathbf{U}\mathbf{U}^* & = & & \implies & \mathbf{U} \text{ is } \textit{normal} \\ \mathbf{U}\mathbf{U}^* & = & \mathbf{U}^*\mathbf{U} & = & & \implies & \mathbf{U}^* \text{ is } \textit{normal} \\ \mathbf{U}^{-1} & = & \mathbf{U}^* & & & \implies & \mathbf{U}^{-1} \text{ is } \textit{normal} \end{array}$$

[45] Rudin (1991) pages 313–314 ⟨Theorem 12.13⟩, Knapp (2005a) page 45 ⟨Proposition 2.6⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

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 O.25 page 343.
3. Proof that $\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I} \implies \mathcal{N}(\mathbf{U}) = \mathcal{N}(\mathbf{U}^*) = \mathcal{N}(\mathbf{U}^{-1})$:

$$\begin{aligned}
\mathcal{N}(\mathbf{U}^*) &= \mathcal{N}(\mathbf{U}) && \text{because } \mathbf{U} \text{ and } \mathbf{U}^* \text{ are both } \textit{normal} \text{ and by Theorem O.21 page 338}\\
&= \mathcal{I}(\mathbf{U})^\perp && \text{by Theorem O.14 page 334}\\
&= X^\perp && \text{by above result}\\
&= \{\mathbb{0}\} && \text{by Proposition I.6 page 257}
\end{aligned}$$

4. Proof that $\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I} \implies \left\|\mathbf{U}^{-1}\right\| = \left\|\mathbf{U}^*\right\| = \|\mathbf{U}\| = 1$:
Because $\mathbf{U}$, $\mathbf{U}^*$, and $\mathbf{U}^{-1}$ are all isometric and by Theorem O.24 page 341.

✏

Example O.5. Examples of *Fredholm integral operators* include

EX							
1.	**Fourier Transform**	$[\tilde{\mathsf{F}}\mathsf{x}](f)$	$=$	$\int_{t\in\mathbb{R}} \mathsf{x}(t)e^{-i2\pi f t}\,\mathrm{d}t$	$\kappa(t,f)$	$=$	$e^{-i2\pi f t}$
2.	**Inverse Fourier Transform**	$[\tilde{\mathbf{F}}^{-1}\tilde{\mathsf{x}}](t)$	$=$	$\int_{f\in\mathbb{R}} \tilde{\mathsf{x}}(f)e^{i2\pi f t}\,\mathrm{d}f$	$\kappa(f,t)$	$=$	$e^{i2\pi f t}$
3.	**Laplace operator**	$[\mathbf{L}\mathsf{x}](s)$	$=$	$\int_{t\in\mathbb{R}} \mathsf{x}(t)e^{-st}\,\mathrm{d}t$	$\kappa(t,s)$	$=$	e^{-st}

Example O.6 (Translation operator). Let $\boldsymbol{X} = \boldsymbol{L}^2_{\mathbb{R}}$ and $\mathbf{T} \in \boldsymbol{X}^{\boldsymbol{X}}$ be defined as

$$\mathbf{T}\mathsf{f}(x) \triangleq \mathsf{f}(x-1) \qquad \forall \mathsf{f}\in\boldsymbol{L}^2_{\mathbb{R}} \qquad \text{(translation operator)}$$

EX					
1.	$\mathbf{T}^{-1}\mathsf{f}(x)$	$=$	$\mathsf{f}(x+1)$	$\forall \mathsf{f}\in\boldsymbol{L}^2_{\mathbb{R}}$	(inverse translation operator)
2.	$\mathbf{T}^*$	$=$	$\mathbf{T}^{-1}$		($\mathbf{T}$ is invertible)
3.	$\mathbf{T}^*\mathbf{T}$	$=$	$\mathbf{T}\mathbf{T}^* = \mathbf{I}$		($\mathbf{T}$ is unitary)

✎PROOF:

1. Proof that $\mathbf{T}^{-1}\mathsf{f}(x) = \mathsf{f}(x+1)$:

$$\begin{aligned}
\mathbf{T}^{-1}\mathbf{T} &= \mathbf{I}\\
\mathbf{T}\mathbf{T}^{-1} &= \mathbf{I}
\end{aligned}$$

2. Proof that $\mathbf{T}$ is unitary:

$$\begin{aligned}
\langle \mathbf{T}\mathsf{f}(x) \,|\, \mathsf{g}(x)\rangle &= \langle \mathsf{f}(x-1) \,|\, \mathsf{g}(x)\rangle && \text{by definition of } \mathbf{T}\\
&= \int_x \mathsf{f}(x-1)\mathsf{g}^*(x)\,\mathrm{d}x\\
&= \int_x \mathsf{f}(x)\mathsf{g}^*(x+1)\,\mathrm{d}x\\
&= \langle \mathsf{f}(x) \,|\, \mathsf{g}(x+1)\rangle\\
&= \left\langle \mathsf{f}(x) \,\middle|\, \underbrace{\mathbf{T}^{-1}}_{\mathbf{T}^*}\mathsf{g}(x) \right\rangle && \text{by 1.}
\end{aligned}$$

✏

Example O.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.	$\mathbf{D}^{-1}\mathsf{f}(x)$	$=$	$\frac{1}{\sqrt{2}}\mathsf{f}\left(\frac{1}{2}x\right)$	$\forall \mathsf{f}\in\boldsymbol{L}^2_{\mathbb{R}}$	(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)

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

1. Proof that $\mathbf{D}^{-1}\mathsf{f}(x) = \frac{1}{\sqrt{2}}\mathsf{f}\left(\frac{1}{2}x\right)$:

$$\begin{aligned}\mathbf{D}^{-1}\mathbf{D} &= \mathbf{I}\\ \mathbf{D}\mathbf{D}^{-1} &= \mathbf{I}\end{aligned}$$

2. Proof that $\mathbf{D}$ is unitary:

$$\begin{aligned}\langle \mathbf{D}\mathsf{f}(x) \,|\, \mathsf{g}(x)\rangle &= \left\langle \sqrt{2}\mathsf{f}(2x) \,|\, \mathsf{g}(x)\right\rangle && \text{by definition of } \mathbf{D}\\ &= \int_x \sqrt{2}\mathsf{f}(2x)\mathsf{g}^*(x)\,\mathrm{d}x\\ &= \int_{u\in\mathbb{R}} \sqrt{2}\mathsf{f}(u)\mathsf{g}^*\left(\frac{1}{2}u\right)\frac{1}{2}\,\mathrm{d}u && \text{let } u \triangleq 2x \quad\Longrightarrow\quad \mathrm{d}x = \frac{1}{2}\,\mathrm{d}u\\ &= \int_{u\in\mathbb{R}} \mathsf{f}(u)\left[\frac{1}{\sqrt{2}}\mathsf{g}\left(\frac{1}{2}u\right)\right]^*\mathrm{d}u\\ &= \left\langle \mathsf{f}(x) \,|\, \frac{1}{\sqrt{2}}\mathsf{g}\left(\frac{1}{2}x\right)\right\rangle\\ &= \left\langle \mathsf{f}(x) \,|\, \underbrace{\mathbf{D}^{-1}}_{\mathbf{D}^*}\mathsf{g}(x)\right\rangle && \text{by 1.}\end{aligned}$$

✏

Example O.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

$$\mathbf{D}^{-1}\,(\!(x_n)\!)_{n\in\mathbb{Z}} \triangleq (\!(x_{n+1})\!)_{n\in\mathbb{Z}}\,.$$

$$\begin{aligned}\langle \mathbf{D}\,(\!(x_n)\!) \,|\, (\!(y_n)\!)\rangle &= \langle (\!(x_{n-1})\!) \,|\, (\!(y_n)\!)\rangle && \text{by definition of } \mathbf{D}\\ &= \sum_n \boldsymbol{x}_{n-1}\,\boldsymbol{y}_n^*\\ &= \sum_n \boldsymbol{x}_n\,\boldsymbol{y}_{n+1}^*\\ &= \left\langle (\!(x_n)\!) \,|\, (\!(y_{n+1})\!)\right\rangle\\ &= \left\langle (\!(x_n)\!) \,|\, \underbrace{\mathbf{D}^{-1}}_{\mathbf{D}^*}(\!(y_n)\!)\right\rangle\end{aligned}$$

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 O.9 (Fourier transform). Let $\tilde{\mathbf{F}}$ be the *Fourier Transform* and $\tilde{\mathbf{F}}^{-1}$ the *inverse Fourier Transform* operator

$$[\tilde{\mathbf{F}}\mathsf{x}](f) \triangleq \int_t \mathsf{x}(t)\underbrace{e^{-i2\pi ft}}_{\kappa(t,f)}\,\mathrm{d}t \qquad \left[\tilde{\mathbf{F}}^{-1}\tilde{\mathsf{x}}\right](t) \triangleq \int_f \tilde{\mathsf{x}}(f)\underbrace{e^{i2\pi ft}}_{\kappa^*(t,f)}\,\mathrm{d}f\,.$$

EX $\tilde{\mathbf{F}}^* = \tilde{\mathbf{F}}^{-1}$ (the Fourier Transform operator $\tilde{\mathbf{F}}$ is unitary)

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✎PROOF:

$$\begin{aligned}
\langle \tilde{\mathbf{F}}\mathsf{x} \mid \tilde{\mathsf{y}} \rangle &= \left\langle \int_t \mathsf{x}(t) e^{-i2\pi f t}\, \mathrm{d}t \mid \tilde{\mathsf{y}}(f) \right\rangle \\
&= \int_t \mathsf{x}(t) \left\langle e^{-i2\pi f t} \mid \tilde{\mathsf{y}}(f) \right\rangle \mathrm{d}t \\
&= \int_t \mathsf{x}(t) \int_f e^{-i2\pi f t} \tilde{\mathsf{y}}^*(f)\, \mathrm{d}f\, \mathrm{d}t \\
&= \int_t \mathsf{x}(t) \left[\int_f e^{i2\pi f t} \tilde{\mathsf{y}}(f)\, \mathrm{d}f \right]^* \mathrm{d}t \\
&= \left\langle \mathsf{x}(t) \mid \int_f \tilde{\mathsf{y}}(f) e^{i2\pi f t}\, \mathrm{d}f \right\rangle \\
&= \left\langle \mathsf{x} \mid \underbrace{\tilde{\mathbf{F}}^{-1}}_{\tilde{\mathbf{F}}^*} \tilde{\mathsf{y}} \right\rangle
\end{aligned}$$

This implies that $\tilde{\mathbf{F}}$ is unitary ($\tilde{\mathbf{F}}^* = \tilde{\mathbf{F}}^{-1}$). ✏

Example O.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:

$$\begin{aligned}
\mathbf{R}^* &= \mathbf{R}^H \\
&= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^H && \text{by definition of } \mathbf{R} \\
&= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} && \text{by definition of Hermetian transpose operator } H \\
&= \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} && \text{by Theorem ?? page ??} \\
&= \mathbf{R}_{-\theta} && \text{by definition of } \mathbf{R} \\
&= \mathbf{R}^{-1} && \text{by 1.}
\end{aligned}$$

✏

O.5 Operator order

Definition O.12. [47] *Let* $\mathbf{P} \in Y^X$ *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 $\mathbf{P} \geq 0$.

[46] Noble and Daniel (1988), page 311
[47] Michel and Herget (1993) page 429 ⟨Definition 7.4.12⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

Theorem O.27. [48]

THM

$$\underbrace{\mathbf{P} \geq 0 \textit{ and } \mathbf{Q} \geq 0}_{\mathbf{P} \textit{ and } \mathbf{Q} \textit{ are both positive}} \quad \Longrightarrow \quad \begin{cases} (\mathbf{P}+\mathbf{Q}) & \geq & 0 & & (\mathbf{P}+\mathbf{Q}) \textit{ is positive} \\ \mathbf{A}^*\mathbf{P}\mathbf{A} & \geq & 0 & \forall \mathbf{A} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{X}) & (\mathbf{A}^*\mathbf{P}\mathbf{A} \textit{ is positive} \\ \mathbf{A}^*\mathbf{A} & \geq & 0 & \forall \mathbf{A} \in \mathcal{B}(\boldsymbol{X},\boldsymbol{X}) & (\mathbf{A}^*\mathbf{A} \textit{ is positive} \end{cases}$$

✎PROOF:

$\langle (\mathbf{P}+\mathbf{Q})\boldsymbol{x} \mid \boldsymbol{x}\rangle = \langle \mathbf{P}\boldsymbol{x} \mid \boldsymbol{x}\rangle + \langle \mathbf{Q}\boldsymbol{x} \mid \boldsymbol{x}\rangle$	by additive property of $\langle \triangle \mid \triangledown \rangle$ (Definition K.1 page 271)
$\geq \langle \mathbf{P}\boldsymbol{x} \mid \boldsymbol{x}\rangle$	by left hypothesis
≥ 0	by left hypothesis
$\langle \mathbf{A}^*\mathbf{P}\mathbf{A}\boldsymbol{x} \mid \boldsymbol{x}\rangle = \langle \mathbf{P}\mathbf{A}\boldsymbol{x} \mid \mathbf{A}\boldsymbol{x}\rangle$	by definition of adjoint (Proposition O.3 page 332)
$= \langle \mathbf{P}\boldsymbol{y} \mid \boldsymbol{y}\rangle$	where $\boldsymbol{y} \triangleq \mathbf{A}\boldsymbol{x}$
≥ 0	by left hypothesis
$\langle \mathbf{I}\boldsymbol{x} \mid \boldsymbol{x}\rangle = \langle \boldsymbol{x} \mid \boldsymbol{x}\rangle$	by definition of $\mathbf{I}$ (Definition O.2 page 319)
≥ 0	by non-negative property of $\langle \triangle \mid \triangledown \rangle$ (Definition K.1 page 271)
$\Longrightarrow \mathbf{I}$ is positive	
$\langle \mathbf{A}^*\mathbf{A}\boldsymbol{x} \mid \boldsymbol{x}\rangle = \langle \mathbf{A}^*\mathbf{I}\mathbf{A}\boldsymbol{x} \mid \boldsymbol{x}\rangle$	by definition of $\mathbf{I}$ (Definition O.2 page 319)
≥ 0	by two previous results

✏

Definition O.13. [49] *Let* $\mathbf{A}, \mathbf{B} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ *be* BOUNDED *operators.*

DEF

$\mathbf{A} \geq \mathbf{B}$ *("$\mathbf{A}$ is greater than or equal to $\mathbf{B}$") if*
$\mathbf{A} - \mathbf{B} \geq 0$ *("$(\mathbf{A} - \mathbf{B})$ is positive")*

[48] Michel and Herget (1993) page 429
[49] Michel and Herget (1993) page 429

COPYRIGHT © 2019 DANIEL J. GREENHOE

Copyright © 2019 Daniel J. Greenhoe

APPENDIX P

FOURIER TRANSFORM

❝The analytical equations ... extend to all general phenomena. There cannot be a language more universal and more simple, more free from errors and from obscurities, ... mathematical analysis is as extensive as nature itself; ... it preserves every principle which it has once acquired; it grows and strengthens itself incessantly in the midst of the many variations and errors of the human mind. Its chief attribute is clearness; it has no marks to express confused notions. It brings together phenomena the most diverse, and discovers the hidden analogies which unite them.❞

Joseph Fourier (1768–1830) [1]

P.1 Definitions

This chapter deals with the *Fourier Transform* in the space of *Lebesgue square-integrable functions* $\boldsymbol{L}^2_{(\mathbb{R},\mathscr{B},\mu)}$, where $\mathbb{R}$ is the set of real numbers, $\mathscr{B}$ is the set of *Borel sets* on $\mathbb{R}$, μ is the standard *Borel measure* on $\mathbb{R}$, and

$$\boldsymbol{L}^2_{(\mathbb{R},\mathscr{B},\mu)} \triangleq \left\{ \mathsf{f} \in \mathbb{R}^{\mathbb{R}} \,\middle|\, \int_{\mathbb{R}} |\mathsf{f}|^2 \, \mathrm{d}\mu < \infty \right\}.$$

Furthermore, $\langle \triangle \mid \triangledown \rangle$ is the *inner product* induced by the operator $\int_{\mathbb{R}} \mathrm{d}\mu$ such that

$$\langle \mathsf{f} \mid \mathsf{g} \rangle \triangleq \int_{\mathbb{R}} \mathsf{f}(x)\mathsf{g}^*(x) \, \mathrm{d}x,$$

and $\left(\boldsymbol{L}^2_{(\mathbb{R},\mathscr{B},\mu)}, \langle \triangle \mid \triangledown \rangle \right)$ is a *Hilbert space*.

Definition P.1. *Let κ be a* FUNCTION *in $\mathbb{C}^{\mathbb{R}^2}$.*

DEF

The function κ is the ***Fourier kernel*** *if* $\quad \kappa(x,\omega) \triangleq e^{i\omega x} \quad \forall_{x,\omega \in \mathbb{R}}$

Definition P.2. [2] *Let $\boldsymbol{L}^2_{(\mathbb{R},\mathscr{B},\mu)}$ be the space of all* LEBESGUE SQUARE-INTEGRABLE FUNCTIONS.

---

[1] quote: Fourier (1878), pages 7–8 ⟨Preliminary Discourse⟩
image: http://en.wikipedia.org/wiki/File:Fourier2.jpg, public domain

[2] Bachman et al. (2000) page 363, Chorin and Hald (2009) page 13, Loomis and Bolker (1965), page 144, Knapp (2005b) pages 374–375, Fourier (1822), Fourier (1878) page 336?

DEF *The **Fourier Transform** operator $\tilde{\mathbf{F}}$ is defined as*

$$\left[\tilde{\mathbf{F}}\mathsf{f}\right](\omega) \triangleq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathsf{f}(x)\, e^{-i\omega x}\ \mathrm{d}x \qquad \forall \mathsf{f} \in L^2_{(\mathbb{R},\mathscr{B},\mu)}$$

*This definition of the Fourier Transform is also called the **unitary Fourier Transform**.*

Remark P.1 (**Fourier transform scaling factor**). [3] If the Fourier transform operator $\tilde{\mathbf{F}}$ and inverse Fourier transform operator $\tilde{\mathbf{F}}^{-1}$ are defined as

$$\tilde{\mathbf{F}}\mathsf{f}(x) \triangleq \mathsf{F}(\omega) \triangleq A \int_{\mathbb{R}} \mathsf{f}(x) e^{-i\omega x}\ \mathrm{d}x \qquad \text{and} \qquad \tilde{\mathbf{F}}^{-1}\tilde{\mathsf{f}}(\omega) \triangleq B \int_{\mathbb{R}} \mathsf{F}(\omega) e^{i\omega x}\ \mathrm{d}\omega$$

then A and B can be any constants as long as $AB = \frac{1}{2\pi}$. The Fourier transform is often defined with the scaling factor A set equal to 1 such that $\left[\tilde{\mathbf{F}}\mathsf{f}(x)\right](\omega) \triangleq \int_{\mathbb{R}} \mathsf{f}(x)\, e^{-i\omega x}\ \mathrm{d}x$. In this case, the inverse Fourier transform operator $\tilde{\mathbf{F}}^{-1}$ is either defined as

- $\left[\tilde{\mathbf{F}}^{-1}\mathsf{f}(x)\right](f) \triangleq \int_{\mathbb{R}} \mathsf{f}(x)\, e^{i2\pi f x}\ \mathrm{d}x$ (using oscillatory frequency free variable f) or
- $\left[\tilde{\mathbf{F}}^{-1}\mathsf{f}(x)\right](\omega) \triangleq \frac{1}{2\pi} \int_{\mathbb{R}} \mathsf{f}(x)\, e^{i\omega x}\ \mathrm{d}x$ (using angular frequency free variable ω).

In short, the 2π has to show up somewhere, either in the argument of the exponential ($e^{-i2\pi f t}$) or in front of the integral ($\frac{1}{2\pi}\int \cdots$). One could argue that it is unnecessary to burden the exponential argument with the 2π factor ($e^{-i2\pi f t}$), and thus could further argue in favor of using the angular frequency variable ω thus giving the inverse operator definition $\left[\tilde{\mathbf{F}}^{-1}\mathsf{f}(x)\right](\omega) \triangleq \frac{1}{2\pi}\int_{\mathbb{R}} \mathsf{f}(x)\, e^{-i\omega x}\ \mathrm{d}x$. But this causes a new problem. In this case, the Fourier operator $\tilde{\mathbf{F}}$ is not *unitary* (see Theorem P.2 page 350)—in particular, $\tilde{\mathbf{F}}\tilde{\mathbf{F}}^* \neq \mathbf{I}$, where $\tilde{\mathbf{F}}^*$ is the *adjoint* of $\tilde{\mathbf{F}}$; but rather, $\tilde{\mathbf{F}}\left(\frac{1}{2\pi}\tilde{\mathbf{F}}^*\right) = \left(\frac{1}{2\pi}\tilde{\mathbf{F}}^*\right)\tilde{\mathbf{F}} = \mathbf{I}$. But if we define the operators $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{F}}^{-1}$ to both have the scaling factor $\frac{1}{\sqrt{2\pi}}$, then $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{F}}^{-1}$ are inverses *and* $\tilde{\mathbf{F}}$ is *unitary*—that is, $\tilde{\mathbf{F}}\tilde{\mathbf{F}}^* = \tilde{\mathbf{F}}^*\tilde{\mathbf{F}} = \mathbf{I}$.

P.2 Operator properties

Theorem P.1 (Inverse Fourier transform). [4] *Let $\tilde{\mathbf{F}}$ be the Fourier Transform operator (Definition P.2 page 349). The inverse $\tilde{\mathbf{F}}^{-1}$ of $\tilde{\mathbf{F}}$ is*

THM

$$\left[\tilde{\mathbf{F}}^{-1}\tilde{\mathsf{f}}\right](x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \tilde{\mathsf{f}}(\omega) e^{i\omega x}\ \mathrm{d}\omega \qquad \forall \tilde{\mathsf{f}} \in L^2_{(\mathbb{R},\mathscr{B},\mu)}$$

Theorem P.2. *Let $\tilde{\mathbf{F}}$ be the Fourier Transform operator with inverse $\tilde{\mathbf{F}}^{-1}$ and adjoint $\tilde{\mathbf{F}}^*$.*

THM

$$\tilde{\mathbf{F}}^* = \tilde{\mathbf{F}}^{-1}$$

✎PROOF:

$$\begin{aligned}
\left\langle \tilde{\mathbf{F}}\mathsf{f} \,\middle|\, \mathsf{g} \right\rangle &= \left\langle \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathsf{f}(x)\, e^{-i\omega x}\ \mathrm{d}x \,\middle|\, \mathsf{g}(\omega) \right\rangle && \text{by definition of } \tilde{\mathbf{F}} \text{ page 349} \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathsf{f}(x) \left\langle e^{-i\omega x} \,\middle|\, \mathsf{g}(\omega) \right\rangle\ \mathrm{d}x && \text{by } \textit{additive property} \text{ of } \langle \triangle \mid \triangledown \rangle \text{ page 271} \\
&= \int_{\mathbb{R}} \mathsf{f}(x) \frac{1}{\sqrt{2\pi}} \left\langle \mathsf{g}(\omega) \,\middle|\, e^{-i\omega x} \right\rangle^*\ \mathrm{d}x && \text{by } \textit{conjugate symmetric property} \text{ of } \langle \triangle \mid \triangledown \rangle \text{ page 271}
\end{aligned}$$

[3] Chorin and Hald (2009) page 13, Jeffrey and Dai (2008) pages xxxi–xxxii, Knapp (2005b) pages 374–375
[4] Chorin and Hald (2009) page 13

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \left\langle \mathsf{f}(x) \,\middle|\, \frac{1}{\sqrt{2\pi}} \left\langle \mathsf{g}(\omega) \,\middle|\, e^{-i\omega x} \right\rangle \right\rangle \qquad \text{by definition of } \langle \triangle \mid \triangledown \rangle$$

$$= \left\langle \mathsf{f} \,\middle|\, \underbrace{\tilde{\mathbf{F}}^{-1}}_{\tilde{\mathbf{F}}^*} \mathsf{g} \right\rangle \qquad \text{by Theorem P.1 page 350}$$

✏

The Fourier Transform operator has several nice properties:

- $\tilde{\mathbf{F}}$ is *unitary*[5] (Corollary P.1—next corollary).
- Because $\tilde{\mathbf{F}}$ is unitary, it automatically has several other nice properties (Theorem P.3 page 351).

Corollary P.1. *Let* $\mathbf{I}$ *be the identity operator and let* $\tilde{\mathbf{F}}$ *be the Fourier Transform operator with adjoint* $\tilde{\mathbf{F}}^*$ *and inverse* $\tilde{\mathbf{F}}^{-1}$.

COR

$$\underbrace{\tilde{\mathbf{F}}\tilde{\mathbf{F}}^* = \tilde{\mathbf{F}}^*\tilde{\mathbf{F}} = \mathbf{I}}_{\tilde{\mathbf{F}}^* = \tilde{\mathbf{F}}^{-1}} \qquad (\tilde{\mathbf{F}} \text{ is unitary})$$

✎PROOF: This follows directly from the fact that $\tilde{\mathbf{F}}^* = \tilde{\mathbf{F}}^{-1}$ (Theorem P.2 page 350). ✏

Theorem P.3. *Let* $\tilde{\mathbf{F}}$ *be the Fourier transform operator with adjoint* $\tilde{\mathbf{F}}^*$ *and inverse* $\tilde{\mathbf{F}}$. *Let* $|||\cdot|||$ *be the operator norm with respect to the vector norm* $\|\cdot\|$ *with respect to the Hilbert space* $\left(\mathbb{C}^{\mathbb{R}}, \langle \triangle \mid \triangledown \rangle\right)$. *Let* $\mathcal{R}(\mathbf{A})$ *be the range of an operator* $\mathbf{A}$.

THM

$$\begin{array}{rclcll}
\mathcal{R}(\mathbf{F}\tau) & = & \mathcal{R}(\tilde{\mathbf{F}}^{-1}) & = & L^2_{\mathbb{R}} & \\
|||\tilde{\mathbf{F}}||| & = & |||\tilde{\mathbf{F}}^{-1}||| & = & 1 & \text{(UNITARY)} \\
\left\langle \tilde{\mathbf{F}}\mathsf{f} \,\middle|\, \tilde{\mathbf{F}}\mathsf{g} \right\rangle & = & \left\langle \tilde{\mathbf{F}}^{-1}\mathsf{f} \,\middle|\, \tilde{\mathbf{F}}^{-1}\mathsf{g} \right\rangle & = & \langle \mathsf{f} \mid \mathsf{g} \rangle & \text{(PARSEVAL'S EQUATION)} \\
\left\|\tilde{\mathbf{F}}\mathsf{f}\right\| & = & \left\|\tilde{\mathbf{F}}^{-1}\mathsf{f}\right\| & = & \|\mathsf{f}\| & \text{(PLANCHEREL'S FORMULA)} \\
\left\|\tilde{\mathbf{F}}\mathsf{f} - \tilde{\mathbf{F}}\mathsf{g}\right\| & = & \left\|\tilde{\mathbf{F}}^{-1}\mathsf{f} - \tilde{\mathbf{F}}^{-1}\mathsf{g}\right\| & = & \|\mathsf{f} - \mathsf{g}\| & \text{(ISOMETRIC)}
\end{array}$$

✎PROOF: These results follow directly from the fact that $\tilde{\mathbf{F}}$ is unitary (Corollary P.1 page 351) and from the properties of unitary operators (Theorem O.26 page 343). ✏

Theorem P.4 (Shift relations)**.** *Let* $\tilde{\mathbf{F}}$ *be the Fourier transform operator.*

THM

$$\begin{array}{rcl}
\tilde{\mathbf{F}}[\mathsf{f}(x-u)](\omega) & = & e^{-i\omega u}\left[\tilde{\mathbf{F}}\mathsf{f}(x)\right](\omega) \\
\left[\tilde{\mathbf{F}}\left(e^{ivx}\mathsf{g}(x)\right)\right](\omega) & = & \left[\tilde{\mathbf{F}}\mathsf{g}(x)\right](\omega - v)
\end{array}$$

✎PROOF:

$$\begin{aligned}
\tilde{\mathbf{F}}[\mathsf{f}(x-u)](\omega) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathsf{f}(x-u)e^{-i\omega x}\,\mathrm{d}x && \text{by definition of } \tilde{\mathbf{F}} && \text{(Definition P.2 page 349))} \\
&= \frac{1}{\sqrt{2\pi}} \int_{v\in\mathbb{R}} \mathsf{f}(v)e^{-i\omega(u+v)}\,\mathrm{d}v && \text{where } v \triangleq x-u \implies t = u+v \\
&= e^{-i\omega u}\,\frac{1}{\sqrt{2\pi}} \int_{v\in\mathbb{R}} \mathsf{f}(v)e^{-i\omega v}\,\mathrm{d}v \\
&= e^{-i\omega u}\,\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathsf{f}(x)e^{-i\omega x}\,\mathrm{d}x && \text{by change of variable } t = v \\
&= e^{-i\omega u}\left[\tilde{\mathbf{F}}\mathsf{f}(x)\right](\omega) && \text{by definition of } \tilde{\mathbf{F}} && \text{(Definition P.2 page 349))} \\
\left[\tilde{\mathbf{F}}\left(e^{ivx}\mathsf{g}(x)\right)\right](\omega) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ivx}\mathsf{g}(x)e^{-i\omega x}\,\mathrm{d}x && \text{by definition of } \tilde{\mathbf{F}} && \text{(Definition P.2 page 349))}
\end{aligned}$$

[5] *unitary operators*: Definition O.11 page 342

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

$$= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(x) e^{-i(\omega - v)x} \, dx$$

$$= \left[\tilde{\mathbf{F}} g(x)\right](\omega - v) \qquad \text{by definition of } \tilde{\mathbf{F}} \qquad \text{(Definition P.2 page 349))}$$

✏

Theorem P.5 (Complex conjugate)**.** *Let $\tilde{\mathbf{F}}$ be the Fourier Transform operator and $*$ represent the complex conjugate operation on the set of complex numbers.*

THM

$$\tilde{\mathbf{F}} f^*(-x) = -\left[\tilde{\mathbf{F}} f(x)\right]^* \quad \forall f \in L^2_{(\mathbb{R},\mathscr{B},\mu)}$$

$$f \text{ is real} \implies \tilde{f}(-\omega) = \left[\tilde{f}(\omega)\right]^* \quad \forall \omega \in \mathbb{R} \qquad \text{REALITY CONDITION}$$

✎PROOF:

$$\left[\tilde{\mathbf{F}} f^*(-x)\right](\omega) \triangleq \frac{1}{\sqrt{2\pi}} \int f^*(-x) e^{-i\omega x} \, dx \qquad \text{by definition of } \tilde{\mathbf{F}} \qquad \text{(Definition P.2 page 349)}$$

$$= \frac{1}{\sqrt{2\pi}} \int f^*(u) e^{i\omega u} (-1) \, du \qquad \text{where } u \triangleq -x \implies dx = -\, du$$

$$= -\left[\frac{1}{\sqrt{2\pi}} \int f(u) e^{-i\omega u} \, du\right]^*$$

$$\triangleq -\left[\tilde{\mathbf{F}} f(x)\right]^* \qquad \text{by definition of } \tilde{\mathbf{F}} \qquad \text{(Definition P.2 page 349)}$$

$$\tilde{f}(-\omega) \triangleq \frac{1}{\sqrt{2\pi}} \int f(x) e^{-i(-\omega)x} \, dx \qquad \text{by definition of } \tilde{\mathbf{F}} \qquad \text{(Definition P.2 page 349)}$$

$$= \left[\frac{1}{\sqrt{2\pi}} \int f^*(x) e^{-i\omega x} \, dx\right]^*$$

$$= \left[\frac{1}{\sqrt{2\pi}} \int f(x) e^{-i\omega x} \, dx\right]^* \qquad \text{by } f \text{ is real hypothesis}$$

$$\triangleq \tilde{f}^*(\omega) \qquad \text{by definition of } \tilde{\mathbf{F}} \qquad \text{(Definition P.2 page 349)}$$

✏

P.3 Convolution

Definition P.3. [6]

DEF

*The **convolution operation** is defined as*

$$\left[f \star g\right](x) \triangleq f(x) \star g(x) \triangleq \int_{u \in \mathbb{R}} f(u) g(x-u) \, du \qquad \forall f,g \in L^2_{(\mathbb{R},\mathscr{B},\mu)}$$

Theorem R.2 (next) demonstrates that multiplication in the "time domain" is equivalent to convolution in the "frequency domain" and vice-versa.

Theorem P.6 (convolution theorem)**.** [7] *Let $\tilde{\mathbf{F}}$ be the Fourier Transform operator and $\star$ the convolution operator.*

[6] Bachman (1964), page 6, Bracewell (1978) page 108 ⟨Convolution theorem⟩

[7] Bracewell (1978) page 110

COPYRIGHT © 2019 DANIEL J. GREENHOE

THM

$$\underbrace{\tilde{\mathbf{F}}\big[\mathsf{f}(x)\star\mathsf{g}(x)\big](\omega)}_{\textit{convolution in "time domain"}} = \underbrace{\sqrt{2\pi}\big[\tilde{\mathbf{F}}\mathsf{f}\big](\omega)\,\big[\tilde{\mathbf{F}}\mathsf{g}\big](\omega)}_{\textit{multiplication in "frequency domain"}} \qquad \forall \mathsf{f},\mathsf{g}\in L^2_{(\mathbb{R},\mathscr{B},\mu)}$$

$$\underbrace{\tilde{\mathbf{F}}\big[\mathsf{f}(x)\mathsf{g}(x)\big](\omega)}_{\textit{multiplication in "time domain"}} = \underbrace{\frac{1}{\sqrt{2\pi}}\big[\tilde{\mathbf{F}}\mathsf{f}\big](\omega)\star\big[\tilde{\mathbf{F}}\mathsf{g}\big](\omega)}_{\textit{convolution in "frequency domain"}} \qquad \forall \mathsf{f},\mathsf{g}\in L^2_{(\mathbb{R},\mathscr{B},\mu)}.$$

✎PROOF:

$$\begin{aligned}
\tilde{\mathbf{F}}\big[\mathsf{f}(x)\star\mathsf{g}(x)\big](\omega) &= \tilde{\mathbf{F}}\left[\int_{u\in\mathbb{R}}\mathsf{f}(u)\mathsf{g}(x-u)\,\mathrm{d}u\right](\omega) && \text{by definition of } \star \text{ (Definition P.3 page 352)}\\
&= \int_{u\in\mathbb{R}}\mathsf{f}(u)\big[\tilde{\mathbf{F}}\mathsf{g}(x-u)\big](\omega)\,\mathrm{d}u\\
&= \int_{u\in\mathbb{R}}\mathsf{f}(u)e^{-i\omega u}\,\big[\tilde{\mathbf{F}}\mathsf{g}(x)\big](\omega)\,\mathrm{d}u && \text{by Theorem P.4 page 351}\\
&= \sqrt{2\pi}\underbrace{\left(\frac{1}{\sqrt{2\pi}}\int_{u\in\mathbb{R}}\mathsf{f}(u)e^{-i\omega u}\,\mathrm{d}u\right)}_{[\tilde{\mathbf{F}}\mathsf{f}](\omega)}\big[\tilde{\mathbf{F}}\mathsf{g}\big](\omega)\\
&= \sqrt{2\pi}\big[\tilde{\mathbf{F}}\mathsf{f}\big](\omega)\,\big[\tilde{\mathbf{F}}\mathsf{g}\big](\omega) && \text{by definition of } \tilde{\mathbf{F}} \text{ (Definition P.2 page 349)}\\
\tilde{\mathbf{F}}[\mathsf{f}(x)\mathsf{g}(x)](\omega) &= \tilde{\mathbf{F}}\big[\big(\tilde{\mathbf{F}}^{-1}\tilde{\mathbf{F}}\mathsf{f}(x)\big)\,\mathsf{g}(x)\big](\omega) && \text{by definition of operator inverse (page 319)}\\
&= \tilde{\mathbf{F}}\left[\left(\frac{1}{\sqrt{2\pi}}\int_{v\in\mathbb{R}}\big[\tilde{\mathbf{F}}\mathsf{f}(x)\big](v)e^{ivx}\,\mathrm{d}v\right)\mathsf{g}(x)\right](\omega) && \text{by Theorem P.1 page 350}\\
&= \frac{1}{\sqrt{2\pi}}\int_{v\in\mathbb{R}}\big[\tilde{\mathbf{F}}\mathsf{f}(x)\big](v)\big[\tilde{\mathbf{F}}\big(e^{ivx}\,\mathsf{g}(x)\big)\big](\omega,v)\,\mathrm{d}v\\
&= \frac{1}{\sqrt{2\pi}}\int_{v\in\mathbb{R}}\big[\tilde{\mathbf{F}}\mathsf{f}(x)\big](v)\big[\tilde{\mathbf{F}}\mathsf{g}(x)\big](\omega-v)\,\mathrm{d}v && \text{by Theorem P.4 page 351}\\
&= \frac{1}{\sqrt{2\pi}}\big[\tilde{\mathbf{F}}\mathsf{f}\big](\omega)\star\big[\tilde{\mathbf{F}}\mathsf{g}\big](\omega) && \text{by definition of } \star \text{ (Definition P.3 page 352)}
\end{aligned}$$

✏

P.4 Real valued functions

Figure P.1: Fourier transform components of real-valued signal

Theorem P.7. *Let* $\mathsf{f}(x)$ *be a function in* $L^2_{\mathbb{R}}$ *and* $\tilde{\mathsf{f}}(\omega)$ *the* FOURIER TRANSFORM *of* $\mathsf{f}(x)$.

THM

$$\left\{\begin{array}{l}\mathsf{f}(x)\ \textit{is}\ \text{REAL-VALUED}\\ (\mathsf{f}\in\mathbb{R}^{\mathbb{R}})\end{array}\right\} \implies \left\{\begin{array}{rcll}\tilde{\mathsf{f}}(\omega) &=& \tilde{\mathsf{f}}^*(-\omega) & (\text{HERMITIAN SYMMETRIC})\\ \mathbf{R}_e\left[\tilde{\mathsf{f}}(\omega)\right] &=& \mathbf{R}_e\left[\tilde{\mathsf{f}}(-\omega)\right] & (\text{SYMMETRIC})\\ \mathbf{I}_m\left[\tilde{\mathsf{f}}(\omega)\right] &=& -\mathbf{I}_m\left[\tilde{\mathsf{f}}(-\omega)\right] & (\text{ANTI-SYMMETRIC})\\ |\tilde{\mathsf{f}}(\omega)| &=& |\tilde{\mathsf{f}}(-\omega)| & (\text{SYMMETRIC})\\ \angle\tilde{\mathsf{f}}(\omega) &=& \angle\tilde{\mathsf{f}}(-\omega) & (\text{ANTI-SYMMETRIC}).\end{array}\right\}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

✎PROOF:

$$\begin{array}{lllllll}
\tilde{f}(\omega) & \triangleq & [\tilde{\mathbf{F}}f(x)](\omega) & \triangleq & \left\langle f(x) \,|\, e^{i\omega x}\right\rangle & = & \left\langle f(x) \,|\, e^{i(-\omega)x}\right\rangle^* \triangleq \tilde{f}^*(-\omega) \\
\mathbf{R}_e\left[\tilde{f}(\omega)\right] & = & \mathbf{R}_e\left[\tilde{f}^*(-\omega)\right] & = & \mathbf{R}_e\left[\tilde{f}(-\omega)\right] & & \\
\mathbf{I}_m\left[\tilde{f}(\omega)\right] & = & \mathbf{I}_m\left[\tilde{f}^*(-\omega)\right] & = & -\mathbf{I}_m\left[\tilde{f}(-\omega)\right] & & \\
|\tilde{f}(\omega)| & = & |\tilde{f}^*(-\omega)| & = & |\tilde{f}(-\omega)| & & \\
\angle\tilde{f}(\omega) & = & \angle\tilde{f}^*(-\omega) & = & -\angle\tilde{f}(-\omega) & &
\end{array}$$

✏

P.5 Moment properties

Definition P.4. [8]

DEF *The quantity* M_n *is the* n***th moment*** *of a function* $f(x) \in L^2_{\mathbb{R}}$ *if*

$$\mathsf{M}_n \triangleq \int_{\mathbb{R}} x^n f(x)\,dx \qquad \text{for } n \in \mathbb{W}.$$

Lemma P.1. [9] *Let* M_n *be the* nTH MOMENT *(Definition P.4 page 354) and* $\tilde{f}(\omega) \triangleq \left[\tilde{\mathbf{F}}f\right](\omega)$ *the* FOURIER TRANSFORM *(Definition P.2 page 349) of a function* $f(x)$ *in* $L^2_{\mathbb{R}}$ *(Definition ?? page ??).*

LEM

$$\begin{array}{rcll}
\mathsf{M}_n & = & \sqrt{2\pi}(i)^n\left[\frac{d}{d\omega}\right]^n\tilde{f}(\omega)\Big|_{\omega=0} & \forall n\in\mathbb{W}, f\in L^2_{\mathbb{R}} \\
\left[\frac{d}{d\omega}\right]^n\tilde{f}(\omega)\Big|_{\omega=0} & = & \frac{1}{\sqrt{2\pi}}(-i)^n\mathsf{M}_n & \forall n\in\mathbb{W}, f\in L^2_{\mathbb{R}}
\end{array}$$

✎PROOF:

$$\begin{aligned}
\sqrt{2\pi}(i)^n\left[\left[\frac{d}{d\omega}\right]^n\tilde{f}(\omega)\right]_{\omega=0} &= \sqrt{2\pi}(i)^n\left[\left[\frac{d}{d\omega}\right]^n\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)e^{-i\omega x}\,dx\right]_{\omega=0} && \text{by definition of } \tilde{\mathbf{F}} \text{ (Definition P.2 page 349)} \\
&= (i)^n\int_{\mathbb{R}} f(x)\left[\left[\frac{d}{d\omega}\right]^n e^{-i\omega x}\right]dx\Big|_{\omega=0} \\
&= (i)^n\int_{\mathbb{R}} f(x)\left[(-i)^n x^n e^{-i\omega x}\right]dx\Big|_{\omega=0} \\
&= (-i^2)^n\int_{\mathbb{R}} f(x)x^n\,dx \\
&= \int_{\mathbb{R}} f(x)x^n\,dx \\
&\triangleq \mathsf{M}_n && \text{by definition of } \mathsf{M}_n \text{ (Definition P.4 page 354)}
\end{aligned}$$

✏

Lemma P.2. [10] *Let* M_n *be the* nTH MOMENT *(Definition P.4 page 354) and* $\tilde{f}(\omega) \triangleq \left[\tilde{\mathbf{F}}f\right](\omega)$ *the* FOURIER TRANSFORM *(Definition P.2 page 349) of a function* $f(x)$ *in* $L^2_{\mathbb{R}}$ *(Definition ?? page ??).*

LEM

$$\mathsf{M}_n = 0 \quad\Longleftrightarrow\quad \left[\frac{d}{d\omega}\right]^n\tilde{f}(\omega)\Big|_{\omega=0} = 0 \qquad \forall n\in\mathbb{W}$$

✎PROOF:

[8] Jawerth and Sweldens (1994), pages 16–17, Sweldens and Piessens (1993), page 2, Vidakovic (1999), page 83

[9] Goswami and Chan (1999), pages 38–39

[10] Vidakovic (1999), pages 82–83, Mallat (1999), pages 241–242

COPYRIGHT © 2019 DANIEL J. GREENHOE

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

$$\begin{aligned}
0 &= \langle \mathsf{f}(x) \mid x^n \rangle && \text{by left hypothesis}\\
&= \sqrt{2\pi}(-i)^{-n}\left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \tilde{\mathsf{f}}(\omega)\Big|_{\omega=0} && \text{by Lemma P.1 page 354}\\
\Longrightarrow\ & \left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \tilde{\mathsf{f}}(\omega)\Big|_{\omega=0} = 0
\end{aligned}$$

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

$$\begin{aligned}
0 &= \left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \tilde{\mathsf{f}}(\omega)\Big|_{\omega=0} && \text{by right hypothesis}\\
&= \left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathsf{f}(x)e^{-i\omega x}\,\mathsf{d}x\Big|_{\omega=0} && \text{by definition of } \tilde{\mathsf{f}}(\omega)\\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathsf{f}(x)\left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n e^{-i\omega x}\,\mathsf{d}x\Big|_{\omega=0}\\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathsf{f}(x)\left[(-i)^n x^n e^{-i\omega x}\right]\,\mathsf{d}x\Big|_{\omega=0}\\
&= (-i)^n\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathsf{f}(x)x^n\,\mathsf{d}x\\
&= (-i)^n\frac{1}{\sqrt{2\pi}}\langle \mathsf{f}(x) \mid x^n\rangle && \text{by definition of } \langle\cdot\mid\cdot\rangle \text{ in } \mathbf{L}^2_{\mathbb{R}} \text{ (Definition ?? page ??)}
\end{aligned}$$

✏

Lemma P.3 (Strang-Fix condition). [11] *Let* $\mathsf{f}(x)$ *be a function in* $\mathbf{L}^2_{\mathbb{R}}$ *and* M_n *the* nTH MOMENT *(Definition P.4 page 354) of* $\mathsf{f}(x)$. *Let* $\mathbf{T}$ *be the* TRANSLATION OPERATOR *(Definition ?? page ??).*

LEM

$$\underbrace{\sum_{k\in\mathbb{Z}} \mathbf{T}^k x^n \mathsf{f}(x) = \mathsf{M}_n}_{\text{STRANG-FIX CONDITION in "time"}} \quad\Longleftrightarrow\quad \underbrace{\left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \tilde{\mathsf{f}}(\omega)\Big|_{\omega=2\pi k} = \frac{1}{\sqrt{2\pi}}(-i)^n\bar{\delta}_k\mathsf{M}_n}_{\text{STRANG-FIX CONDITION in "frequency"}}$$

✎PROOF:

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

$$\begin{aligned}
\left[\left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \tilde{\mathsf{f}}(\omega)\right]_{\omega=2\pi k} &= \sum_{k\in\mathbb{Z}}\left[\left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \tilde{\mathsf{f}}(\omega)\right]_{\omega=2\pi k} e^{i2\pi kx}\bar{\delta}_k\\
&= \frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}}\left[\left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \int_{\mathbb{R}} \mathsf{f}(x)e^{-i\omega x}\,\mathsf{d}x\right]_{\omega=2\pi k} e^{i2\pi kx}\bar{\delta}_k && \text{by Definition P.2 page 349}\\
&= \frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}}\left[\int_{\mathbb{R}} \mathsf{f}(x)(-ix)^n e^{-i\omega x}\,\mathsf{d}x\right]_{\omega=2\pi k} e^{i2\pi kx}\bar{\delta}_k\\
&= (-i)^n\frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}}\left[\int_{\mathbb{R}} x^n\mathsf{f}(x)e^{-i\omega x}\,\mathsf{d}x\right]_{\omega=2\pi k} e^{i2\pi kx}\bar{\delta}_k\\
&= (-i)^n\frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}}(x-k)^n\mathsf{f}(x-k)\bar{\delta}_k && \text{by } PSF \text{ (Theorem ?? page ??)}\\
&= \frac{1}{\sqrt{2\pi}}(-i)^n\bar{\delta}_k\mathsf{M}_n && \text{by left hypothesis}
\end{aligned}$$

[11] Jawerth and Sweldens (1994), pages 16–17, Sweldens and Piessens (1993), page 2, Vidakovic (1999), page 83, Mallat (1999), pages 241–243, Fix and Strang (1969)

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

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

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}}(-i)^n \mathsf{M}_n &= \frac{1}{\sqrt{2\pi}} \sum_{k\in\mathbb{Z}} \left[(-i)^n \bar{\delta}_k \mathsf{M}_n\right] e^{-i2\pi kx} && \text{by definition of } \bar{\delta} && \text{(Definition K.3 page 283)}\\
&= \sum_{k\in\mathbb{Z}} \left[\left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \tilde{\mathsf{f}}(\omega)\right]\Bigg|_{\omega=2\pi k} e^{-i2\pi kx} && \text{by right hypothesis}\\
&= \sum_{k\in\mathbb{Z}} \left[\left[\frac{\mathsf{d}}{\mathsf{d}\omega}\right]^n \int_{\mathbb{R}} \mathsf{f}(x)e^{-i\omega x}\,\mathsf{d}x\right]\Bigg|_{\omega=2\pi k} e^{-i2\pi kx}\\
&= \sum_{k\in\mathbb{Z}} \left[\int_{\mathbb{R}} \mathsf{f}(x)(-ix)^n e^{-i\omega x}\,\mathsf{d}x\right]\Bigg|_{\omega=2\pi k} e^{-i2\pi kx}\\
&= (-i)^n \sum_{k\in\mathbb{Z}} \left[\int_{\mathbb{R}} x^n \mathsf{f}(x)e^{-i\omega x}\,\mathsf{d}x\right]\Bigg|_{\omega=2\pi k} e^{-i2\pi kx}\\
&= (-i)^n \sum_{k\in\mathbb{Z}} (x-k)^n \mathsf{f}(x-k) && \text{by } PSF && \text{(Theorem ?? page ??)}
\end{aligned}$$

✏

P.6 Examples

Example P.1 (rectangular pulse). Let $\tilde{\mathsf{f}}(\omega)$ be the *Fourier transform* of a function $\mathsf{f}(x) \in \boldsymbol{L}^2_{\mathbb{R}}$.

EX

$$\mathsf{f}(x) = \begin{cases} c & \text{for } x \in [a : b)\\ 0 & \text{otherwise}\end{cases} \qquad \tilde{\mathsf{f}}(\omega) = \frac{c(b-a)}{\sqrt{2\pi}} e^{-i\left(\frac{a+b}{2}\omega\right)} \left[\frac{\sin\left(\frac{b-a}{2}\omega\right)}{\left(\frac{b-a}{2}\omega\right)}\right]$$

✎PROOF:

$$\begin{aligned}
\tilde{\mathsf{f}}(\omega) &= \tilde{\mathbf{F}}[\mathsf{f}(x)](\omega) && \text{by definition of } \tilde{\mathsf{f}}(\omega)\\
&= e^{-i\left(\frac{a+b}{2}\right)\omega} \tilde{\mathbf{F}}\left[\mathsf{f}\left(x - \frac{a+b}{2}\right)\right](\omega) && \text{by } \textit{shift relation} && \text{(Theorem P.4 page 351)}\\
&= e^{-i\left(\frac{a+b}{2}\right)\omega} \tilde{\mathbf{F}}\left[c\mathbb{1}_{[a:b)}\left(x - \frac{a+b}{2}\right)\right](\omega) && \text{by definition of } \mathsf{f}(x)\\
&= e^{-i\left(\frac{a+b}{2}\right)\omega} \tilde{\mathbf{F}}\left[c\mathbb{1}_{\left[-\frac{b-a}{2}:\frac{b-a}{2}\right)}(x)\right](\omega) && \text{by definition of } \mathbb{1} && \text{(Definition ?? page ??)}\\
&= \frac{1}{\sqrt{2\pi}} e^{-i\left(\frac{a+b}{2}\right)\omega} \int_{\mathbb{R}} c\mathbb{1}_{\left[-\frac{b-a}{2}:\frac{b-a}{2}\right)}(x) e^{-i\omega x}\,\mathsf{d}x && \text{by definition of } \tilde{\mathbf{F}} && \text{(Definition P.2 page 349)}\\
&= \frac{1}{\sqrt{2\pi}} e^{-i\left(\frac{a+b}{2}\right)\omega} \int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} c e^{-i\omega x}\,\mathsf{d}x && \text{by definition of } \mathbb{1} && \text{(Definition ?? page ??)}\\
&= \frac{c}{\sqrt{2\pi}} e^{-i\left(\frac{a+b}{2}\right)\omega} \frac{1}{-i\omega} e^{-i\omega x}\Bigg|_{-\frac{b-a}{2}}^{\frac{b-a}{2}}\\
&= \frac{2c}{\sqrt{2\pi}\omega} e^{-i\left(\frac{a+b}{2}\right)\omega} \left[\frac{e^{i\left(\frac{b-a}{2}\omega\right)} - e^{-i\left(\frac{b-a}{2}\omega\right)}}{2i}\right]
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \frac{c(b-a)}{\sqrt{2\pi}} e^{-i\left(\frac{a+b}{2}\omega\right)} \left[\frac{\sin\left(\frac{b-a}{2}\omega\right)}{\left(\frac{b-a}{2}\omega\right)} \right] \qquad \text{by } \textit{Euler formulas} \qquad \text{(Corollary ?? page ??)}$$

✏

Example P.2 (triangle). Let $\tilde{f}(\omega)$ be the *Fourier transform* of a function $f(x) \in L^2_{\mathbb{R}}$.

EX

$$f(x) = \begin{cases} c\left[1 - \frac{|2x-b-a|}{b-a}\right] & \text{for } x \in [a : b) \\ 0 & \text{otherwise} \end{cases} \qquad \tilde{f}(\omega) = \frac{c(b-a)^2}{4\sqrt{2\pi}} e^{-i\left(\frac{a+b}{2}\omega\right)} \left[\frac{\sin\left(\frac{b-a}{4}\omega\right)}{\left(\frac{b-a}{4}\omega\right)} \right]^2$$

✎PROOF:

$$\begin{aligned}
\tilde{f}(\omega) &= \tilde{\mathbf{F}}[f(x)](\omega) && \text{by definition of } \tilde{f}(\omega) \\
&= e^{-i\left(\frac{a+b}{2}\right)\omega} \tilde{\mathbf{F}}\left[f\left(x - \frac{a+b}{2}\right)\right](\omega) && \text{by } \textit{shift relation} \qquad \text{(Theorem P.4 page 351)} \\
&= \tilde{\mathbf{F}}\left[c\left(1 - \frac{|2x-b-a|}{b-a}\right)\mathbb{1}_{[a:b)}(x)\right](\omega) && \text{by definition of } f(x) \\
&= c\tilde{\mathbf{F}}\left[\mathbb{1}_{\left[\frac{a}{2}:\frac{b}{2}\right)}(x) \star \mathbb{1}_{\left[\frac{a}{2}:\frac{b}{2}\right)}(x)\right](\omega) \\
&= c\sqrt{2\pi}\tilde{\mathbf{F}}\left[\mathbb{1}_{\left[\frac{a}{2}:\frac{b}{2}\right)}\right]\tilde{\mathbf{F}}\left[\mathbb{1}_{\left[\frac{a}{2}:\frac{b}{2}\right)}\right] && \text{by } \textit{convolution theorem} \qquad \text{(Theorem R.2 page 372)} \\
&= c\sqrt{2\pi}\left(\tilde{\mathbf{F}}\left[\mathbb{1}_{\left[\frac{a}{2}:\frac{b}{2}\right)}\right]\right)^2 \\
&= c\sqrt{2\pi}\left(\frac{\left(\frac{b}{2}-\frac{a}{2}\right)}{\sqrt{2\pi}} e^{-i\left(\frac{a+b}{4}\omega\right)} \left[\frac{\sin\left(\frac{b-a}{4}\omega\right)}{\left(\frac{b-a}{4}\omega\right)}\right]\right)^2 && \text{by } \textit{Rectangular pulse} \text{ ex.} \qquad \text{Example P.1 page 356} \\
&= \frac{c(b-a)^2}{4\sqrt{2\pi}} e^{-i\left(\frac{a+b}{2}\omega\right)} \left[\frac{\sin\left(\frac{b-a}{4}\omega\right)}{\left(\frac{b-a}{4}\omega\right)}\right]^2
\end{aligned}$$

✏

Example P.3. Let a function f be defined in terms of the cosine function (Definition ?? page ??) as follows:

EX

$$f(x) \triangleq \begin{cases} \cos^2\left(\frac{\pi}{2}x\right) & \text{for } |x| \le 1 \\ 0 & \text{otherwise} \end{cases} \qquad \tilde{f}(\omega) = \frac{1}{2\sqrt{2\pi}}\left[\underbrace{\frac{2\sin\omega}{\omega}}_{2\,\mathrm{sinc}(\omega)} + \underbrace{\frac{\sin(\omega+\pi)}{(\omega+\pi)}}_{\mathrm{sinc}(\omega+\pi)} + \underbrace{\frac{\sin(\omega-\pi)}{(\omega-\pi)}}_{\mathrm{sinc}(\omega-\pi)}\right]$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✎PROOF: Let $\mathbb{1}_A(x)$ be the *set indicator function* (Definition ?? page ??) on a set A.

$$\begin{aligned}
\tilde{\mathsf{f}}(\omega) &\triangleq \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathsf{f}(x)e^{-i\omega x}\,\mathrm{d}x && \text{by definition of } \tilde{\mathsf{f}}(\omega) \text{ (Definition P.2)}\\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \cos^2\left(\frac{\pi}{2}x\right)\mathbb{1}_{[-1:1]}(x)e^{-i\omega x}\,\mathrm{d}x && \text{by definition of } \mathsf{f}(x)\\
&= \frac{1}{\sqrt{2\pi}}\int_{-1}^{1} \cos^2\left(\frac{\pi}{2}x\right)e^{-i\omega x}\,\mathrm{d}x && \text{by definition of } \mathbb{1} \text{ (Definition ??)}\\
&= \frac{1}{\sqrt{2\pi}}\int_{-1}^{1} \left[\frac{e^{i\frac{\pi}{2}x}+e^{-i\frac{\pi}{2}x}}{2}\right]^2 e^{-i\omega x}\,\mathrm{d}x && \text{by Corollary ?? page ??}\\
&= \frac{1}{4\sqrt{2\pi}}\int_{-1}^{1} \left[2+e^{i\pi x}+e^{-i\pi x}\right]e^{-i\omega x}\,\mathrm{d}x\\
&= \frac{1}{4\sqrt{2\pi}}\int_{-1}^{1} 2e^{-i\omega x}+e^{-i(\omega+\pi)x}+e^{-i(\omega-\pi)x}\,\mathrm{d}x\\
&= \frac{1}{4\sqrt{2\pi}}\left[2\frac{e^{-i\omega x}}{-i\omega}+\frac{e^{-i(\omega+\pi)x}}{-i(\omega+\pi)}+\frac{e^{-i(\omega-\pi)x}}{-i(\omega-\pi)}\right]_{-1}^{1}\\
&= \frac{1}{2\sqrt{2\pi}}\left[2\frac{e^{-i\omega}-e^{+i\omega}}{-2i\omega}+\frac{e^{-i(\omega+\pi)}-e^{+i(\omega+\pi)}}{-2i(\omega+\pi)}+\frac{e^{-i(\omega-\pi)}-e^{+i(\omega-\pi)}}{-2i(\omega-\pi)}\right]_{-1}^{1}\\
&= \frac{1}{2\sqrt{2\pi}}\left[\underbrace{\frac{2\sin\omega}{\omega}}_{2\,\mathrm{sinc}(\omega)}+\underbrace{\frac{\sin(\omega+\pi)}{(\omega+\pi)}}_{\mathrm{sinc}(\omega+\pi)}+\underbrace{\frac{\sin(\omega-\pi)}{(\omega-\pi)}}_{\mathrm{sinc}(\omega-\pi)}\right]
\end{aligned}$$

✏

COPYRIGHT © 2019 DANIEL J. GREENHOE

APPENDIX Q

DISCRETE TIME FOURIER TRANSFORM

Q.1 Definition

Definition Q.1.

DEF *The **discrete-time Fourier transform** $\breve{\mathbf{F}}$ of $(\!(x_n)\!)_{n\in\mathbb{Z}}$ is defined as*

$$\left[\breve{\mathbf{F}}\,(\!(x_n)\!)\right](\omega) \triangleq \sum_{n\in\mathbb{Z}} x_n e^{-i\omega n} \qquad \forall (\!(x_n)\!)_{n\in\mathbb{Z}} \in \ell^2_{\mathbb{R}}$$

If we compare the definition of the *Discrete Time Fourier Transform* (Definition Q.1 page 359) to the definition of the Z-transform (Definition R.4 page 370), we see that the DTFT is just a special case of the more general Z-Transform, with $z = e^{i\omega}$. If we imagine $z \in \mathbb{C}$ as a complex plane, then $e^{i\omega}$ is a unit circle in this plane. The "frequency" ω in the DTFT is the unit circle in the much larger z-plane, as illustrated to the right.

Q.2 Properties

Proposition Q.1 (DTFT periodicity)**.** *Let $\breve{x}(\omega) \triangleq \breve{\mathbf{F}}\left[(\!(x_n)\!)\right](\omega)$ be the* DISCRETE-TIME FOURIER TRANSFORM *(Definition Q.1 page 359) of a sequence $(\!(x_n)\!)_{n\in\mathbb{Z}}$ in $\ell^2_{\mathbb{R}}$.*

PRP $$\underbrace{\breve{x}(\omega) = \breve{x}(\omega + 2\pi n)}_{\text{PERIODIC } \textit{with period } 2\pi} \qquad \forall n \in \mathbb{Z}$$

✎PROOF:

$$\begin{aligned}
\breve{x}(\omega + 2\pi n) &= \sum_{m\in\mathbb{Z}} x_m e^{-i(\omega+2\pi n)m} &&= \sum_{m\in\mathbb{Z}} x_m e^{-i\omega m} \cancelto{1}{e^{-i2\pi n m}} \\
&= \sum_{m\in\mathbb{Z}} x_m e^{-i\omega m} &&= \breve{x}(\omega)
\end{aligned}$$

✏

Theorem Q.1. *Let* $\breve{\mathsf{x}}(\omega) \triangleq \breve{\mathbf{F}}\big[(\!|\mathsf{x}[n]|\!)\big](\omega)$ *be the* DISCRETE-TIME FOURIER TRANSFORM *(Definition Q.1 page 359) of a sequence* $(\!|x_n|\!)_{n\in\mathbb{Z}}$ *in* $\ell^2_{\mathbb{R}}$.

THM

$$\left\{\ \breve{\mathsf{x}}(\omega) \triangleq \breve{\mathbf{F}}\left(\!|\mathsf{x}[n]|\!\right)\ \right\} \implies \left\{\begin{array}{llcll} (1). & \breve{\mathbf{F}}\left(\!|\mathsf{x}[-n]|\!\right) & = & \breve{\mathsf{x}}(-\omega) & \textit{and} \\ (2). & \breve{\mathbf{F}}\left(\!|\mathsf{x}^*[n]|\!\right) & = & \breve{\mathsf{x}}^*(-\omega) & \textit{and} \\ (3). & \breve{\mathbf{F}}\left(\!|\mathsf{x}^*[-n]|\!\right) & = & \breve{\mathsf{x}}^*(\omega) & \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\breve{\mathbf{F}}\left(\!|\mathsf{x}[-n]|\!\right) &\triangleq \sum_{n\in\mathbb{Z}} \mathsf{x}[-n]e^{-i\omega n} && \text{by definition of } \textit{DTFT} && \text{(Definition Q.1 page 359)} \\
&= \sum_{m\in\mathbb{Z}} \mathsf{x}[m]e^{i\omega m} && \text{where } m \triangleq -n \implies n = -m \\
&= \sum_{m\in\mathbb{Z}} \mathsf{x}[m]e^{-i(-\omega)m} \\
&\triangleq \breve{\mathsf{x}}(-\omega) && \text{by left hypothesis}
\end{aligned}$$

$$\begin{aligned}
\breve{\mathbf{F}}\left(\!|\mathsf{x}^*[n]|\!\right) &\triangleq \sum_{n\in\mathbb{Z}} \mathsf{x}^*[n]e^{-i\omega n} && \text{by definition of } \textit{DTFT} && \text{(Definition Q.1 page 359)} \\
&= \left(\sum_{n\in\mathbb{Z}} \mathsf{x}[n]e^{i\omega n}\right)^* && \text{by } \textit{distributive} \text{ property of } *\textbf{-algebras} && \text{(Definition J.3 page 266)} \\
&= \left(\sum_{n\in\mathbb{Z}} \mathsf{x}[n]e^{-i(-\omega)n}\right)^* \\
&\triangleq \breve{\mathsf{x}}^*(-\omega) && \text{by left hypothesis}
\end{aligned}$$

$$\begin{aligned}
\breve{\mathbf{F}}\left(\!|\mathsf{x}^*[-n]|\!\right) &\triangleq \sum_{n\in\mathbb{Z}} \mathsf{x}^*[-n]e^{-i\omega n} && \text{by definition of } \textit{DTFT} && \text{(Definition Q.1 page 359)} \\
&= \left(\sum_{n\in\mathbb{Z}} \mathsf{x}[-n]e^{i\omega n}\right)^* && \text{by } \textit{distributive} \text{ property of } *\textbf{-algebras} && \text{(Definition J.3 page 266)} \\
&= \left(\sum_{m\in\mathbb{Z}} \mathsf{x}[m]e^{-i\omega m}\right)^* && \text{where } m \triangleq -n \implies n = -m \\
&\triangleq \breve{\mathsf{x}}^*(\omega) && \text{by left hypothesis}
\end{aligned}$$

✏

Theorem Q.2. *Let* $\breve{\mathsf{x}}(\omega) \triangleq \breve{\mathbf{F}}\big[(\!|\mathsf{x}[n]|\!)\big](\omega)$ *be the* DISCRETE-TIME FOURIER TRANSFORM *(Definition Q.1 page 359) of a sequence* $(\!|\mathsf{x}[n]|\!)_{n\in\mathbb{Z}}$ *in* $\ell^2_{\mathbb{R}}$.

THM

$$\left\{\begin{array}{lll} (1). & \breve{\mathsf{x}}(\omega) \triangleq \breve{\mathbf{F}}\left(\!|\mathsf{x}[n]|\!\right) & \textit{and} \\ (2). & (\!|\mathsf{x}[n]|\!) \textit{ is } \text{REAL-VALUED} & \end{array}\right\} \implies \left\{\begin{array}{llcllll} (1). & \breve{\mathbf{F}}\left(\!|\mathsf{x}[-n]|\!\right) & = & \breve{\mathsf{x}}(-\omega) & & & \textit{and} \\ (2). & \breve{\mathbf{F}}\left(\!|\mathsf{x}^*[n]|\!\right) & = & \breve{\mathsf{x}}^*(-\omega) & = & \breve{\mathsf{x}}(\omega) & \textit{and} \\ (3). & \breve{\mathbf{F}}\left(\!|\mathsf{x}^*[-n]|\!\right) & = & \breve{\mathsf{x}}^*(\omega) & = & \breve{\mathsf{x}}(-\omega) & \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\breve{\mathbf{F}}\left(\!|\mathsf{x}[-n]|\!\right) &\triangleq \sum_{n\in\mathbb{Z}} \mathsf{x}[-n]e^{-i\omega n} && \text{by definition of } \textit{DTFT} && \text{(Definition Q.1 page 359)} \\
&= \sum_{m\in\mathbb{Z}} \mathsf{x}[m]e^{i\omega m} && \text{where } m \triangleq -n \implies n = -m \\
&= \sum_{m\in\mathbb{Z}} \mathsf{x}[m]e^{-i(-\omega)m}
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$\triangleq \tilde{\mathsf{x}}(-\omega) \qquad \text{by left hypothesis}$$

$$\begin{aligned}
\boxed{\tilde{\mathsf{x}}^*(-\omega)} &= \boxed{\breve{\mathbf{F}}\left(\!\left(\mathsf{x}^*[n]\right)\!\right)} && \text{by Theorem Q.1 page 360}\\
&= \breve{\mathbf{F}}\left(\!\left(\mathsf{x}[n]\right)\!\right) && \text{by } \textit{real-valued} \text{ hypothesis}\\
&= \boxed{\tilde{\mathsf{x}}(\omega)} && \text{by definition of } \tilde{\mathsf{x}}(\omega) \qquad \text{(Definition Q.1 page 359)}
\end{aligned}$$

$$\begin{aligned}
\boxed{\tilde{\mathsf{x}}^*(\omega)} &= \boxed{\breve{\mathbf{F}}\left(\!\left(\mathsf{x}^*[-n]\right)\!\right)} && \text{by Theorem Q.1 page 360}\\
&= \breve{\mathbf{F}}\left(\!\left(\mathsf{x}[-n]\right)\!\right) && \text{by } \textit{real-valued} \text{ hypothesis}\\
&= \boxed{\tilde{\mathsf{x}}(-\omega)} && \text{by result (1)}
\end{aligned}$$

✍

Proposition Q.2. *Let* $\breve{\mathsf{x}}(z)$ *be the* Z-TRANSFORM *(Definition R.4 page 370) and* $\breve{\mathsf{x}}(\omega)$ *the* DISCRETE-TIME FOURIER TRANSFORM *(Definition Q.1 page 359) of* $(\!(x_n)\!)$.

PRP

$$\underbrace{\left\{\sum_{n\in\mathbb{Z}} x_n = c\right\}}_{\text{(1) time domain}} \iff \underbrace{\left\{\breve{\mathsf{x}}(z)\Big|_{z=1} = c\right\}}_{\text{(2) } z \text{ domain}} \iff \underbrace{\left\{\breve{\mathsf{x}}(\omega)\Big|_{\omega=0} = c\right\}}_{\text{(3) frequency domain}} \qquad \forall (\!(x_n)\!)_{n\in\mathbb{Z}} \in \ell^2_{\mathbb{R}}, c\in\mathbb{R}$$

✎PROOF:

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

$$\begin{aligned}
\breve{\mathsf{x}}(z)\Big|_{z=1} &= \sum_{n\in\mathbb{Z}} x_n z^{-n}\Big|_{z=1} && \text{by definition of } \breve{\mathsf{x}}(z) \text{ (Definition R.4 page 370)}\\
&= \sum_{n\in\mathbb{Z}} x_n && \text{because } z^n = 1 \text{ for all } n\in\mathbb{Z}\\
&= c && \text{by hypothesis (1)}
\end{aligned}$$

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

$$\begin{aligned}
\breve{\mathsf{x}}(\omega)\Big|_{\omega=0} &= \sum_{n\in\mathbb{Z}} x_n e^{-i\omega n}\Big|_{\omega=0} && \text{by definition of } \breve{\mathsf{x}}(\omega) \qquad \text{(Definition Q.1 page 359)}\\
&= \sum_{n\in\mathbb{Z}} x_n z^{-n}\Big|_{z=1}\\
&= \breve{\mathsf{x}}(z)\Big|_{z=1} && \text{by definition of } \breve{\mathsf{x}}(z) \qquad \text{(Definition R.4 page 370)}\\
&= c && \text{by hypothesis (2)}
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

3. Proof that (3) $\Longrightarrow$ (1):

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} x_n &= \sum_{n\in\mathbb{Z}} x_n e^{-i\omega n}\Big|_{\omega=0} \\
&= \check{\mathsf{x}}(\omega) && \text{by definition of } \check{\mathsf{x}}(\omega) && \text{(Definition Q.1 page 359)} \\
&= c && \text{by hypothesis (3)}
\end{aligned}$$

✏

Proposition Q.3. *If the coefficients are **real**, then the magnitude response (MR) is **symmetric**.*

✎PROOF:

$$\begin{aligned}
\left|\tilde{\mathsf{h}}(-\omega)\right| &\triangleq \left|\check{\mathsf{h}}(z)\right|_{z=e^{-i\omega}} && \triangleq \left|\sum_{m\in\mathbb{Z}} \mathsf{x}[m]z^{-m}\right|_{z=e^{-i\omega}} \\
&= \left|\sum_{m\in\mathbb{Z}} \mathsf{x}[m]e^{i\omega m}\right| && = \left|\left(\sum_{m\in\mathbb{Z}} \mathsf{x}^*[m]e^{-i\omega m}\right)^*\right| \\
&= \underbrace{\left|\left(\sum_{m\in\mathbb{Z}} \mathsf{x}[m]e^{-i\omega m}\right)^*\right|}_{\text{if } x[m] \text{ is } real} && = \left|\sum_{m\in\mathbb{Z}} \mathsf{x}[m]e^{-i\omega m}\right| \\
&\triangleq \left|\check{\mathsf{h}}(z)\right|_{z=e^{-i\omega}} && \triangleq \left|\tilde{\mathsf{h}}(\omega)\right|
\end{aligned}$$

✏

Proposition Q.4. [1]

PRP

$$\underbrace{\sum_{n\in\mathbb{Z}}(-1)^n x_n = c}_{\text{(1) in "time"}} \iff \underbrace{\check{\mathsf{x}}(z)|_{z=-1} = c}_{\text{(2) in "z domain"}} \iff \underbrace{\check{\mathsf{x}}(\omega)|_{\omega=\pi} = c}_{\text{(3) in "frequency"}}$$

$$\iff \underbrace{\left(\sum_{n\in\mathbb{Z}} h_{2n}, \sum_{n\in\mathbb{Z}} h_{2n+1}\right) = \left(\frac{1}{2}\left(\sum_{n\in\mathbb{Z}} h_n + c\right), \frac{1}{2}\left(\sum_{n\in\mathbb{Z}} h_n - c\right)\right)}_{\text{(4) sum of even, sum of odd}}$$

$\forall c\in\mathbb{R}, (x_n)_{n\in\mathbb{Z}}, (y_n)_{n\in\mathbb{Z}} \in \ell^2_{\mathbb{R}}$

✎PROOF:

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

$$\begin{aligned}
\check{\mathsf{x}}(z)|_{z=-1} &= \sum_{n\in\mathbb{Z}} x_n z^{-n}\Big|_{z=-1} \\
&= \sum_{n\in\mathbb{Z}} (-1)^n x_n \\
&= c && \text{by (1)}
\end{aligned}$$

[1] Chui (1992) page 123

COPYRIGHT © 2019 DANIEL J. GREENHOE

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

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} x_n e^{-i\omega n}\Big|_{\omega=\pi} &= \sum_{n\in\mathbb{Z}} (-1)^n x_n \\
&= \sum_{n\in\mathbb{Z}} (-1)^{-n} x_n && = \sum_{n\in\mathbb{Z}} z^{-n} x_n\Big|_{z=-1} \\
&= c && \text{by (2)}
\end{aligned}$$

3. Proof that (3) $\Longrightarrow$ (1):

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} (-1)^n x_n &= \sum_{n\in\mathbb{Z}} (-1)^{-n} x_n \\
&= \sum_{n\in\mathbb{Z}} e^{-i\omega n} x_n\Big|_{\omega=\pi} \\
&= c && \text{by (3)}
\end{aligned}$$

4. Proof that (2) $\Longrightarrow$ (4):

   (a) Define $A \triangleq \sum_{n\in\mathbb{Z}} h_{2n}$ $\qquad B \triangleq \sum_{n\in\mathbb{Z}} h_{2n+1}$.

   (b) Proof that $A - B = c$:

$$\begin{aligned}
c &= \sum_{n\in\mathbb{Z}} (-1)^n x_n && \text{by (2)} \\
&= \underbrace{\sum_{n\in\mathbb{Z}_e} (-1)^n x_n}_{\text{even terms}} + \underbrace{\sum_{n\in\mathbb{Z}_o} (-1)^n x_n}_{\text{odd terms}} \\
&= \sum_{n\in\mathbb{Z}} (-1)^{2n} x_{2n} + \sum_{n\in\mathbb{Z}} (-1)^{2n+1} x_{2n+1} \\
&= \underbrace{\sum_{n\in\mathbb{Z}} x_{2n}}_{A} - \underbrace{\sum_{n\in\mathbb{Z}} x_{2n+1}}_{B} \\
&\triangleq A - B && \text{by definitions of } A \text{ and } B
\end{aligned}$$

   (c) Proof that $A + B = \sum_{n\in\mathbb{Z}} x_n$:

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} x_n &= \sum_{n \text{ even}} x_n + \sum_{n \text{ odd}} x_n \\
&= \underbrace{\sum_{n\in\mathbb{Z}} x_{2n}}_{A} + \underbrace{\sum_{n\in\mathbb{Z}} x_{2n+1}}_{B} \\
&= A + B && \text{by definitions of } A \text{ and } B
\end{aligned}$$

   (d) This gives two simultaneous equations:

$$\begin{aligned}
A - B &= c \\
A + B &= \sum_{n\in\mathbb{Z}} x_n
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

(e) Solutions to these equations give

$$\sum_{n\in\mathbb{Z}} x_{2n} \triangleq A \qquad = \frac{1}{2}\left(\sum_{n\in\mathbb{Z}} x_n + c\right)$$

$$\sum_{n\in\mathbb{Z}} x_{2n+1} \triangleq B \qquad = \frac{1}{2}\left(\sum_{n\in\mathbb{Z}} x_n - c\right)$$

5. Proof that (2) $\Longleftarrow$ (4):

$$\begin{aligned}
\sum_{n\in\mathbb{Z}}(-1)^n x_n &= \underbrace{\sum_{n\in\mathbb{Z}_e}(-1)^n x_n}_{\text{even terms}} + \underbrace{\sum_{n\in\mathbb{Z}_o}(-1)^n x_n}_{\text{odd terms}} \\
&= \sum_{n\in\mathbb{Z}}(-1)^{2n} x_{2n} + \sum_{n\in\mathbb{Z}}(-1)^{2n+1} x_{2n+1} \\
&= \sum_{n\in\mathbb{Z}} x_{2n} - \sum_{n\in\mathbb{Z}} x_{2n+1} \\
&= \frac{1}{2}\left(\sum_{n\in\mathbb{Z}} x_n + c\right) - \frac{1}{2}\left(\sum_{n\in\mathbb{Z}} x_n - c\right) && \text{by (3)} \\
&= c
\end{aligned}$$

✏

Lemma Q.1. *Let* $\tilde{\mathsf{f}}(\omega)$ *be the* DTFT *(Definition Q.1 page 359) of a sequence* $(\!(x_n)\!)_{n\in\mathbb{Z}}$.

LEM

$$\underbrace{(\!(x_n \in \mathbb{R})\!)_{n\in\mathbb{Z}}}_{\text{REAL-VALUED } sequence} \quad \Longrightarrow \quad \underbrace{|\check{\mathsf{x}}(\omega)|^2 = |\check{\mathsf{x}}(-\omega)|^2}_{\text{EVEN}} \qquad \forall (\!(x_n)\!)_{n\in\mathbb{Z}} \in \ell^2_{\mathbb{R}}$$

✎PROOF:

$$\begin{aligned}
|\check{\mathsf{x}}(\omega)|^2 &= |\check{\mathsf{x}}(z)|^2\big|_{z=e^{i\omega}} \\
&= \check{\mathsf{x}}(z)\check{\mathsf{x}}^*(z)\big|_{z=e^{i\omega}} \\
&= \left[\sum_{n\in\mathbb{Z}} x_n z^{-n}\right]\left[\sum_{m\in\mathbb{Z}} x_m z^{-n}\right]^*\Bigg|_{z=e^{i\omega}} \\
&= \left[\sum_{n\in\mathbb{Z}} x_n z^{-n}\right]\left[\sum_{m\in\mathbb{Z}} x_m^*(z^*)^{-m}\right]\Bigg|_{z=e^{i\omega}} \\
&= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} x_n x_m^* z^{-n}(z^*)^{-m}\Bigg|_{z=e^{i\omega}} \\
&= \sum_{n\in\mathbb{Z}}\left[|x_n|^2 + \sum_{m>n} x_n x_m^* z^{-n}(z^*)^{-m} + \sum_{m<n} x_n x_m^* z^{-n}(z^*)^{-m}\right]\Bigg|_{z=e^{i\omega}} \\
&= \sum_{n\in\mathbb{Z}}\left[|x_n|^2 + \sum_{m>n} x_n x_m e^{i\omega(m-n)} + \sum_{m<n} x_n x_m e^{i\omega(m-n)}\right] \\
&= \sum_{n\in\mathbb{Z}}\left[|x_n|^2 + \sum_{m>n} x_n x_m e^{i\omega(m-n)} + \sum_{m>n} x_n x_m e^{-i\omega(m-n)}\right] \\
&= \sum_{n\in\mathbb{Z}}\left[|x_n|^2 + \sum_{m>n} x_n x_m \left(e^{i\omega(m-n)} + e^{-i\omega(m-n)}\right)\right]
\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \sum_{n\in\mathbb{Z}} \left[|x_n|^2 + \sum_{m>n} x_n x_m 2\cos[\omega(m-n)] \right]$$
$$= \sum_{n\in\mathbb{Z}} |x_n|^2 + 2\sum_{n\in\mathbb{Z}}\sum_{m>n} x_n x_m \cos[\omega(m-n)]$$

Since cos is real and even, then $|\check{\mathrm{x}}(\omega)|^2$ must also be real and even. ✏

Theorem Q.3 (inverse DTFT). [2] *Let $\check{\mathrm{x}}(\omega)$ be the* DISCRETE-TIME FOURIER TRANSFORM *(Definition Q.1 page 359) of a sequence $(\!(x_n)\!)_{n\in\mathbb{Z}} \in \ell^2_{\mathbb{R}}$. Let $\check{\mathrm{x}}^{-1}$ be the inverse of $\check{\mathrm{x}}$.*

THM
$$\underbrace{\left\{ \check{\mathrm{x}}(\omega) \triangleq \sum_{n\in\mathbb{Z}} x_n e^{-i\omega n} \right\}}_{\check{\mathrm{x}}(\omega) \triangleq \check{\mathbf{F}}(\!(x_n)\!)} \implies \underbrace{\left\{ x_n = \frac{1}{2\pi}\int_{\alpha-\pi}^{\alpha+\pi} \check{\mathrm{x}}(\omega) e^{i\omega n}\,\mathrm{d}\omega \quad \forall_{\alpha\in\mathbb{R}} \right\}}_{(\!(x_n)\!) = \check{\mathbf{F}}^{-1}\check{\mathbf{F}}(\!(x_n)\!)} \quad \forall_{(\!(x_n)\!)_{n\in\mathbb{Z}}\in\ell^2_{\mathbb{R}}}$$

✎PROOF:

$$\frac{1}{2\pi}\int_{\alpha-\pi}^{\alpha+\pi} \check{\mathrm{x}}(\omega)e^{i\omega n}\,\mathrm{d}\omega = \frac{1}{2\pi}\int_{\alpha-\pi}^{\alpha+\pi} \underbrace{\left[\sum_{m\in\mathbb{Z}} x_m e^{-i\omega m}\right]}_{\check{\mathrm{x}}(\omega)} e^{i\omega n}\,\mathrm{d}\omega \qquad \text{by definition of } \check{\mathrm{x}}(\omega)$$
$$= \frac{1}{2\pi}\int_{\alpha-\pi}^{\alpha+\pi} \sum_{m\in\mathbb{Z}} x_m e^{-i\omega(m-n)}\,\mathrm{d}\omega$$
$$= \frac{1}{2\pi}\sum_{m\in\mathbb{Z}} x_m \int_{\alpha-\pi}^{\alpha+\pi} e^{-i\omega(m-n)}\,\mathrm{d}\omega$$
$$= \frac{1}{2\pi}\sum_{m\in\mathbb{Z}} x_m \left[2\pi\bar{\delta}_{m-n}\right]$$
$$= x_n$$

✏

Theorem Q.4 (orthonormal quadrature conditions). [3] *Let $\check{\mathrm{x}}(\omega)$ be the* DISCRETE-TIME FOURIER TRANSFORM *(Definition Q.1 page 359) of a sequence $(\!(x_n)\!)_{n\in\mathbb{Z}} \in \ell^2_{\mathbb{R}}$. Let $\bar{\delta}_n$ be the* KRONECKER DELTA FUNCTION *at n (Definition K.3 page 283).*

THM
$$\sum_{m\in\mathbb{Z}} x_m y^*_{m-2n} = 0 \iff \check{\mathrm{x}}(\omega)\check{\mathrm{y}}^*(\omega) + \check{\mathrm{x}}(\omega+\pi)\check{\mathrm{y}}^*(\omega+\pi) = 0 \qquad \forall_{n\in\mathbb{Z}}, \forall_{(\!(x_n)\!),(\!(y_n)\!)\in\ell^2_{\mathbb{R}}}$$
$$\sum_{m\in\mathbb{Z}} x_m x^*_{m-2n} = \bar{\delta}_n \iff |\check{\mathrm{x}}(\omega)|^2 + |\check{\mathrm{x}}(\omega+\pi)|^2 = 2 \qquad \forall_{n\in\mathbb{Z}}, \forall_{(\!(x_n)\!),(\!(y_n)\!)\in\ell^2_{\mathbb{R}}}$$

✎PROOF: Let $z \triangleq e^{i\omega}$.

[2] J.S.Chitode (2009a) page 3-95 ⟨(3.6.2)⟩
[3] Daubechies (1992) pages 132–137 ⟨(5.1.20),(5.1.39)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

1. Proof that $2\sum_{n\in\mathbb{Z}}\left[\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n}\right] e^{-i2\omega n} = \check{x}(\omega)\check{y}^*(\omega) + \check{x}(\omega+\pi)\check{y}^*(\omega+\pi)$:

$$\begin{aligned}
&2\sum_{n\in\mathbb{Z}}\left[\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n}\right] e^{-i2\omega n}\\
&= 2\sum_{k\in\mathbb{Z}} x_k \sum_{n\in\mathbb{Z}} y^*_{k-2n} z^{-2n}\\
&= 2\sum_{k\in\mathbb{Z}} x_k \sum_{n \text{ even}} y^*_{k-n} z^{-n}\\
&= \sum_{k\in\mathbb{Z}} x_k \sum_{n\in\mathbb{Z}} y^*_{k-n} z^{-n}\left(1+e^{i\pi n}\right)\\
&= \sum_{k\in\mathbb{Z}} x_k \sum_{n\in\mathbb{Z}} y^*_{k-n} z^{-n} + \sum_{k\in\mathbb{Z}} x_k \sum_{n\in\mathbb{Z}} y^*_{k-n} z^{-n} e^{i\pi n}\\
&= \sum_{k\in\mathbb{Z}} x_k \sum_{m\in\mathbb{Z}} y^*_{m} z^{-(k-m)} + \sum_{k\in\mathbb{Z}} x_k \sum_{m\in\mathbb{Z}} y^*_{m} e^{-i(\omega+\pi)(k-m)} && \text{where } m \triangleq k-n\\
&= \sum_{k\in\mathbb{Z}} x_k z^{-k} \sum_{m\in\mathbb{Z}} y^*_{m} z^{m} + \sum_{k\in\mathbb{Z}} x_k e^{-i(\omega+\pi)k} \sum_{m\in\mathbb{Z}} y^*_{m} e^{+i(\omega+\pi)m}\\
&= \sum_{k\in\mathbb{Z}} x_k e^{-i\omega k}\left[\sum_{m\in\mathbb{Z}} y_m e^{-i\omega m}\right]^* + \sum_{k\in\mathbb{Z}} x_k e^{-i(\omega+\pi)k}\left[\sum_{m\in\mathbb{Z}} y_m e^{-i(\omega+\pi)m}\right]^*\\
&\triangleq \check{x}(\omega)\check{y}^*(\omega) + \check{x}(\omega+\pi)\check{y}^*(\omega+\pi)
\end{aligned}$$

2. Proof that $\sum_{m\in\mathbb{Z}} x_m y^*_{m-2n} = 0 \implies \check{x}(\omega)\check{y}^*(\omega) + \check{x}(\omega+\pi)\check{y}^*(\omega+\pi) = 0$:

$$\begin{aligned}
0 &= 2\sum_{n\in\mathbb{Z}}\left[\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n}\right] e^{-i2\omega n} && \text{by left hypothesis}\\
&= \check{x}(\omega)\check{y}^*(\omega) + \check{x}(\omega+\pi)\check{y}^*(\omega+\pi) && \text{by item (1)}
\end{aligned}$$

3. Proof that $\sum_{m\in\mathbb{Z}} x_m y^*_{m-2n} = 0 \impliedby \check{x}(\omega)\check{y}^*(\omega) + \check{x}(\omega+\pi)\check{y}^*(\omega+\pi) = 0$:

$$\begin{aligned}
2\sum_{n\in\mathbb{Z}}\left[\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n}\right] e^{-i2\omega n} &= \check{x}(\omega)\check{y}^*(\omega) + \check{x}(\omega+\pi)\check{y}^*(\omega+\pi) && \text{by item (1)}\\
&= 0 && \text{by right hypothesis}
\end{aligned}$$

Thus by the above equation, $\sum_{n\in\mathbb{Z}}\left[\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n}\right] e^{-i2\omega n} = 0$. The only way for this to be true is if $\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n} = 0$.

4. Proof that $\sum_{m\in\mathbb{Z}} x_m x^*_{m-2n} = \bar{\delta}_n \implies |\check{x}(\omega)|^2 + |\check{x}(\omega'+\pi)|^2 = 2$:
Let $g_n \triangleq x_n$.

$$\begin{aligned}
2 &= 2\sum_{n\in\mathbb{Z}} \bar{\delta}_{n\in\mathbb{Z}} e^{-i2\omega n}\\
&= 2\sum_{n\in\mathbb{Z}}\left[\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n}\right] e^{-i2\omega n} && \text{by left hypothesis}\\
&= \check{x}(\omega)\check{y}^*(\omega) + \check{x}(\omega+\pi)\check{y}^*(\omega+\pi) && \text{by item (1)}
\end{aligned}$$

5. Proof that $\sum_{m\in\mathbb{Z}} x_m x^*_{m-2n} = \bar{\delta}_n \impliedby |\check{x}(\omega)|^2 + |\check{x}(\omega'+\pi)|^2 = 2$:
Let $g_n \triangleq x_n$.

$$\begin{aligned}
2\sum_{n\in\mathbb{Z}}\left[\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n}\right] e^{-i2\omega n} &= \check{x}(\omega)\check{y}^*(\omega) + \check{x}(\omega+\pi)\check{y}^*(\omega+\pi) && \text{by item (1)}\\
&= 2 && \text{by right hypothesis}
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe

Thus by the above equation, $\sum_{n\in\mathbb{Z}}\left[\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n}\right] e^{-i2\omega n} = 1$. The only way for this to be true is if $\sum_{k\in\mathbb{Z}} x_k y^*_{k-2n} = \bar{\delta}_n$.

✏

Q.3 Derivatives

Theorem Q.5. [4] *Let* $\check{x}(\omega)$ *be the* DTFT *(Definition Q.1 page 359) of a sequence* $(\!(x_n)\!)_{n\in\mathbb{Z}}$.

THM

(A)	$\left[\frac{d}{d\omega}\right]^n \check{x}(\omega)\Big\vert_{\omega=0}$	$=$	0	$\Longleftrightarrow$	$\sum_{k\in\mathbb{Z}} k^n x_k$	$=$	0	*(B)*	$\forall n\in\mathbb{W}$
(C)	$\left[\frac{d}{d\omega}\right]^n \check{x}(\omega)\Big\vert_{\omega=\pi}$	$=$	0	$\Longleftrightarrow$	$\sum_{k\in\mathbb{Z}} (-1)^k k^n x_k$	$=$	0	*(D)*	$\forall n\in\mathbb{W}$

✎PROOF:

1. Proof that $(A) \Longrightarrow (B)$:

$$\begin{aligned}
0 &= \left[\frac{d}{d\omega}\right]^n \check{x}(\omega)\Big\vert_{\omega=0} && \text{by hypothesis (A)}\\
&= \left[\frac{d}{d\omega}\right]^n \sum_{k\in\mathbb{Z}} x_k e^{-i\omega k}\Bigg\vert_{\omega=0} && \text{by definition of } \check{x}(\omega) \text{ (Definition Q.1 page 359)}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[\frac{d}{d\omega}\right]^n e^{-i\omega k}\Bigg\vert_{\omega=0}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[(-i)^n k^n e^{-i\omega k}\right]\Bigg\vert_{\omega=0}\\
&= (-i)^n \sum_{k\in\mathbb{Z}} k^n x_k
\end{aligned}$$

2. Proof that $(A) \Longleftarrow (B)$:

$$\begin{aligned}
\left[\frac{d}{d\omega}\right]^n \check{x}(\omega)\Big\vert_{\omega=0} &= \left[\frac{d}{d\omega}\right]^n \sum_{k\in\mathbb{Z}} x_k e^{-i\omega k}\Bigg\vert_{\omega=0} && \text{by definition of } \tilde{g}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[\left[\frac{d}{d\omega}\right]^n e^{-i\omega k}\right]\Bigg\vert_{\omega=0}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[(-i)^n k^n e^{-i\omega k}\right]\Bigg\vert_{\omega=0}\\
&= (-i)^n \sum_{k\in\mathbb{Z}} k^n x_k\\
&= 0 && \text{by hypothesis (B)}
\end{aligned}$$

[4] Vidakovic (1999), pages 82–83, Mallat (1999), pages 241–242

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

3. Proof that $(C) \implies (D)$:

$$
\begin{aligned}
0 &= \left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n \check{\mathrm{x}}(\omega)\Big|_{\omega=\pi} && \text{by hypothesis (C)}\\
&= \left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n \sum_{k\in\mathbb{Z}} x_k e^{-i\omega k}\Bigg|_{\omega=\pi} && \text{by definition of } \check{\mathrm{x}} \text{ (Definition Q.1 page 359)}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n e^{-i\omega k}\Bigg|_{\omega=\pi}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[(-i)^n k^n e^{-i\omega k}\right]\Bigg|_{\omega=\pi}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[(-i)^n k^n (-1)^k\right]\\
&= (-i)^n \sum_{k\in\mathbb{Z}} (-1)^k k^n x_k
\end{aligned}
$$

4. Proof that $(C) \impliedby (D)$:

$$
\begin{aligned}
\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n \check{\mathrm{x}}(\omega)\Big|_{\omega=\pi} &= \left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n \sum_{k\in\mathbb{Z}} x_k e^{-i\omega k}\Bigg|_{\omega=\pi} && \text{by definition of } \check{\mathrm{x}} \text{ (Definition Q.1 page 359)}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n e^{-i\omega k}\Bigg|_{\omega=\pi}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[(-i)^n k^n e^{-i\omega k}\right]\Bigg|_{\omega=\pi}\\
&= \sum_{k\in\mathbb{Z}} x_k \left[(-i)^n k^n (-1)^k\right]\\
&= (-i)^n \sum_{k\in\mathbb{Z}} (-1)^k k^n x_k\\
&= 0 && \text{by hypothesis (D)}
\end{aligned}
$$

✏

Copyright © 2019 Daniel J. Greenhoe

APPENDIX R

OPERATIONS ON SEQUENCES

R.1 Convolution operator

Definition R.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.*

DEF

*A function f in X^Y is a **sequence** over X if $Y = \mathbb{Z}$.*
A sequence may be denoted in the form $(\!(x_n)\!)_{n\in\mathbb{Z}}$ or simply as $(\!(x_n)\!)$.

Definition R.2. [2] *Let $(\mathbb{F}, +, \cdot, 0, 1)$ be a* FIELD *(Definition F.5 page 208).*

DEF

*The **space of all absolutely square summable sequences** $\ell^2_{\mathbb{F}}$ over $\mathbb{F}$ is defined as*

$$\ell^2_{\mathbb{F}} \triangleq \left\{ (\!(x_n)\!)_{n\in\mathbb{Z}} \mid \sum_{n\in\mathbb{Z}} |x_n|^2 < \infty \right\}$$

The space $\ell^2_{\mathbb{R}}$ is an example of a *separable Hilbert space*. In fact, $\ell^2_{\mathbb{R}}$ is the *only* separable Hilbert space in the sense that all separable Hilbert spaces are isomorphically equivalent. For example, $\ell^2_{\mathbb{R}}$ is isomorphic to $L^2_{\mathbb{R}}$, the *space of all absolutely square Lebesgue integrable functions*.

Definition R.3.

DEF

*The **convolution** operation $\star$ is defined as*

$$(\!(x_n)\!) \star (\!(y_n)\!) \triangleq \left(\left(\sum_{m\in\mathbb{Z}} x_m y_{n-m} \right) \right)_{n\in\mathbb{Z}} \qquad \forall (\!(x_n)\!)_{n\in\mathbb{Z}}, (\!(y_n)\!)_{n\in\mathbb{Z}} \in \ell^2_{\mathbb{R}}$$

Proposition R.1. *Let $\star$ be the* CONVOLUTION OPERATOR *(Definition R.3 page 369).*

PRP

$$(\!(x_n)\!) \star (\!(y_n)\!) = (\!(y_n)\!) \star (\!(x_n)\!) \qquad \forall (\!(x_n)\!)_{n\in\mathbb{Z}}, (\!(y_n)\!)_{n\in\mathbb{Z}} \in \ell^2_{\mathbb{R}} \qquad (\star \text{ is COMMUTATIVE})$$

[1] Bromwich (1908), page 1, Thomson et al. (2008) page 23 ⟨Definition 2.1⟩, Joshi (1997) page 31
[2] Kubrusly (2011) page 347 ⟨Example 5.K⟩

✎PROOF:

$$
\begin{aligned}
[x \star y](n) &\triangleq \sum_{m\in\mathbb{Z}} x_m y_{n-m} && \text{by Definition R.3 page 369}\\
&= \sum_{k\in\mathbb{Z}} x_{n-k} y(k) && \text{where } k = n-m \iff m = n-k\\
&= \sum_{k\in\mathbb{Z}} x_{n-k} y(k) && \text{by change commutivity of addition}\\
&= \sum_{m\in\mathbb{Z}} x_{n-m} y_m && \text{by change of variables}\\
&= \sum_{m\in\mathbb{Z}} y_m x_{n-m} && \text{by commutative property of the field over } \mathbb{C}\\
&\triangleq (y \star x)_n && \text{by Definition R.3 page 369}
\end{aligned}
$$

✍

Proposition R.2. *Let $\star$ be the* CONVOLUTION OPERATOR *(Definition R.3 page 369). Let $\ell^2_{\mathbb{R}}$ be the set of* ABSOLUTELY SUMMABLE *sequences (Definition R.2 page 369).*

PRP

$$
\left\{ \begin{array}{lll} (A). & \mathsf{x}(n) \in \ell^2_{\mathbb{R}} & \text{and} \\ (B). & \mathsf{y}(n) \in \ell^2_{\mathbb{R}} & \end{array} \right\} \implies \left\{ \sum_{k\in\mathbb{Z}} \mathsf{x}[k]\mathsf{y}[n+k] = \mathsf{x}[-n] \star \mathsf{y}(n) \right\}
$$

✎PROOF:

$$
\begin{aligned}
\sum_{k\in\mathbb{Z}} \mathsf{x}[k]\mathsf{y}[n+k] &= \sum_{-p\in\mathbb{Z}} \mathsf{x}[-p]\mathsf{y}[n-p] && \text{where } p \triangleq -k && \implies k = -p\\
&= \sum_{p\in\mathbb{Z}} \mathsf{x}[-p]\mathsf{y}[n-p] && \text{by } \textit{absolutely summable} \text{ hypothesis} && \text{(Definition R.2 page 369)}\\
&= \sum_{p\in\mathbb{Z}} \mathsf{x}'[p]\mathsf{y}[n-p] && \text{where } \mathsf{x}'[n] \triangleq \mathsf{x}[-n] && \implies \mathsf{x}[-n] = \mathsf{x}'[n]\\
&\triangleq \mathsf{x}'[n] \star \mathsf{y}[n] && \text{by definition of } \textit{convolution} \star && \text{(Definition R.3 page 369)}\\
&\triangleq \mathsf{x}[-n] \star \mathsf{y}[n] && \text{by definition of } \mathsf{x}'[n]
\end{aligned}
$$

✍

R.2 Z-transform

Definition R.4. [3]

DEF

*The **z-transform** $\mathbf{Z}$ of $(\!(x_n)\!)_{n\in\mathbb{Z}}$ is defined as*

$$
\left[\mathbf{Z}\left(\!(x_n)\!\right)\right](z) \triangleq \underbrace{\sum_{n\in\mathbb{Z}} x_n z^{-n}}_{\textit{Laurent series}} \qquad \forall (\!(x_n)\!) \in \ell^2_{\mathbb{R}}
$$

Theorem R.1. *Let $X(z) \triangleq \mathbf{Z}\mathsf{x}[n]$ be the* Z-TRANSFORM *of $\mathsf{x}[n]$.*

THM

$$
\left\{ \check{\mathsf{x}}(z) \triangleq \mathbf{Z}\left(\!(\mathsf{x}[n])\!\right) \right\} \implies \left\{ \begin{array}{llll}
(1). & \mathbf{Z}\left(\!(\alpha\mathsf{x}[n])\!\right) = \alpha\check{\mathsf{x}}(z) & \forall (\!(x_n)\!) \in \ell^2_{\mathbb{R}} & \text{and}\\
(2). & \mathbf{Z}\left(\!(\mathsf{x}[n-k])\!\right) = z^{-k}\check{\mathsf{x}}(z) & \forall (\!(x_n)\!) \in \ell^2_{\mathbb{R}} & \text{and}\\
(3). & \mathbf{Z}\left(\!(\mathsf{x}[-n])\!\right) = \check{\mathsf{x}}\left(\frac{1}{z}\right) & \forall (\!(x_n)\!) \in \ell^2_{\mathbb{R}} & \text{and}\\
(4). & \mathbf{Z}\left(\!(\mathsf{x}^*[n])\!\right) = \check{\mathsf{x}}^*\left(z^*\right) & \forall (\!(x_n)\!) \in \ell^2_{\mathbb{R}} & \text{and}\\
(5). & \mathbf{Z}\left(\!(\mathsf{x}^*[-n])\!\right) = \check{\mathsf{x}}^*\left(\frac{1}{z^*}\right) & \forall (\!(x_n)\!) \in \ell^2_{\mathbb{R}} &
\end{array} \right\}
$$

[3] *Laurent series*: Abramovich and Aliprantis (2002) page 49

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$
\begin{aligned}
\alpha\mathbb{Z}\check{\mathsf{x}}(z) &\triangleq \alpha\mathbf{Z}\left(\!\left(\mathsf{x}[n]\right)\!\right) && \text{by definition of } \check{\mathsf{x}}(z) \\
&\triangleq \alpha\sum_{n\in\mathbb{Z}}\mathsf{x}[n]z^{-n} && \text{by definition of } \mathbf{Z} \text{ operator} \\
&\triangleq \sum_{n\in\mathbb{Z}}(\alpha\mathsf{x}[n])z^{-n} && \text{by } \textit{distributive} \text{ property} \\
&\triangleq \mathbf{Z}\left(\!\left(\alpha\mathsf{x}[n]\right)\!\right) && \text{by definition of } \mathbf{Z} \text{ operator} \\
z^{-k}\check{\mathsf{x}}(z) &= z^{-k}\mathbf{Z}\left(\!\left(\mathsf{x}[n]\right)\!\right) && \text{by definition of } \check{\mathsf{x}}(z) && \text{(left hypothesis)} \\
&\triangleq z^{-k}\sum_{n=-\infty}^{n=+\infty}\mathsf{x}[n]z^{-n} && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)} \\
&= \sum_{n=-\infty}^{n=+\infty}\mathsf{x}[n]z^{-n-k} \\
&= \sum_{m-k=-\infty}^{m-k=+\infty}\mathsf{x}[m-k]z^{-m} && \text{where } m\triangleq n+k && \implies n=m-k \\
&= \sum_{m=-\infty}^{m=+\infty}\mathsf{x}[m-k]z^{-m} \\
&= \sum_{n=-\infty}^{n=+\infty}\mathsf{x}[n-k]z^{-n} && \text{where } n\triangleq m \\
&\triangleq \mathbf{Z}\left(\!\left(\mathsf{x}[n-k]\right)\!\right) && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)} \\
\mathbf{Z}\left(\!\left(\mathsf{x}^*[n]\right)\!\right) &\triangleq \sum_{n\in\mathbb{Z}}\mathsf{x}^*[n]z^{-n} && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)} \\
&\triangleq \left(\sum_{n\in\mathbb{Z}}\mathsf{x}[n](z^*)^{-n}\right)^* && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)} \\
&\triangleq \check{\mathsf{x}}^*(z^*) && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)} \\
\mathbf{Z}\left(\!\left(\mathsf{x}[-n]\right)\!\right) &\triangleq \sum_{n\in\mathbb{Z}}\mathsf{x}[-n]z^{-n} && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)} \\
&= \sum_{-m\in\mathbb{Z}}\mathsf{x}[m]z^{m} && \text{where } m\triangleq -n && \implies n=-m \\
&= \sum_{m\in\mathbb{Z}}\mathsf{x}[m]z^{m} && \text{by } \textit{absolutely summable} \text{ property} && \text{(Definition R.2 page 369)} \\
&= \sum_{m\in\mathbb{Z}}\mathsf{x}[m]\left(\frac{1}{z}\right)^{-m} && \text{by } \textit{absolutely summable} \text{ property} && \text{(Definition R.2 page 369)} \\
&\triangleq \check{\mathsf{x}}\left(\frac{1}{z}\right) && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)} \\
\mathbf{Z}\left(\!\left(\mathsf{x}^*[-n]\right)\!\right) &\triangleq \sum_{n\in\mathbb{Z}}\mathsf{x}^*[-n]z^{-n} && \text{by definition of } \mathbf{Z} && \text{(Definition R.4 page 370)} \\
&= \sum_{-m\in\mathbb{Z}}\mathsf{x}^*[m]z^{m} && \text{where } m\triangleq -n && \implies n=-m \\
&= \sum_{m\in\mathbb{Z}}\mathsf{x}^*[m]z^{m} && \text{by } \textit{absolutely summable} \text{ property} && \text{(Definition R.2 page 369)} \\
&= \sum_{m\in\mathbb{Z}}\mathsf{x}^*[m]\left(\frac{1}{z}\right)^{-m} && \text{by } \textit{absolutely summable} \text{ property} && \text{(Definition R.2 page 369)} \\
&= \left(\sum_{m\in\mathbb{Z}}\mathsf{x}[m]\left(\frac{1}{z^*}\right)^{-m}\right)^* && \text{by } \textit{absolutely summable} \text{ property} && \text{(Definition R.2 page 369)}
\end{aligned}
$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Figure R.1: Direct form 1 order 2 IIR filters

$$\triangleq \check{x}^*\left(\frac{1}{z^*}\right) \qquad \text{by definition of } \mathbf{Z} \qquad \text{(Definition R.4 page 370)}$$

✏

Theorem R.2 (convolution theorem)**.** *Let $\star$ be the convolution operator* (*Definition R.3 page 369*)*.*

THM

$$\mathbf{Z}\underbrace{\left(\langle\!\langle x_n\rangle\!\rangle \star \langle\!\langle y_n\rangle\!\rangle\right)}_{\textit{sequence convolution}} = \underbrace{\left(\mathbf{Z}\langle\!\langle x_n\rangle\!\rangle\right)\left(\mathbf{Z}\langle\!\langle y_n\rangle\!\rangle\right)}_{\textit{series multiplication}} \qquad \forall \langle\!\langle x_n\rangle\!\rangle_{n\in\mathbb{Z}}, \langle\!\langle y_n\rangle\!\rangle_{n\in\mathbb{Z}} \in \ell^2_{\mathbb{R}}$$

✎PROOF:

$$\begin{aligned}
[\mathbf{Z}(x\star y)](z) &\triangleq \mathbf{Z}\left(\sum_{m\in\mathbb{Z}} x_m y_{n-m}\right) && \text{by Definition R.3 page 369}\\
&\triangleq \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} x_m y_{n-m} z^{-n} && \text{by Definition R.4 page 370}\\
&= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} x_m y_{n-m} z^{-n}\\
&= \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} x_m y_{n-m} z^{-n}\\
&= \sum_{m\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} x_m y_k z^{-(m+k)} && \text{where } k = n-m \iff n = m+k\\
&= \left[\sum_{m\in\mathbb{Z}} x_m z^{-m}\right]\left[\sum_{k\in\mathbb{Z}} y_k z^{-k}\right]\\
&\triangleq \left(\mathbf{Z}\langle\!\langle x_n\rangle\!\rangle\right)\left(\mathbf{Z}\langle\!\langle y_n\rangle\!\rangle\right) && \text{by Definition R.4 page 370}
\end{aligned}$$

✏

R.3 From z-domain back to time-domain

$$\check{y}(z) = b_0X(z) + b_1z^{-1}X(z) + b_2z^{-2}X(z) - a_1z^{-1}Y(z) + a_2z^{-2}Y(z)$$

$$y[n] = b_0x[n] + b_1x[n-1] + b_2x[n-2] - a_1y[n-1] - a_2y[n-2]$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

Example R.1. See Figure R.1 (page 372)

$$\frac{3z^2+5z+7}{2z^2+10z+12} = \frac{3z^2+5z+7}{2\left(z^2+5z+6\right)} = \frac{\left(\frac{3}{2}z^2+\frac{5}{2}z+\frac{7}{2}\right)}{z^2+5z+6} = \frac{\left(\frac{3}{2}+\frac{5}{2}z^{-1}+\frac{7}{2}z^{-2}\right)}{1+5z^{-1}+6z^{-2}}$$

R.4 Zero locations

The system property of *minimum phase* is defined in Definition R.5 (next) and illustrated in Figure R.2 (page 373).

Figure R.2: Minimum Phase filter

Definition R.5. [4] *Let* $\check{x}(z) \triangleq \mathbf{Z}(\langle x_n \rangle)$ *be the* Z TRANSFORM *(Definition R.4 page 370) of a sequence* $\langle x_n \rangle_{n\in\mathbb{Z}}$ *in* $\ell^2_{\mathbb{R}}$. *Let* $\langle z_n \rangle_{n\in\mathbb{Z}}$ *be the* ZEROS *of* $\check{x}(z)$.

DEF *The sequence* $\langle x_n \rangle$ *is* ***minimum phase*** *if*

$$\underbrace{|z_n| < 1 \qquad \forall n \in \mathbb{Z}}_{\check{x}(z) \text{ has all its ZEROS inside the unit circle}}$$

The impulse response of a minimum phase filter has most of its energy concentrated near the beginning of its support, as demonstrated next.

Theorem R.3 (Robinson's Energy Delay Theorem). [5] *Let* $\mathsf{p}(z) \triangleq \sum_{n=0}^{N} a_n z^{-n}$ *and* $\mathsf{q}(z) \triangleq \sum_{n=0}^{N} b_n z^{-n}$ *be polynomials.*

THM

$$\left\{\begin{array}{ll} \mathsf{p} & \text{is MINIMUM PHASE} \\ \mathsf{q} & \text{is NOT minimum phase} \end{array}\right\} \text{ and } \Longrightarrow \underbrace{\sum_{n=0}^{m-1} |a_n|^2}_{\text{"energy" of the first } m \text{ coefficients of } \mathsf{p}(z)} \geq \underbrace{\sum_{n=0}^{m-1} |b_n|^2}_{\text{"energy" of the first } m \text{ coefficients of } \mathsf{q}(z)} \qquad \forall 0 \leq m \leq N$$

But for more *symmetry*, put some zeros inside and some outside the unit circle.

Example R.2. An example of a minimum phase polynomial is the Daubechies-4 scaling function. The minimum phase polynomial causes most of the energy to be concentrated near the origin, making it very *asymmetric*. In contrast, the Symlet-4 has a design very similar to that of Daubechies-4, but the selected zeros are not all within the unit circle in the complex z plane. This results in a

[4] Farina and Rinaldi (2000) page 91, Dumitrescu (2007) page 36
[5] Dumitrescu (2007) page 36, Robinson (1962), Robinson (1966) ⟨???⟩, Claerbout (1976), pages 52–53

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Figure R.3: Daubechies-4 and Symlet-4 scaling functions pole-zero plots

scaling function that is more symmetric and less contrated near the origin. Both scaling functions are illustrated in Figure R.3 (page 374).

R.5 Pole locations

Definition R.6.

DEF *A filter (or system or operator)* **H** *is* ***causal*** *if its current output does not depend on future inputs.*

Definition R.7.

DEF *A filter (or system or operator)* **H** *is* ***time-invariant*** *if the mapping it performs does not change with time.*

Definition R.8.

DEF *An operation* **H** *is* ***linear*** *if any output* y_n *can be described as a linear combination of inputs* x_n *as in*

$$y_n = \sum_{m\in\mathbb{Z}} h(m)x(n-m) .$$

For a filter to be *stable*, place all the poles *inside* the unit circle.

Theorem R.4. *A causal LTI filter is* ***stable*** *if all of its poles are* ***inside*** *the unit circle.*

Example R.3. Stable/unstable filters are illustrated in Figure R.4 (page 375).

True or False? This filter has no poles:

Copyright © 2019 Daniel J. Greenhoe

Figure R.4: Pole-zero plot stable/unstable causal LTI filters (Example R.3 page 374)

$H(z) = b_0 + b_1 z^{-1} + b_2 z^{-2}$

$$H(z) = b_0 + b_1 z^{-1} + b_2 z^{-2} = \frac{z^2}{z^2} \times \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1} = \frac{b_0 z^2 + b_1 z^1 + b_2}{z^2}$$

R.6 Mirroring for real coefficients

If you want real coefficients, choose poles and zeros in conjugate pairs (next).

Proposition R.3.

PRP $\left\{ \begin{array}{l} \text{ZEROS } \textit{and} \text{ POLES} \\ \textit{occur in} \text{ CONJUGATE PAIRS} \end{array} \right\} \Longrightarrow \left\{ \begin{array}{l} \text{COEFFICIENTS} \\ \textit{are} \text{ REAL.} \end{array} \right\}$

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

Figure R.5: Conjugate pair structure yielding real coefficients

✎PROOF:

$$\begin{aligned}(z-p_1)(z-p_1^*) &= [z-(a+ib)][z-(a-ib)]\\ &= z^2+[-a+ib-ib-a]z-[ib]^2\\ &= z^2-2az+b^2\end{aligned}$$

✏

Example R.4. See Figure R.5 (page 376).

$$\begin{aligned}H(z) &= G\frac{[z-z_1][z-z_2]}{[z-p_1][z-p_2]} = G\frac{[z-(1+i)][z-(1-i)]}{[z-(-2/3+i1/2)][z-(-2/3-i1/2)]}\\ &= G\frac{z^2-z[(1-i)+(1+i)]+(1-i)(1+i)}{z^2-z[(-2/3+i1/2)+(-2/3+i1/2)]+(-2/3+i1/2)(-2/3+i1/2)}\\ &= G\frac{z^2-2z+2}{z^2-4/3z+(4/3+1/4)} = G\frac{z^2-2z+2}{z^2-4/3z+19/12}\end{aligned}$$

R.7 Rational polynomial operators

A digital filter is simply an operator on $\ell^2_{\mathbb{R}}$. If the digital filter is a causal LTI system, then it can be expressed as a rational polynomial in z as shown next.

Lemma R.1. *A causal LTI operator* **H** *can be expressed as a rational expression* $\check{\mathsf{h}}(z)$.

$$\begin{aligned}\check{\mathsf{h}}(z) &= \frac{b_0+b_1z^{-1}+b_2z^{-2}+\cdots+b_Nz^{-N}}{1+a_1z^{-1}+a_2z^{-2}+\cdots+a_Nz^{-N}}\\ &= \frac{\sum_{n=0}^{N} b_nz^{-n}}{1+\sum_{n=1}^{N} a_nz^{-n}}\end{aligned}$$

A filter operation $\check{\mathsf{h}}(z)$ can be expressed as a product of its roots (poles and zeros).

$$\begin{aligned}\check{\mathsf{h}}(z) &= \frac{b_0+b_1z^{-1}+b_2z^{-2}+\cdots+b_Nz^{-N}}{1+a_1z^{-1}+a_2z^{-2}+\cdots+a_Nz^{-N}}\\ &= \alpha\frac{(z-z_1)(z-z_2)\cdots(z-z_N)}{(z-p_1)(z-p_2)\cdots(z-p_N)}\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE

where α is a constant, z_i are the zeros, and p_i are the poles. The poles and zeros of such a rational expression are often plotted in the z-plane with a unit circle about the origin (representing $z = e^{i\omega}$). Poles are marked with $\times$ and zeros with $\bigcirc$. An example is shown in Figure R.6 page 377. Notice that in this figure the zeros and poles are either real or occur in complex conjugate pairs.

Figure R.6: Pole-zero plot for rational expression with real coefficients

R.8 Filter Banks

Conjugate quadrature filters (next definition) are used in *filter banks*. If $\check{x}(z)$ is a *low-pass filter*, then the conjugate quadrature filter of $\check{y}(z)$ is a *high-pass filter*.

Definition R.9. [6] *Let* $(\!(x_n)\!)_{n\in\mathbb{Z}}$ *and* $(\!(y_n)\!)_{n\in\mathbb{Z}}$ *be* SEQUENCES *(Definition R.1 page 369) in* $\ell^2_{\mathbb{R}}$ *(Definition R.2 page 369).*

DEF

The sequence $(\!(y_n)\!)$ *is a* ***conjugate quadrature filter*** *with shift* N *with respect to* $(\!(x_n)\!)$ *if*

$$y_n = \pm(-1)^n x^*_{N-n}$$

A CONJUGATE QUADRATURE FILTER *is also called a* ***CQF*** *or a* ***Smith-Barnwell filter****.*
Any triple $\left((\!(x_n)\!), (\!(y_n)\!), N\right)$ *in this form is said to satisfy the* ***conjugate quadrature filter condition*** *or the* ***CQF condition****.*

Theorem R.5 (CQF theorem)**.** [7] *Let* $\check{y}(\omega)$ *and* $\check{x}(\omega)$ *be the* DTFT*s (Definition Q.1 page 359) of the sequences* $(\!(y_n)\!)_{n\in\mathbb{Z}}$ *and* $(\!(x_n)\!)_{n\in\mathbb{Z}}$*, respectively, in* $\ell^2_{\mathbb{R}}$ *(Definition R.2 page 369).*

THM

$$\underbrace{y_n = \pm(-1)^n x^*_{N-n}}_{\text{(1) CQF in "time"}} \iff \check{y}(z) = \pm(-1)^N z^{-N}\check{x}^*\left(\frac{-1}{z^*}\right) \qquad \text{(2) CQF in "z-domain"}$$

$$\iff \check{y}(\omega) = \pm(-1)^N e^{-i\omega N}\check{x}^*(\omega+\pi) \qquad \text{(3) CQF in "frequency"}$$

$$\iff x_n = \pm(-1)^N(-1)^n y^*_{N-n} \qquad \text{(4) "reversed" CQF in "time"}$$

$$\iff \check{x}(z) = \pm z^{-N}\check{y}^*\left(\frac{-1}{z^*}\right) \qquad \text{(5) "reversed" CQF in "z-domain"}$$

$$\iff \check{x}(\omega) = \pm e^{-i\omega N}\check{y}^*(\omega+\pi) \qquad \text{(6) "reversed" CQF in "frequency"}$$

$\forall N\in\mathbb{Z}$

✎PROOF:

[6] Strang and Nguyen (1996) page 109, Haddad and Akansu (1992) pages 256–259 ⟨section 4.5⟩, Vaidyanathan (1993) page 342 ⟨(7.2.7), (7.2.8)⟩, Smith and Barnwell (1984a), Smith and Barnwell (1984b), Mintzer (1985)

[7] Strang and Nguyen (1996) page 109, Mallat (1999) pages 236–238 ⟨(7.58),(7.73)⟩, Haddad and Akansu (1992) pages 256–259 ⟨section 4.5⟩, Vaidyanathan (1993) page 342 ⟨(7.2.7), (7.2.8)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

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

$$\begin{aligned}
\check{y}(z) &= \sum_{n\in\mathbb{Z}} y_n z^{-n} && \text{by definition of } z\text{-transform} \quad \text{(Definition R.4 page 370)}\\
&= \sum_{n\in\mathbb{Z}} \underbrace{(\pm)(-1)^n x^*_{N-n}}_{CQF} z^{-n} && \text{by (1)}\\
&= \pm \sum_{m\in\mathbb{Z}} (-1)^{N-m} x^*_m z^{-(N-m)} && \text{where } m \triangleq N-n \implies \quad n = N-m\\
&= \pm(-1)^N z^{-N} \sum_{m\in\mathbb{Z}} (-1)^{-m} x^*_m \left(z^{-1}\right)^{-m}\\
&= \pm(-1)^N z^{-N} \sum_{m\in\mathbb{Z}} x^*_m \left(-\frac{1}{z}\right)^{-m}\\
&= \pm(-1)^N z^{-N} \left[\sum_{m\in\mathbb{Z}} x_m \left(-\frac{1}{z^*}\right)^{-m}\right]^*\\
&= \pm(-1)^N z^{-N} \check{x}^*\left(\frac{-1}{z^*}\right) && \text{by definition of } z\text{-transform} \quad \text{(Definition R.4 page 370)}
\end{aligned}$$

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

$$\begin{aligned}
\check{y}(z) &= \pm(-1)^N z^{-N} \check{x}^*\left(\frac{-1}{z^*}\right) && \text{by (2)}\\
&= \pm(-1)^N z^{-N} \left[\sum_{m\in\mathbb{Z}} x_m \left(\frac{-1}{z^*}\right)^{-m}\right]^* && \text{by definition of } z\text{-transform} \quad \text{(Definition R.4 page 370)}\\
&= \pm(-1)^N z^{-N} \left[\sum_{m\in\mathbb{Z}} x^*_m \left(-z^{-1}\right)^{-m}\right] && \text{by definition of } z\text{-transform} \quad \text{(Definition R.4 page 370)}\\
&= \sum_{m\in\mathbb{Z}} (\pm)(-1)^{N-m} x^*_m z^{-(N-m)}\\
&= \sum_{m\in\mathbb{Z}} (\pm)(-1)^n x^*_{N-n} z^{-n} && \text{where } n = N-m \implies \quad m \triangleq N-n\\
&\implies x_n = \pm(-1)^n x^*_{N-n}
\end{aligned}$$

3. Proof that (1) $\implies$ (3):

$$\begin{aligned}
\check{y}(\omega) &\triangleq \check{x}(z)\Big|_{z=e^{i\omega}} && \text{by definition of } DTFT \quad \text{(Definition Q.1 page 359)}\\
&= \left[\pm(-1)^N z^{-N} \check{x}\left(\frac{-1}{z^*}\right)\right]_{z=e^{i\omega}} && \text{by (2)}\\
&= \pm(-1)^N e^{-i\omega N} \check{x}\left(e^{i\pi} e^{\imath\omega}\right)\\
&= \pm(-1)^N e^{-i\omega N} \check{x}\left(e^{i(\omega+\pi)}\right)\\
&= \pm(-1)^N e^{-i\omega N} \check{x}(\omega+\pi) && \text{by definition of } DTFT \quad \text{(Definition Q.1 page 359)}
\end{aligned}$$

4. Proof that (1) $\implies$ (6):

$$\begin{aligned}
\check{x}(\omega) &= \sum_{n\in\mathbb{Z}} y_n e^{-i\omega n} && \text{by definition of } DTFT \quad \text{(Definition Q.1 page 359)}\\
&= \sum_{n\in\mathbb{Z}} \underbrace{\pm(-1)^n x^*_{N-n}}_{CQF} e^{-i\omega n} && \text{by (1)}\\
&= \sum_{m\in\mathbb{Z}} \pm(-1)^{N-m} x^*_m e^{-i\omega(N-m)} && \text{where } m \triangleq N-n \implies \quad n = N-m\\
&= \pm(-1)^N e^{-i\omega N} \sum_{m\in\mathbb{Z}} (-1)^m x^*_m e^{i\omega m}
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe

$$= \pm(-1)^N e^{-i\omega N} \sum_{m\in\mathbb{Z}} e^{i\pi m} x_m^* e^{i\omega m}$$
$$= \pm(-1)^N e^{-i\omega N} \sum_{m\in\mathbb{Z}} x_m^* e^{i(\omega+\pi)m}$$
$$= \pm(-1)^N e^{-i\omega N} \left[\sum_{m\in\mathbb{Z}} x_m e^{-i(\omega+\pi)m}\right]^*$$
$$= \pm(-1)^N e^{-i\omega N} \check{x}^*(\omega+\pi) \qquad \text{by definition of } DTFT \qquad \text{(Definition Q.1 page 359)}$$

5. Proof that (1) $\Longleftarrow$ (3):

$$y_n = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \check{y}(\omega)e^{i\omega n}\,d\omega \qquad \text{by } \textit{inverse DTFT} \qquad \text{(Theorem Q.3 page 365)}$$
$$= \frac{1}{2\pi}\int_{-\pi}^{+\pi} \underbrace{\pm(-1)^N e^{-iN\omega}\check{x}^*(\omega+\pi)}_{\text{right hypothesis}} e^{i\omega n}\,d\omega \qquad \text{by right hypothesis}$$
$$= \pm(-1)^N \frac{1}{2\pi}\int_{-\pi}^{+\pi} \check{x}^*(\omega+\pi)e^{i\omega(n-N)}\,d\omega \qquad \text{by right hypothesis}$$
$$= \pm(-1)^N \frac{1}{2\pi}\int_{0}^{2\pi} \check{x}^*(v)e^{i(v-\pi)(n-N)}\,dv \qquad \text{where } v \triangleq \omega+\pi \implies \omega = v-\pi$$
$$= \pm(-1)^N e^{-i\pi(n-N)} \frac{1}{2\pi}\int_{0}^{2\pi} \check{x}^*(v)e^{iv(n-N)}\,dv$$
$$= \pm(-1)^N \underbrace{(-1)^N}_{e^{i\pi N}}\underbrace{(-1)^n}_{e^{-i\pi n}} \left[\frac{1}{2\pi}\int_{0}^{2\pi} \check{x}(v)e^{iv(N-n)}\,dv\right]^*$$
$$= \pm(-1)^n x_{N-n}^* \qquad \text{by } \textit{inverse DTFT} \qquad \text{(Theorem Q.3 page 365)}$$

6. Proof that (1) $\Longleftrightarrow$ (4):

$$\begin{aligned}
y_n = \pm(-1)^n x_{N-n}^* &\iff (\pm)(-1)^n y_n = (\pm)(\pm)(-1)^n(-1)^n x_{N-n}^* \\
&\iff \pm(-1)^n y_n = x_{N-n}^* \\
&\iff \left(\pm(-1)^n y_n\right)^* = \left(x_{N-n}^*\right)^* \\
&\iff \pm(-1)^n y_n^* = x_{N-n} \\
&\iff x_{N-n} = \pm(-1)^n y_n^* \\
&\iff x_m = \pm(-1)^{N-m} y_{N-m}^* \qquad \text{where } m \triangleq N-n \implies n = N-m \\
&\iff x_m = \pm(-1)^{N-m} y_{N-m}^* \\
&\iff x_m = \pm(-1)^N(-1)^m y_{N-m}^* \\
&\iff x_n = \pm(-1)^N(-1)^n y_{N-n}^* \qquad \text{by change of free variables}
\end{aligned}$$

7. Proofs for (5) and (6): not included. See proofs for (2) and (3).

✍

Theorem R.6. [8] *Let* $\check{y}(\omega)$ *and* $\check{x}(\omega)$ *be the* DTFTs *(Definition Q.1 page 359) of the sequences* $(\!(y_n)\!)_{n\in\mathbb{Z}}$ *and* $(\!(x_n)\!)_{n\in\mathbb{Z}}$*, respectively, in* $\ell^2_{\mathbb{R}}$ *(Definition R.2 page 369).*

THM *Let* $y_n = \pm(-1)^n x_{N-n}^*$ (CQF CONDITION, *Definition R.9 page 377*)*. Then*

$$\left\{\begin{aligned}
&(A) \quad \left[\frac{d}{d\omega}\right]^n \check{y}(\omega)\Big|_{\omega=0} = 0 &&\iff \left[\frac{d}{d\omega}\right]^n \check{x}(\omega)\Big|_{\omega=\pi} = 0 \quad (B)\\
&&&\iff \sum_{k\in\mathbb{Z}} (-1)^k k^n x_k = 0 \quad (C)\\
&&&\iff \sum_{k\in\mathbb{Z}} k^n y_k = 0 \quad (D)
\end{aligned}\right\} \quad \forall n\in\mathbb{W}$$

[8] Vidakovic (1999), pages 82–83, Mallat (1999), pages 241–242

✎PROOF:

1. Proof that (A) $\Longrightarrow$ (B):

$$
\begin{aligned}
0 &= \left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n \breve{\mathrm{y}}(\omega)\Big|_{\omega=0} && \text{by (A)} \\
&= \left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n (\pm)(-1)^N e^{-i\omega N}\breve{\mathrm{x}}^*(\omega+\pi)\Big|_{\omega=0} && \text{by } \textit{CQF theorem} \quad \text{(Theorem R.5 page 377)} \\
&= (\pm)(-1)^N \sum_{\ell=0}^{n}\binom{n}{\ell}\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^{\ell}\left[e^{-i\omega N}\right]\cdot\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^{n-\ell}\left[\breve{\mathrm{x}}^*(\omega+\pi)\right]\Bigg|_{\omega=0} && \text{by } \textit{Leibnitz GPR} \quad \text{(Lemma ?? page ??)} \\
&= (\pm)(-1)^N \sum_{\ell=0}^{n}\binom{n}{\ell}-iN^{\ell}e^{-i\omega N}\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^{n-\ell}\left[\breve{\mathrm{x}}^*(\omega+\pi)\right]\Bigg|_{\omega=0} \\
&= (\pm)(-1)^N \cancelto{1}{e^{-i0N}} \sum_{\ell=0}^{n}\binom{n}{\ell}-iN^{\ell}\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^{n-\ell}\left[\breve{\mathrm{x}}^*(\omega+\pi)\right]\Bigg|_{\omega=0}
\end{aligned}
$$

$$
\begin{aligned}
\Longrightarrow \quad \breve{\mathrm{x}}^{(0)}(\pi) &= 0 \\
\Longrightarrow \quad \breve{\mathrm{x}}^{(1)}(\pi) &= 0 \\
\Longrightarrow \quad \breve{\mathrm{x}}^{(2)}(\pi) &= 0 \\
\Longrightarrow \quad \breve{\mathrm{x}}^{(3)}(\pi) &= 0 \\
\Longrightarrow \quad \breve{\mathrm{x}}^{(4)}(\pi) &= 0 \\
\vdots \quad & \quad \vdots \\
\Longrightarrow \quad \breve{\mathrm{x}}^{(n)}(\pi) &= 0 \quad \text{for } n = 0,1,2,\ldots
\end{aligned}
$$

2. Proof that (A) $\Longleftarrow$ (B):

$$
\begin{aligned}
0 &= \left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n \breve{\mathrm{x}}(\omega)\Big|_{\omega=\pi} && \text{by (B)} \\
&= \left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^n (\pm)e^{-i\omega N}\breve{\mathrm{y}}^*(\omega+\pi)\Big|_{\omega=\pi} && \text{by } \textit{CQF theorem} \quad \text{(Theorem R.5 page 377)} \\
&= (\pm)\sum_{\ell=0}^{n}\binom{n}{\ell}\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^{\ell}\left[e^{-i\omega N}\right]\cdot\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^{n-\ell}\left[\breve{\mathrm{y}}^*(\omega+\pi)\right]\Bigg|_{\omega=\pi} && \text{by } \textit{Leibnitz GPR} \quad \text{(Lemma ?? page ??)} \\
&= (\pm)\sum_{\ell=0}^{n}\binom{n}{\ell}(-iN)^{\ell}e^{-i\omega N}\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^{n-\ell}\left[\breve{\mathrm{y}}^*(\omega+\pi)\right]\Bigg|_{\omega=\pi} \\
&= (\pm)\cancelto{1}{e^{-i\pi N}}\sum_{\ell=0}^{n}\binom{n}{\ell}-iN^{\ell}\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^{n-\ell}\left[\breve{\mathrm{y}}^*(\omega+\pi)\right]\Bigg|_{\omega=\pi} \\
&= (\pm)(-1)^N\sum_{\ell=0}^{n}\binom{n}{\ell}-iN^{\ell}\left[\frac{\mathrm{d}}{\mathrm{d}\omega}\right]^{n-\ell}\left[\breve{\mathrm{y}}^*(\omega+\pi)\right]\Bigg|_{\omega=\pi}
\end{aligned}
$$

$$
\begin{aligned}
\Longrightarrow \quad \breve{\mathrm{y}}^{(0)}(0) &= 0 \\
\Longrightarrow \quad \breve{\mathrm{y}}^{(1)}(0) &= 0 \\
\Longrightarrow \quad \breve{\mathrm{y}}^{(2)}(0) &= 0 \\
\Longrightarrow \quad \breve{\mathrm{y}}^{(3)}(0) &= 0 \\
\Longrightarrow \quad \breve{\mathrm{y}}^{(4)}(0) &= 0 \\
\vdots \quad & \quad \vdots \\
\Longrightarrow \quad \breve{\mathrm{y}}^{(n)}(0) &= 0 \\
\Longrightarrow \quad \breve{\mathrm{y}}^{(n)}(0) &= 0 \quad \text{for } n = 0,1,2,\ldots
\end{aligned}
$$

3. Proof that (B) $\Longleftrightarrow$ (C): by Theorem Q.5 page 367

4. Proof that (A) $\Longleftrightarrow$ (D): by Theorem Q.5 page 367

COPYRIGHT © 2019 DANIEL J. GREENHOE

$\times$

$= \quad 1$

$$\frac{(z-z_1)(z-z_2)(z-z_3)(z-z_4)}{(z-p_1)(z-p_2)(z-p_3)(z-p_4)} \times \frac{(z-p_1)(z-p_2)(z-p_3)(z-p_4)}{(z-z_1)(z-z_2)(z-z_3)(z-z_4)} = 1$$

5. Proof that (CQF) $\not\Leftarrow$ (A): Here is a counterexample: $\check{y}(\omega) = 0$.

✏

R.9 Inverting non-minimum phase filters

Minimum phase filters are easy to invert: each *zero* becomes a *pole* and each *pole* becomes a *zero*.

unstable

$\times$

all-pass

$=$

stable!

$$\begin{aligned}
|A(z)|_{z=e^{i\omega}} &= \frac{1}{r}\left|\frac{z-re^{i\phi}}{z-\frac{1}{r}e^{i\phi}}\right|_{z=e^{i\omega}} &&= \left|\frac{z-re^{i\phi}}{rz-e^{i\phi}}\right|_{z=e^{i\omega}} \\
&= \left|e^{i\phi}\left(\frac{e^{-i\phi}z-r}{rz-e^{i\phi}}\right)\right|_{z=e^{i\omega}} &&= \left|z\left(\frac{e^{-i\phi}-rz^{-1}}{rz-e^{i\phi}}\right)\right|_{z=e^{i\omega}} \\
&= \left|-z\left(\frac{rz^{-1}-e^{-i\phi}}{rz-e^{i\phi}}\right)\right|_{z=e^{i\omega}} &&= \left|\underbrace{e^{i\pi}}_{-1}e^{i\omega}\left(\frac{re^{-i\omega}-e^{-i\phi}}{re^{i\omega}-e^{i\phi}}\right)\right| \\
&= \left|\frac{1}{e^{-iv}}\left(\frac{re^{-i\omega}-e^{-i\phi}}{\left(re^{i\omega}-e^{i\phi}\right)^*}\right)\right| &&= \left|\frac{re^{-i\omega}-e^{-i\phi}}{re^{-i\omega}-e^{-i\phi}}\right| \\
&= 1
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

Copyright © 2019 Daniel J. Greenhoe

APPENDIX S

B-SPLINES

S.1 Definitions

Definition S.1. *Let X be a set.*

DEF *The **step function** $\sigma \in \mathbb{R}^{\mathbb{R}}$ is defined as*

$$\sigma(x) \triangleq \mathbb{1}_{[0:\infty)}(x) \qquad \forall x \in \mathbb{R}.$$

Lemma S.1. *Let $\sigma(x)$ be the* STEP FUNCTION *(Definition S.1 page 383).*

LEM $$\{g(x) > 0\} \quad \Longrightarrow \quad \left\{\sigma\left[g(x)f(x)\right] = \sigma[f(x)]\right\} \qquad \forall f,g \in \mathbb{R}^{\mathbb{R}}$$

✎PROOF:

$$\begin{aligned}
\sigma\left[g(x)f(x)\right] &\triangleq \mathbb{1}_{[0:\infty)}\left[g(x)f(x)\right] && \text{by definition of } \sigma(x) && \text{(Definition S.1 page 383)}\\
&\triangleq \begin{cases} 1 & \text{for } g(x)f(x) \geq 0 \\ 0 & \text{otherwise} \end{cases} && \text{by definition of } \mathbb{1}_A(x) && \text{(Definition ?? page ??)}\\
&= \begin{cases} 1 & \text{for } f(x) \geq 0 \\ 0 & \text{otherwise} \end{cases} && \text{by } g(x) > 0 \text{ hypothesis} \\
&\triangleq \mathbb{1}_{[0:\infty)}[f(x)] && \text{by definition of } \mathbb{1}_A(x) && \text{(Definition ?? page ??)}\\
&\triangleq \sigma[f(x)] && \text{by definition of } \sigma(x) && \text{(Definition S.1 page 383)}
\end{aligned}$$

✏

Definition S.2. [1] *Let $\mathbb{1}$ be the* SET INDICATOR *function (Definition ?? page ??). Let $f(x) \star g(x)$ represent the* CONVOLUTION *operation (Definition P.3 page 352).*

DEF *The n**th order cardinal B-spline** $\mathsf{N}_n(x)$ for $n \in \mathbb{W}$ is defined as*

$$\mathsf{N}_n(x) \triangleq \left\{ \begin{array}{ll} \mathbb{1}_{[0:1)}(x) & \text{for } n = 0 \\ \mathsf{N}_{n-1}(x) \star \mathsf{N}_0(x) & \text{for } n \in \mathbb{W}\backslash 0 \end{array} \right\} \qquad \forall x \in \mathbb{R}$$

Lemma S.2. [2]

[1] Chui (1992) page 85 ⟨(4.2.1)⟩, Christensen (2008) page 140, Chui (1988) page 1

[2] Christensen (2008) page 140, Chui (1992) page 85 ⟨(4.2.1)⟩, Chui (1988) page 1, Prasad and Iyengar (1997) page 145

LEM

$$\mathsf{N}_n(x) = \int_{\tau=0}^{\tau=1} \mathsf{N}_{n-1}(x-\tau)\,\mathsf{d}\tau \qquad \forall n \in \{1,2,3,\ldots\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{N}_n(x) &\triangleq \mathsf{N}_{n-1}(x) \star \mathsf{N}_0(x) && \text{by definition of } \mathsf{N}_n(x) && \text{(Definition S.2 page 383)}\\
&\triangleq \int_{\mathbb{R}} \mathsf{N}_{n-1}(x-\tau)\mathsf{N}_0(\tau)\,\mathsf{d}\tau && \text{by definition of convolution operation } \star && \text{(Definition P.3 page 352)}\\
&\triangleq \int_{\mathbb{R}} \mathsf{N}_{n-1}(x-\tau)\mathbb{1}_{[0:1)}(\tau)\,\mathsf{d}\tau && \text{by definition of } \mathsf{N}_0(x) && \text{(Definition S.2 page 383)}\\
&= \int_{[0:1)} \mathsf{N}_{n-1}(x-\tau)\,\mathsf{d}\tau && \text{by definition of } \mathbb{1}_A(x) && \text{(Definition ?? page ??)}\\
&= \int_{[0:1]} \mathsf{N}_{n-1}(x-\tau)\,\mathsf{d}\tau\\
&\triangleq \int_{0}^{1} \mathsf{N}_{n-1}(x-\tau)\,\mathsf{d}\tau
\end{aligned}$$

✏

Lemma S.3. *Let* $\mathsf{f}(x)$ *be a* FUNCTION *in* $\mathbb{R}^{\mathbb{R}}$. *Let* $\mathsf{F}(x)$ *be the* ANTI-DERIVATIVE *of* $\mathsf{f}(x)$. *Let* $\sigma(x)$ *be the* STEP FUNCTION *(Definition S.1 page 383)*.

LEM

$$\begin{aligned}
&\int_{y=a}^{y=b} \mathsf{f}(x-y)\sigma(x-y)\,\mathsf{d}y\\
&= \left\{\begin{array}{ll} -\displaystyle\int_{y=x-a}^{y=x-b} \mathsf{f}(y)\,\mathsf{d}y & \textit{for } x \geq b\\ -\displaystyle\int_{y=x-a}^{y=0} \mathsf{f}(y)\,\mathsf{d}y & \textit{for } a \leq x \leq b\\ 0 & \textit{for } x \leq a \end{array}\right\} = \left\{\begin{array}{ll} \mathsf{F}(x-a)-\mathsf{F}(x-b) & \textit{for } x \geq b\\ \mathsf{F}(x-a)-\mathsf{F}(0) & \textit{for } a \leq x \leq b\\ 0 & \textit{for } x \leq a \end{array}\right\}\\
&= [\mathsf{F}(x-a)-\mathsf{F}(0)]\sigma(x-a) + [\mathsf{F}(0)-\mathsf{F}(x-b)]\sigma(x-b)
\end{aligned}$$

✎PROOF:

$$\begin{aligned}
\int_{y=a}^{y=b} \mathsf{f}(x-y)\sigma(x-y)\,\mathsf{d}y &= \left\{\begin{array}{ll} \int_{y=a}^{y=b} \mathsf{f}(x-y)\,\mathsf{d}y & \text{for } x \geq b\\ \int_{y=a}^{y=x} \mathsf{f}(x-y)\,\mathsf{d}y & \text{for } a \leq x \leq b\\ 0 & \text{for } x \leq a \end{array}\right\} && \text{by definition of } \sigma \text{ (Definition S.1 page 383)}\\
&= \left\{\begin{array}{ll} -\int_{u=x-a}^{u=x-b} \mathsf{f}(u)\,\mathsf{d}u & \text{for } x \geq b\\ -\int_{u=x-a}^{u=0} \mathsf{f}(u)\,\mathsf{d}u & \text{for } a \leq x \leq b\\ 0 & \text{for } x \leq a \end{array}\right\} && \text{where } u \triangleq x-y \implies y = x-u\\
&= \left\{\begin{array}{ll} -\int_{y=x-a}^{y=x-b} \mathsf{f}(y)\,\mathsf{d}y & \text{for } x \geq b\\ -\int_{y=x-a}^{y=0} \mathsf{f}(y)\,\mathsf{d}y & \text{for } a \leq x \leq b\\ 0 & \text{for } x \leq a \end{array}\right\} && \text{by change of dummy variable } (u \to y)\\
&= \left\{\begin{array}{ll} \mathsf{F}(x-a)-\mathsf{F}(x-b) & \text{for } x \geq b\\ \mathsf{F}(x-a)-\mathsf{F}(0) & \text{for } a \leq x \leq b\\ 0 & \text{for } x \leq a \end{array}\right\} && \text{by } \textit{Fundamental Theorem of Calculus}\\
&= [\mathsf{F}(x-a)-\mathsf{F}(x-b)]\sigma(x-b) + [\mathsf{F}(x-a)-\mathsf{F}(0)][\sigma(x-a)-\sigma(x-b)]\\
&= [\mathsf{F}(x-a)-\mathsf{F}(0)]\sigma(x-a) + [\mathsf{F}(x-a)-\mathsf{F}(x-b)-\mathsf{F}(x-a)+\mathsf{F}(0)]\sigma(x-b)\\
&= [\mathsf{F}(x-a)-\mathsf{F}(0)]\sigma(x-a) + [\mathsf{F}(0)-\mathsf{F}(x-b)]\sigma(x-b)
\end{aligned}$$

✏

COPYRIGHT © 2019 DANIEL J. GREENHOE

Figure S.1: some low order B-splines (Example S.1 page 385)

Lemma S.4. *Let* $\sigma(x)$ *be the* STEP FUNCTION *(Definition S.1 page 383).*

LEM
$$\int_{\tau=0}^{\tau=1} (x-\tau-k)^n \sigma(x-\tau-k)\,d\tau = \frac{1}{n+1}\left[(x-k)^{n+1}\sigma(x-k) - (x-k-1)^{n+1}\sigma(x-k-1)\right]$$

✎PROOF:

$$\begin{aligned}
&\int_{\tau=0}^{\tau=1} (x-\tau-k)^n \sigma(x-\tau-k)\,d\tau \\
&= \int_{y=k}^{y=k+1} (x-y)^n \sigma(x-y)\,dy && \text{where } y \triangleq \tau+k \implies \tau = y-k \\
&= [F(x-k)-F(0)]\sigma(x-k) + [F(0)-F(x-k-1)]\sigma(x-k-1) && \text{by Lemma S.3 (page 384), where } f(x) \triangleq x^n \\
&= \frac{\left[(x-k)^{n+1}-0\right]\sigma(x-k) + \left[0-(x-k-1)^{n+1}\right]\sigma(x-k-1)}{n+1} && \text{because } F(x) \triangleq \int f(x)\,dx = \frac{x^{n+1}}{n+1}+c \\
&= \frac{1}{n+1}\left[(x-k)^{n+1}\sigma(x-k) - (x-k-1)^{n+1}\sigma(x-k-1)\right]
\end{aligned}$$

✏

Example S.1. [3] Let $\sigma(x)$ be the *step function* (Definition S.1 page 383). Let $\binom{n}{k}$ be the *binomial coefficient* (Definition ?? page ??). The 0*th order B-spline* (Definition S.2 page 383) $N_0(x)$ can be expressed as follows:

EX
$$N_0(x) = \left\{\begin{array}{ll} 1 & \text{for } x \in [0:1) \\ 0 & \text{otherwise} \end{array}\right\} = \left\{\sum_{k=0}^{1} (-1)^k \binom{1}{k} (x-k)^0 \sigma(x-k) \quad \forall x \in \mathbb{R}\right\}$$

The B-spline $N_0(x)$ is illustrated in Figure S.1 (page 385).

[3] Schumaker (2007) page 136 ⟨Table 1⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

✎PROOF:

$$\begin{aligned}
\mathsf{N}_0(x) &= \mathbb{1}_{[0:1)}(x) && \text{by definition of } \mathsf{N}_0(x) && \text{(Definition S.2 page 383)}\\
&= \sigma(x) - \sigma(x-1) && \text{by definition of } \sigma(x) && \text{(Definition S.1 page 383)}\\
&= \left[\binom{1}{0}\sigma(x) - \binom{1}{1}\sigma(x-1)\right] && \text{by definition of } \textit{binomial coefficient } \binom{n}{k} && \text{(Definition ?? page ??)}\\
&= \sum_{k=0}^{1}(-1)^k\binom{1}{k}(x-k)^0\sigma(x-k) && \text{by definition of } \sum \text{ operator}
\end{aligned}$$

✏

Example S.2. [4] Let $\sigma(x)$ be the *step function*. Let $\binom{n}{k}$ be the *binomial coefficient*.
The 1*st order B-spline* $\mathsf{N}_1(x)$ can be expressed as follows:

EX
$$\mathsf{N}_1(x) = \left\{\begin{array}{rl} x & \text{for } x \in [0:1]\\ -x+2 & \text{for } x \in [1:2]\\ 0 & \text{otherwise}\end{array}\right\} = \left\{\sum_{k=0}^{2}(-1)^k\binom{2}{k}(x-k)\sigma(x-k) \quad \forall x\in\mathbb{R}\right\}$$

The B-spline $\mathsf{N}_1(x)$ is illustrated in Figure S.1 (page 385).

✎PROOF:

$$\begin{aligned}
\mathsf{N}_1(x) &= \int_{\tau=0}^{\tau=1} \mathsf{N}_0(x-\tau)\,\mathrm{d}\tau && \text{by Lemma S.2 page 383}\\
&= \int_{\tau=0}^{\tau=1} \sum_{k=0}^{1}(-1)^k\binom{1}{k}(x-\tau-k)^0\sigma(x-\tau-k)\,\mathrm{d}\tau && \text{by Example S.1 page 385}\\
&= \sum_{k=0}^{1}(-1)^k\binom{1}{k}\int_{\tau=0}^{\tau=1}(x-\tau-k)^0\sigma(x-\tau-k)\,\mathrm{d}\tau && \text{by } \textit{linearity} \text{ of } \int \mathrm{d}\tau \text{ operator}\\
&= \sum_{k=0}^{1}(-1)^k\binom{1}{k}\frac{1}{0+1}\left[(x-k)^{0+1}\sigma(x-k) - (x-k-1)^{0+1}\sigma(x-k-1)\right] && \text{by Lemma S.4 page 385}\\
&= \begin{pmatrix} 1\{(x-0)\sigma(x-0) - (x-1)\sigma(x-1)\}\\ -1\{(x-1)\sigma(x-1) - (x-2)\sigma(x-2)\}\end{pmatrix}\\
&= x\sigma(x) - 2(x-1)\sigma(x-1) + (x-2)\sigma(x-2)\\
&= \sum_{k=0}^{2}(-1)^k\binom{2}{k}(x-k)\sigma(x-k) && \text{by def. of } \binom{n}{k} \text{ (Definition ?? page ??)}\\
&= \left\{\begin{array}{ll} x & \text{for } x \in [0:1)\\ -x+2 & \text{for } x \in [1:2]\\ 0 & \text{otherwise}\end{array}\right\} && \text{by def. of } \sigma(x) \text{ (Definition S.1 page 383)}
\end{aligned}$$

✏

Example S.3. [5] Let $\sigma(x)$ be the *step function*. Let $\binom{n}{k}$ be the *binomial coefficient*.
The 2*nd order B-spline* $\mathsf{N}_2(x)$ can be expressed as follows:

EX
$$\mathsf{N}_2(x) = \frac{1}{2}\left\{\begin{array}{rl} x^2 & \text{for } x \in [0:1)\\ -2x^2+6x-3 & \text{for } x \in [1:2]\\ x^2-6x+9 & \text{for } x \in [2:3]\\ 0 & \text{otherwise}\end{array}\right\} = \left\{\frac{1}{2}\sum_{k=0}^{3}(-1)^k\binom{3}{k}(x-k)^2\sigma(x-k) \quad \forall x\in\mathbb{R}\right\}$$

The B-spline $\mathsf{N}_2(x)$ is illustrated in Figure S.1 (page 385).

[4] Christensen (2008) page 148 ⟨Exercise 6.2⟩, Christensen (2010) page 212 ⟨Exercise 10.2⟩, Heil (2011) pages 142–143 ⟨Definition 4.22 (The Schauder System)⟩, Schumaker (2007) page 136 ⟨Table 1⟩, Stoer and Bulirsch (2002) page 124

[5] Christensen (2008) page 148 ⟨Exercise 6.2⟩, Christensen (2010) page 212 ⟨Exercise 10.2⟩, Schumaker (2007) page 136 ⟨Table 1⟩, Stoer and Bulirsch (2002) page 124

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

$$
\begin{aligned}
\mathsf{N}_2(x) &= \int_{\tau=0}^{\tau=1} \mathsf{N}_1(x-\tau)\,\mathrm{d}\tau && \text{by Lemma S.2 page 383}\\
&= \int_{\tau=0}^{\tau=1} \sum_{k=0}^{2}(-1)^k\binom{2}{k}(x-\tau-k)\sigma(x-\tau-k)\,\mathrm{d}\tau && \text{by Example S.2 page 386}\\
&= \sum_{k=0}^{2}(-1)^k\binom{2}{k}\int_{\tau=0}^{\tau=1}(x-\tau-k)\sigma(x-\tau-k)\,\mathrm{d}\tau && \text{by \textit{linearity} of } \int \mathrm{d}\tau \text{ operator}\\
&= \sum_{k=0}^{2}(-1)^k\binom{2}{k}\frac{1}{1+1}\left[(x-k)^{1+1}\sigma(x-k)-(x-k-1)^{1+1}\sigma(x-k-1)\right] && \text{by Lemma S.4 page 385}\\
&= \frac{1}{2}\begin{pmatrix} 1 & \{(x-0)^2\sigma(x-0)-(x-1)^2\sigma(x-1)\}\\ -2 & \{(x-1)^2\sigma(x-1)-(x-2)^2\sigma(x-2)\}\\ +1 & \{(x-2)^2\sigma(x-2)-(x-3)^2\sigma(x-3)\}\end{pmatrix}\\
&= \frac{1}{2}\left[x^2\sigma(x)-3(x-1)^2\sigma(x-1)+3(x-2)^2\sigma(x-2)-(x-3)^2\sigma(x-3)\right]\\
&= \frac{1}{2}\sum_{k=0}^{3}(-1)^k\binom{3}{k}(x-k)^2\sigma(x-k) && \text{by def. of } \binom{n}{k} \text{ (Definition ?? page ??)}\\
&= \frac{1}{2}\left\{\begin{array}{rl} x^2 & \text{for } x\in[0:1)\\ -2x^2+6x-3 & \text{for } x\in[1:2]\\ x^2-6x+9 & \text{for } x\in[2:3]\\ 0 & \text{otherwise}\end{array}\right\} && \text{by def. of } \sigma(x) \text{ (Definition S.1 page 383)}
\end{aligned}
$$

The final steps of this proof can be calculated "by hand" or by using the free and open source software package *Maxima* along with the script file listed in Section **??** (page **??**). ✏

S.2 Algebraic properties

Theorem S.1 (next) presents a closed form expression for an *nth order B-spline* $\mathsf{N}_n(x)$ based on the definition of $\mathsf{N}_n(x)$ given in Definition S.2 (page 383). Alternatively, Theorem S.1 could serve as the definition and Definition S.2 as a property.

Theorem S.1. [6] *Let* $\mathsf{N}_n(x)$ *be the* nTH ORDER B-SPLINE *(Definition S.2 page 383)*.
Let $\sigma(x)$ *be the* STEP FUNCTION *(Definition S.1 page 383)*.

THM

$$\mathsf{N}_n(x) = \frac{1}{n!}\sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k}(x-k)^n\sigma(x-k) \qquad \forall n\in\{0,1,2,\ldots\}=\mathbb{W}$$

✎PROOF: Proof follows by induction:

1. base case (choose one):
   - Proof for $n=0$ case: by Example S.1 (page 385).
   - Proof for $n=1$ case: by Example S.2 (page 386).
   - Proof for $n=2$ case: by Example S.3 (page 386).

[6] Christensen (2008) page 142 ⟨Theorem 6.1.3⟩, Chui (1992) page 84 ⟨(4.1.12)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

2. inductive step—proof that n case $\implies$ $n+1$ case:

$$
\begin{aligned}
&\mathsf{N}_{n+1}(x) \\
&= \int_0^1 \mathsf{N}_n(x-\tau)\,\mathrm{d}\tau && \text{by Lemma S.2 page 383} \\
&= \int_0^1 \frac{1}{n!}\sum_{k=0}^{k=n+1}(-1)^k\binom{n+1}{k}(x-\tau-k)^n\sigma(x-\tau-k)\,\mathrm{d}\tau && \text{by induction hypothesis} \\
&= \frac{1}{n!}\sum_{k=0}^{k=n+1}(-1)^k\binom{n+1}{k}\int_0^1(x-\tau-k)^n\sigma(x-\tau-k)\,\mathrm{d}\tau && \text{by \textit{linearity} of } \int \mathrm{d}\tau \text{ operator} \\
&= \frac{1}{n!}\sum_{k=0}^{k=n+1}(-1)^k\binom{n+1}{k}\frac{1}{n+1}\left[(x-k)^{n+1}\sigma(x-k)-(x-k-1)^{n+1}\sigma(x-k-1)\right] && \text{by Lemma S.4 page 385} \\
&= \frac{1}{(n+1)!}\sum_{k=0}^{k=n+1}(-1)^k\binom{n+1}{k}\left[(x-k)^{n+1}\sigma(x-k)-(x-k-1)^{n+1}\sigma(x-k-1)\right] \\
&= \frac{1}{(n+1)!}\left[\sum_{k=0}^{k=n+1}(-1)^k\binom{n+1}{k}(x-k)^{n+1}\sigma(x-k)-\sum_{k=0}^{k=n+1}(-1)^k\binom{n+1}{k}(x-k-1)^{n+1}\sigma(x-k-1)\right] \\
&= \frac{1}{(n+1)!}\left[\sum_{k=0}^{k=n+1}(-1)^k\binom{n+1}{k}(x-k)^{n+1}\sigma(x-k)-\sum_{m=1}^{m=n+2}(-1)^{m-1}\binom{n+1}{m-1}(x-m)^{n+1}\sigma(x-m)\right] \\
&\quad \text{where } m \triangleq k+1 \implies k=m-1 \\
&= \frac{1}{(n+1)!}\left(\begin{array}{l}\displaystyle\sum_{k=0}^{k=n+1}(-1)^k\binom{n+1}{k}(x-k)^{n+1}\sigma(x-k) \\ \displaystyle-\sum_{m=1}^{m=n+2}(-1)^{m-1}\left[\binom{n+2}{m}-\binom{n+1}{m}\right](x-m)^{n+1}\sigma(x-m)\end{array}\right) && \begin{array}{l}\text{by \textit{Pascal's identity} /} \\ \textit{Stifel formula} \\ \text{(Theorem ?? page ??)}\end{array} \\
&= \frac{1}{(n+1)!}\left(\begin{array}{l}\displaystyle\sum_{m=1}^{m=n+2}(-1)^m\binom{n+2}{m}(x-m)^{n+1}\sigma(x-m) \\ \displaystyle-\sum_{m=1}^{m=n+2}(-1)^m\binom{n+1}{m}(x-m)^{n+1}\sigma(x-m) \\ \displaystyle+\sum_{k=0}^{k=n+1}(-1)^k\binom{n+1}{k}(x-k)^{n+1}\sigma(x-k)\end{array}\right) && \text{note } (-1)^{m-1}=-(-1)^m \\
&= \frac{1}{(n+1)!}\left(\begin{array}{l|ll}\displaystyle\sum_{m=0}^{m=n+2}(-1)^m\tbinom{n+2}{m}(x-m)^{n+1}\sigma(x-m) & \text{(A)} & \text{desired } n+1 \text{ case} \\ \displaystyle-\quad(-1)^0\tbinom{n+2}{0}(x-0)^{n+1}\sigma(x-0) & \text{(B)} & \text{cancelled by (F)} \\ \displaystyle-\sum_{m=1}^{m=n+1}(-1)^m\tbinom{n+1}{m}(x-m)^{n+1}\sigma(x-m) & \text{(C)} & \text{cancelled by (E)} \\ \displaystyle-\quad(-1)^{n+2}\tbinom{n+1}{n+2}(x-n-2)^{n+1}\sigma(x-n-2) & \text{(D)} & \tbinom{n+1}{n+2}=0 \text{ by Proposition ?? page ??} \\ \displaystyle+\sum_{k=1}^{k=n+1}(-1)^k\tbinom{n+1}{k}(x-k)^{n+1}\sigma(x-k) & \text{(E)} & \text{cancelled by (C)} \\ \displaystyle+\quad(-1)^0\tbinom{n+1}{0}(x-0)^{n+1}\sigma(x-0) & \text{(F)} & \tbinom{n+2}{0}=\tbinom{n+1}{0}=1, \text{ so (F) is cancelled by (B)}\end{array}\right) \\
&= \frac{1}{(n+1)!}\sum_{m=0}^{m=n+2}(-1)^m\binom{n+2}{m}(x-m)^{n+1}\sigma(x-m) && (n+1 \text{ case})
\end{aligned}
$$

Copyright © 2019 Daniel J. Greenhoe

✏

Example S.4. [7] Let $\mathsf{N}_3(x)$ be the 3*rd order B-spline* (Definition S.2 page 383).[8]

EX

$$\mathsf{N}_3(x) = \frac{1}{6}\left\{\begin{array}{rl} x^3 & \text{for } 0 \le x \le 1 \\ -3x^3+12x^2-12x+4 & \text{for } 1 \le x \le 2 \\ 3x^3-24x^2+60x-44 & \text{for } 2 \le x \le 3 \\ -x^3+12x^2-48x+64 & \text{for } 3 \le x \le 4 \\ 0 & \text{otherwise} \end{array}\right\}$$

✎PROOF: This expression can be calculated "by hand" using Theorem S.1 (page 387) or by using the free and open source software package *Maxima* along with the script file listed in Section **??** (page **??**). ✏

Example S.5. Let $\mathsf{N}_4(x)$ be the 4*th order B-spline* (Definition S.2 page 383).

EX

$$\mathsf{N}_4(x) = \frac{1}{24}\left\{\begin{array}{rl} x^4 & \text{for } 0 \le x \le= 1 \\ -4x^4+20x^3-30x^2+20x-5 & \text{for } 1 \le x \le= 2 \\ 6x^4-60x^3+210x^2-300x+155 & \text{for } 2 \le x \le= 3 \\ -4x^4+60x^3-330x^2+780x-655 & \text{for } 3 \le x \le= 4 \\ x^4-20x^3+150x^2-500x+625 & \text{for } 4 \le x \le= 5 \\ 0 & \text{otherwise} \end{array}\right\}$$

✎PROOF: This expression can be calculated "by hand" using Theorem S.1 (page 387) or by using the free and open source software package *Maxima* along with the script file listed in Section **??** (page **??**). ✏

Example S.6. Let $\mathsf{N}_5(x)$ be the 5*th order B-spline* (Definition S.2 page 383).

EX

$$\mathsf{N}_5(x) = \frac{1}{120}\left\{\begin{array}{rl} x^5 & \text{for } 0 \le x \le= 1 \\ -5x^5+30x^4-60x^3+60x^2-30x+6 & \text{for } 1 \le x \le= 2 \\ 10x^5-120x^4+540x^3-1140x^2+1170x-474 & \text{for } 2 \le x \le= 3 \\ -10x^5+180x^4-1260x^3+4260x^2-6930x+4386 & \text{for } 3 \le x \le= 4 \\ 5x^5-120x^4+1140x^3-5340x^2+12270x-10974 & \text{for } 4 \le x \le= 5 \\ -x^5+30x^4-360x^3+2160x^2-6480x+7776 & \text{for } 5 \le x \le= 6 \\ 0 & \text{otherwise} \end{array}\right\}$$

The 5*th order B-spline* $\mathsf{N}_5(x)$ is illustrated in Figure S.1 (page 385).

✎PROOF: This expression can be calculated "by hand" using Theorem S.1 (page 387) or by using the free and open source software package *Maxima* along with the script file listed in Section **??** (page **??**). ✏

Example S.7. Let $\mathsf{N}_6(x)$ be the 6*th order B-spline* (Definition S.2 page 383).

EX

$$\mathsf{N}_6(x) = \frac{1}{720}\left\{\begin{array}{rl} x^6 & \text{for } 0 \le x \le 1 \\ -6x^6+42x^5-105x^4+140x^3-105x^2+42x-7 & \text{for } 1 \le x \le 2 \\ 15x^6-210x^5+1155x^4-3220x^3+4935x^2-3990x+1337 & \text{for } 2 \le x \le 3 \\ -20x^6+420x^5-3570x^4+15680x^3-37590x^2+47040x-24178 & \text{for } 3 \le x \le 4 \\ 15x^6-420x^5+4830x^4-29120x^3+96810x^2-168000x+119182 & \text{for } 4 \le x \le 5 \\ -6x^6+210x^5-3045x^4+23380x^3-100065x^2+225750x-208943 & \text{for } 5 \le x \le 6 \\ x^6-42x^5+735x^4-6860x^3+36015x^2-100842x+117649 & \text{for } 6 \le x \le 7 \\ 0 & \text{otherwise} \end{array}\right\}$$

The 6*th order B-spline* $\mathsf{N}_6(x)$ is illustrated in Figure S.1 (page 385).

✎PROOF: This expression can be calculated "by hand" using Theorem S.1 (page 387) or by using the free and open source software package *Maxima* along with the script file listed in Section **??** (page **??**). ✏

[7] Schumaker (2007) page 136 ⟨Table 1⟩, Shizgal (2015) page 92 ⟨(2.199)⟩, Szabó and Horváth (2004) page 146 ⟨(4)⟩, Wei and Billings (2006) page 578 ⟨Table 1⟩, Maleknejad et al. (2013) ⟨(9)⟩

[8] For help with plotting *B-splines*, see APPENDIX T (page 415).

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Example S.8. Let $\mathsf{N}_7(x)$ be the 7*th order B-spline* (Definition S.2 page 383).

EX

$$7!\mathsf{N}_7(x) = 5040\mathsf{N}_7(x) =$$
$$\left\{\begin{array}{rrrrrrrrrl}
x^7 & & & & & & & & & \text{for } 0 \le x \le 1 \\
-7x^7 & +56x^6 & -168x^5 & +280x^4 & -280x^3 & +168x^2 & -56x & +8 & & \text{for } 1 \le x \le 2 \\
21x^7 & -336x^6 & +2184x^5 & -7560x^4 & +15400x^3 & -18648x^2 & +12488x & -3576 & & \text{for } 2 \le x \le 3 \\
-35x^7 & +840x^6 & -8400x^5 & +45360x^4 & -143360x^3 & +267120x^2 & -273280x & +118896 & & \text{for } 3 \le x \le 4 \\
35x^7 & -1120x^6 & +15120x^5 & -111440x^4 & +483840x^3 & -1238160x^2 & +1733760x & -1027984 & & \text{for } 4 \le x \le 5 \\
-21x^7 & +840x^6 & -14280x^5 & +133560x^4 & -741160x^3 & +2436840x^2 & -4391240x & +3347016 & & \text{for } 5 \le x \le 6 \\
7x^7 & -336x^6 & +6888x^5 & -78120x^4 & +528920x^3 & -2135448x^2 & +4753336x & -4491192 & & \text{for } 6 \le x \le 7 \\
-x^7 & +56x^6 & -1344x^5 & +17920x^4 & -143360x^3 & +688128x^2 & -1835008x & +2097152 & & \text{for } 7 \le x \le 8 \\
 & & & & & & & & 0 & \text{otherwise}
\end{array}\right\}$$

The 7*th order B-spline* $\mathsf{N}_7(x)$ is illustrated in Figure S.1 (page 385).

✎PROOF: This expression can be calculated "by hand" using Theorem S.1 (page 387) or by using the free and open source software package *Maxima* along with the script file listed in Section **??** (page **??**). ✏

Example S.9. [9] The $(n+1)^2$ coefficients of the order $n, n-1, \ldots, 0$ monomials of each *B-spline* $\mathsf{N}_n(x)$ multiplied by $n!$ induce an *integer sequence*
$\mathbb{x} \triangleq (\!(1,1,0,-1,2,1,0,0,-2,6,-3,1,-6,9,1,0,0,0,-3,12,-12,4,3,-24,60,-44,-1,12,-48,64,\ldots)\!)$
as more fully listed in Table S.1 (page 414). In this sequence $\mathbb{x} \triangleq (\!(x_0, x_1, x_2, \ldots)\!)$, the coefficients for the *order n B-spline* $\mathsf{N}_n(x)$ begin at the sequence index value

$$p \triangleq \sum_{k=0}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1) \qquad \text{and end at index value } p + (n+1)^2 - 1.$$

For example, the coefficients for $\mathsf{N}_3(x)$ begin at index value $p \triangleq 0+1+4+9 = 14$ and end at index value $p + 4^2 - 1 = 29$. Using these coefficients gives the following expression for $\mathsf{N}_3(x)$:

$$\mathsf{N}_3(x) = \left[\begin{array}{rrrr|l}
1 & 0 & 0 & 0 & \text{for } 0 \le x < 1 \\
-3 & 12 & -12 & 4 & \text{for } 1 \le x < 2 \\
3 & -24 & 60 & -44 & \text{for } 2 \le x < 3 \\
-1 & 12 & -48 & 64 & \text{for } 3 \le x < 4
\end{array}\right]\left[\begin{array}{c} x^3 \\ x^2 \\ x \\ 1 \end{array}\right] = \left\{\begin{array}{rrrrl}
x^3 & & & & \text{for } 0 \le x < 1 \\
-3x^3 & +12x^2 & -12x & +4 & \text{for } 1 \le x < 2 \\
3x^3 & -24x^2 & +60x & -44 & \text{for } 2 \le x < 3 \\
-x^3 & +12x^2 & -48x & +64 & \text{for } 3 \le x < 4 \\
 & & & 0 & \text{otherwise}
\end{array}\right.$$

…which agrees with the result presented in Example S.4 (page 389).

✎PROOF:

1. The coefficients for the sequence $\mathbb{x}$ may be computed with assistance from *Maxima* together with the script file listed in Section **??** (page **??**).

2. Proof that $\sum_{k=0}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1)$: The summation is a *power sum*. The relation may be proved using *induction*.[10]

   (a) Base case: n=0 case …

$$\begin{aligned}
\sum_{k=0}^{k=0} k^2 &= 0 \\
&= \frac{0(0+1)(2\cdot 0+1)}{6} \\
&= \left.\frac{n(n+1)(2n+1)}{6}\right|_{n=0}
\end{aligned}$$

[9] Greenhoe (2017b)
[10] Greenhoe (2017a), pages 186–187 ⟨Proposition 11.2 (Power Sums)⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

(b) Base case: n=1 case …

$$\begin{aligned}\sum_{k=0}^{k=1} k^2 &= 0+1\\ &= \frac{1(1+1)(2\cdot 1+1)}{6}\\ &= \left.\frac{n(n+1)(2n+1)}{6}\right|_{n=1}\end{aligned}$$

(c) inductive step—proof that n case $\implies$ $n+1$ case:

$$\begin{aligned}\sum_{k=0}^{n+1} k^2 &= \left(\sum_{k=0}^{n} k^2\right) + \left(n+1\right)^2\\ &= \left(\frac{n(n+1)(2n+1)}{6}\right) + \left(n+1\right)^2 && \text{by } n \text{ case hypothesis}\\ &= \left(n+1\right)\left(\frac{n(2n+1)+6(n+1)}{6}\right)\\ &= \left(n+1\right)\left(\frac{2n^2+7n+6}{6}\right)\\ &= \left(n+1\right)\left(\frac{(n+2)(2n+3)}{6}\right)\\ &= \frac{(n+1)[(n+1)+1][2(n+1)+1]}{6}\end{aligned}$$

✏

Theorem S.2. [11]

THM $\frac{\mathsf{d}}{\mathsf{d}x}\mathsf{N}_n(x) = \mathsf{N}_{n-1}(x) - \mathsf{N}_{n-1}(x-1) \qquad \forall n \in \mathbb{N}, \forall x \in \mathbb{R}$

✎PROOF:

1. Proof using Lemma S.2 (page 383) and the *Fundamental Theorem of Calculus*:

$$\begin{aligned}\frac{\mathsf{d}}{\mathsf{d}x}\mathsf{N}_n(x) &= \frac{\mathsf{d}}{\mathsf{d}x}\int_0^1 \mathsf{N}_{n-1}(x-\tau)\,\mathsf{d}\tau && \text{by Lemma S.2 page 383}\\ &= \frac{\mathsf{d}}{\mathsf{d}x}\int_{x-u=0}^{x-u=1} \mathsf{N}_{n-1}(u)(-1)\,\mathsf{d}u && \text{where } u \triangleq x-\tau \implies \tau = x-u\\ &= \frac{\mathsf{d}}{\mathsf{d}x}\int_{u=x-1}^{u=x} \mathsf{N}_{n-1}(u)\,\mathsf{d}u\\ &= \left\{\frac{\mathsf{d}}{\mathsf{d}x}\int \mathsf{N}_{n-1}(u)\,\mathsf{d}u\Big|_{u=x}\right\} - \left\{\frac{\mathsf{d}}{\mathsf{d}x}\int \mathsf{N}_{n-1}(u)\,\mathsf{d}u\Big|_{u=x-1}\right\} && \text{by } \textit{Fundamental Theorem of Calculus}^{12}\\ &= \left\{\mathsf{N}_{n-1}(x)\frac{\mathsf{d}}{\mathsf{d}x}(x)\right\} - \left\{\mathsf{N}_{n-1}(x-1)\frac{\mathsf{d}}{\mathsf{d}x}(x-1)\right\} && \text{by } \textit{Chain Rule}^{13}\\ &= \mathsf{N}_{n-1}(x) - \mathsf{N}_{n-1}(x-1)\end{aligned}$$

[11] Höllig (2003) page 25 ⟨3.2⟩, Schumaker (2007) page 121 ⟨Theorem 4.16⟩
[12] Hijab (2011) page 163 ⟨Theorem 4.4.3⟩
[13] Hijab (2011) pages 73–74 ⟨Theorem 3.1.2⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

2. Proof using Lemma S.2 (page 383) and *induction*:

   (a) Base case …proof for $n = 1$ case:

$$\begin{aligned}
\mathsf{N}_0(x) - \mathsf{N}_0(x-1) &= \underbrace{\sigma(x) - \sigma(x-1)}_{\mathsf{N}_0(x)} - \underbrace{[\sigma(x-1) - \sigma(x-2)]}_{\mathsf{N}_0(x-1)} && \text{by Example S.1 page 385}\\
&= \sigma(x) - 2\sigma(x-1) + \sigma(x-2)\\
&= \sum_{k=0}^{2}(-1)^k\binom{2}{k}\sigma(x-k)\\
&= \frac{\mathsf{d}}{\mathsf{d}x}\sum_{k=0}^{2}(-1)^k\binom{2}{k}(x-k)\sigma(x-k)\\
&= \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{N}_1(x) && \text{by Example S.2 page 386}
\end{aligned}$$

   (b) Base case …proof for $n = 2$ case:

$$\begin{aligned}
&\mathsf{N}_1(x) - \mathsf{N}_1(x-1)\\
&= \underbrace{x\sigma(x) - 2(x-1)\sigma(x-1) + (x-2)\sigma(x-2)}_{\mathsf{N}_1(x)}\\
&\quad - \underbrace{[(x-1)\sigma(x-1) - 2(x-2)\sigma(x-2) + (x-3)\sigma(x-3)]}_{\mathsf{N}_1(x-1)} && \text{by Example S.2 page 386}\\
&= x\sigma(x) + [-2x+2-x+1]\sigma(x-1) + [x-2+2x-4]\sigma(x-2) + [-x+3]\sigma(x-3)\\
&= x\sigma(x) + [-3x+3]\sigma(x-1) + [3x-6]\sigma(x-2) + [-x+3]\sigma(x-3)\\
&= \frac{\mathsf{d}}{\mathsf{d}x}\left\{\begin{array}{l}\frac{1}{2}x^2\sigma(x) + \left[-\frac{3}{2}x^2 + 3x - \frac{1}{2}\right]\sigma(x-1) + \left[\frac{3}{2}x^2 - 6x + 3\right]\sigma(x-2)\\ \qquad + \left[-\frac{1}{2}x^2 + 3x - \frac{5}{2}\right]\sigma(x-3)\end{array}\right\}\\
&= \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{N}_2(x) && \text{by Example S.3 page 386}
\end{aligned}$$

   (c) Proof that n case $\implies$ $n+1$ case:

$$\begin{aligned}
\frac{\mathsf{d}}{\mathsf{d}x}\mathsf{N}_{n+1}(x) &= \frac{\mathsf{d}}{\mathsf{d}x}\int_0^1 \mathsf{N}_n(x-\tau)\,\mathsf{d}\tau && \text{by Lemma S.2 page 383}\\
&= \int_0^1 \frac{\partial}{\partial x}\mathsf{N}_n(x-\tau)\,\mathsf{d}\tau && \text{by \textit{Leibniz Integration Rule} (Theorem ?? page ??)}\\
&= \int_0^1 \left[\mathsf{N}_{n-1}(x-\tau) - \mathsf{N}_{n-1}(x-1-\tau)\right]\mathsf{d}\tau && \text{by left hypothesis}\\
&= \int_0^1 \mathsf{N}_{n-1}(x-\tau)\,\mathsf{d}\tau - \int_0^1 \mathsf{N}_{n-1}(x-1-\tau)\,\mathsf{d}\tau\\
&= \mathsf{N}_n(x) - \mathsf{N}_n(x-1) && \text{by Lemma S.2 page 383}
\end{aligned}$$

✏

Theorem S.3 (*B-spline recursion*). [14] *Let* $\mathsf{N}_n(x)$ *be the* nTH ORDER B-SPLINE *(Definition S.2 page 383)*.

THM
$$\mathsf{N}_n(x) = \frac{x}{n}\mathsf{N}_{n-1}(x) + \frac{n+1-x}{n}\mathsf{N}_{n-1}(x-1) \quad \forall n \in \{1,2,3,\ldots\},\ \forall x \in \mathbb{R}$$

[14] Christensen (2008) page 140 ⟨Theorem 6.1.1⟩, Höllig (2003) page 27 ⟨3.4⟩, Schumaker (2007) page 120 ⟨Theorem 4.15⟩, de Boor (2001) page 90 ⟨B-Spline Property (i)⟩, Chui (1988) page 2 ⟨Theorem 1.1⟩, Wojtaszczyk (1997) page 53 ⟨Theorem 3.7⟩, Cox (1972), de Boor (1972)

COPYRIGHT © 2019 DANIEL J. GREENHOE

✎PROOF:

1. Base case …proof for $n = 1$ case:

$$\begin{aligned}\frac{x}{1}\mathsf{N}_0(x) + \frac{1+1-x}{1}\mathsf{N}_0(x-1) &= \frac{x}{1}\underbrace{[\sigma(x)-\sigma(x-1)]}_{\mathsf{N}_0(x)} + \frac{1+1-x}{1}\underbrace{[\sigma(x-1)-\sigma(x-2)]}_{\mathsf{N}_0(x-1)}\\ &= x\sigma(x) + [-x-x+2]\sigma(x-1) + [x-2]\sigma(x-2)\\ &= \mathsf{N}_1(x) \qquad \text{by Example S.2 page 386}\end{aligned}$$

2. Induction step …proof that n case $\implies$ $n+1$ case:

$$\begin{aligned}&\frac{x}{n+1}\mathsf{N}_n(x) + \frac{n+2-x}{n+1}\mathsf{N}_n(x-1) + c_1\\ &= \int \frac{\mathsf{d}}{\mathsf{d}x}\left\{\frac{x}{n+1}\mathsf{N}_n(x) + \frac{n+2-x}{n+1}\mathsf{N}_n(x-1)\right\}\mathsf{d}x\\ &= \int \underbrace{\frac{1}{n+1}\mathsf{N}_n(x) + \frac{x}{n+1}\frac{\mathsf{d}}{\mathsf{d}x}\mathsf{N}_n(x)}_{\frac{\mathsf{d}}{\mathsf{d}x}\frac{x}{n+1}\mathsf{N}_n(x)} + \underbrace{\frac{-1}{n+1}\mathsf{N}_n(x-1) + \frac{n+2-x}{n}\frac{\mathsf{d}}{\mathsf{d}x}\mathsf{N}_n(x-1)}_{\frac{\mathsf{d}}{\mathsf{d}x}\frac{n+2-x}{n+1}\mathsf{N}_n(x-1)}\,\mathsf{d}x\\ &\qquad \text{by } \textit{product rule}\\ &= \int \frac{1}{n+1}\underbrace{\left[\frac{x}{n}\mathsf{N}_{n-1}(x) + \frac{n+1-x}{n}\mathsf{N}_{n-1}(x-1)\right]}_{\text{by } n \text{ hypothesis}} + \frac{x}{n+1}\underbrace{\left[\mathsf{N}_{n-1}(x) - \mathsf{N}_{n-1}(x-1)\right]}_{\text{by Theorem S.2 page 391}}\\ &\qquad - \underbrace{\left[\frac{x-1}{n^2+n}\mathsf{N}_{n-1}(x-1) + \frac{n-x+2}{n(n+1)}\mathsf{N}_{n-1}(x-2)\right]}_{\text{by induction hypothesis}}\\ &\qquad + \frac{n+2-x}{n+1}\underbrace{\left[\mathsf{N}_{n-1}(x-1) - \mathsf{N}_{n-1}(x-2)\right]}_{\text{by Theorem S.2 page 391}}\mathsf{d}x\\ &= \int \left[\frac{x}{n(n+1)} + \frac{x}{n+1}\right]\mathsf{N}_{n-1}(x) + \left[\frac{n-x+1}{n(n+1)} - \frac{x-1}{n(n+1)} + \frac{n+2-2x}{n+1}\right]\mathsf{N}_{n-1}(x-1)\\ &\qquad + \left[\frac{-n-2+x}{n(n+1)} + \frac{-n-2+x}{n+1}\right]\mathsf{N}_{n-1}(x-2)\,\mathsf{d}x\\ &= \int \left[\frac{x+nx}{n(n+1)}\right]\mathsf{N}_{n-1}(x) + \left[\frac{n+2-2x+n(n+2-2x)}{n(n+1)}\right]\mathsf{N}_{n-1}(x-1)\\ &\qquad + \left[\frac{-n-2+x+n(-n-2+x)}{n(n+1)}\right]\mathsf{N}_{n-1}(x-2)\,\mathsf{d}x\\ &= \int \left[\frac{x}{n}\right]\mathsf{N}_{n-1}(x) + \left[\frac{n+2-2x}{n}\right]\mathsf{N}_{n-1}(x-1) + \left[\frac{-n-2+x}{n}\right]\mathsf{N}_{n-1}(x-2)\,\mathsf{d}x\\ &= \int \underbrace{\left[\frac{x}{n}\right]\mathsf{N}_{n-1}(x) + \left[\frac{n+1-x}{n}\right]\mathsf{N}_{n-1}(x-1)}_{\mathsf{N}_n(x)}\\ &\qquad - \underbrace{\left[\frac{x-1}{n}\right]\mathsf{N}_{n-1}(x-1) - \left[\frac{n+2-x}{n}\right]\mathsf{N}_{n-1}(x-2)}_{\mathsf{N}_{n-1}(x-1)}\,\mathsf{d}x\\ &= \int \mathsf{N}_n(x) - \mathsf{N}_n(x-1)\,\mathsf{d}x \qquad \text{by } n \text{ hypothesis}\\ &= \int \frac{\mathsf{d}}{\mathsf{d}x}\mathsf{N}_{n+1}(x)\,\mathsf{d}x \qquad \text{by Theorem S.2 page 391}\\ &= \mathsf{N}_{n+1}(x) + c_2\end{aligned}$$

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Proof that $c_1 = c_2$: By item (2) (page 394), $\mathsf{N}_n(x) = 0$ for $x < 0$. Therefore, $c_1 = c_2$.

✏

Theorem S.4 (*B-spline general form*). [15] *Let* $\mathsf{N}_n(x)$ *be the* nTH ORDER B-SPLINE *(Definition S.2 page 383). Let* $\mathsf{supp}\,\mathsf{f}$ *be the* SUPPORT *of a function* $\mathsf{f} \in \mathbb{R}^{\mathbb{R}}$.

THM

1.	$\mathsf{N}_n(x)$	$\geq$	0	$\forall n \in \mathbb{W}, \quad \forall x \in \mathbb{R}$	(NON-NEGATIVE)
2.	$\mathsf{supp}\,\mathsf{N}_n(x)$	$=$	$[0 : n+1]$	$\forall n \in \mathbb{W}$	(CLOSED SUPPORT)
3.	$\int_{\mathbb{R}} \mathsf{N}_n(x)\,\mathsf{d}x$	$=$	1	$\forall n \in \mathbb{W}$	(UNIT AREA)
4.	$\mathsf{N}_n\left(\frac{n+1}{2} - x\right)$	$=$	$\mathsf{N}_n\left(\frac{n+1}{2} + x\right)$	$\forall n \in \mathbb{N} \quad \forall x \in \mathbb{R}$	(SYMMETRIC *about* $x = \frac{n+1}{2}$)

✎PROOF:

1. Proof that $\mathsf{N}_n(x) \geq 0$ (proof by induction):

   (a) base case…proof that $\mathsf{N}_0(x) \geq 0$:

$$\begin{aligned} \mathsf{N}_0(x) &\triangleq \mathbb{1}_{[0:1)}(x) && \text{by definition of } \mathsf{N}_0(x) && \text{(Definition S.2 page 383)} \\ &\geq 0 && \text{by definition of } \mathbb{1}_A(x) && \text{(Definition ?? page ??)} \end{aligned}$$

   (b) inductive step—proof that $\{\mathsf{N}_n(x) \geq 0\} \implies \{\mathsf{N}_{n+1}(x) \geq 0\}$:

$$\begin{aligned} \mathsf{N}_{n+1}(x) &= \int_{\tau=0}^{\tau=1} \mathsf{N}_n(x-\tau)\,\mathsf{d}\tau && \text{by Lemma S.2 page 383} \\ &\geq 0 && \text{by induction hypothesis } (\mathsf{N}_n(x) \geq 0) \end{aligned}$$

2. Proof that $\mathsf{supp}\,\mathsf{N}_n(x) = [0 : n+1]$ (proof by induction):

   (a) Base case …proof that $\mathsf{supp}\,\mathsf{N}_0 = [0 : 1]$:

$$\begin{aligned} \mathsf{supp}\,\mathsf{N}_0 &\triangleq \mathsf{supp}\,\mathbb{1}_{[0:1)} && \text{by definition of } \mathsf{N}_0(x) && \text{(Definition S.2 page 383)} \\ &= \{[0 : 1)\}^- && \text{by definition of } \textit{support} \text{ operator} \\ &= [0 : 1] && \text{by definition of } \textit{closure} \text{ operator} \end{aligned}$$

   (b) Induction step …proof that $\{\mathsf{supp}\,\mathsf{N}_n = [0 : n+1]\} \implies \{\mathsf{supp}\,\mathsf{N}_{n+1} = [0 : n+2]\}$:

$$\begin{aligned} \mathsf{supp}\,\mathsf{N}_{n+1}(x) &= \mathsf{supp} \int_{\tau=0}^{\tau=1} \mathsf{N}_n(x-\tau)\,\mathsf{d}\tau && \text{by Lemma S.2 page 383} \\ &= \mathsf{supp} \int_{[0:1]} \mathsf{N}_n(x-\tau)\,\mathsf{d}\tau && \text{by def. of } \textit{Lebesgue integration} \\ &= \left\{x \in \mathbb{R} \,|\, (x-\tau) \in [0 : n+1] \text{ for some } \tau \in [0 : 1]\right\}^- && \text{by induction hypothesis} \\ &= [0 : n+1] \cup [0+1 : n+1+1]^- \\ &= [0 : n+2]^- \\ &= [0 : n+2] && \text{by property of } \textit{closure} \text{ operator} \end{aligned}$$

3. Proof that $\int_{\mathbb{R}} \mathsf{N}_n(x)\,\mathsf{d}x = 1$ (proof by induction):

[15] Christensen (2008) page 140 ⟨Theorem 6.1.1⟩, Höllig (2003) page 27 ⟨3.4⟩, Schumaker (2007) page 120 ⟨Theorem 4.15⟩, de Boor (2001) page 90 ⟨B-Spline Property (i)⟩, Chui (1988) page 2 ⟨Theorem 1.1⟩, Wojtaszczyk (1997) page 53 ⟨Theorem 3.7⟩, Cox (1972), de Boor (1972), Prasad and Iyengar (1997) pages 145–146 ⟨Theorem 6.2⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

(a) Base case …proof that $\int_{\mathbb{R}} \mathsf{N}_0(x) = 1$:

$$\begin{aligned}
\int_{\mathbb{R}} \mathsf{N}_0(x)\,\mathrm{d}x &= \int_{\mathbb{R}} \mathbb{1}_{[0:1)}\,\mathrm{d}x && \text{by definition of } \mathsf{N}_0(x) && \text{(Definition S.2 page 383)}\\
&= \int_{[0:1)} 1\,\mathrm{d}x && \text{by definition of } \mathbb{1}_A(x) && \text{(Definition ?? page ??)}\\
&= \int_{[0:1]} 1\,\mathrm{d}x && \text{by property of } \textit{Lebesgue integration}\\
&= 1
\end{aligned}$$

(b) Induction step …proof that $\left\{\int_{\mathbb{R}} \mathsf{N}_n(x) = 1\right\} \implies \left\{\int_{\mathbb{R}} \mathsf{N}_{n+1} = 1\right\}$:

$$\begin{aligned}
\int_{\mathbb{R}} \mathsf{N}_{n+1}(x)\,\mathrm{d}x &= \int_{\mathbb{R}}\int_0^1 \mathsf{N}_n(x-\tau)\,\mathrm{d}\tau\,\mathrm{d}x && \text{by Lemma S.2 page 383}\\
&= \int_0^1\int_{\mathbb{R}} \mathsf{N}_n(x-\tau)\,\mathrm{d}x\,\mathrm{d}\tau\\
&= \int_0^1\int_{\mathbb{R}} \mathsf{N}_n(u)\,\mathrm{d}u\,\mathrm{d}\tau && \text{where } u \triangleq x-\tau \implies \tau = x-u\\
&= \int_0^1 1\,\mathrm{d}\tau && \text{by induction hypothesis}\\
&= 1
\end{aligned}$$

4. Proof that $\mathsf{N}_n(x)$ is *symmetric* for $n \in \{1, 2, 3, \ldots\}$:

(a) Note that $\mathsf{N}_0(x)$ $(n = 0)$ is *not symmetric* (in particular it fails at $x = 1/2$) because
$\mathsf{N}_0\left(\frac{0+1}{2} - \frac{1}{2}\right) = \mathsf{N}_0(0) = 1 \neq 0 = \mathsf{N}_1(1) = \mathsf{N}_0\left(\frac{0+1}{2} + \frac{1}{2}\right)$

(b) Base case …proof for $n = 1$ case:

$$\begin{aligned}
\mathsf{N}_1\left(\frac{1+1}{2} - x\right) &= \mathsf{N}_1(1-x)\\
&= \left\{\begin{array}{rl} (1-x) & \text{for } 1-x \in [0:1]\\ -(1-x)+2 & \text{for } 1-x \in [1:2]\\ 0 & \text{otherwise}\end{array}\right\} && \text{by Example S.2 page 386}\\
&= \left\{\begin{array}{rl} -x+1 & \text{for } -x \in [-1:0]\\ x+1 & \text{for } -x \in [0:1]\\ 0 & \text{otherwise}\end{array}\right\}\\
&= \left\{\begin{array}{rl} x+1 & \text{for } x \in [-1:0]\\ -x+1 & \text{for } x \in [0:1]\\ 0 & \text{otherwise}\end{array}\right\}\\
&= \left\{\begin{array}{rl} (1+x) & \text{for } 1+x \in [0:1]\\ -(1+x)+2 & \text{for } 1+x \in [1:2]\\ 0 & \text{otherwise}\end{array}\right\}\\
&= \mathsf{N}_1(1+x) && \text{by Example S.2 page 386}\\
&= \mathsf{N}_1\left(\frac{1+1}{2} + x\right)
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe

(c) Induction step …proof that $n-1$ case $\implies$ n case:

$$
\begin{aligned}
&\mathsf{N}_n\left(\frac{n+1}{2}+x\right)\\
&=\frac{\frac{n+1}{2}+x}{n}\mathsf{N}_{n-1}\left(\frac{n+1}{2}+x\right)+\frac{n+1-\left(\frac{n+1}{2}+x\right)}{n}\mathsf{N}_{n-1}\left(\frac{n+1}{2}+x-1\right) && \text{by Theorem S.3 page 392}\\
&=\frac{n+1-\left(\frac{n+1}{2}-x\right)}{n}\mathsf{N}_{n-1}\left(\frac{n}{2}+\left[x+\frac{1}{2}\right]\right)+\frac{\frac{n+1}{2}-x}{n}\mathsf{N}_{n-1}\left(\frac{n}{2}+\left[x-\frac{1}{2}\right]\right)\\
&=\frac{n+1-\left(\frac{n+1}{2}-x\right)}{n}\mathsf{N}_{n-1}\left(\frac{n}{2}-\left[x+\frac{1}{2}\right]\right)+\frac{\frac{n+1}{2}-x}{n}\mathsf{N}_{n-1}\left(\frac{n}{2}-\left[x-\frac{1}{2}\right]\right) && \text{by induction hypothesis}\\
&=\frac{n+1-\left(\frac{n+1}{2}-x\right)}{n}\mathsf{N}_{n-1}\left(\left[\frac{n+1}{2}-x\right]-1\right)+\frac{\frac{n+1}{2}-x}{n}\mathsf{N}_{n-1}\left(\frac{n+1}{2}-x\right)\\
&=\mathsf{N}_n\left(\frac{n+1}{2}-x\right) && \text{by Theorem S.3 page 392}
\end{aligned}
$$

✏

S.3 Projection properties

In the case where $(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$ is to be used as a basis in some subspace of $\boldsymbol{L}^2_{\mathbb{R}}$, one may want to *project* a function $\mathsf{f}(x)$ onto a basis function $\mathsf{N}_n(x-k)$. This is especially true when $(\!(\mathsf{N}_n(x-k))\!)$ is *orthogonal*; but in the case of *B-splines* this is only true when $n=0$ (Theorem S.8 page 406). Nevertheless, projection of a function onto $\mathsf{N}_n(x-k)$, or the projection of $\mathsf{N}_n(x)$ onto another basis fucntion (such as the complex exponential in the case of *Fourier analysis* as in Lemma S.5 page 398), is still useful. Projection in an *inner product space* is typically performed using the *inner product* $\langle \mathsf{f}(x) \mid \mathsf{N}_n(x-k)\rangle$; and in the space $\boldsymbol{L}^2_{\mathbb{R}}$, this inner product is typically defined as an *integral* such that

$$\langle \mathsf{f}(x) \mid \mathsf{N}_n(x-k)\rangle \triangleq \int_{\mathbb{R}} \mathsf{f}(x)\mathsf{N}_n(x-k)\,\mathrm{d}x.$$

As it turns out, there is a way to compute this inner product that only involves the function $\mathsf{f}(x)$ and the order parameter n (next theorem).

Theorem S.5. [16] *Let* f *be a* CONTINUOUS *function in* $\boldsymbol{L}^2_{\mathbb{R}}$ *and* $\mathsf{f}^{(n)}$ *the* nTH DERIVATIVE *of* $\mathsf{f}(x)$.

THM

(1). $$\int_{\mathbb{R}} \mathsf{f}(x)\mathsf{N}_n(x)\,\mathrm{d}x = \int_{[0:1)^{n+1}} \mathsf{f}(x_1+x_2+\cdots+x_{n+1})\,\mathrm{d}x_1\,\mathrm{d}x_2\cdots\mathrm{d}x_{n+1}$$

(2). $$\int_{\mathbb{R}} \mathsf{f}^{(n)}(x)\mathsf{N}_n(x)\,\mathrm{d}x = \sum_{k=0}^{n+1}(-1)^{n-k+1}\binom{n+1}{k}\mathsf{f}(k)$$

✎PROOF:

1. Proof for (1) (proof by induction):

   (a) Base case …proof for $n=0$ case:

$$\int_{\mathbb{R}} \mathsf{f}(x)\mathsf{N}_0(x)\,\mathrm{d}x = \int_{[0:1)} \mathsf{f}(x)\,\mathrm{d}x \qquad \text{by definition of } \mathsf{N}_0(x) \qquad \text{(Definition S.2 page 383)}$$

[16] Chui (1992) page 85 ⟨(4.2.2), (4.2.3)⟩, Christensen (2008) page 140 ⟨Theorem 6.1.1⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

(b) Inductive step—proof that n case $\implies$ $n+1$ case:

$$\begin{aligned}
&\int_{\mathbb{R}} \mathsf{f}(x)\mathsf{N}_{n+1}(x)\,\mathrm{d}x \\
&= \int_{\mathbb{R}} \left[\int_0^1 \mathsf{N}_n(x-\tau)\,\mathrm{d}\tau\right]\mathsf{f}(x)\,\mathrm{d}x && \text{by Lemma S.2 page 383} \\
&= \int_{[0:1)}\int_{\mathbb{R}} \mathsf{N}_n(x-\tau)\mathsf{f}(x)\,\mathrm{d}x\,\mathrm{d}\tau \\
&= \int_{[0:1)}\int_{\mathbb{R}} \mathsf{N}_n(u)\mathsf{f}(u+\tau)\,\mathrm{d}u\,\mathrm{d}\tau && \text{where } u \triangleq x-\tau \implies x = u+\tau \\
&= \int_{[0:1)}\int_{[0:1)^{n+1}} \mathsf{f}(u_1+u_2+\cdots+u_{n+1}+\tau)\,\mathrm{d}u_1\ \mathrm{d}u_2\ \cdots\ \mathrm{d}u_{n+1}\ \mathrm{d}\tau && \text{by induction hypothesis} \\
&= \int_{[0:1)^{n+2}} \mathsf{f}(u_1+u_2+\cdots+u_{n+1}+u_{n+2})\,\mathrm{d}u_1\ \mathrm{d}u_2\ \cdots\ \mathrm{d}u_{n+2}\ \mathrm{d}\tau \\
&= \int_{[0:1)^{n+2}} \mathsf{f}(x_1+x_2+\cdots+x_{n+1}+x_{n+2})\,\mathrm{d}x_1\ \mathrm{d}x_2\ \cdots\ \mathrm{d}x_{n+2} && \text{by change of variables } u_k \to x_k
\end{aligned}$$

2. Proof for (2):

$$\begin{aligned}
\int_{\mathbb{R}} \mathsf{f}^{(n)}(x)\mathsf{N}_n(x)\,\mathrm{d}x &= \int_{[0:1)^{n+1}} \mathsf{f}^{(n)}\left(\sum_{k=1}^{n+1} x_k\right)\mathrm{d}x_1\ \mathrm{d}x_2\cdots\ \mathrm{d}x_{n+1} && \text{by (1)} \\
&= \sum_{k=0}^{n+1}(-1)^{n-k+1}\binom{n+1}{k}\mathsf{f}(k) && \text{by Theorem ?? page ??}
\end{aligned}$$

✏

Figure S.2: Zero locations for B-cardinal spline $N_n(x)$ scaling coefficients

S.4 Fourier analysis

Simply put, no matter what new and fancy basis sequences are discovered, the *Fourier transform* never goes out of style. This is largely because the *kernel* of the Fourier transform—the *complex exponential* function—has two properties that makes it extremely special:

- The complex exponential is an eigenvalue of any *linear time invariant* (*LTI*) operator (Theorem ?? page ??).
- The complex exponential generates a *continuous point spectrum* for the *differential operator*.

Thus, we might expect the projection of the *B-spline* function $\mathsf{N}_n(x)$ onto the complex exponential (essentially the *Fourier transform* of $\mathsf{N}_n(x)$,…next lemma) to be useful. Such a hunch would be confirmed because it is useful for proving that

- the sequence $(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$ is a *Riesz basis* (Lemma S.6 page 401, Theorem S.8 page 406) and
- the sequence $(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$ is a *multiresolution analysis* (Theorem S.10 page 409).

Lemma S.5. [17] *Let* $\tilde{\mathbf{F}}$ *be the* FOURIER TRANSFORM *operator* (Definition P.2 page 349).

LEM

$$\tilde{\mathbf{F}}\mathsf{N}_n(\omega) = \frac{1}{\sqrt{2\pi}}\left(\frac{1-e^{-i\omega}}{i\omega}\right)^{n+1} = \frac{1}{\sqrt{2\pi}}e^{-i\frac{(n+1)\omega}{2}}\left(\frac{\sin(\omega/2)}{\omega/2}\right)^{n+1} \triangleq \frac{1}{\sqrt{2\pi}}e^{-i\frac{(n+1)\omega}{2}}\left(\operatorname{sinc}\frac{\omega}{2}\right)^{n+1}$$

✎PROOF:

1. Proof using Theorem S.5 page 396:

$$\begin{aligned}
\tilde{\mathbf{F}}\mathsf{N}_n(\omega) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathsf{N}_n(x)e^{-i\omega x}\,\mathrm{d}x && \text{by definition of } \tilde{\mathbf{F}} \quad \text{(Definition P.2 page 349)}\\
&= \frac{1}{\sqrt{2\pi}}\int_{[0:1)^{n+1}} e^{-i\omega(x_1+x_2+\cdots+x_{n+1})}\,\mathrm{d}x_1\,\mathrm{d}x_2\cdots,\,\mathrm{d}x_{n+1} && \text{by Theorem S.5}\\
&= \frac{1}{\sqrt{2\pi}}\prod_{k=1}^{n+1}\left(\int_{[0:1)} e^{-i\omega x_k}\,\mathrm{d}x_k\right) && \text{because } e^{x+y}=e^xe^y\\
&= \frac{1}{\sqrt{2\pi}}\left(\int_0^1 e^{-i\omega x}\,\mathrm{d}x\right)^{n+1} = \frac{1}{\sqrt{2\pi}}\left(\frac{e^{-i\omega x}}{-i\omega}\bigg|_0^1\right)^{n+1}\\
&= \frac{1}{\sqrt{2\pi}}\left(\frac{1-e^{-i\omega}}{i\omega}\right)^{n+1} = \frac{1}{\sqrt{2\pi}}\left[e^{-i\frac{\omega}{2}}\left(\frac{e^{i\frac{\omega}{2}}-e^{-i\frac{\omega}{2}}}{i\omega}\right)\right]^{n+1}\\
&= \frac{1}{\sqrt{2\pi}}\left[e^{-i\frac{\omega}{2}}\left(\frac{2i\sin\left(\frac{\omega}{2}\right)}{\frac{2i\omega}{2}}\right)\right]^{n+1} && \text{by } \textit{Euler formulas} \quad \text{(Corollary ?? page ??)}\\
&= \frac{1}{\sqrt{2\pi}}e^{-i\frac{(n+1)\omega}{2}}\left(\frac{\sin(\omega/2)}{\omega/2}\right)^{n+1}
\end{aligned}$$

2. Proof using *rectangular pulse* example (Example P.1 page 356) and *Convolution Theorem* (Theorem R.2 page 372):

$$\begin{aligned}
\tilde{\mathbf{F}}\mathsf{N}_n(\omega) &= \left[\sqrt{2\pi}\right]^n\left[\tilde{\mathbf{F}}\mathsf{N}_0\right]^{n+1} && \text{by } \textit{Convolution Theorem} \quad \text{(Theorem R.2 page 372)}\\
&= \frac{1}{\sqrt{2\pi}}\left[\sqrt{2\pi}\frac{c(b-a)}{\sqrt{2\pi}}e^{-i\left(\frac{a+b}{2}\omega\right)}\left(\frac{\sin\left(\frac{b-a}{2}\omega\right)}{\left(\frac{b-a}{2}\omega\right)}\right)\right]^{n+1} && \text{by } \textit{rectangular pulse} \text{ example with } a=0,\ b=c=1 \quad \text{(Example P.1 page 356)}\\
&= \frac{1}{\sqrt{2\pi}}\left[\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}e^{-i\left(\frac{\omega}{2}\right)}\left(\frac{\sin\left(\frac{\omega}{2}\right)}{(\omega/2)}\right)\right]^{n+1}\\
&= \frac{1}{\sqrt{2\pi}}e^{-i\left(\frac{(n+1)\omega}{2}\right)}\left(\frac{\sin\left(\frac{\omega}{2}\right)}{(\omega/2)}\right)^{n+1}
\end{aligned}$$

✏

Example S.10. The *Fourier transform magnitude* $\left|\left[\tilde{\mathbf{F}}\mathsf{N}_0\right](\omega)\right|$ of the *0 order B-spline* $\mathsf{N}_0(x)$ is illustrated to the right.

[17] Christensen (2008) page 142 ⟨Corollary 6.1.2⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

Example S.11. The *Fourier transform magnitude* $\left|\left[\tilde{\mathbf{F}}\mathsf{N}_1\right](\omega)\right|$ of the *1st order B-spline* $\mathsf{N}_1(x)$ is illustrated to the right.

Example S.12. The *Fourier transform magnitude* $\left|\left[\tilde{\mathbf{F}}\mathsf{N}_2\right](\omega)\right|$ of the *2nd order B-spline* $\mathsf{N}_2(x)$ is illustrated to the right.

S.5 Basis properties

S.5.1 Uniqueness properties

Coefficients of a *basis sequence* are not always *unique*. Take for example a very trivial sequence (α_1, α_2) in which the coefficients are summed. If $\mathsf{f}(x) \triangleq \alpha_1 + \alpha_2$ and $\mathsf{g}(x) \triangleq \beta_1 + \beta_2$,

$$\begin{aligned} &\text{then} \quad \{(\alpha_1,\alpha_2) = (\beta_1,\beta_2)\} \quad &&\Longrightarrow \quad \mathsf{f}(x) = \mathsf{g}(x) \\ &\text{but} \quad \mathsf{f}(x) = \mathsf{g}(x) &&\not\Longrightarrow \quad \{(\alpha_1,\alpha_2) = (\beta_1,\beta_2)\}, \end{aligned}$$

because for example if $(\alpha_1, \alpha_2) = (1, 2)$ and $(\beta_1, \beta_2) = (-6, 9)$, then $\mathsf{f}(x) = \mathsf{g}(x)$, but $(\alpha_1, \alpha_2) \neq (\beta_1, \beta_2)$. This example demonstrates that the "if and only if" condition $\Longleftrightarrow$ does not hold and coefficients are not unique in all *basis sequences*. But arguably a minimal requirement for any practical basis sequence is that the coefficients are *unique* (the "if and only if" condition $\Longleftrightarrow$ holds). And indeed, in a *B-spline* basis sequence $(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$, the coefficients $(\!(\alpha_k)\!)_{k\in\mathbb{Z}}$ are *unique*, as demonstrated by Theorem S.6 (next).

Theorem S.6. [18] *Let* $\mathsf{N}_n(x)$ *be the* nTH-ORDER B-SPLINE *(Definition S.2 page 383). Let*

$$\mathsf{f}(x) \triangleq \sum_{k\in\mathbb{Z}} \alpha_k \mathsf{N}_n(x-k) \quad \textit{and} \quad \mathsf{g}(x) \triangleq \sum_{k\in\mathbb{Z}} \beta_k \mathsf{N}_n(x-k).$$

THM

$$\underbrace{\left\{\ \mathsf{f}(x) = \mathsf{g}(x) \quad \forall x \in \mathbb{R}\ \right\} \quad \Longleftrightarrow \quad \left\{(\!(\alpha_k)\!)_{k\in\mathbb{Z}} = (\!(\beta_k)\!)_{k\in\mathbb{Z}}\right\}}_{\textit{coefficients are UNIQUE}}$$

✎PROOF:

1. Proof that $\Longleftarrow$ condition holds:

$$\begin{aligned} \mathsf{f}(x) &\triangleq \sum_{k\in\mathbb{Z}} \alpha_k \mathsf{N}_n(x-k) && \text{by definition of } \mathsf{f}(x) \\ &= \sum_{k\in\mathbb{Z}} \beta_k \mathsf{N}_n(x-k) && \text{by right hypothesis} \\ &\triangleq \mathsf{g}(x) && \text{by definition of } \mathsf{g}(x) \end{aligned}$$

2. Proof that $\Longrightarrow$ condition holds (proof by contradiction):

   (a) Suppose it does *not* hold.

[18] Wojtaszczyk (1997) page 55 ⟨Theorem 3.11⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

(b) Then there exists sequences $(\!(\alpha_k)\!)_{k\in\mathbb{Z}}$ and $(\!(\beta_k)\!)_{k\in\mathbb{Z}}$ such that
$$(\!(\alpha_k)\!) - (\!(\beta_k)\!) \triangleq (\!(alpha_k - \beta_k)\!) \neq (\!(0,0,0,\cdots)\!)$$
but also such that $\mathsf{f}(x) - \mathsf{g}(x) = 0\ \forall x \in \mathbb{R}$.

(c) If this were possible, then

$$\begin{aligned}
0 &= \mathsf{f}(x) - \mathsf{g}(x) \\
&= \sum_{m\in\mathbb{Z}} \alpha_m \mathsf{N}_n(x-m) - \sum_{m\in\mathbb{Z}} \beta_m \mathsf{N}_n(x-m) \\
&= \sum_{m\in\mathbb{Z}} \big(\alpha_m - \beta_m\big)\mathsf{N}_n(x-m) \\
&= \sum_{m=0}^{m=n} \big(\alpha_m - \beta_m\big)\frac{1}{n!}\left[\sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k}(x-k)^n\sigma(x-k)\right] && \text{by Theorem S.1 page 387}
\end{aligned}$$

(d) But this is *impossible* because $\mathsf{N}(x)$ is *non-negative* (Theorem S.4 page 394).

(e) Therefore, there is a contradiction, and the $\Longrightarrow$ condition *does* hold.

✏

S.5.2 Partition of unity properties

In the case in which a sequence of *B-splines* $(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$ is to be used as a *basis* for some subspace of $\boldsymbol{L}^2_{\mathbb{R}}$, arguably one of the most important properties for the sequence to have is the *partition of unity* property such that $\sum_{k\in\mathbb{Z}} \mathsf{N}_n(x-k) = 1$. This allows for convenient representation of the most basic functions, such as constants.[19] As it turns out, B-splines *do* have this property (next theorem).

Theorem S.7 (*B-spline partition of unity*). [20] *Let* $\mathsf{N}_n(x)$ *be the* nTH ORDER B-SPLINE *(Definition S.2 page 383)*.

THM $$\sum_{k\in\mathbb{Z}} \mathsf{N}_n(x-k) = 1 \quad \forall n \in \mathbb{W} \qquad \text{(PARTITION OF UNITY)}$$

✎PROOF:

1. lemma: $\sum_{k\in\mathbb{Z}} \mathsf{N}_0(x-k) = 1$. Proof:

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} \mathsf{N}_0(x-k) &= \sum_{k\in\mathbb{Z}} \mathbb{1}_{[0:1)}(x-k) && \text{by definition of } \mathsf{N}_0(x) && \text{(Definition S.2 page 383)} \\
&= 1 && \text{by definition of } \mathbb{1}_A(x) && \text{(Definition ?? page ??)}
\end{aligned}$$

2. Proof for this theorem follows from the $n = 0$ case ((1) lemma page 400), the definition of $\mathsf{N}_n(x)$ (Definition S.2 page 383), and Corollary ?? (page ??).

3. Alternatively, this theorem can be proved by *induction*:

   (a) Base case ($n = 0$ case): by (1) lemma.

[19] Jawerth and Sweldens (1994) page 8

[20] Christensen (2008) page 140 ⟨Theorem 6.1.1⟩, Höllig (2003) page 27 ⟨3.4⟩, Schumaker (2007) page 120 ⟨Theorem 4.15⟩, de Boor (2001) page 90 ⟨B-Spline Property (i)⟩, Chui (1988) page 2 ⟨Theorem 1.1⟩, Wojtaszczyk (1997) page 53 ⟨Theorem 3.7⟩, Cox (1972), de Boor (1972)

COPYRIGHT © 2019 DANIEL J. GREENHOE

Figure S.3: *auto-power spectrum* $\tilde{S}_n(\omega)$ plots of *B-splines* $\mathsf{N}_n(x)$ (Lemma S.6 page 401) For C and LaTeX source code to generate such a plot, see Section **??** (page **??**).

(b) Inductive step—proof that $\sum_{k\in\mathbb{Z}} \mathsf{N}_n(x-k) = 1 \implies \sum_{k\in\mathbb{Z}} \mathsf{N}_{n+1}(x-k) = 1$:

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} \mathsf{N}_{n+1}(x-k) &= \sum_{k\in\mathbb{Z}} \int_{\tau=0}^{\tau=1} \mathsf{N}_n(x-k-\tau)\,\mathrm{d}\tau && \text{by Lemma S.2 page 383} \\
&= \sum_{k\in\mathbb{Z}} \int_{x-u=0}^{x-u=1} \mathsf{N}_n(u-k)(-1)\,\mathrm{d}u && \text{where } u \triangleq x-\tau \implies \tau = x-u \\
&= \sum_{k\in\mathbb{Z}} \int_{u=x-1}^{u=x} \mathsf{N}_n(u-k)\,\mathrm{d}u \\
&= \int_{u=x-1}^{u=x} \left(\sum_{k\in\mathbb{Z}} \mathsf{N}_n(u-k)\right) \mathrm{d}u \\
&= \int_{u=x-1}^{u=x} 1\,\mathrm{d}u && \text{by induction hypothesis} \\
&= 1
\end{aligned}$$

✏

Example S.13. The *partition of unity* property for the 0 *order B-spline* $\mathsf{N}_0(x)$ (Example S.1 page 385) is illustrated to the right.

Example S.14. The *partition of unity* property for the 1*st order B-spline* $\mathsf{N}_1(x)$ (Example S.2 page 386) is illustrated to the right.

Example S.15. The *partition of unity* property for the 2*nd order B-spline* $\mathsf{N}_2(x)$ (Example S.3 page 386) is illustrated to the right.

S.5.3 Riesz basis properties

Lemma S.6. *Let* $\mathsf{N}_n(x)$ *be the* nTH ORDER B-SPLINE *(Definition S.2 page 383).*
Let $\tilde{S}_n(\omega) \triangleq 2\pi \sum_{k\in\mathbb{Z}} \left|\tilde{\mathsf{F}}\mathsf{N}_n(\omega - 2\pi k)\right|^2$ *be the* AUTO-POWER SPECTRUM *(Definition **??** page **??**) of* $\mathsf{N}_n(x)$.

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

LEM				
	(1). $0 < \tilde{S}_n(\omega) \le 1 \quad \forall \omega \in \mathbb{R} \quad , \quad \forall n \in \mathbb{W}$		(3). $\tilde{S}_n(0) = 1 \quad \forall n \in \mathbb{W}$	
	(2). $\tilde{S}_n(\omega) = 1 \quad \forall \omega \in \mathbb{R} \quad , \quad \text{for } n = 0$		(4). $\tilde{S}_n(\pi) \le \frac{1}{3} \quad \forall n \in \mathbb{W}\setminus\{0\}$	

(*Note: see illustration in Figure S.3 page 401.*)

✎PROOF:

1. lemma: $\tilde{S}_n(\omega) = \sum_{k\in\mathbb{Z}} \left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi - \frac{\omega}{2}}\right]^{2(n+1)}$. Proof:

$$\begin{aligned}
\tilde{S}_n(\omega) &\triangleq 2\pi \sum_{k\in\mathbb{Z}} \left|\tilde{\mathbf{F}}\mathsf{N}_n(\omega - 2\pi k)\right|^2 && \text{by Definition ?? page ??}\\
&= 2\pi \sum_{k\in\mathbb{Z}} \left|\frac{1}{\sqrt{2\pi}} e^{-i\frac{(n+1)(\omega-2\pi k)}{2}} \left(\frac{\sin\left(\frac{\omega-2\pi k}{2}\right)}{\frac{\omega-2\pi k}{2}}\right)^{n+1}\right|^2 && \text{by Lemma S.5 page 398}\\
&= \sum_{k\in\mathbb{Z}} \left[\frac{\sin\left(\frac{\omega-2\pi k}{2}\right)}{\frac{\omega-2\pi k}{2}}\right]^{2(n+1)}\\
&= \sum_{k\in\mathbb{Z}} \left[\frac{\sin\left(\frac{\omega}{2} - k\pi\right)}{k\pi - \frac{\omega}{2}}\right]^{2(n+1)}\\
&= \sum_{k\in\mathbb{Z}} \left[\frac{(-1)^k \sin\left(\frac{\omega}{2}\right)}{k\pi - \frac{\omega}{2}}\right]^{2(n+1)}\\
&= \sum_{k\in\mathbb{Z}} \left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi - \frac{\omega}{2}}\right]^{2(n+1)}
\end{aligned}$$

2. lemma (one sided series form):

$$\begin{aligned}
\tilde{S}_n(\omega) &= \sum_{k\in\mathbb{Z}} \left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi - \frac{\omega}{2}}\right]^{2(n+1)} && \text{by (1) lemma}\\
&= \left[\frac{\sin(\omega/2)}{(\omega/2)}\right]^{2(n+1)} + \left[\frac{\sin(\omega/2)}{(\pi/2)}\right]^{2(n+1)} \left(\sum_{k=1}^{\infty} \left[\frac{1}{2k - \frac{\omega}{\pi}}\right]^{2(n+1)} + \sum_{k=1}^{\infty} \left[\frac{1}{2k + \frac{\omega}{\pi}}\right]^{2(n+1)}\right)
\end{aligned}$$

3. lemma: $\tilde{S}_n(\omega)$ is *continuous* for all $\omega \in \mathbb{R}$.
   Proof: $\sin(\omega/2)$ and $\omega/2$ are *continuous*, so $\tilde{S}_n(\omega)$ is *continuous* as well.

4. lemma: $\tilde{S}_n(\omega)$ is *periodic* with period 2π (and so we only need to examine $\tilde{S}_n(\omega)$ for $\omega \in [0 : 2\pi)$). Proof of *periodicity*: This follows directly from Proposition **??** (page **??**).

5. lemma: $\tilde{S}_n(-\omega) = \tilde{S}_n(\omega)$ (*symmetric* about 0) and $\tilde{S}_n(\pi - \omega) = \tilde{S}_n(\pi + \omega)$ (*symmetric* about π). Proof: This follows directly from Proposition **??** (page **??**).

COPYRIGHT © 2019 DANIEL J. GREENHOE

6. Proof that $\tilde{S}_n(0) = 1$:

$$\begin{aligned}
\tilde{S}_n(0) &= \lim_{\omega\to 0} \tilde{S}_n(\omega) && \text{by (3) lemma}\\
&= \lim_{\omega\to 0}\left[\left[\frac{\sin(\omega/2)}{(\omega/2)}\right]^{2(n+1)} + \left[\frac{\sin(\omega/2)}{(\pi/2)}\right]^{2(n+1)}\left(\sum_{k=1}^{\infty}\left[\frac{1}{2k-\frac{\omega}{\pi}}\right]^{2(n+1)} + \sum_{k=1}^{\infty}\left[\frac{1}{2k+\frac{\omega}{\pi}}\right]^{2(n+1)}\right)\right] && \text{by (2) lemma}\\
&= \lim_{\omega\to 0}\left[\frac{\cos\left(\frac{\omega}{2}\right)}{-\frac{1}{2}}\right]^{2(n+1)} + 0 && \text{by } \textit{l'Hôpital's rule}\\
&= (-1)^{2(n+1)} = 1
\end{aligned}$$

7. Proof that $\tilde{S}_n(\pi)$ converges to some value > 0:

(a) Proof that $\tilde{S}_n(\pi) > 0$:

$$\begin{aligned}
\tilde{S}_n(\pi) &= \left[\frac{\sin(\pi/2)}{(\pi/2)}\right]^{2(n+1)} + \left[\frac{\sin(\pi/2)}{(\pi/2)}\right]^{2(n+1)}\left(\sum_{k=1}^{\infty}\left[\frac{1}{2k-\frac{\pi}{\pi}}\right]^{2(n+1)} + \sum_{k=1}^{\infty}\left[\frac{1}{2k+\frac{\pi}{\pi}}\right]^{2(n+1)}\right) && \text{by (2) lemma}\\
&= \left(\frac{2}{\pi}\right)^{2(n+1)}\left[1 + \left(\frac{1}{1}\right)^{2(n+1)} + \left(\frac{1}{3}\right)^{2(n+1)} + \left(\frac{1}{3}\right)^{2(n+1)} + \left(\frac{1}{5}\right)^{2(n+1)} + \left(\frac{1}{5}\right)^{2(n+1)} + \cdots\right]\\
&= 2\left(\frac{2}{\pi}\right)^{2(n+1)} \underbrace{\sum_{k=1}^{\infty}\left[\frac{1}{2k-1}\right]^{2(n+1)}}_{\textit{Dirichlet Lambda function } \lambda(2n+2)}\\
&> 0 \qquad \text{because } x^2 > 0 \text{ for all } x \in \mathbb{R}\setminus\{0\}
\end{aligned}$$

(b) Proof that $\tilde{S}_n(\pi)$ *converges*:

$$\begin{aligned}
\tilde{S}_n(\pi) &= 2\left(\frac{2}{\pi}\right)^{2(n+1)}\sum_{k=1}^{\infty}\left(\frac{1}{2k-1}\right)^{2(n+1)} && \text{by item (7a)}\\
&\le 2\left(\frac{2}{\pi}\right)^{2(n+1)}\sum_{k=1}^{\infty}\left(\frac{1}{k}\right)^{2(n+1)}\\
&\le 2\left(\frac{2}{\pi}\right)^{2(n+1)}\sum_{k=1}^{\infty}\left(\frac{1}{k}\right)^{2}\\
&\implies \textit{convergence} && \text{by } \textit{comparison test}
\end{aligned}$$

(c) Tighter bounds for $\tilde{S}_n(\pi)$ for certain values of $n \in \{0,1,2,3,4\}$:

$$\begin{aligned}
\tilde{S}_n(\pi) &= 2\left(\frac{2}{\pi}\right)^{2(n+1)}\sum_{k=1}^{\infty}\left(\frac{1}{2k-1}\right)^{2(n+1)} && \text{by item (7a)}\\
&= 2\left(\frac{2}{\pi}\right)^{2(n+1)} U_{2(n+1)} && \text{by Jolley (1961), pages 56–57 } \langle(307)\rangle\\
&= 2\left(\frac{2}{\pi}\right)^{2(n+1)}\left[\frac{\pi^{2(n+1)}\alpha_{n+1}}{(4)[(2n+2)!]}\right] && \text{by Jolley (1961), pages 56–57 } \langle(307)\rangle\\
&= \frac{2^{2n+1}\alpha_{n+1}}{(2n+2)!}
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

$$= \left\{ \begin{array}{lll} \frac{2^1(1)}{2!} & \text{for } n=0 & (\alpha_1 = 1) \\ \frac{2^5(1)}{4!} & \text{for } n=1 & (\alpha_2 = 1) \\ \frac{2^5(3)}{6!} & \text{for } n=2 & (\alpha_3 = 3) \\ \frac{2^7(17)}{8!} & \text{for } n=3 & (\alpha_4 = 17) \\ \frac{2^9(155)}{10!} & \text{for } n=4 & (\alpha_5 = 155) \end{array} \right\} \qquad \text{by Jolley (1961), page 234 } \langle(1130)\rangle$$

$$= \left\{ \begin{array}{ll} 1 & \text{for } n=0 \\ \frac{1}{3} & \text{for } n=1 \\ \frac{2}{15} & \text{for } n=2 \\ \frac{17}{315} & \text{for } n=3 \\ \frac{62}{2835} & \text{for } n=4 \end{array} \right\} = \left\{ \begin{array}{ll} 1 & \text{for } n=0 \\ 0.3333333333333\cdots & \text{for } n=1 \\ 0.1333333333333\cdots & \text{for } n=2 \\ 0.0539682539682\cdots & \text{for } n=3 \\ 0.0218694885361\cdots & \text{for } n=4 \end{array} \right\}$$

(d) Being important for the $n = 0$ case, note that[21]

$$\underbrace{\sum_{k=1}^{\infty}\left(\frac{1}{2k-1}\right)^2}_{\textit{Dirichlet Lambda function } \lambda(2)} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots = \frac{\pi^2}{8}$$

(e) Proof that $\tilde{S}_n(\pi) \le \frac{1}{3}$: because $\tilde{S}_n(\pi) = \frac{1}{3}$ for $n = 1$ (item (7c) page 403) and because $\tilde{S}_n(\pi)$ is decreasing for increasing n.

8. lemma: $\tilde{S}_n(\omega)$ *converges* to some value $> 0\ \forall\omega \in \mathbb{R}$. Proof:

(a) For $\omega = 0$, $\tilde{S}_n(\omega) = 1$ by item (6).

(b) Proof that $\tilde{S}_n(\omega) > 0$ for $\omega \in (0 : 2\pi)$:

$$\tilde{S}_n(\omega) = \left[\frac{\sin(\omega/2)}{(\omega/2)}\right]^{2(n+1)} + \left[\frac{\sin(\omega/2)}{(\pi/2)}\right]^{2(n+1)}\left(\sum_{k=1}^{\infty}\left[\frac{1}{2k-\frac{\omega}{\pi}}\right]^{2(n+1)} + \sum_{k=1}^{\infty}\left[\frac{1}{2k+\frac{\omega}{\pi}}\right]^{2(n+1)}\right) \qquad \text{by (2) lemma}$$

$$> 0$$

(c) Proof that $\tilde{S}_n(\omega)$ *converges*:

i. lemma: $\displaystyle\sum_{k=1}^{\infty}\left[\frac{1}{2k \pm \frac{\omega}{\pi}}\right]^{2(n+1)}$ *converges*. Proof:

$$\begin{aligned}
\lim_{b\to\infty}\int_1^b \left[\frac{1}{2y \pm \frac{\omega}{\pi}}\right]^{2(n+1)} \mathrm{d}y &= \lim_{b\to\infty}\int_1^b \left[2y \pm \frac{\omega}{\pi}\right]^{-2n-2} \mathrm{d}y \\
&= \lim_{b\to\infty} \left.\frac{\left[2y \pm \frac{\omega}{\pi}\right]^{-2n-1}}{2(-2n-1)}\right|_1^b \\
&= \lim_{b\to\infty}\left(\frac{-1}{2(2n+1)}\right)\left[\frac{1}{\left[2b \pm \frac{\omega}{\pi}\right]^{2n+1}} - \frac{1}{\left[2 \pm \frac{\omega}{\pi}\right]^{2n+1}}\right] \\
&= 0 + \frac{1}{2(2n+1)\left[2 \pm \frac{\omega}{\pi}\right]^{2n+1}} \\
&< \infty \qquad \forall\omega \in [0 : 2\pi) \\
\Longrightarrow \sum_{k=1}^{\infty}\left[\frac{1}{2k \pm \frac{\omega}{\pi}}\right]^{2(n+1)} &\ \textit{converges} \qquad \text{by } \textit{integral test}
\end{aligned}$$

[21] Nahin (2011) page 153, Bailey et al. (2013) page 334 ⟨Catalan's Constant⟩, Bailey et al. (2011) ⟨(15)⟩, Wells (1987) page 36 ⟨Dictionary entry for π: pages 31–37⟩, Heinbockel (2010) page 94 ⟨(2.27) *Dirichlet Lambda function*⟩

Copyright © 2019 Daniel J. Greenhoe

ii. completion of proof using (8(c)i) lemma …

$$\begin{aligned}
&\tilde{S}_n(\omega)\\
&= \left[\frac{\sin(\omega/2)}{(\omega/2)}\right]^{2(n+1)} + \left[\frac{\sin(\omega/2)}{(\pi/2)}\right]^{2(n+1)}\left(\sum_{k=1}^{\infty}\left[\frac{1}{2k-\frac{\omega}{\pi}}\right]^{2(n+1)} + \sum_{k=1}^{\infty}\left[\frac{1}{2k+\frac{\omega}{\pi}}\right]^{2(n+1)}\right) && \text{by (2) lemma}\\
&\Longrightarrow \tilde{S}_n(\omega) \textit{ converges } \forall\omega \in (0:2\pi) && \text{by (8(c)i) lemma}
\end{aligned}$$

9. lemma (an expression for $\tilde{S}'_n(\omega)$):

$$\begin{aligned}
\tilde{S}'_n(\omega) &\triangleq \frac{\mathrm{d}}{\mathrm{d}\omega}\tilde{S}_n(\omega)\\
&= \frac{\mathrm{d}}{\mathrm{d}\omega}\sum_{k\in\mathbb{Z}}\left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi-\frac{\omega}{2}}\right]^{2(n+1)} && \text{by (1) lemma page 402}\\
&= \sum_{k\in\mathbb{Z}}\frac{\mathrm{d}}{\mathrm{d}\omega}\left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi-\frac{\omega}{2}}\right]^{2(n+1)} && \text{by } \textit{linearity} \text{ of } \frac{\mathrm{d}}{\mathrm{d}\omega} \text{ operator}\\
&= \sum_{k\in\mathbb{Z}}2(n+1)\left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi-\frac{\omega}{2}}\right]^{2n+1}\frac{\mathrm{d}}{\mathrm{d}\omega}\left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi-\frac{\omega}{2}}\right] && \text{by } \textit{power rule}\\
&= 2(n+1)\sum_{k\in\mathbb{Z}}\left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi-\frac{\omega}{2}}\right]^{2n+1}\left[\frac{\frac{1}{2}\cos\left(\frac{\omega}{2}\right)\left(k\pi-\frac{\omega}{2}\right)-\sin\left(\frac{\omega}{2}\right)\left(-\frac{1}{2}\right)}{\left(k\pi-\frac{\omega}{2}\right)^2}\right] && \text{by } \textit{quotient rule}\\
&= (n+1)\sum_{k\in\mathbb{Z}}\left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi-\frac{\omega}{2}}\right]^{2n+1}\left[\frac{\cos\left(\frac{\omega}{2}\right)\left(k\pi-\frac{\omega}{2}\right)+\sin\left(\frac{\omega}{2}\right)}{\left(k\pi-\frac{\omega}{2}\right)^2}\right]
\end{aligned}$$

10. lemma: $\tilde{S}'_n(0) = \tilde{S}'_n(\pi) = 0$. Proof: This follows from Proposition **??** (page **??**). Here is alternate proof:

$$\begin{aligned}
\tilde{S}'_n(0) &= \lim_{\omega\to 0}\tilde{S}'_n(\omega)\\
&= \lim_{\omega\to 0}(n+1)\sum_{k\in\mathbb{Z}}\left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi-\frac{\omega}{2}}\right]^{2n+1}\left[\frac{\cos\left(\frac{\omega}{2}\right)\left(k\pi-\frac{\omega}{2}\right)+\sin\left(\frac{\omega}{2}\right)}{\left(k\pi-\frac{\omega}{2}\right)^2}\right] && \text{by (9) lemma}\\
&= \lim_{\omega\to 0}(n+1)\left[\frac{\sin\left(\frac{\omega}{2}\right)}{-\frac{\omega}{2}}\right]^{2n+1}\left[\frac{\cos\left(\frac{\omega}{2}\right)\left(-\frac{\omega}{2}\right)+\sin\left(\frac{\omega}{2}\right)}{\left(-\frac{\omega}{2}\right)^2}\right]\\
&= (n+1)\lim_{\omega\to 0}\left[\frac{\sin\left(\frac{\omega}{2}\right)}{-\frac{\omega}{2}}\right]^{2n+1}\lim_{\omega\to 0}\left[\frac{\cos\left(\frac{\omega}{2}\right)\left(-\frac{\omega}{2}\right)+\sin\left(\frac{\omega}{2}\right)}{\left(-\frac{\omega}{2}\right)^2}\right]\\
&= (n+1)[-1]^{2n+1}\lim_{\omega\to 0}\left[\frac{-\frac{1}{2}\sin\left(\frac{\omega}{2}\right)\left(-\frac{\omega}{2}\right)+\cos\left(\frac{\omega}{2}\right)\left(-\frac{1}{2}\right)+\cos\left(\frac{\omega}{2}\right)\left(\frac{1}{2}\right)}{-\frac{2}{2}\left(-\frac{\omega}{2}\right)}\right] && \text{by } \textit{l'Hôpital's rule}\\
&= (1)(0)\\
&= 0
\end{aligned}$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

$$\begin{aligned}
\tilde{S}_n'(\pi) &= (n+1)\sum_{k\in\mathbb{Z}}\left[\frac{\sin\left(\frac{\pi}{2}\right)}{k\pi-\frac{\pi}{2}}\right]^{2n+1}\left[\frac{\cos\left(\frac{\pi}{2}\right)\left(k\pi-\frac{\pi}{2}\right)+\sin\left(\frac{\pi}{2}\right)}{\left(k\pi-\frac{\pi}{2}\right)^2}\right] && \text{by (9) lemma}\\
&= (n+1)\sum_{k\in\mathbb{Z}}\left[\frac{1}{k\pi-\frac{\pi}{2}}\right]^{2n+1}\left[\frac{0\left(k\pi-\frac{\pi}{2}\right)+1}{\left(k\pi-\frac{\pi}{2}\right)^2}\right]\\
&= (n+1)\left(\frac{2}{\pi}\right)^{2n+3}\sum_{k\in\mathbb{Z}}\left[\frac{1}{2k-1}\right]^{2n+3}\\
&= (n+1)\left(\frac{2}{\pi}\right)^{2n+3}\left[\left(\frac{1}{1}\right)^{2n+3}+\left(\frac{1}{-1}\right)^{2n+3}+\left(\frac{1}{3}\right)^{2n+3}+\left(\frac{1}{-3}\right)^{2n+3}+\ldots\right]\\
&= (n+1)\left(\frac{2}{\pi}\right)^{2n+3}\sum_{k=1}^{\infty}(-1)^{k+1}\alpha_k \qquad \text{where } \alpha_k \triangleq \left\{\begin{array}{ll}\left(\frac{1}{k}\right)^{2n+3} & \text{for } k \text{ odd}\\ \left(\frac{1}{k-1}\right)^{2n+3} & \text{for } k \text{ even}\end{array}\right\}\\
&= 0 \qquad \text{because } \lim_{k\to\infty}\alpha_k = 0 \text{ and by } \textit{Alternating Series Test}
\end{aligned}$$

11. lemma: $\tilde{S}_n(\omega)$ is *decreasing* with respect to $\omega \in [0 : \pi]$. Proof:

$$\begin{aligned}
\tilde{S}_n'(\omega) &= (n+1)\sum_{k\in\mathbb{Z}}\left[\frac{\sin\left(\frac{\omega}{2}\right)}{k\pi-\frac{\omega}{2}}\right]^{2n+1}\left[\frac{\cos\left(\frac{\omega}{2}\right)\left(k\pi-\frac{\omega}{2}\right)+\sin\left(\frac{\omega}{2}\right)}{\left(k\pi-\frac{\omega}{2}\right)^2}\right] && \text{by (9) lemma page 405}\\
&= \underbrace{(n+1)\left(\sin\frac{\omega}{2}\right)^{2n+1}}_{\geq 0 \text{ for } \omega\in[0:2\pi)}\underbrace{\sum_{k\in\mathbb{Z}}\left[\frac{1}{k\pi-\frac{\omega}{2}}\right]^{2n+2}}_{>0}\left[\underbrace{\left(\cos\frac{\omega}{2}\right)}_{\text{sign change at } \omega=\pi}+\frac{\overbrace{\sin\frac{\omega}{2}}^{>0 \text{ for } \omega\in(0:2\pi)}}{\underbrace{k\pi-\frac{\omega}{2}}_{\textit{decreasing} \text{ w.r.t. } \omega\in\mathbb{R}}}\right]
\end{aligned}$$

12. lemma: $\tilde{S}_n(\omega)$ is *increasing* with respect to $\omega \in [\pi : 2\pi)$. Proof: This is true because $\tilde{S}_n(\omega)$ is *decreasing* in $[0 : \pi]$ ((11) lemma) and because $\tilde{S}_n(\omega)$ is *symmetric* about $\omega = \pi$ ((5) lemma).

13. Proof that $0 < \tilde{S}_n(\omega) \leq 1$:
    (a) $\tilde{S}_n(\omega) > 0$ by (8) lemma and
    (b) $\tilde{S}_n(0) = 1$ by item (6) and
    (c) $\tilde{S}_n(\omega)$ is *decreasing* from $\omega = 0$ to $\omega = \pi$ by (11) lemma and
    (d) $\tilde{S}_n(\omega)$ is *increasing* from $\omega = \pi$ to $\omega = 2\pi$ by (12) lemma and
    (e) $\tilde{S}_n(2\pi) = 1$ because $\tilde{S}_n(2\pi) = \tilde{S}_n(0)$ by (4) lemma.

✏

Theorem S.8. [22]

THM					
1.	$(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$ *is a*	RIESZ BASIS	*for* $\operatorname{span}(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$		$\forall n \in \mathbb{W}$
2.	$(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$ *is an*	ORTHONORMAL BASIS	*for* $\operatorname{span}(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$	$\Longleftrightarrow$	$n = 0$

✎PROOF:

[22] Wojtaszczyk (1997) page 56 ⟨Proposition 3.12⟩, Prasad and Iyengar (1997) page 148 ⟨Theorem 6.3⟩, Forster and Massopust (2009) page 66 ⟨Theorem 2.25⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

1. Proof that $(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$ is a *Riesz basis* for $\mathsf{span}(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$:

$$\begin{aligned}
&0 < \tilde{S}_n(\omega) \le 1 && \text{by Lemma S.6 page 401 (1)}\\
&\implies (\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}} \text{ is a } \textit{Riesz basis} \text{ for } \mathsf{span}(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}} && \text{by Theorem ?? page ??}
\end{aligned}$$

2. Proof that $\{n=0\} \iff (\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$ is an *orthonormal basis* for $\mathsf{span}(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$:

$$\begin{aligned}
n=0 &\iff \tilde{S}_n(\omega) = 1 && \text{by Lemma S.6 page 401 (2), (4)}\\
&\iff (\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}} \text{ is an } \textit{orthonormal basis} \text{ for } \mathsf{span}(\!(\mathsf{N}_n(x-k))\!) && \text{by Theorem ?? page ??}
\end{aligned}$$

✏

S.6 Mutiresolution properties

S.6.1 Introduction

In 1989, Stéphane G. Mallat introduced the *Multiresolution Analysis* (MRA) structure (Definition ?? page ??) An MRA is very powerful because it can be used to approximate functions at incrementally increasing "scales" of resolution, and furthermore induces a *wavelet*. In fact, the MRA has become the dominate wavelet construction method. Moreover, P.G. Lemarié has proved that all wavelets with *compact support* are generated by an MRA. [23]

S.6.2 B-spline dyadic decomposition

One key feature of an MRA is *dyadic decomposition* such that $\mathsf{N}_n(x) = \sum_k \alpha_n \mathsf{N}_n(2x-n)$ for some sequence $(\!(\alpha_n)\!)$. As it turns out, *B-splines* also have this property (next theorem).

Theorem S.9 (*B-spline dyadic decomposition*). [24] *Let* $\mathsf{N}_n(x)$ *be the* nTH ORDER B-SPLINE.

THM
$$\mathsf{N}_n(x) = \frac{1}{2^n}\sum_{k=0}^{n+1}\binom{n+1}{k}\mathsf{N}_n(2x-k) \qquad \forall n\in\mathbb{W},\ \forall x\in\mathbb{R}$$

✎PROOF:

1. Base case ...proof for $n=0$ case:

$$\begin{aligned}
\mathsf{N}_0(x) &= \mathbb{1}_{[0:1)}(x) && \text{by definition of } \mathbb{1}_A(x) \quad \text{(Definition ?? page ??)}\\
&= \mathbb{1}_{[0:\frac{1}{2})}(x) + \mathbb{1}_{[\frac{1}{2}:1)}(x)\\
&= \mathbb{1}_{[2\times 0:2\times\frac{1}{2})}(2x) + \mathbb{1}_{[2\times\frac{1}{2}-1:2\times 1-1)}(2x-1)\\
&= \mathbb{1}_{[0:1)}(2x) + \mathbb{1}_{[0:1)}(2x-1)\\
&= \frac{1}{2^0}\sum_{k=0}^{0+1}\binom{0+1}{k}\mathsf{N}_n(2x-k)
\end{aligned}$$

[23] Mallat (1999) page 240, Definition ?? (page ??)
[24] Prasad and Iyengar (1997) pages 151–152 ⟨proof using Fourier transform⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
CC BY-NC-ND

2. Induction step…proof that n case $\implies n+1$ case:

$$
\begin{aligned}
\mathsf{N}_{n+1}(x) &= \int_0^1 \mathsf{N}_n(x-\tau)\,\mathrm{d}\tau && \text{by Lemma S.2 page 383}\\
&= \int_0^1 \frac{1}{2^n}\sum_{k=0}^{n+1}\binom{n+1}{k}\mathsf{N}_n(2x-2\tau-k)\,\mathrm{d}\tau && \text{by induction hypothesis}\\
&= \frac{1}{2^n}\sum_{k=0}^{n+1}\binom{n+1}{k}\int_{\tau=0}^{\tau=1}\mathsf{N}_n(2x-2\tau-k)\,\mathrm{d}\tau && \text{by \textit{linearity} of } \textstyle\sum \text{ operator}\\
&= \frac{1}{2^n}\sum_{k=0}^{n+1}\binom{n+1}{k}\int_{u=0}^{u=2}\mathsf{N}_n(2x-u-k)\frac{1}{2}\,\mathrm{d}u && \text{where } u \triangleq 2\tau \implies \tau = \tfrac{1}{2}u\\
&= \frac{1}{2^{n+1}}\sum_{k=0}^{n+1}\binom{n+1}{k}\left[\int_{u=0}^{u=1}\mathsf{N}_n(2x-k-u)\,\mathrm{d}u + \int_{u=1}^{u=2}\mathsf{N}_n(2x-k-u)\,\mathrm{d}u\right]\\
&= \frac{1}{2^{n+1}}\sum_{k=0}^{n+1}\binom{n+1}{k}\left[\int_{u=0}^{u=1}\mathsf{N}_n(2x-k-u)\,\mathrm{d}u + \int_{v=0}^{v=1}\mathsf{N}_n(2x-k-v-1)\,\mathrm{d}v\right] && \begin{array}{l}\text{where } v \triangleq u-1\\ \implies u = v+1\end{array}\\
&= \frac{1}{2^{n+1}}\sum_{k=0}^{n+1}\binom{n+1}{k}\left[\mathsf{N}_n(2x-k)+\mathsf{N}_n(2x-k-1)\right] && \text{by Lemma S.2 page 383}\\
&= \frac{1}{2^{n+1}}\left[\sum_{k=0}^{n+1}\binom{n+1}{k}\mathsf{N}_n(2x-k) + \sum_{m=1}^{n+2}\binom{n+1}{m-1}\mathsf{N}_n(2x-m)\right] && \text{where } m \triangleq k+1 \implies k = m-1\\
&= \frac{1}{2^{n+1}}\left[\underbrace{\sum_{k=1}^{n+1}\left[\binom{n+1}{k}+\binom{n+1}{k-1}\right]\mathsf{N}_n(2x-k)}_{\text{common indices of above two summations}} + \underbrace{\binom{n+1}{0}\mathsf{N}_n(2x-0)}_{k=0\text{ term}} + \underbrace{\binom{n+2}{n+2}\mathsf{N}_n(2x-n-2)}_{m=n+2\text{ term}}\right]\\
&= \frac{1}{2^{n+1}}\left[\underbrace{\sum_{k=1}^{n+1}\binom{n+2}{k}\mathsf{N}_n(2x-k)}_{\text{by \textit{Stifel formual} (Theorem ?? page ??)}} + \underbrace{\binom{n+2}{0}\mathsf{N}_n(2x-0)}_{\text{because } \binom{n+1}{0}=1=\binom{n+2}{0}} + \binom{n+2}{n+2}\mathsf{N}_n(2x-n-2)\right]\\
&= \frac{1}{2^{n+1}}\sum_{k=0}^{n+2}\binom{n+2}{k}\mathsf{N}_n(2x-k)
\end{aligned}
$$

✍

Example S.16. [25]The 0 *order B-spline dyadic decomposition*

$$\mathsf{N}_0(x) = \frac{1}{1}\sum_{k=0}^{k=1}\binom{1}{k}\mathsf{N}_0(2x-k)$$

is illustrated to the right.

[25] Strang (1989) page 615 ⟨Box function⟩, Strang and Nguyen (1996) page 441 ⟨Box function⟩

Copyright © 2019 Daniel J. Greenhoe

Example S.17. [26]The 1*st order B-spline dyadic decomposition*

$$\mathsf{N}_1(x) = \frac{1}{2}\sum_{k=0}^{k=2}\binom{2}{k}\mathsf{N}_1(2x-k)$$

is illustrated to the right.

Example S.18. The 2*nd order B-spline dyadic decomposition*

$$\mathsf{N}_2(x) = \frac{1}{4}\sum_{k=0}^{k=3}\binom{3}{k}\mathsf{N}_2(2x-k)$$

is illustrated to the right.

S.6.3 B-spline MRA scaling functions

Theorem S.10. *Let f* $N_n(x)$ *be the* nTH ORDER B-SPLINE *(Definition S.2 page 383)*.
Let $V_j \triangleq \operatorname{span}\left(\left(\mathsf{N}_n\left(2^j x - k\right)\right)\right)_{k\in\mathbb{Z}}$.

THM $\left(\left(V_j\right)\right)_{j\in\mathbb{Z}}$ *is a* MULTIRESOLUTION ANALYSIS *on* $L^2_{\mathbb{R}}$ *with* SCALING FUNCTION $\phi(x) \triangleq \mathsf{N}_n(x)$

✎PROOF:

1. lemma: $(\!(\mathsf{N}_n(x-k))\!)_{k\in\mathbb{Z}}$ is a *Riesz sequence*. Proof: by Theorem S.8 (page 406).
2. lemma: $\exists\,(\!(h_k)\!)$ such that $\mathsf{N}_n(x) = \sum_{k\in\mathbb{Z}} h_k \mathsf{N}_n(2x-k)$. Proof: by Theorem S.9 (page 407). In fact, note that $h_k = \frac{1}{2^n\sqrt{2}}\binom{n+1}{k}$
3. lemma: $\tilde{\mathbf{F}}\mathsf{N}_n(\omega)$ is *continuous* at 0. Proof:

$$\begin{aligned}\tilde{\mathbf{F}}\mathsf{N}_n(\omega) &= \frac{1}{\sqrt{2\pi}}e^{-i\frac{(n+1)\omega}{2}}\left(\operatorname{sinc}\frac{\omega}{2}\right)^{n+1} && \text{by Lemma S.5 page 398}\\ &\Longrightarrow \textit{continuous} \text{ at } 0 && \text{by known property of sinc function}\end{aligned}$$

4. lemma: $\tilde{\phi}(0) \neq 0$. Proof:

$$\begin{aligned}\tilde{\mathbf{F}}\mathsf{N}_n(0) &= \frac{1}{\sqrt{2\pi}}e^{-i\frac{(n+1)\omega}{2}}\left(\operatorname{sinc}\frac{\omega}{2}\right)^{n+1}\Bigg|_{\omega=0} && \text{by Lemma S.5 page 398}\\ &= 1\cdot\frac{1}{1/2} = 2 && \text{by } \textit{l'Hôpital's rule}\\ &\neq 0\end{aligned}$$

5. The completion of this proof follows directly from (1) lemma, (2) lemma, (3) lemma, (4) lemma, and Theorem **??** (page **??**).

✏

[26] Strang (1989) page 615 ⟨Hat function⟩, Strang and Nguyen (1996) page 442 ⟨Hat function⟩, Heil (2011) page 380 ⟨Fig. 12.10⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Figure S.4: *dilation equation* demonstrations for selected B-splines (Example S.19 page 410)

S.6.4 B-spline MRA coefficient sequences

Because each *B-spline* $\mathsf{N}_n(x)$ is the *scaling function* for an *MRA* (Theorem S.10 page 409), each *B-spline* also satisfies the *dilation equation* (Theorem **??** page **??**) such that

$$\mathsf{N}_n(x) = \sqrt{2}\sum_{k\in\mathbb{Z}} h_k \mathsf{N}(2x-k) \quad \text{where} \quad h_k = \left\{\begin{array}{ll} \frac{\sqrt{2}}{2^{n+1}}\binom{n+1}{k} & \text{for } n = 0,1,\ldots,n \\ 0 & \text{otherwise} \end{array}\right\}$$

The resulting *sequence* $(\!(h_k)\!)_{k\in\mathbb{Z}}$ is the *order n B-spline MRA coefficient sequence* induced by the *order n B-spline MRA scaling sequence* $\phi(x) \triangleq \mathsf{N}_n(x)$.[27]

Example S.19. See Figure S.4 (page 410) for some *dilation equation* demonstrations of selected B-splines.

Theorem S.11 (B-spline scaling coefficients)**.** *Let* $\left(L^2_{\mathbb{R}}, (\!(V_j)\!), \phi, (\!(h_n)\!)\right)$ *be an* MRA SYSTEM *(Definition* **??** *page* **??***). Let* $\mathsf{N}_n(x)$ *be a* nTH ORDER B-SPLINE *(Definition S.2 page 383).*

THM

$$\underbrace{\phi(x) \triangleq \mathsf{N}_n(x)}_{\text{(1) B-spline scaling function}} \implies (\!(h_k)\!) = \left\{\begin{array}{ll} \frac{\sqrt{2}}{2^{n+1}}\binom{n}{k} & \text{for } k = 0,1,\ldots,n \\ 0 & \text{otherwise} \end{array}\right. \qquad \text{(2) scaling sequence in "time"}$$

$$\iff \check{\mathsf{h}}(z)\Big|_{z\triangleq e^{i\omega}} = \frac{\sqrt{2}}{2^n}\left(1+z^{-1}\right)^{n+1}\Big|_{z\triangleq e^{i\omega}} \qquad \text{(3) scaling sequence in "z domain"}$$

$$\iff \tilde{\mathsf{h}}(\omega) = 2\sqrt{2}e^{-i\frac{n+1}{2}\omega}\left[\cos\left(\frac{\omega}{2}\right)\right]^{n+1} \qquad \text{(4) scaling sequence in "frequency"}$$

✎PROOF:

1. Proof that (1) $\implies$ (3): By Theorem S.10 page 409 we know that $\mathsf{N}_n(x)$ is a *scaling function* (Definition **??** page **??**). So then we know that we can use Lemma **??** page **??**.

$$\begin{aligned}
\tilde{\mathsf{h}}(\omega) &= \sqrt{2}\frac{\tilde{\phi}(2\omega)}{\tilde{\phi}(\omega)} && \text{by Lemma ?? page ??} \\
&= \sqrt{2}\frac{\tilde{\mathsf{N}}_n(2\omega)}{\tilde{\mathsf{N}}_n(\omega)} && \text{by (1)} \\
&= \sqrt{2}\frac{\frac{1}{\sqrt{2\pi}}\left(\frac{1-e^{-i2\omega}}{2i\omega}\right)^{n+1}}{\frac{1}{\sqrt{2\pi}}\left(\frac{1-e^{-i\omega}}{i\omega}\right)^{n+1}} && \text{by Lemma S.5 page 398}
\end{aligned}$$

[27] For *Octave/MatLab* code useful for plotting a function given a sequence of coefficients $(\!(h_k)\!)$, see Section **??** (page **??**).

COPYRIGHT © 2019 DANIEL J. GREENHOE

$$= \frac{\sqrt{2}}{2^{n+1}}\left(\frac{1-z^{-2}}{1-z^{-1}}\right)^{n+1}\Bigg|_{z=e^{i\omega}}$$
$$= \frac{\sqrt{2}}{2^{n+1}}\left[\left(\frac{1-z^{-2}}{1-z^{-1}}\right)\left(\frac{1+z^{-1}}{1+z^{-1}}\right)\right]^{n+1}\Bigg|_{z=e^{i\omega}}$$
$$= \frac{\sqrt{2}}{2^{n+1}}\left(\frac{\left(1-z^{-2}\right)\left(1+z^{-1}\right)}{1-z^{-2}}\right)^{n+1}\Bigg|_{z=e^{i\omega}}$$
$$= \frac{\sqrt{2}}{2^{n}}\left(1+z^{-1}\right)^{n+1}\Bigg|_{z\triangleq e^{i\omega}}$$

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

$$\check{\mathsf{h}}(z)\Big|_{z\triangleq e^{i\omega}} = \frac{\sqrt{2}}{2^{n}}\left(1+z^{-1}\right)^{n+1}\Bigg|_{z\triangleq e^{i\omega}} \qquad \text{by (3)}$$
$$= \frac{\sqrt{2}}{2^{n}}\left(\sum_{k=0}^{n+1}\binom{n}{k}z^{-k}\right)\Bigg|_{z\triangleq e^{i\omega}} \qquad \text{by } \textit{binomial theorem}$$
$$\Longleftrightarrow \qquad h_k = \frac{\sqrt{2}}{2^{n+1}}\binom{n}{k} \qquad \text{by definition of } Z \textit{ transform} \text{ (Definition R.4 page 370)}$$

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

$$\tilde{\mathsf{h}}(\omega) = \check{\mathsf{h}}(z)\Big|_{z\triangleq e^{i\omega}} \qquad \text{by definition of } DTFT \text{ (Definition Q.1 page 359)}$$
$$= \frac{\sqrt{2}}{2^{n}}\left(1+z^{-1}\right)^{n+1}\Bigg|_{z\triangleq e^{i\omega}} \qquad \text{by (3)}$$
$$= \frac{\sqrt{2}}{2^{n}}\left(1+e^{-i\omega}\right)^{n+1} \qquad \text{by definition of } z$$
$$= \frac{\sqrt{2}}{2^{n}}\left[e^{-i\frac{1}{2}\omega}\left(e^{i\frac{\omega}{2}}+e^{-i\frac{\omega}{2}}\right)\right]^{n+1}$$
$$= \frac{\sqrt{2}}{2^{n}}e^{-i\frac{n+1}{2}\omega}\left[2\cos\left(\frac{\omega}{2}\right)\right]^{n+1}$$
$$= 2\sqrt{2}e^{-i\frac{n+1}{2}\omega}\left[\cos\left(\frac{\omega}{2}\right)\right]^{n+1}$$

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

$$\check{\mathsf{h}}(z)\Big|_{z\triangleq e^{i\omega}} = \check{\mathsf{h}}\left(e^{i\omega}\right)$$
$$= \tilde{\mathsf{h}}(\omega)$$
$$= 2\sqrt{2}e^{-i\frac{n+1}{2}\omega}\left[\cos\left(\frac{\omega}{2}\right)\right]^{n+1} \qquad \text{by (4)}$$
$$= \frac{\sqrt{2}}{2^{n}}e^{-i\frac{n+1}{2}\omega}\left[2\cos\left(\frac{\omega}{2}\right)\right]^{n+1}$$
$$= \frac{\sqrt{2}}{2^{n}}\left[e^{-i\frac{1}{2}\omega}\left(e^{i\frac{\omega}{2}}+e^{-i\frac{\omega}{2}}\right)\right]^{n+1}$$
$$= \frac{\sqrt{2}}{2^{n}}\left(1+e^{-i\omega}\right)^{n+1}$$
$$= \frac{\sqrt{2}}{2^{n}}\left(1+z^{-1}\right)^{n+1}\Bigg|_{z\triangleq e^{i\omega}}$$

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

✏

Example S.20 (2 coefficient case). [28] Let $\left(L^2_{\mathbb{R}}, (\!(V_j)\!), \phi, (\!(h_n)\!)\right)$ be an *MRA system* (Definition ?? page ??).

EX

$$\underbrace{\left\{\begin{array}{ll} 1. & \operatorname{supp}\phi(x) = [0:1] \quad \text{and} \\ 2. & (\!(\phi(x-k))\!) \text{ forms a} \\ & \textit{partition of unity} \end{array}\right\}}_{\text{(A)}} \iff \underbrace{h_n = \left\{\begin{array}{ll} \frac{\sqrt{2}}{2} & \text{for } n=0 \\ \frac{\sqrt{2}}{2} & \text{for } n=1 \\ 0 & \text{otherwise} \end{array}\right\}}_{\text{(B)}} \iff \underbrace{\left\{\phi(x) = \mathsf{N}_0(x)\right\}}_{\text{(C)}}$$

✎PROOF:

1. Proof that (A) $\Longrightarrow$ (B):

   (a) lemma: Only h_0 and h_1 are *non-zero*; All other coefficients h_k are 0. Proof: This follows from $\operatorname{supp}\phi(x) = [0:1]$ (Definition ?? page ??) and by Theorem ?? page ??.

   (b) lemma (equations for $(\!(h_k)\!)$): Because $(\!(h_k)\!)$ is a *scaling coefficient sequence* (Definition ?? page ??), it must satisfy the *admissibility equation* (Theorem ?? page ??). And because $(\!(\phi(x-k))\!)$ forms a *partition of unity*, it must satisfy the equations given by Theorem ?? (page ??). (1a) lemma and these two constraints yield two simultaneous equations and two unknowns:

   $$\begin{array}{rcll} h_0 + h_1 & = & \sqrt{2} & (\textit{admissibility condition}) \\ h_0 - h_1 & = & 0 & (\textit{partition of unity / zero at } -1 \textit{ / vanishing 0th moment}) \end{array}$$

   (c) lemma: The equations provided by (1b) lemma can be expressed in matrix algebra form as follows…

   $$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}$$

   (d) lemma: The *inverse* $\mathbf{A}^{-1}$ of $\mathbf{A}$ can be expressed as demonstrated below…

   $$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 2 & 0 & 1 & 1 \\ 1 & -1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 2 & 0 & 1 & 1 \\ 0 & -1 & -\frac{1}{2} & \frac{1}{2} \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} \end{array}\right]$$
   $$\Longrightarrow \mathbf{A}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

   (e) Proof for the values of $(\!(h_k)\!)$ (B):

   $$\begin{aligned} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \mathbf{A}^{-1}\mathbf{A}\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} &= \mathbf{A}^{-1}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} && \text{by (1c) lemma} \\ &= \mathbf{A}^{-1}\begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix} && \text{by (1c) lemma} \\ &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix} && \text{by (1d) lemma} \\ &= \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}$$

[28] Haar (1910), Wojtaszczyk (1997) pages 14–15 ⟨"Sources and comments"⟩

COPYRIGHT © 2019 DANIEL J. GREENHOE

2. Proof that (B) $\Longrightarrow$ (C):

$$
\begin{aligned}
\text{(B)} \Longrightarrow \phi(x) &= \sum_{k\in\mathbb{Z}} h_k\sqrt{2}\phi(2x-k) && \textit{dilation equation} && \text{(Theorem ?? page ??)}\\
&= \sum_{k=0}^{k=1}\left(\frac{\sqrt{2}}{2}\right)\sqrt{2}\phi(2x-k) && \text{by item (1e) page 412}\\
&= \sum_{k=0}^{k=1}\phi(2x-k)\\
&= \sum_{k=0}^{k=1}\binom{1}{k}\phi(2x-k) && \text{by definition of } \binom{n}{k} && \text{(Definition ?? page ??)}\\
&\Longrightarrow \text{(D)} && \text{by } \textit{B-spline dyadic decomposition} && \text{(Theorem S.9 page 407)}
\end{aligned}
$$

3. Proof that (B) $\Longleftarrow$ (C):

$$
\begin{aligned}
\text{(C)} \Longrightarrow \mathsf{N}_0(x) &= \sum_{k=0}^{k=1}\binom{1}{k}\mathsf{N}_0(2x-k) && \text{by } \textit{B-spline dyadic decomposition} && \text{(Theorem S.9 page 407)}\\
&= \sum_{k=0}^{k=1}\left(\frac{\sqrt{2}}{2}\right)\sqrt{2}\mathsf{N}_0(2x-k) && \text{by definition of } \binom{n}{k} && \text{(Definition ?? page ??)}\\
&= \sum_{k\in\mathbb{Z}} h_k\sqrt{2}\mathsf{N}_0(2x-k) && \text{by definition of } \binom{n}{k} && \text{(Definition ?? page ??)}\\
&\Longrightarrow \text{(B)}
\end{aligned}
$$

4. Proof that (A) $\Longleftarrow$ (C):
   1. Proof that (C) $\Longrightarrow$ $\operatorname{supp}\phi(x) = [0:1]$: by Theorem S.4 (page 394)
   2. Proof that (C) $\Longrightarrow$ $(\!(\phi(x-k))\!)$ forms a *partition of unity*: by Theorem S.7 (page 400)

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

EX

n=0,	(÷0!)	1;							
n=1,	(÷1!)	1,	0;						
		−1,	2;						
n=2,	(÷2!)	1,	0,	0;					
		−2,	6,	−3;					
		1,	−6,	9;					
n=3,	(÷3!)	1,	0,	0,	0;				
		−3,	12,	−12,	4;				
		3,	−24,	60,	−44;				
		−1,	12,	−48,	64;				
n=4,	(÷4!)	1,	0,	0,	0,	0;			
		−4,	20,	−30,	20,	−5;			
		6,	−60,	210,	−300,	155;			
		−4,	60,	−330,	780,	−655;			
		1,	−20,	150,	−500,	625;			
n=5,	(÷5!)	1,	0,	0,	0,	0,	0;		
		−5,	30,	−60,	60,	−30,	6;		
		10,	−120,	540,	−1140,	1170,	−474;		
		−10,	180,	−1260,	4260,	−6930,	4386;		
		5,	−120,	1140,	−5340,	12270,	−10974;		
		−1,	30,	−360,	2160,	−6480,	7776;		
n=6,	(÷6!)	1,	0,	0,	0,	0,	0,	0;	
		−6,	42,	−105,	140,	−105,	42,	−7;	
		15,	−210,	1155,	−3220,	4935,	−3990,	1337;	
		−20,	420,	−3570,	15680,	−37590,	47040,	−24178;	
		15,	−420,	4830,	−29120,	96810,	−168000,	119182;	
		−6,	210,	−3045,	23380,	−100065,	225750,	−208943;	
		1,	−42,	735,	−6860,	36015,	−100842,	117649;	
n=7,	(÷7!)	1,	0,	0,	0,	0,	0,	0,	0;
		−7,	56,	−168,	280,	−280,	168,	−56,	8;
		21,	−336,	2184,	−7560,	15400,	−18648,	12488,	−3576;
		−35,	840,	−8400,	45360,	−143360,	267120,	−273280,	118896;
		35,	−1120,	15120,	−111440,	483840,	−1238160,	1733760,	−1027984;
		−21,	840,	−14280,	133560,	−741160,	2436840,	−4391240,	3347016;
		7,	−336,	6888,	−78120,	528920,	−2135448,	4753336,	−4491192;
		−1,	56,	−1344,	17920,	−143360,	688128,	−1835008,	2097152;
n=8,	(÷8!)	1,	0,	0,	0,	0,	0,	0,	0,	0;
		−8,	72,	−252,	504,	−630,	504,	−252,	72,	−9
		28,	−504,	3780,	−15624,	39690,	−64008,	64260,	−36792,	9207
		−56,	1512,	−17388,	111384,	−436590,	1079064,	−1650348,	1432872,	−541917
		70,	−2520,	39060,	−340200,	1821330,	−6146280,	12800340,	−15082200,	7715619
		−56,	2520,	−49140,	541800,	−3691170,	15903720,	−42324660,	63667800,	−41503131
		28,	−1512,	35532,	−474264,	3929310,	−20674584,	67410252,	−124449192,	99584613
		−8,	504,	−13860,	217224,	−2121210,	13208328,	−51179940,	112731192,	−107948223
		1,	−72,	2268,	−40824,	459270,	−3306744,	14880348,	−38263752,	43046721

Table S.1: Coefficients of the *B-splines* $\mathsf{N}_n(x)$ multiplied by $n!$ (Example S.9 page 390)

Copyright © 2019 Daniel J. Greenhoe

APPENDIX T

SOURCE CODE

The free and open source software package Maxima has been used to compute some of the algebraic expressions for *B-spline*s used in APPENDIX S (page 383):

```
/*=====================================================================*
 * Daniel J. Greenhoe                                                  *
 * Maxima script file                                                  *
 * To execute this script, start Maxima in a command window            *
 * in the subdirectory containing this file (e.g. c:\math\maxima\)     *
 * and then after the (%i...) prompt enter                             *
 *   batchload("bspline.max")$                                         *
 * Data produced will be written to the file "bsplineout.txt".         *
 * reference: http://maxima.sourceforge.net/documentation.html         *
 *=====================================================================*/
/*-----------------------------------------
 * initialize script
 *-----------------------------------------*/
reset();
kill(all);
load(orthopoly);
display2d:false; /* 2-dimensional display */
writefile("bsplineout.txt");
/*-----------------------------------------
 * n = B-spline order parameter
 * may be set to any value in {1,2,3,...}
 *-----------------------------------------*/
n:2;
print("=====================================================================");
print("Daniel J. Greenhoe");
print("Output file for nth order B-spline Nn(x) calculation, n=",n," .");
print("Output produced using Maxima running the script file bspline.max");
print("=====================================================================");
Nnx:(1/n!)*sum((-1)^k*binomial(n+1,k)*(x-k)^n*unit_step(x-k),k,0,n+1);
print("---------------------------------------------------------------");
print("              n+1      k (n+1)       n                    ");
print(" n! Nn(x) = SUM (-1)  (    ) (x-k)  step(x-k)  ,n=",n);
print("              k=0       ( k )                            ");
print("              ",n+1,"      k (",n+1,")      ",n);
print(n,"! Nn(x) = SUM (-1)  (    ) (x-k)  step(x-k)");
print("              k=0       ( k )");
print("            = ",expand(n!*Nnx));
print("---------------------------------------------------------------");
assume(x<=0);    print(n!,"N(x)= ",expand(n!*Nnx)," for x<=0");  forget(x<=0);
for i:0 thru n step 1 do(
  assume(x>i,x<(i+1)),
  print(n!,"N(x)= ",expand(n!*Nnx)," for ",i,"<x<",i+1),
  tex(expand(n!*Nnx),"djg.tex"),/*write output in TeX format to file "djg.tex"*/
  forget(x>i,x<(i+1))
  );
assume(x>(n+1)); print(n!,"N(x)= ",expand(n!*Nnx)," for x>",n+1); forget(x>(n+1));
```

```
print("-----------------------------------------------------------");
print(" values at some specific points x:                          ");
print("-----------------------------------------------------------");
y:Nnx,x=(n+1)/2;print("N(",(n+1)/2,")= ",y," (center value)");
y:Nnx,x=(n+2)/2;print("N(",(n+2)/2,")= ",y);
y:Nnx,x=(n+3)/2;print("N(",(n+3)/2,")= ",y);
/*-----------------------------------------
 * close output file
 *-----------------------------------------*/
closefile();
```

Once the polynomial expressions for a *B-spline* have been calculated, they can be plotted within a LaTeX environment using the pst-plot package along with a LaTeX source file such as the following:[1]

```
%==========================================================================
% Daniel J. Greenhoe
% LaTeX file
% N_3(x) B-spline
% nominal unit = 10mm
%==========================================================================
\begin{pspicture}(-1,-0.5)(5,1)
  %-----------------------------------------
  % parameters
  %-----------------------------------------
  \psset{plotpoints=64,labelsep=1pt}%
  %-----------------------------------------
  % axes
  %-----------------------------------------
  \psaxes[linewidth=0.75pt,linecolor=axis,yAxis=false,ticks=x,labels=x]{<->}(0,0)(-1,0)(5,1)% x axis
  \psaxes[linewidth=0.75pt,linecolor=axis,xAxis=false,ticks=x,labels=x]{->}(0,0)(-1,0)(5,1)%  y axis
  %-----------------------------------------
  % annotation
  %-----------------------------------------
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](2,0)(2,0.667)%
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](0,0.667)(2,0.667)%
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](1,0)(1,0.1667)%
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](3,0)(3,0.1667)%
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](0,0.1667)(3,0.1667)%
  \uput[180](0,0.667){$\sfrac{2}{3}$}%
  \uput[180](0,0.1667){$\sfrac{1}{6}$}%
  %-----------------------------------------
  % function plot
  %-----------------------------------------
  \psplot{0}{1}{+1 x 3 exp mul                                          6 div}% for 0<=x<=1
  \psplot{1}{2}{-3 x 3 exp mul +12 x 2 exp mul add -12 x mul add  +4 add 6 div}% for 1<=x<=2
  \psplot{2}{3}{+3 x 3 exp mul -24 x 2 exp mul add +60 x mul add -44 add 6 div}% for 2<=x<=3
  \psplot{3}{4}{-1 x 3 exp mul +12 x 2 exp mul add -48 x mul add +64 add 6 div}% for 3<=x<=4
  \psline(0,0)(-0.5,0)\psline[linestyle=dotted](-0.5,0)(-0.75,0)%              % for x<=0
  \psline(4,0)(4.5,0)\psline[linestyle=dotted](4.5,0)(4.75,0)%                 % for x>=4
\end{pspicture}%
```

Alternatively, one can plot $\mathsf{N}_3(x)$ more or less directly from the equation given in Theorem S.1 (page 387) without first calculating the polynomial expressions:

```
%==========================================================================
% Daniel J. Greenhoe
% LaTeX file
% N_3(x) B-spline
% nominal unit = 10mm
%==========================================================================
\begin{pspicture}(-1,-0.5)(5,1)
  %-----------------------------------------
  % parameters
  %-----------------------------------------
  \psset{plotpoints=64,labelsep=1pt}%
```

[1] For help with PostScript® math operators, see Adobe (1999), pages 508–509 ⟨Arithmetic and Math Operators⟩.

Copyright © 2019 Daniel J. Greenhoe

```
  %-------------------------------------
  % axes
  %-------------------------------------
  \psaxes[linewidth=0.75pt,linecolor=axis,yAxis=false,ticks=x,labels=x]{<->}(0,0)(-1,0)(5,1)% x axis
  \psaxes[linewidth=0.75pt,linecolor=axis,xAxis=false,ticks=x,labels=x]{->}(0,0)(-1,0)(5,1)%  y axis
  %-------------------------------------
  % annotation
  %-------------------------------------
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](2,0)(2,0.667)%
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](0,0.667)(2,0.667)%
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](1,0)(1,0.1667)%
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](3,0)(3,0.1667)%
  \psline[linestyle=dashed,linewidth=0.75pt,linecolor=red](0,0.1667)(3,0.1667)%
  \uput[180](0,0.667){$\sfrac{2}{3}$}%
  \uput[180](0,0.1667){$\sfrac{1}{6}$}%
  %-----------------------------------------------------------------------------
  % for n=3
  %        1  n+1        (n+1)                    1   4          ( 4 )
  % N_n(x)--- SUM (-1)^k (    ) (x-k)^n s(x-k) = --- SUM  (-1)^k (    ) (x-k)^3 s(x-k)
  %        n! k=1        ( k )                    3! k=1         ( k )
  %
  % where s(x) = 0 for x<0 and 1 for x>=0  (step function)
  %-----------------------------------------------------------------------------
  \psplot{0}{1}  {1 x 0 sub 3 exp mul   6 div}%
  \psplot{1}{2}  {1 x 0 sub 3 exp mul   4 x 1 sub 3 exp mul sub   6 div}%
  \psplot{2}{3}  {1 x 0 sub 3 exp mul   4 x 1 sub 3 exp mul sub   6 x 2 sub 3 exp mul add   6 div}%
  \psplot{3}{4}  {1 x 0 sub 3 exp mul   4 x 1 sub 3 exp mul sub   6 x 2 sub 3 exp mul add   4 x 3 sub
            3 exp mul sub   6 div}%
  \psplot{4}{4.5}{1 x 0 sub 3 exp mul   4 x 1 sub 3 exp mul sub   6 x 2 sub 3 exp mul add   4 x 3 sub
            3 exp mul sub   1 x 4 sub 3 exp mul add 6 div}%
  %    N_3(x) = [(4choose0)(x-0)^3    - (4choose1)(x-1)^3       + (4choose2)(x-2)^3       -
            (4choose3)(x-3)^3       + (4choose4)(x-4)^3       ]/3!
  %           = [         1 (x-0)^3    -         4 (x-1)^3       +         6 (x-2)^3       -         4
            (x-3)^3       +         1 (x-4)^3       ]/6
  \psline(0,0)(-0.5,0)%
  \psline[linestyle=dotted](-0.5,0)(-0.75,0)%
  \psline[linestyle=dotted](4.5,0)(4.75,0)%
\end{pspicture}%
```


Copyright © 2019 Daniel J. Greenhoe

Copyright © 2019 Daniel J. Greenhoe

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 [2]

❝ 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. [3]

ancient library of Alexandria

The Book Worm by *Carl Spitzweg*, circa 1850 [4]

❝ 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 [5]

[2] quote: Simmons (2007), page 187.
image: http://en.wikipedia.org/wiki/Image:Niels_Henrik_Abel.jpg, public domain
[3] quote: Machiavelli (1961), page 139?.
image: http://commons.wikimedia.org/wiki/File:Santi_di_Tito_-_Niccolo_Machiavelli%27s_portrait_headcrop.jpg, public domain
[4] http://en.wikipedia.org/wiki/File:Ancientlibraryalex.jpg, public domain http://en.wikipedia.org/wiki/File:Carl_Spitzweg_021.jpg,

Copyright © 2019 Daniel J. Greenhoe
CC BY-NC-ND

[5] quote: Kenko (circa 1330)
image: http://en.wikipedia.org/wiki/Yoshida_Kenko

Copyright © 2019 Daniel J. Greenhoe

BIBLIOGRAPHY

M. S. Abdaheer. *Digital Signal Processing*. University Science Press, New Delhi, 2009.

M. Abom. A note on random errors in frequency response estimators. *Journal of Sound and Vibration*, 107(2):355–358, 1986. doi: 10.1016/0022-460X(86)90244-0. URL http://adsabs.harvard.edu/abs/1986JSV...107..355A.

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.

Adobe. *PostScript Language Reference*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1999. ISBN 0-201-37922-8. URL http://wwwimages.adobe.com/content/dam/Adobe/en/devnet/postscript/pdfs/PLRM.pdf.

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*.

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.

Randall J. Allemang, R. Zimmerman, and David L. Brown. Technlqucs of reducing noise in frquency response measurements. *ASME Winter Annual Meeting*, 1979. session on Application of System Identlfication Techniques.

Randall J. Allemang, David L. Brown, and Robert W. Rost. *Experimental Modal Analysis And Dynamic Component Synthesis*, 2(AFWAL-TR-87-3069), December 1987. URL https://apps.dtic.mil/dtic/tr/fulltext/u2/a195145.pdf.

Theodore T. Allen. *Introduction to Engineering Statistics and Lean Six Sigma: Statistical Quality Control and Design of Experiments and Systems*. Springer, 3 edition, 2018. ISBN 9781447174202.

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.

Theodore Wilbur Anderson. *An introduction to multivariate statistical analysis*. Wiley series in probability and mathematical statistics. Wiley, 1 edition, 1958. URL https://books.google.com/books?id=7YpqAAAAMAAJ.

Theodore Wilbur Anderson. *An Introduction to Multivariate Statistical Analysis*, volume 114 of *Wiley Series in Probability and Statistics—Applied Probability and Statistics Section Series*. Wiley, 2 edition, 1984. ISBN 9780471889878. URL http://books.google.com/books?vid=ISBN9780471889878.

John E. Angus. The probability integral transform and related results. *SIAM Review*, 36(4):652–654, February 25 1994. doi: 10.1137/1036146. URL http://dx.doi.org/10.1137/1036146.

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.

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).

George Bachman. *Elements of Abstract Harmonic Analysis*. Academic paperbacks. Academic Press, New York, 1964. URL http://books.google.com/books?id=ZP8-AAAAIAAJ.

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.

David H. Bailey, Jonathan M. Borwein, and Andrew Mattingly. The computation of previously inaccessible digits of π^2 and catalan's constant (2013). page 23, March 24 2011. URL https://pdfs.semanticscholar.org/0b15/22a1b2f76729e19be2be1bca9f57031a4db9.pdf.

Copyright © 2019 Daniel J. Greenhoe

David H. Bailey, Jonathan M. Borwein, and Andrew Mattingly. 20. the computation of previously inaccessible digits of π^2 and catalan's constant (2013). In *Pi: The Next Generation: A Sourcebook on the Recent History of Pi and Its Computation*, pages 324–339. Springer, 2013. ISBN 9783319323770. URL http://books.google.com/books?vid=ISBN3319323776.

Om Prakash Bajpai. *Foundations of statistics*. Asia Publishing House, 1967. URL https://books.google.com/books?id=QrPuAAAAMAAJ.

Per Bak. *How Nature Works: the science of self-organized criticality*. Springer Science & Business Media, 2013. ISBN 9781475754261. URL http://books.google.com/books?vid=ISBN1475754264.

N Balakrishnan and Chin-Diew Lai. *Continuous Bivariate Distributions*. Springer Science & Business Media, 2 edition, May 31 2009. ISBN 9780387096148. URL http://books.google.com/books?vid=ISBN0387096140.

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.

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.

Eric Temple Bell. *Men of Mathematics*. Simon & Schuster, New York, 1986. ISBN 9780671628185. URL http://books.google.com/books?vid=ISBN0671628186.

J. S. Bendat. Statistical errors in measurement of coherence functions and input/output quantities. *Journal of Sound and Vibration*, 59(3):405–421, August 08 1978. doi: 10.1016/S0022-460X(78)80007-8. URL https://doi.org/10.1016/S0022-460X(78)80007-8.

Julius S. Bendat and Allan G. Piersol. *Measurement and Analysis of Random Data*. John Wiley & Sons, 1966.

Julius S. Bendat and Allan G. Piersol. *Engineering Applications of Correlation and Spectral Analysis*. John Wiley & Sons, 1980. ISBN 9780471058878. URL http://www.amazon.com/dp/0471058874.

Julius S. Bendat and Allan G. Piersol. *Engineering Applications of Correlation and Spectral Analysis*. Wiley-Interscience, 2 edition, 1993. ISBN 9780471570554. URL http://www.amazon.com/dp/0471570559.

Julius S. Bendat and Allan G. Piersol. *Random Data: Analysis and Measurement Procedures*, volume 729 of *Wiley Series in Probability and Statistics*. John Wiley & Sons, 4 edition, 2010. ISBN 9781118210826. URL http://books.google.com/books?vid=ISBN1118210824.

Sterling Khazag Berberian. *Introduction to Hilbert Space*. Oxford University Press, New York, 1961. URL http://books.google.com/books?vid=ISBN0821819127.

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.

M. Bertero and P. Boccacci. *Introduction to Inverse Problems in Imaging*. CRC Press, 1998. ISBN 9781439822067. URL http://books.google.com/books?vid=ISBN9781439822067.

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.

Patrick Billingsley. *Probability And Measure*. Wiley series in probability and mathematical statistics. Wiley, 3 edition, 1995. ISBN 978-0471007104. URL http://books.google.com/books?vid=ISBN0471007102.

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.

William M. Bolstad. *Introduction to Bayesian Statistics*. Wiley, 2 edition, 2007. ISBN 9780470141151. URL http://books.google.com/books?vid=ISBN9780470141151.

A. Böttcher and B. Silbermann. *Introduction to Large Truncated Toeplitz Matrices*. Springer, 1999. ISBN 9780387985701. URL https://d-nb.info/955562643/04.

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 9780470525487. 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.

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.

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.

Copyright © 2019 Daniel J. Greenhoe

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.

Charles L. Byrne. *Signal Processing: A Mathematical Approach*. AK Peters Series. A K Peters, 2005. ISBN 1568812426.

James A. Cadzow. *Foundations of Digital Signal Processing and Data Analysis*. Macmillan Publishing Company, New York, 1987. ISBN 0023180102. URL http://frontweb.vuse.vanderbilt.edu/vuse_web/directory/facultybio.asp?FacultyID=9.

Peter E. Caines. *Linear Stochastic Systems*, volume 100 of *Wiley Series in Probability and Statistics*. J. Wiley, 1988. ISBN 9780471081012. URL http://books.google.com/books?vid=ISBN0471081019.

Peter E. Caines. *Linear Stochastic Systems*, volume 77 of *Classics in Applied Mathematics*. SIAM, 2018. ISBN 9781611974713. URL http://books.google.com/books?vid=ISBN1611974712.

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*.

James J. Callahan. *Advanced Calculus: A Geometric View*. Springer Science & Business Media, September 9 2010. ISBN 9781441973320. URL http://books.google.com/books?vid=ISBN144197332X.

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.

Thomas G. Carne and C. R. Dohrmann. Improving experimental frequency response function matrices for admittance modeling. In *Proceedings of the XXIV International Modal Analysis Conference (IMAC)*, St. Louis, Missouri, USA, 2006. URL https://www.researchgate.net/publication/265790751.

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.

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).

Arthur Cayley. A memoir on the theory of matrices. *Philosophical Transactions of the Royal Society of London*, 148:17–37, 1858. ISSN 1364-503X. URL http://www.jstor.org/view/02610523/ap000059/00a00020/0.

Si-Wei Chen, Xue-Song Wang, and Motoyuki Sato. Polinsar complex coherence estimation based on covariance matrix similarity test. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11):4699–4710, November 2012. doi: 10.1109/TGRS.2012.2192937. URL https://www.researchgate.net/profile/Si_Wei_Chen/publication/281271737.

Sung C. Choi. *Introductory applied statistics in science*. Prentice-Hall, 1978. ISBN 9780135016190. URL http://books.google.com/books?vid=ISBN9780135016190.

Alexandre J. Chorin and Ole H. Hald. *Stochastic Tools in Mathematics and Science*, volume 1 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2 edition, 2009. ISBN 978-1-4419-1001-1. URL http://books.google.com/books?vid=ISBN9781441910011.

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.

Ole Christensen. *Functions, Spaces, and Expansions: Mathematical Tools in Physics and Engineering*. Applied and Numerical Harmonic Analysis. Birkhäuser, Boston/Basel/Berlin, 2010. ISBN 978-0-8176-4979-1. doi: 10.1007/978-0-8176-4980-7. URL http://books.google.com/books?vid=ISBN0817649794.

Charles K. Chui. *Multivariate Splines*, volume 54 of *CBMS-NSF regional conference series in applied mathematics*. SIAM, 1988. ISBN 1611970172. URL http://books.google.com/books?vid=ISBN1611970172.

Charles K. Chui. *An Introduction to Wavelets*. Academic Press, San Diego, California, USA, January 3 1992. ISBN 9780121745844. URL http://books.google.com/books?vid=ISBN0121745848.

Donato Michele Cifarelli and Eugenio Regazzini. De finetti's contribution to probability and statistics. *Statistical Science*, 11(4):253–282, 1996. URL http://www.jstor.org/stable/2246020.

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.

Jon F. Claerbout. *Fundamentals of Geophysical Data Processing with Applications to Petroleum Prospecting*. International series in the earth and planetary sciences, Tab Mastering Electronics Series. McGraw-Hill, New York, 1976. ISBN 9780070111172. URL http://sep.stanford.edu/sep/prof/.

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.

Peter M. Clarkson. *Optimal and Adaptive Signal Processing*. Electronic Engineering Systems Series. CRC Press, 1993. ISBN 0849386098.

Copyright © 2019 Daniel J. Greenhoe

R. E. Cobb and L. D. Mitchell. Estimation of uncorrelated content in experimentally measured frequency response functions using three measurement channels. *Mechanical Systems and Signal Processing*, 4(6):449–461, November 1990. URL https://doi.org/10.1016/0888-3270(90)90045-M.

Richard E. Cobb. *Confidence Bands, Measurement Noise, and Multiple Input - Multiple Output Measurements Using the Three-Channel Frequency Response Function Estimator*. Ph.d. dissertation, May 1988. URL http://hdl.handle.net/10919/53675. "Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fululllment of the requirements for the degree of Doctor of Philosophy in Mechanical Engineering".

Paul M. Cohn. *Basic Algebra; Groups, Rings and Fields*. Springer, December 6 2002. ISBN 1852335874. URL http://books.google.com/books?vid=isbn1852335874.

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.

R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Interscience Publishers, New York, 1930. URL http://www.worldcat.org/isbn/0471504475.

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.

M. G. Cox. The numerical evaluation of b-splines. *IMA Journal of Applied Mathematics*, 10(2):134–149, 1972. doi: 10.1093/imamat/10.2.134. URL http://imamat.oxfordjournals.org/content/10/2/134.abstract.

Guy L. Curry and Richard M. Feldman. *Manufacturing Systems Modeling and Analysis*. Technology & Engineering. Springer Science & Business Media, 2 edition, 2010. ISBN 9783642166181. URL http://books.google.com/books?vid=ISBN3642166180.

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.

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.

Bruno de Finetti. Sui passaggi al limite nel calcolo delle probabilità. *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere*, 63:155–166, 1930a.

Bruno de Finetti. A proposito dell'estensione del teorema delle probabilita totali alle classi numerabili. *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere*, 63:901–905, 1930b.

Bruno de Finetti. Ancora sull'estensione alle classi numerabili del teorema delle probabilità totali. *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere*, 63:1063–1069, 1930c.

Johan de Witt. *Elementa curvarum linearum (Elements of Curves)*. 1659. Elementa curvarum linearum was an appendix to de Witt's translation of of Descartes' *La Géométrie*.

Carl R. de Boor. On calculating with b-splines. *Journal of Approximation Theory*, pages 50–62, 1972.

Carl R. de Boor. *A Practical Guide to Splines*, volume 27 of *Applied Mathematical Sciences*. Springer, revised edition, 2001. ISBN 0387953663. URL http://books.google.com/books?vid=ISBN0387953663.

F.M. Dekking, C. Kraaikamp, H.P. Lopuhaa, and L.E. Meester. *A Modern Introduction to Probability and Statistics: Understanding Why and How*. Springer Texts in Statistics. Springer Science & Business Media, 2006. ISBN 9781846281686. URL http://books.google.com/books?vid=ISBN1846281687.

René Descartes. *La géométrie*. 1637. URL http://historical.library.cornell.edu/math/math_D.html.

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.

René Descartes. *The Geometry of Rene Descartes*. Courier Dover Publications, June 1 1954. ISBN 0486600688. URL http://books.google.com/books?vid=isbn0486600688. orginally published by Open Court Publishing, Chicago, 1925; translation of La géométrie.

Luc Devroye. *Non-Uniform Random Variate Generation*. Springer-Verlag, New York, 1986. ISBN 0387963057. URL http://www.eirene.de/Devroye.pdf.

Elena Deza and Michel-Marie Deza. *Dictionary of Distances*. Elsevier Science, Amsterdam, 2006. ISBN 0444520872. URL http://books.google.com/books?vid=ISBN0444520872.

Jean Alexandre Dieudonné. *Foundations of Modern Analysis*. Academic Press, New York, 1969. ISBN 1406727911. URL http://books.google.com/books?vid=ISBN1406727911.

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.

Bogdan Dumitrescu. *Positive Trigonometric Polynomials and Signal Processing Applications*. Signals and Communication Technology. Springer, 2007. ISBN 978-1-4020-5124-1. URL gen.lib.rus.ec/get?md5=5346e169091b2d928d8333cd053300f9.

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.

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.

Copyright © 2019 Daniel J. Greenhoe

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.

Jacques Dutka. On gauss' priority in the discovery of the method of least squares. *Archive for History of Exact Sciences*, 49(4):355–370, December 1995. URL https://www.jstor.org/stable/41134010.

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.

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.

D. J. Ewins. *Modal Testing: Theory and Practice*. Research Studies Press, "reprinted with amendments april 1986" edition, April 1986. ISBN 0863800173. URL http://www.amazon.com/dp/086380036X.

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.

Lorenzo Farina and Sergio Rinaldi. *Positive Linear Systems: Theory and Applications*. Pure and applied mathematics. John Wiley & Sons, 1 edition, July 3 2000. ISBN 9780471384564. URL http://books.google.com/books?vid=ISBN0471384569.

Richard M. Feldman and Ciriaco Valdez-Flores. *Applied Probability and Stochastic Processes*. Technology & Engineering. Springer Science & Business Media, 2 edition, 2010. ISBN 9783642051586. URL http://books.google.com/books?vid=ISBN3642051588.

Carlos A. Felippa. *Matrix Calculus*. University of Colorado at Boulder, August 18 1999. URL http://caswww.colorado.edu/courses.d/IFEM.d/.

William Feller. *An Introduction to Probability Theory and its Applications Volume I*. Wiley series in probability and mathematical statistics. John Wiley & Sons, 3, revised edition, 1970. ISBN 9780471257080. URL http://www.amazon.com/dp/0471257087.

R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transacations of the Royal Society*, January 1922. URL https://doi.org/10.1098/rsta.1922.0009.

G.L. Fix and G. Strang. Fourier analysis of the finite element method in ritz-galerkin theory. *Studies in Applied Mathematics*, 48:265–273, 1969.

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.

Brigitte Forster and Peter Massopust, editors. *Four Short Courses on Harmonic Analysis: Wavelets, Frames, Time-Frequency Methods, and Applications to Signal and Image Analysis*. Applied and Numerical Harmonic Analysis. Springer, November 19 2009. ISBN 9780817648909. URL http://books.google.com/books?vid=ISBN0817648909.

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.

Jean-Baptiste-Joseph Fourier. *Théorie Analytique de la Chaleur (The Analytical Theory of Heat)*. Chez Firmin Didot, pere et fils, Paris, 1822. URL http://books.google.com/books?vid=04X2vlqZx7hydlQUWEq&id=TDQJAAAAIAAJ.

Jean-Baptiste-Joseph Fourier. *The Analytical Theory of Heat (Théorie Analytique de la Chaleur)*. Cambridge University Press, Cambridge, February 20 1878. URL http://www.archive.org/details/analyticaltheory00fourrich. 1878 English translation of the original 1822 French edition. A 2003 Dover edition is also available: isbn 0486495310.

Maurice Fréchet. Sur l'extension du théorème des probabilités totales au cas d'une suite infinie d'evénéments. *Ren- diconti del Reale Istituto Lombardo di Scienze e Lettere*, 63:899–900, 1930a.

Maurice Fréchet. Sur l'extension du théorème des probabilités totales au cas d'une suite infinie d'evénéments. *Ren- diconti del Reale Istituto Lombardo di Scienze e Lettere*, 63:1059–1062, 1930b.

Erik Ivar Fredholm. Sur une nouvelle méthode pour la résolution du problème de dirichlet (on a new method for the resolution of the problem of dirichlet). */:Oefversigt af Kongl. Sv. Vetenskaps-Academiens Förhandlingar*, 57:39–66, 1900.

Erik Ivar Fredholm. Sur une classe d'equations fonctionnelles (on a class of functional equations). *Acta Mathematica*, 27(1):365–390, December 1903. ISSN 0001-5962. doi: 10.1007/BF02421317. URL http://www.springerlink.com/content/c41371137837p252/.

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.

Copyright © 2019 Daniel J. Greenhoe

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.

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.

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.

Israel Gohberg, Seymour Goldberg, and Marinus A. Kaashoek. *Basic Classes of Linear Operators*. Birkhäuser, Basel, 1 edition, 2003. ISBN 3764369302. URL http://books.google.com/books?vid=ISBN3764369302.

Steve Goldman. *Vibration Spectrum Analysis: A Practical Approach*. Industrial Press Inc., 1999. ISBN 9780831130886. URL http://books.google.com/books?vid=ISBN0831130881.

Jaideva C. Goswami and Andrew K. Chan. *Fundamentals of Wavelets; Theory, Algorithms, and Applications*. John Wiley & Sons, Inc., 1999. ISBN 0-471-19748-3. URL http://vadkudr.boom.ru/Collection/fundwave_contents.html.

H. G. D. Goyder. Foolproof methods for frequency response measurements. Technical Report ADP003677, United Kingdom Atomic Energy Authority London (England), 1984. URL https://apps.dtic.mil/dtic/tr/fulltext/u2/a143301.pdf#page=118.

I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, Series, and Products*. 1980. ISBN 0-12-294760-6. URL http://www.worldcat.org/isbn/0122947606.

Alexander Graham. *Kronecker Products and Matrix Calculus: With Applications*. Ellis Horwood Series; Mathematics and its Applications. Ellis Horwood Limited, Chichester, 1981. ISBN 0-85312-391-8. URL http://books.google.com/books?vid=ISBN0853123918.

Alexander Graham. *Kronecker Products and Matrix Calculus: With Applications*. Ellis Horwood Series; Mathematics and its Applications. Ellis Horwood Limited, 2018. ISBN 9780486824178. URL http://books.google.com/books?vid=ISBN9780486824178.

Robert M. Gray. Toeplitz and circulant matrices: A review. Technical Report AD0727139, Stanford University California Stanford Electronics Labs, Norwell, Massachusetts, June 1971. URL https://apps.dtic.mil/docs/citations/AD0727139.

Robert M. Gray. Toeplitz and circulant matrices: A review, January 31 2006. URL https://ee.stanford.edu/~gray/toeplitz.pdf.

R.J Greechie. Orthomodular lattices admitting no states. *Journal of Combinatorial Theory, Series A*, 10(2):119–132, March 1971. URL http://dx.doi.org/10.1016/0097-3165(71)90015-X.

Daniel J. Greenhoe. *Structure and Analysis of Mathematical Spaces*, volume 2 of *Mathematical Structure and Design series*. version 0.30 edition, January 2017a. URL https://www.researchgate.net/publication/312201065.

Daniel J. Greenhoe. The sequence a(n,m) of the m polynomial coefficients of the n-th order b-spline scaled by n!, read by rows, with n in 0,1,2,... and m in 1,2,3,...,(n+1)2. *Online Encyclopedia of Integer Sequences*, July 2017b. URL https://oeis.org/A289358.

Daniel J. Greenhoe. The effects of the assorted cross-correlation definitions. *ResearchGate.com*, June 2019. doi: 10.13140/RG.2.2.15881.70240. URL https://www.researchgate.net/publication/333882235.

Ulf Grenander and Gábor Szegö. *Toeplitz Forms and Their Applications*. California monographs in mathematical sciences. University of California Press, 1958. URL https://www.worldcat.org/title/toeplitz-forms-and-their-applications/oclc/648601989.

Ulf Grenander and Gábor Szegö. *Toeplitz Forms and Their Applications*. Chelsea Publishing Company, 2 edition, 1984. ISBN 9780828403214. URL http://books.google.com/books?vid=ISBN9780828403214.

Stanley Gudder. *Stochastic Methods in Quantum Mechanics*. North Holland, 1979. ISBN 0444002995. URL http://books.google.com/books?vid=ISBN0444002995.

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.

Alfréd Haar. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 69:331–371, September 1910. ISSN 1432-1807. doi: 10.1007/BF01456326.

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.

Paul R. Haddad and Ali N. Akansu. *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets*. Acedemic Press, October 1 1992. ISBN 0323138365. URL http://books.google.com/books?vid=ISBN0323138365.

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. *Finite Dimensional Vector Spaces*. Springer-Verlag, New York, 2 edition, 1958. ISBN 0-387-90093-4. URL http://books.google.com/books?vid=isbn0387900934.

Copyright © 2019 Daniel J. Greenhoe

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.

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, 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.

S. Haykin and S. Kesler. *Prediction-Error Filtering and Maximum-Entropy Spectral Estimation*, volume 34 of *Topics in Applied Physics*, pages 9–72. Springer-Verlag, 1 edition, 1979.

S. Haykin and S. Kesler. *Prediction-Error Filtering and Maximum-Entropy Spectral Estimation*, volume 34 of *Topics in Applied Physics*, pages 9–72. Springer-Verlag, "second corrected and updated edition" edition, 1983.

Simon Haykin. *Communication Systems*. John Wiley & Sons, 3 edition, 1994. ISBN 0471571768. URL http://www.amazon.com/dp/0471571768.

Simon Haykin. *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, 4 edition, September 24 2001. ISBN 978-0130901262. URL http://books.google.com/books?vid=isbn0130901261.

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.

Carl W. Helstrom. *Probability and Stochastic Processes for Engineers*. Maxwell Macmillan international editions in engineering. Macmillan, 2 edition, 1991. ISBN 9780023535710. URL http://books.google.com/books?vid=ISBN0023535717.

Omar Hijab. *Introduction to Calculus and Classical Analysis*. Undergraduate Texts in Mathematics. Springer, 3 edition, 2011. ISBN 9781441994875. URL http://books.google.com/books?vid=ISBN1441994874.

David Hilbert. Grundzüge einer allgemeinen theorie der linearen integralgleichungen (fundamentals of a general theory of the linear integral equations). *Mathematisch-Physikalische Klasse (Mathematical-Physical Class)*, pages 49–91, March 1904. URL http://dz-srv1.sub.uni-goettingen.de/sub/digbib/loader?did=D57552. Report 1 of 6.

David Hilbert. Grundzüge einer allgemeinen theorie der linearen integralgleichungen (fundamentals of a general theory of the linear integral equations). *Mathematisch-Physikalische Klasse (Mathematical-Physical Class)*, pages 157–228, March 1906. URL http://dz-srv1.sub.uni-goettingen.de/sub/digbib/loader?did=D58133. Report 4 of 6.

David Hilbert. *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen (Fundamentals of a general theory of the linear integral equations)*, volume 26 of *Fortschritte der mathematischen Wissenschaften in Monographien (Progress of the mathematical sciences in Monographs)*. B.G. Teubner, Leipzig und Berlin, 1912. URL http://www.worldcat.org/oclc/13468199.

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.

Otto Hölder. üeber einen mittelwerthssatz. *Göttingen Nachrichten*, pages 38–47, 1889. URL http://www.digizeitschriften.de/dms/img/?PPN=GDZPPN00252421X.

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.".

Klaus Höllig. *Finite Element Methods With B-Splines*, volume 26 of *Frontiers in Applied Mathematics*. SIAM, 2003. ISBN 0-89871-533-4. URL http://books.google.com/books?vid=ISBN0898715334.

Lars Hörmander. *Notions of Convexity*. Birkhäuser, Boston, 1994. ISBN 0817637990. URL http://books.google.com/books?vid=ISBN0817637990.

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.

K Husimi. Studies on the foundations of quantum mechanics i. *Proceedings of the Physico-Mathematical Society of Japan*, 19:766–789, 1937.

Copyright © 2019 Daniel J. Greenhoe

Emmanuel C. Ifeachor and Barrie W. Jervis. *Digital Signal Processing: A Practical Approach*. Electronic systems engineering series. Prentice Hall, 1993. ISBN 020154413X. URL http://www.amazon.com/dp/020154413X.

Emmanuel C. Ifeachor and Barrie W. Jervis. *Digital Signal Processing: A Practical Approach*. Electronic systems engineering series. Prentice Hall, 2002. ISBN 9780201596199. URL http://www.amazon.com/dp/0201596199.

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.

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.

Bjorn Jawerth and Wim Sweldens. An overview of wavelet based multiresolutional analysis. *SIAM Review*, 36:377–412, September 1994. URL http://cm.bell-labs.com/who/wim/papers/papers.html#overview.

Alan Jeffrey and Hui Hui Dai. *Handbook of Mathematical Formulas and Integrals*. Handbook of Mathematical Formulas and Integrals Series. Academic Press, 4 edition, January 18 2008. ISBN 9780080556840. URL http://books.google.com/books?vid=ISBN0080556841.

Gwilym M. Jenkins and Donald G. Watts. *Spectral analysis and its applications*. Holden-Day series in time series analysis. Holden-Day, 1968. URL https://archive.org/details/spectralanalysis00jenk_0.

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/.

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. *Applied Discrete Structures*. New Age International, New Delhi, 1997. ISBN 8122408265. URL http://books.google.com/books?vid=ISBN8122408265.

J.S.Chitode. *Signals And Systems*. Technical Publications, 2009a. ISBN 9788184316780. URL http://books.google.com/books?vid=ISBN818431678X.

J.S.Chitode. *Digital Signal Processing*. Technical Publications, 2009b. ISBN 9788184316469. URL http://books.google.com/books?vid=ISBN9788184316469.

Shizuo Kakutani and George W. Mackey. Ring and lattice characterizations of complex hilbert space. *Bulletin of the American Mathematical Societiy*, 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/.

Gudrun Kalmbach. *Measures and Hilbert Lattices*. World Scientific, Singapore, 1986. ISBN 9789971500092. URL http://books.google.com/books?vid=ISBN9971500094.

Nidal S. Kamel and K. S. Sim. Image signal-to-noise ratio and noise variance estimation using autoregressive model. *The Journal of Scanning Microscopes*, 26:277–281, September 15 2004. URL https://onlinelibrary.wiley.com/doi/pdf/10.1002/sca.4950260605.

Steven M. Kay. *Modern Spectral Estimation: Theory and Application*. Prentice-Hall signal processing series. Prentice Hall, 1988. ISBN 9788131733561. URL http://books.google.com/books?vid=ISBN8131733564.

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.

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.

John Francis Kenney. *Mathematics of Statistics, Volume 2*. D. Van Nostrand Company, New York, second edition, second printing edition, 1947.

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.

K.S. Kim and Nidal S. Kamel. Image signal-to-noise ratio and noise variance estimation using autoregressive model. *The Journal of Scanning Microscopes*, 26(6):135–139, September 24 2004. URL https://doi.org/10.1002/sca.4950260605.

Anthony W Knapp. *Advanced Real Analysis*. Cornerstones. Birkhäuser, Boston, Massachusetts, USA, 1 edition, July 29 2005a. ISBN 0817643826. URL http://books.google.com/books?vid=ISBN0817643826.

Anthony W Knapp. *Basic Real Analysis*. Cornerstones. Birkhäuser, Boston, Massachusetts, USA, 1 edition, July 29 2005b. ISBN 0817632506. URL http://books.google.com/books?vid=ISBN0817632506.

Copyright © 2019 Daniel J. Greenhoe

Andrei N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin, 1933a.

Andrei Nikolaevich Kolmogorov. *Foundations of the theory of probability*. Chelsea Publishing Company, New Yourk, 2 edition, 1933b. ISBN B0006AUOGO. URL http://statweb.stanford.edu/~cgates/PERSI/Courses/Phil166-266/Kolmogorov-Foundations.pdf. 1956 2nd edition English translation of A. N. Kolmogorov's 1933 "Grundbegriffe der Wahrscheinlichkeitsrechnung".

Lambert Herman Koopmans. volume 22 of *Probability and mathematical statistics*. Academic Press, 2 edition, 1995. ISBN 9780124192515.

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.

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.

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.

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.

Q. Leclere, Bert Roozen, and Celine Sandier. On the use of the hs estimator for the experimental assessment of transmissibility matrices. *Mechanical Systems and Signal Processing*, 43(1-2):237–245, 2014. URL https://hal.archives-ouvertes.fr/hal-00911916/document. <hal-00911916>.

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.

J. Leuridan, J. De Vis, and Van der Auweraer H. F. Lembregts D. "a comparison of some frequency response function measurement techniques. pages 908–918, 1986. URL https://www.researchgate.net/publication/283994172.

Zach Liang and George C. Lee. *Random Vibration: Mechanical, Structural, and Earthquake Engineering Applications*. Advances in Earthquake Engineering. Taylor & Francis, illustrated edition, 2015. ISBN 9781498702348. URL http://www.amazon.com/dp/1498702341.

Elliott H. Lieb and Michael Loss. *Analysis*, volume 14 of *Graduate studies in mathematics*. American Mathematical Soceity, Providence, Rhode Island, USA, 2 edition, 2001. ISBN 0821827839. URL http://books.google.com/books?vid=ISBN0821827839.

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.

Anders Lindquist and Giorgio Picci. *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification*, volume 1 of *Series in Contemporary Mathematics*. Springer, 2015. ISBN 9783662457504. URL http://books.google.com/books?vid=ISBN3662457504.

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.

Lynn H. Loomis and Ethan D. Bolker. *Harmonic analysis*. Mathematical Association of America, 1965. URL http://books.google.com/books?id=MEfvAAAAMAAJ.

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.

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.

Khosrow Maleknejad, Monireh Nosrati Sahlan, and Azadeh Ostad. Numerical solution of fractional integro-differential equation by using cubic b-spline wavelets. In *Proceedings of the World Congress on Engineering*, volume I, London, July 3–5 2013. URL https://pdfs.semanticscholar.org/bbe0/55f40768bbefe9e6476bdce701effb9243f7.pdf.

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.

R. Viswanathan Mandyam D. Srinath, P.K. Rajasekaran. *Introduction to Statistical Signal Processing with Applications*. Prentice Hall Inc, Upper Saddle River, 1996. ISBN 013125295X. URL http://engr.smu.edu/ee/mds/.

Dimitris G. Manolakis, Vinay K. Ingle, and Stephen M. Kogon. *Statistical and Adaptive Signal Processing*. Mc-Graw Hill, Boston, 2000. ISBN 0-07-040051-2.

Dimitris G. Manolakis, Vinay K. Ingle, and Stephen M. Kogon. *Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering, and Array Processing*. Artech House signal processing library. Artech House, illustrated, reprint edition, 2005. ISBN 9781580536103. URL http://www.amazon.com/dp/1580536107.

Copyright © 2019 Daniel J. Greenhoe

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.

MatLab. cpsd: Cross power spectral density. *MatLab*, accessed 2018 November 26 2018a. URL https://www.mathworks.com/help/signal/ref/cpsd.html.

MatLab. xcorr: Cross-correlation. *MatLab*, accessed 2018 November 26 2018b. URL https://www.mathworks.com/help/signal/ref/xcorr.html.

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.

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.

James Mercer. Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society of London*, 209:415–446, 1909. ISSN 02643952. URL http://www.jstor.org/stable/91043.

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.

Gregory K. Miller. *Probability: Modeling and Applications to Random Processes*. Wiley, August 25 2006. URL http://www.amazon.com/dp/0471458929.

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.

Fred Mintzer. Filters for distortion-free two-band multi-rate filter banks. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32, 1985.

Larry D. Mitchell. An improved method for the fast fourier transform (fft) calculation of the frequency response function. In M. Kunt and F. De Coulon, editors, *Proceedings of EUSIPCO-80, First European Signal Processing Conference*, Signal processing, theories and applications, pages 159–160, Lausanne, Switzerland, September 16–19 1980. Elsevier. URL http://www.worldcat.org/oclc/63416846. citation from Cobb(1988).

Larry D. Mitchell. Improved methods for the fast fourier transform (fft) calculation of the frequency response function. *Journal of Mechanical Design*, 104(2):277–279, April 01 1982. doi: 10.1115/1.3256336. URL http://mechanicaldesign.asmedigitalcollection.asme.org/article.aspx?articleid=1451670.

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.

Todd K. Moon and Wynn C. Stirling. *Mathematical Methods and Algorithms for Signal Processing*. Prentice Hall, Upper Saddle River, 2000. ISBN 0-201-36186-8. URL http://books.google.com/books?vid=isbn0201361868.

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.

Paul J. Nahin. *Dr. Euler's Fabulous Formula: Cures Many Mathematical Ills*. Princeton Science Library. Princeton University Press, reprint, revised edition, April 25 2011. ISBN 9781400838479. URL http://books.google.com/books?vid=ISBN1400838479.

Louis Narens. *Theories of Probability: An examination of logical and qualitative foundations*, volume 2 of *Advanced series on mathematical psychology*. World Scientific, January 2007. ISBN 9789812708014. URL http://www.amazon.com/dp/9812708014.

Louis Narens. Alternative probability theories for cognitive psychology. *Topics in Cognitive Science*, 6(1):114–120, January 2014. ISSN 1756-8757. doi: 10.1111/tops.12071. URL http://onlinelibrary.wiley.com/doi/10.1111/tops.12071/full.

Oliver Nelles. *Nonlinear System Identification*. Springer, New York, 2001. ISBN 9783540673699.

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.

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.

Brad Osgood. The fourier transform and its applications; ee261: Standford university course. university course notes, 2002. URL http://www.stanford.edu/class/ee261/.

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.

Anthanasios Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Series in Electrical Engineering. McGraw-Hill Book Company, New York, 2 edition, 1984. ISBN 0070484686. URL http://books.google.com/books?vid=ISBN0070484686.

Anthanasios Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, 3 edition, 1991. ISBN 0070484775. URL http://books.google.com/books?vid=ISBN0070484775.

Copyright © 2019 Daniel J. Greenhoe

Anthanasios Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill series in electrical and computer engineering. Tata McGraw-Hill, New York, 4 edition, 2002. ISBN 9780071226615. URL http://books.google.com/books?vid=ISBN0071226613.

Athanasios Papoulis. *Probability & Statistics*. Prentice-Hall international editions. Prentice Hall, illustrated edition, 1990. ISBN 9780137116980. URL http://books.google.com/books?vid=ISBN0137116985.

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.

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.

Rik Pintelon and Johan Schoukens. *System Identification: A Frequency Domain Approach*. John Wiley & Sons, 2 edition, 2012. ISBN 9781118287392. URL http://books.google.com/books?vid=ISBN1118287398.

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).

R. L. Plackett. Studies in the history of probability and statistics. xxix: The discovery of the method of least squares. *Biometrika*, 59(2):239–251, August 1972. doi: 10.2307/2334569. URL https://www.jstor.org/stable/2334569. https://doi.org/10.1093/biomet/59.2.239.

Henri Poincaré. *Calcul Des Probabilités*. De L'ecole Polytechnique, Paris, 1912. URL http://www.archive.org/details/calculdeprobabil00poinrich.

Karl Raimund Popper. *Conjectures and refutations: the growth of scientific knowledge*. Basic Books, 1962. URL https://books.google.com/books?id=OBgNAQAAIAAJ.

Karl Raimund Popper. *Conjectures and refutations: the growth of scientific knowledge*. Routledge, 2 edition, 1963. ISBN 9781135971304.

Lakshman Prasad and Sundararaja S. Iyengar. *Wavelet Analysis with Applications to Image Processing*. CRC Press LLC, Boca Raton, 1997. ISBN 978-0849331695. URL http://books.google.com/books?vid=ISBN0849331692. Library TA1637.P7 1997.

John G. Proakis. *Digital Communications*. McGraw Hill, 4 edition, 2001. ISBN 0-07-232111-3. URL http://www.mhhe.com/.

John G. Proakis, editor. *Algorithms for Statistical Signal Processing*. Prentice Hall, 2002. ISBN 9780130622198. URL http://books.google.com/books?vid=ISBN9780130622198.

John G. Proakis and Dimitris G. Manolakis. *Digital Signal Processing: Principles, Algorithms, and Applications*. Prentice Hall international editions. Prentice Hall, 3 edition, 1996. ISBN 9780133737622. URL http://books.google.com/books?vid=ISBN0133737624.

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.

K. Deergha Rao and M. N. S. Swamy. *Digital Signal Processing: Theory and Practice*. Springer, 2018. ISBN 9789811080814. URL http://books.google.com/books?vid=ISBN9789811080814.

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. 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.

Enders A. Robinson. *Random Wavelets and Cybernetic Systems*, volume 9 of *Griffins Statistical Monographs & Courses*. Lubrecht & Cramer Limited, London, June 1962. ISBN 0852640757. URL http://books.google.com/books?vid=ISBN0852640757.

Enders A. Robinson. Multichannel z-transforms and minimum delay. *Geophyics*, 31(3):482–500, June 1966. doi: 10.1190/1.1439788. URL http://dx.doi.org/10.1190/1.1439788.

R. Tyrrell Rockafellar. *Convex Analysis*. Princeton landmarks in mathematics. Princeton University Press, 1970. ISBN 9780691015866. URL http://books.google.com/books?vid=ISBN0691015864.

Murray Rosenblatt. Remarks on some non-parametric estimates of a density function. *Annals of Mathematical Statistics*, 27(3):832–837, September 1956. URL https://link.springer.com/content/pdf/10.1007/978-1-4419-8339-8_13.pdf. https://projecteuclid.org/download/pdf_1/euclid.aoms/1177728190.

Sheldon Ross. *A First Course in Probability*. Prentice Hall, Upper Saddle River, New Jersey, USA 07458, 5 edition, 1998. ISBN 0-13-746314-6. URL http://www.amazon.com/dp/0137463146.

George G. Roussas. *An Introduction to Probability and Statistical Inference*. Academic Press, 2 edition, October 21 2014. ISBN 9780128004371. URL http://books.google.com/books?vid=ISBN0128004371.

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.

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.

Jr. S. Lawrence Marple. *Digital Spectral Analysis with Applications*. Prentice Hall, 1 edition, 1987. ISBN 9780132141499. URL http://books.google.com/books?vid=ISBN9780132141499.

Copyright © 2019 Daniel J. Greenhoe

Jr. S. Lawrence Marple. *Digital Spectral Analysis*. Dover Books on Electrical Engineering. Courier Dover Publications, 2019. ISBN 9780486780528. URL http://books.google.com/books?vid=ISBN9780486780528.

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.

S. Salivahanan and A. Vallavaraj. *Digital Signal Processing*. Tata McGraw-Hill Education, 2001. ISBN 9780074639962. URL http://books.google.com/books?vid=ISBN9780074639962.

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.

Louis L. Scharf. *Statistical Signal Processing*. Addison-Wesley Publishing Company, Reading, MA, 1991. ISBN 0-201-19038-9.

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.

Larry L. Schumaker. *Spline Functions: Basic Theory*. Cambridge Mathematical Library. Cambridge University Press, 3 edition, August 2007. ISBN 978-0-521-70512-7. URL http://books.google.com/books?vid=ISBN0521705126.

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.

Jun Shao. *Mathematical Statistics*. Springer Texts in Statistics. Springer Science & Business Media, 2003. ISBN 9780387953823. URL http://books.google.com/books?vid=ISBN0387953825.

Kihong Shin and Joseph Hammond. *Fundamentals of Signal Processing for Sound and Vibration Engineers*. John Wiley & Sons, 2008. ISBN 9780470725641.

Bernard Shizgal. *Spectral Methods in Chemistry and Physics: Applications to Kinetic Theory and Quantum Mechanics*. Scientific Computation. Springer, January 7 2015. ISBN 9789401794541. URL http://books.google.com/books?vid=ISBN9401794545.

Bernard. W. Silverman. *Density Estimation for Statistics and Data Analysis*, volume 26 of *Monographs on Statistics & Applied Probability*. Chapman & Hall/CRC, illustrated, reprint edition, 1986. ISBN 9780412246203. URL http://books.google.com/books?vid=ISBN9780412246203.

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.

Herbert Ellsworth Slaught and Nels Johann Lennes. *Elementary Algebra*. Allyn and Bacon, 1915. URL https://archive.org/details/elementaryalgebr00slaurich/.

M.J.T. Smith and T.P. Barnwell. A procedure for designing exact reconstruction filter banks for tree-structured subband coders. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 9:421–424, 1984a. T.P. Barnwell is T.P. Barnwell III.

M.J.T. Smith and T.P. Barnwell. The design of digital filters for exact reconstruction in subband coding. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34(3):434–441, June 1984b. ISSN 0096-3518. doi: 10.1109/TASSP.1986.1164832. T.P. Barnwell is T.P. Barnwell III.

D. E. Smylie, G. K. C. Clarke, and T. J. Ulrych. *Analysis of Irregularities in the Earth's Rotation*, volume 13 of *Geophyics*, pages 391–340. Academic Press, 1973. ISBN 9780323148368. URL http://books.google.com/books?vid=ISBN9780323148368.

J. Laurie Snell. A conversation with joe doob. *Statistical Science*, 12(4):301–311, 1997. URL http://www.jstor.org/stable/2246220. similar article on web at http://www.dartmouth.edu/~chance/Doob/conversation.html.

J. Laurie Snell. Obituary: Joseph leonard doob. *Journal of Applied Probability*, 42(1):247–256, 2005. URL http://projecteuclid.org/euclid.jap/1110381384.

Houshang H. Sohrab. *Basic Real Analysis*. Birkhäuser, Boston, 1 edition, 2003. ISBN 978-0817642112. URL http://books.google.com/books?vid=ISBN0817642110.

H. W. Sorenson. Least-squares estimation: from gauss to kalman. *IEEE Spectrum*, pages 63–68, July 1970. URL https://www2.pv.infn.it/~fontana/download/lect/Sorenson.pdf.

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.

Copyright © 2019 Daniel J. Greenhoe

Stephen M. Stigler. Gauss and the invention of least squares. *The Annals of Statistics*, 9(3):465–474, May 1981. URL https://projecteuclid.org/euclid.aos/1176345451. https://www.jstor.org/stable/2240811.

Josef Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, volume 12 of *Texts in Applied Mathematics*. Springer Science & Business Media, 3 edition, August 21 2002. ISBN 9780387954523. URL http://books.google.com/books?vid=ISBN038795452X.

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.

Gilbert Strang. Wavelets and dilation equations: A brief introduction. *SIAM Review*, 31(4): 614–627, December 1989. ISSN 0036-1445. URL https://pdfs.semanticscholar.org/d118/a25a91488b4b37457cdc506472819de90523.pdf.

Gilbert Strang and Truong Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Wellesley, MA, 1996. ISBN 9780961408879. URL http://books.google.com/books?vid=ISBN0961408871.

Alan Stuart and J. Keith Ord. *Kendall's Advanced Theory of Statistics Volume 2 Classical Inference and Relationship*. Hodder & Stoughton, 5 edition, 1991. ISBN 9780340560235. URL http://books.google.com/books?vid=ISBN9780340560235.

Yuri M. Suhov, Michael Kelbert, and Mark Kelbert. *Probability and Statistics by Example: Volume 1, Basic Probability and Statistics*, volume 1 of *Probability and Statistics by Example*. Cambridge University Press, illustrated, reprint edition, October 13 2005. ISBN 9780521847667.

Wim Sweldens and Robert Piessens. Wavelet sampling techniques. In *1993 Proceedings of the Statistical Computing Section*, pages 20–29. American Statistical Association, August 1993. URL http://citeseer.ist.psu.edu/18531.html.

Tamás Szabó and Gábor Horváth. An efficient hardware implementation of feed-forward neural networks. *Applied Intelligence*, 1:143–158, 2004. URL https://www.researchgate.net/profile/Gabor_Horvath/publication/263107833.

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.

J. T. L. Thong, K. S. Sim, and J. C. H. Phang. Single-image signal-to-noise ratio estimation. *The Journal of Scanning Microscopes*, 23:328–336, December 07 2001. URL https://doi.org/10.1002/sca.4950230506.

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.

Harry L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I*. Wiley-Interscience, reprint edition edition, September 27 2001. ISBN 0471095176. URL http://ece.gmu.edu/faculty_info/van.html.

Harry L. Van Trees. *Otimum Array Processing; Part IV of Detection, Estimation, and Modulation Theory*. Wiley-Interscience, New York, 2002. ISBN 0-471-09390-4. URL http://ece.gmu.edu/faculty_info/van.html.

Kishor S. Trivedi. *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*. John Wiley & Sons, 2 edition, 2016. ISBN 9781119314202.

Howard G. Tucker. *An Introduction to Probability and Mathematical Statistics*. Academic press textbooks in mathematics. Academic Press, 1 edition, 1965.

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.

P.P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice Hall Signal Processing Series. Prentice Hall, 1993. ISBN 0136057187. URL http://books.google.com/books?vid=ISBN0136057187.

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.

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.

Saeed V. Vaseghi. *Advanced Digital Signal Processing and Noise Reduction*. John Wiley & Sons, 2 edition, 2000. ISBN 0471626929. URL http://books.google.com/books?vid=ISBN0471626929.

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.

Wolfgang von der Linden, Volker Dose, and Udo Von Toussaint. *Bayesian Probability Theory: Applications in the Physical Sciences*. Cambridge University Press, 2014. ISBN 9781107035904. URL http://books.google.com/books?vid=ISBN1107035902.

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.

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.

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.

Copyright © 2019 Daniel J. Greenhoe

H. L. Wei and S. A. Billings. An efficient nonlinear cardinal b-spline model for high tide forecasts at the venice lagoon. *Nonlinear Processes in Geophysics*, 13:577–584, 2006. URL https://www.nonlin-processes-geophys.net/13/577/2006/.

David Wells. *The Penguin Dictionary of Curious and Interesting Numbers*. revised edition, April 7 1987. ISBN 9780140080292. URL http://www.amazon.com/dp/0140080295.

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.

P. R. White, M. H. Tan, and J. K. Hammond. Analysis of the maximum likelihood, total least squares and principal component approaches for frequency response function estimation. *Journal of Sound and Vibration*, 290(3–5):676–689, March 2006. URL https://doi.org/10.1016/j.jsv.2005.04.029.

Alfred L. Wicks and Harvard Vold. The h_s frequency response function estimator. In *Proceedings of the 4th International Modal Analysis Conference*, pages 897–899, Los Angeles Airport Marriott Hotel, California, USA, February 3–6 1986. URL https://www.tib.eu/en/search/id/TIBKAT%3A246489537/Proceedings-of-the-4th-International-Modal-Analysis/.

Harold Widom. *Toeplitz Matrices*, volume 3 of *Studies in mathematics (Mathematical Association of America)*. Mathematical Association of America; distributed by Prentice-Hall, 1965. ISBN 9780883851036. URL https://www.worldcat.org/title/studies-in-real-and-complex-analysis/oclc/506377.

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.

Samuel Stanley Wilks. *Mathematical Statistics*. Wiley series in probability and mathematical statistics. Wiley, 1 edition, 1963a.

Samuel Stanley Wilks. *Mathematical Statistics*. Wiley series in probability and mathematical statistics. Wiley, "second printing with corrections" edition, 1963b. URL https://books.google.com/books?id=-sJLAAAAMAAJ.

Peter Willett, Peter F. Swaszek, and Rick S. Blum. The good, bad, and ugly: Distributed detection of a known signal in dependent gaussian noise. *IEEE Transactions on Signal Processing*, 48(12):3266–3279, December 2000.

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.

W. J. Yan and W. X. Ren. Operational modal parameter identification from power spectrum density transmissibility. *Computer-Aided Civil and Infrastructure Engineering*, 27(3):202–217, 2012. URL https://doi.org/10.1111/j.1467-8667.2011.00735.x.

Frank Yates and Kenneth Mather. Ronald aylmer fisher, 1890–1962. *Biographical Memoirs of Fellows of the Royal Society of London*, 9:90–120, November 1963. URL https://doi.org/10.1098/rsbm.1963.0006. print ISSN:0080-4606, online ISSN:1748-8494, https://pdfs.semanticscholar.org/6ccd/4303f68f3543fc794c376ce00395947a79b6.pdf.

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.

Yu Zhang, Tong Zhu, and Jing Zhou. Understanding power spectrum density transmissibility. *Shock and Vibration*, 2016(8479721):9, 2016. URL http://dx.doi.org/10.1155/2016/8479721.

Mu-Qiao Zheng, She-Ping Hu, and Li Li. Improved methods for the estimate of frequency response functions. pages 917–921, 2002. URL https://pdfs.semanticscholar.org/edfc/0572a727633e729204bc9a8cbce4eb3211e5.pdf.

Yun-Lai Zhou and Magd Abdel Wahab. Republished paper. rapid early damage detection using transmissibility with distance measure analysis under unknown excitation in long-term health monitoring. *Journal of Vibroengineering*, 20(1):823–831, 2018. doi: 10.21595/jve.2016.19718. URL https://doi.org/10.21595/jve.2016.19718.

Copyright © 2019 Daniel J. Greenhoe

REFERENCE INDEX

Abramovich and Aliprantis (2002), 208, 370
Aliprantis and Burkinshaw (2006), 214, 216–219
Abdaheer (2009), 99, 107
Abom (1986), 114
Aliprantis and Burkinshaw (1998), 228, 230, 234, 261, 271, 272, 274–277, 279, 280, 284, 287–289, 314, 320, 324, 326–328
Adobe (1999), 416
Akhiezer and Glazman (1993), 210, 263
Allemang et al. (1979), 113
Allemang et al. (1987), 114
Allen (2018), 6
Amann and Escher (2008), 316
Amemiya and Araki (1966), 264
Amir (1986), 279, 294–296
Anderson (1958), 188, 195
Anderson (1984), 188, 195
Angus (1994), 27
Apostol (1975), 297
Aubin and Frankowska (2009), 299
Aubin (2011), 299
Autonne (1901), 342
Autonne (1902), 342
Bachman (1964), 352
Bachman and Narici (1966), 221, 223, 225, 234, 262, 334, 336
Bachman et al. (2000), 226, 229, 238, 349
Bailey et al. (2011), 404
Bailey et al. (2013), 404
Bajpai (1967), 100
Bak (2013), 125
Balakrishnan and Lai (2009), 27
Banach (1922), 209, 212, 213, 287, 288
Banach (1932b), 287
Banach (1932a), 226, 287, 290
Barvinok (2002), 298, 299
Beer (1993), 299
Bell (1986), 314
Bendat and Piersol (1966), 67
Bendat (1978), 113
Bendat and Piersol (1980), 67, 129
Bendat and Piersol (1993), 113, 129
Bendat and Piersol (2010), 67, 113, 114, 129
?, 46
Berberian (1961), 211, 212, 221, 234, 238, 252, 257, 260, 261, 269, 271, 278, 305, 320–323, 339
Berkson (1963), 263
Bertero and Boccacci (1998), 336
Besso (1879), 308
Bienaymé (1840), 308
Billingsley (1995), 166
Birkhoff and Neumann (1936), 264
Blumenthal (1970), 303
Bollobás (1999), 191–193, 230, 271, 275, 277, 278, 284, 293, 299, 300, 306–308, 311, 328, 329
Bolstad (2007), 68
Bouniakowsky (1859), 314
Boyer and Merzbach (2011), 314
Bracewell (1978), 234, 352
Bromwich (1908), 369
Bullen (2003), 307, 308, 311–313, 315
Bunyakovsky (1859), 272
Burago et al. (2001), 264
Byrne (2005), 247
Cadzow (1987), 11, 12, 46
Caines (1988), 8
Caines (2018), 8
Callahan (2010), 28
Cardano (1545), 207
Carne and Dohrmann (2006), 116
Carothers (2000), 312, 313, 315
Carothers (2005), 223, 226
Cauchy (1821), 272, 311, 314
Cayley (1858), 191
Chen et al. (2012), 134
J.S.Chitode (2009b), 107
J.S.Chitode (2009a), 365
Choi (1978), 67, 68
Chorin and Hald (2009), 349, 350
Christensen (2003), 221, 226, 234, 239, 244, 247
Christensen (2008), 239, 244, 383, 386, 387, 392, 394, 396, 398, 400
Christensen (2010), 386
Chui (1988), 383, 392, 394, 400
Chui (1992), 362, 383, 387, 396
Cifarelli and Regazzini (1996), 166
Çinlar and Vanderbei (2013), 300
Claerbout (1976), 373
Clarkson (1993), 67
Cobb (1988), 111, 114, 118
Cobb and Mitchell (1990), 118
Cohn (2002), 208
Conway (1990), 275

Courant and Hilbert (1930), 194
Cover and Thomas (1991), 6, 302
Cox (1972), 392, 394, 400
Curry and Feldman (2010), 175
Daubechies (1992), 227, 243, 365
Day (1973), 279
de Witt (1659), 193
de Boor (1972), 392, 394, 400
de Boor (2001), 392, 394, 400
de Finetti (1930a), 166
de Finetti (1930b), 166
de Finetti (1930c), 166
Dekking et al. (2006), 6
Descartes (1954), 191
Descartes (1637), 191
Devroye (1986), 27, 31
Deza and Deza (2006), 263
Daubechies et al. (1986), 244
Dieudonné (1969), 337
Drljević (1989), 283, 296
Duffin and Schaeffer (1952), 244
Dumitrescu (2007), 373
Dunford and Schwartz (1957), 294, 296, 330
Durbin (2000), 207, 208
Duthie (1942), 297
Dutka (1995), 71
Edelman and Jamison (1985), 303
Edelman (1986), 303
Edwards (1995), 230
Eidelman et al. (2004), 322
Enflo (1973), 226
Ewen (1950), viii
Ewen (1961), viii
Ewins (1986), 112, 134
Fabian et al. (2010), 232
Farina and Rinaldi (2000), 373
Feldman and Valdez-Flores (2010), 175
Felippa (1999), 195
Feller (1970), 169
Fisher (1922), 81
Fix and Strang (1969), 355
Folland (1995), 265, 266, 269, 270
Forster and Massopust (2009), 406
Fourier (1820), 305
Fourier (1878), 349
Fourier (1822), 349
Fréchet (1930a), 166
Fréchet (1930b), 166
Fredholm (1900), 18, 193
Fredholm (1903), 18, 193
Frobenius (1968), 337
Frobenius (1878), 337
Fuchs (1995), 265
Gelfand (1941), 269
Gelfand and Neumark (1943b), 268–270
Gelfand and Neumark (1943a), 269, 270
Gelfand and Naimark (1964), 266, 267, 269
Giles (1987), 289, 290
Giles (2000), 226, 230, 288, 290, 330, 332
Gohberg et al. (2003), 194
Goldman (1999), 113
Goswami and Chan (1999), 354
Goyder (1984), 114, 118
Gradshteyn and Ryzhik (1980), 202
Graham (2018), 195, 199
Graham (1981), 195, 199–201
Gray (1971), 23
Gray (2006), 23
Greechie (1971), 166
Greenhoe (2017a), 390
Greenhoe (2019), 13
Grenander and Szegö (1958), 23
Grenander and Szegö (1984), 23
Gudder (1979), 264
Gudder (2005), 264
Gupta (1998), 225
Haar (1910), 412
Haaser and Sullivan (1991), 209, 210, 225, 252, 271–273, 276
Haddad and Akansu (1992), 377
Hall and Knight (1894), 316
Halmos (1948), 209
Halmos (1958), 252, 334
Halmos (1998b), 264
Halmos (1998a), 267, 279, 333
Hamel (1905), 211, 223
Han et al. (2007), 232, 247
Hardy (1929), 305
Hardy et al. (1952), 307, 312–314
Haykin and Kesler (1979), 23
Haykin and Kesler (1983), 23
Haykin (1994), 28
Hazewinkel (2000), 265
Heijenoort (1967), viii
Heil (2011), 221, 223, 225, 226, 234, 239, 244, 247, 275, 319, 386, 409
Heinbockel (2010), 404
Helstrom (1991), 46
Hijab (2011), 391
Hilbert (1904), 191
Hilbert (1906), 192, 193
Hilbert (1912), 191, 193
Hilbert et al. (1927), 320
Hölder (1889), 312
Holland (1970), 264
Höllig (2003), 391, 392, 394, 400
Hörmander (1994), 303
Horn and Johnson (1990), 76, 204, 324
Housman (1936), viii
Husimi (1937), 264
Ifeachor and Jervis (1993), 73
Ifeachor and Jervis (2002), 73, 107
Igari (1996), 242
Istrăţescu (1987), 216, 271, 278, 279
Iturrioz (1985), 256, 262, 264
James (1945), 283, 294, 296
James (1947), 296
Jawerth and Sweldens (1994), 354, 355, 400
Jeffrey and Dai (2008), 350
Jenkins and Watts (1968), 105
Jensen (1906), 6, 299, 300, 311
Jolley (1961), 403, 404
Jordan and von Neumann (1935), 277–279
Jørgensen et al. (2008), 244
Joshi (1997), 369
Kalmbach (1986), 166
Kamel and Sim (2004), 137
Kay (1988), 46, 67, 73, 98
Keener (1988), 18, 20, 21, 193, 336
Kenko (circa 1330), 420
Kenney (1947), 100
Khamsi and Kirk (2001), 303
Kim and Kamel (2004), 137
Knapp (2005a), 343
Knapp (2005b), 349, 350
Kolmogorov (1933b), 166
Kolmogorov (1933a), 166
Koopmans (1995), 46
Krein and Krasnoselski (1947), 263
Krein et al. (1948), 263
Kubrusly (2001), 209, 211, 221, 223, 225, 228, 234, 238, 256, 257, 260–262, 281, 282, 284, 320, 334, 339, 340
Kubrusly (2011), 257, 369
Kurdila and Zabarankin (2005), 221, 300
Lax (2002), 336
Lay (1982), 300
Leclere et al. (2014), 117, 123

COPYRIGHT © 2019 DANIEL J. GREENHOE

Leibniz (1679), 319
Leuridan et al. (1986), 111, 113
Liang and Lee (2015), 134
Lieb and Loss (2001), 204
von der Linden et al. (2014), 176
Lindenstrauss and Tzafriri (1977), 226
Lindquist and Picci (2015), 8
Loomis (1953), 281
Loomis (1955), 264
Loomis and Bolker (1965), 349
Machiavelli (1961), 419
Maeda and Maeda (1970), 298
Maleknejad et al. (2013), 388
Maligranda (1995), 276, 312, 313
Mallat (1989), 243
Mallat (1999), 243, 354, 355, 367, 377, 379, 407
Manolakis et al. (2005), 46
S. Lawrence Marple (1987), 13
S. Lawrence Marple (2019), 13
Massera and Schäffer (1958), 263
MatLab (2018a), 11
MatLab (2018b), 11
Mazur and Ulam (1932), 330
Mazur (1938), 269
Menger (1928), 303
Menini and Oystaeyen (2004), 316
Mercer (1909), 194
Michel and Herget (1993), 18, 191, 193, 194, 208, 211–213, 221, 223, 225, 252–254, 256, 257, 267, 268, 287, 290, 319, 321, 323, 326, 331, 332, 335, 337, 340–342, 346, 347
Manolakis et al. (2000), 78
Miller (2006), 175
Milovanović and Milovanović (1979), 307
Minkowski (1910), 276, 313
Mintzer (1985), 377
Mitchell (1980), 114
Mitchell (1982), 114, 130
Mitrinović et al. (2010), 300
Moon and Stirling (2000), 8, 188
Mulholland (1950), 307
Nahin (2011), 404
Narens (2014), 166
Nelles (2001), 65, 77–79
Noble and Daniel (1988), 346
Greenhoe (2017b), 390
Oikhberg and Rosenthal (2007), 292, 330
Ore (1935), 297
Osgood (2002), 181
Paine (2000), vi
Pap (1995), 166
Papoulis (1984), 11, 12, 18, 28, 30, 31, 46
Papoulis (1990), 6, 28, 33, 167, 168, 177, 184
Papoulis (2002), 46
Papoulis (1991), 17, 28, 30–32, 37–39, 41, 43, 47, 52, 152, 166, 168, 169, 176, 183
Peano (1888b), 209, 212, 213, 271
Peano (1888a), 212, 213
Pečarić et al. (1992), 307
Pedersen (2000), 191, 229, 230
Pintelon and Schoukens (2012), 114, 118
Plackett (1972), 71
Poincaré (1912), 182
Popper (1962), 111
Popper (1963), 111
Prasad and Iyengar (1997), 383, 394, 406, 407
Proakis and Manolakis (1996), 107
Proakis (2002), 107
Proakis (2001), 28, 30, 31, 184–186, 188
Rana (2002), 308
Rao and Swamy (2018), 107
Rickart (1960), 266–269
Riesz (1934), 279
Robinson (1962), 373
Robinson (1966), 373
Rockafellar (1970), 300
Rosenblatt (1956), 67
Ross (1998), 6, 7, 183–185
Roussas (2014), 27
Rudin (1991), 324, 326, 327, 329, 331–334, 336–338, 342, 343
Rudin (1976), 310
Rynne and Youngson (2008), 225
Sakai (1998), 333
Salivahanan and Vallavaraj (2001), 107
Sasaki (1954), 264
Scharf (1991), 195, 197, 200
Schauder (1927), 226
Schauder (1928), 226
Schumaker (2007), 385, 386, 388, 391, 392, 394, 400
Schur (1909), 342
Scharz (1885), 314
Schwarz (1885), 272
Searcóid (2002), 221
Shao (2003), 6, 68
Shin and Hammond (2008), 111, 113, 114, 117, 118, 123, 130, 132
Shizgal (2015), 388
Silverman (1986), 67
Simmons (2007), 419
Simon (2011), 299
Singer (1970), 226
Slaught and Lennes (1915), 124
Smith and Barnwell (1984a), 377
Smith and Barnwell (1984b), 377
Smylie et al. (1973), 23
Snell (1997), 175
Snell (2005), 175
Sohrab (2003), 225
Sorenson (1970), 71
Mandyam D. Srinath (1996), 65, 87, 91, 92
Steele (2004), 274
Steen (1973), 191, 342
Stigler (1981), 71
Stoer and Bulirsch (2002), 386
Stone (1932), 320, 332, 336
Strang and Nguyen (1996), 377, 408, 409
Strang (1989), 408, 409
Stuart and Ord (1991), 67, 100
Suhov et al. (2005), 6
Sweldens and Piessens (1993), 354, 355
Szabó and Horváth (2004), 388
Thomson et al. (2008), 369
Thong et al. (2001), 137
Tolsted (1964), 312, 313
Trivedi (2016), 175
Tucker (1965), 100
Udriste (1994), 300
Ulam (1991), 330
Vaidyanathan (1993), 377
Väisälä (2003), 330
Trees (2001), 143
Trees (2002), 195, 197, 200
Vaseghi (2000), 105
van de Vel (1993), 298, 303
Vidakovic (1999), 243, 354, 355, 367, 379
von Neumann (1929), 193, 332, 336
von Neumann (1960), 264
Walter and Shen (2001), 230
Weber (1893), 208
Wedderburn (1907), 253
Wei and Billings (2006), 388

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND

Wells (1987), 404
Weyl (1940), 230
White et al. (2006), 114, 117, 132
Wicks and Vold (1986), 117, 132
Widom (1965), 23
Wilks (1963a), 100, 102
Wilks (1963b), 4, 7, 11, 100
Willett et al. (2000), 160
Wojtaszczyk (1997), 239, 241–243, 392, 394, 399, 400, 406, 412
Yan and Ren (2012), 113
Yates and Mather (1963), 11
Young (1912), 312
Young (1980), 232
Young (2001), 223, 225–227, 232, 236, 237, 239, 244
Zhang et al. (2016), 113
Zheng et al. (2002), 114, 115, 117, 137
Zhou and Wahab (2018), 113

COPYRIGHT © 2019 DANIEL J. GREENHOE

SUBJECT INDEX

0 order B-spline, 401
0 order B-spline dyadic decomposition, 408
0th order B-spline, 385
1st order B-spline, 386, 401
1st order B-spline dyadic decomposition, 409
2nd order B-spline, 386, 401
2nd order B-spline dyadic decomposition, 409
3rd order B-spline, 388
4th order B-spline, 389
5th order B-spline, 389
6th order B-spline, 389
7th order B-spline, 389, 390
C^* algebra, **269, 270**
C^*-algebra, 270
$\Gamma(b)$, 183
$*$-algebra, **12**, **15**, **37–39**, **266**, 266–268, 333, **360**
$\langle \triangle \mid \triangledown \rangle$, 271
αth-order modified Bessel function of the first kind, **187**
nth derivative, 396
nth moment, **354**, 354, 355
nth order B-spline, 387, 392, 394, 400, 401, 407, 409, 410
nth order cardinal B-spline, **383**
nth-order B-spline, 399
N-tuple, 211, 287
ϕ-mean, **306**
(ESD), 46
*-algebra, 72
*-algebras, 333
LaTeX, vi
TeX-Gyre Project, vi
XeLaTeX, vi
attention markers, 148, 152, 281, 330
    problem, 191, 324, 329, 336, 338
σ-additive, 166, 172
σ-additivity, 166
σ-algebra, 167, 172, 174
0 order B-spline, 398
1st order B-spline, 399
2 coefficient case, **412**
2nd order B-spline, 399
4th central moment, 101

inverse, 142

Abel, 314
Abel, Niels Henrik, 419
abelian group, 210
absolute value, x, **208**, **216**, 217–219, 272, 287, 289
absolutely summable, 15, 370, 371
ADC, 113
additive, 20, 166–169, 171, 259, 271, 272, 281, 283, 284, 320, 323
additive identity, 12, 68, 211–213, 321
additive identity element, 211, 212
additive inverse, 212, 213, 321
additive property, 350
additivity, 20, 166, 167, 248, 275, 332
adjoint, 268, **329**, 329, **332**, 333, 350
admissibility condition, 412
admissibility equation, 412
Adobe Systems Incorporated, vi
affine, 192, **299**, 330
Affine equations, **199**
affine equations, 73
algebra, 207, **208**, 209, **265**, 265, 266
algebra of sets, xi, 173
algebraic, 209
algebras
    C^*-algebra, 269
    $*$-algebra, 266
algebras of sets, 173
Alternating Series Test, 406
AM-GM inequality, 311
analog-to-digital converter, 113
AND, xi
anti-derivative, 384
anti-symmetric, 13, 14, 50, 262, 353
anti-symmetric , 50
antiautomorphic, 15, 266, 267, 272, 333
antilinear, 267
antitone, 168, 171, 257, 259, 262
aperature, 263
Aristotle, 11, 17
arithmetic, 100
arithmetic mean, 97, 99, 114, **311**
arithmetic mean estimate, 67
arithmetic mean geometric mean inequality, 311
Arithmetic Mean transfer function estimate, **114**
associates, 210
associative, 5, 207, 208, 210, 212, 305, 306, 322, 323, 342
asymmetric, 373
asymptotically unbiased, 105
auto-correlation, **12**, 13, 19, 50
auto-correlation function, **18**
auto-correlation operator,

19, 21
auto-power spectrum, 401
auto-spectral density, **14**, 16
autocorrelation, 336
Avant-Garde, vi
average, **98**, 98–101, **311**
average power, **46**
AWGN, 81
Axiom of Continuity, 166

B-spline, 36, 388, 390, 396, 397, 399, 400, 407, 410, 414–416
B-spline dyadic decomposition, 407, 413
B-spline general form, 394
B-spline partition of unity, 400
B-spline recursion, 392
B-spline scaling coefficients, **410**
B-splines, 401
Bak, Per, 125
ball, 298
Banach algebra, **269**
Banach space, 226, 227
Bartlett Method, 107
basis, 21, 22, 223–225, 232, 400
  frame, 244
  orthogonal, 231
  orthonormal, 231
  Riesz, 239
  tight frame, 244
basis sequence, 399
Battle-Lemarié orthogonalization, **243**
Bayesian, 65
Bayesian estimate, **66**
Bessel's correction, 100
Bessel's Equality, **229**, 230, 231, 235
Bessel's Inequality, 237
Bessel's inequality, **230**
Best Approximation Theorem, **230**, 230
biased, 100, 101, 114
biased sample variance, **100**, 100
bijection, 237
bijective, xi, 237, 330
bilinear functionals, 278
binomial coefficient, 385, 386
Binomial Theorem, 5, 68, 102
binomial theorem, 411
biorthogonal, **239**, 241, 242
Birkhoff orthogonality, 294, **296**, 296
Blackman-Tukey Method, 107
Bohr, Harald, 305
Boole's inequality, 169
Boolean, 167, 169
Boolean algebra, 172
Boolean lattice, 166, 169
Borel algebra on $\mathbb{R}$, 172
Borel measure, 349
Borel sets, 349
bounded, xi, 168, 227, **327**, 336, 347
bounded bijective, 240
bounded lattice, 169
bounded linear operator, 343
bounded linear operators, **327**, 329, 331, 332, 334, 335, 337, 338, 340–343
bounded operator, 327

C star algebra, **269**
calculus, 209
Cardano, Gerolamo, 207
Carl Spitzweg, 419
Cartesian product, x
Cauchy, 261
Cauchy Schwartz inequality, 120, 122
Cauchy-Bunyakovsky-Schwarz Inequality, 272
Cauchy-Schwartz inequality, 9, 99
Cauchy-Schwarz Inequality, 95, 99, **272**, 272, 275, 276
Cauchy-Schwarz inequality, 245, 315
Cauchy-Schwarz Inequality for inner-product spaces, 315
Cauchy-Schwarz inequality for inner-product spaces, 315
Cauchy-Schwarz Inequality for sequences, **314**, 315
Cauchy-Schwarz inequality for sequences, 315
causal, **374**, 374
cdf, 27, **176**, 176
Chain Rule, 391
chain rule, 29, 201
channel capacity, 144
channel coding, 144
channel statistics, 144
characteristic function, x
Chebyshev's inequality, 7
chi-square distribution, **184**
Chi-squared distribution, 184
chi-squared distribution, 184
Chi-squared distribution with n degrees of freedom, **185**, 185
chi-squared distribution with n degrees of freedom, 185
closed, 259, 261
closed interval, **297**, **298**, 298
Closed Set Theorem, 259
closed sets, 254
closed support, 394
closure, 225, 394
coefficient functionals, **226**, 226
coefficients, 375
coherence, 133, 134
colored, 19
communication system, 128, 141
communications additive noise cross-correlation, 57
communications additive noise model, **112**, 125
Communications additive noise system, 57
communications LTI additive noise model, **129**, **137**
commutative, 12, 208, 210, 217, 322, 369, 370
commutative group, 210
commutative ring, 207, **208**, 208
compact, 193
compact support, 407
comparable, 297
comparison test, 403
complement, x
complemented, 262
complete, **225**
complete metric space, 225
complete set, 225
complex, 122
complex coherence, **134**
complex exponential, 397
complex linear space, **210**
complex transmissibility, **116**, 116, 134
complex-valued, 116
concave, **298**, 298, **299**, 299, **300**, 300, 307, 308
conditional cumulative distribution function, **176**, 176
conditional probability, 67, 83, 169
conditional probability density function, **176**, 176
conditional probability of x given y, **168**
conjuage symmetric, 18
conjugate linear, 266, 267, 333
conjugate pairs, 375
conjugate quadrature filter, **377**, 377
conjugate quadrature filter

COPYRIGHT © 2019 DANIEL J. GREENHOE

condition, **377**
Conjugate quadrature filters, 377
conjugate symmetric, 14, 15, 18, 50, 194, 271, 272, 280, 281, 284
conjugate symmetric property, 350
conjugate symmetry, 20
conjugate-symmetric, 15
consistent, 98, 99, 101, 105
constant, 12, 45, 67
constraint, 142
Continuity, 290
continuity, 177, 178
continuous, xi, 27, 170, 177, 194, 269, 275, 290, 306, 308, 322, 396, 402, 409
Continuous data whitening, 94
continuous point spectrum, 397
continuous spectrum, **192**
contradiction, 212
convergence, 403
convergence in probability, 21
convergent, 46, 227
converges, 403–405
converges absolutely, 194
Convex, 300
convex, 5, 6, 252, 253, 261, 290, 292, 298, **299**, 299, **300**, 300, 307, 308
functional, 299
strictly, 299
convex function, 299, 300
convex set, **298**, 298–300
convexity, 261
convolution, 33, 34, 37, 38, 40, 47, 59, 184, 352, **369**, 369, 370, 383
convolution operation, **352**
convolution operator, 369, 370
Convolution Theorem, 40, 398
convolution theorem, **352**, 357, **372**
coordinate functionals, **226**
coordinates, **223**
correlated, 54, 112, 113
cost function, 65, 72, 125
countably infinite, 193, 226
counting measure, xi
CQF, **377**, 377, 378
CQF condition, **377**, 379
CQF theorem, **377**, 380
Cramér-Rao bound, 92
Cramér-Rao lower bound, 91
criterion, 142
critical point, **28**, 28, 30, 31
cross spectral density, **14**
cross-correlation, **12**, 13, **18**, 18
CS Inequality, 248
CSD, **14**
cummulative distribution function, **176**, 176
cumulative distribution function, 27, 170

de Morgan, 263
decision region, 159, 160
decreasing, 307, 406
definitions
C^* algebra, **269**, **270**
$*$-algebra, **12**, **15**, **37–39**, **266**, **360**
nth order cardinal B-spline, **383**
absolute value, **216**
algebra, **208**, **265**
Banach algebra, **269**
biorthogonal, **239**
Birkhoff orthogonality, **296**
bounded, **327**
bounded linear operators, **327**
C star algebra, **269**
closed interval, **297**, **298**
coefficient functionals, **226**
commutative ring, **208**
complete, **225**
complex linear space, **210**
conditional probability of x given y, **168**
continuous spectrum, **192**
convex set, **298**
coordinate functionals, **226**
coordinates, **223**
CQF, **377**
difference, **212**
disjoint, **218**
disjoint complement, **219**
eigenspace, **191**
epigraph, **299**
equal, **319**
equivalent, **227**
even/odd die toss, **171**
event, **167**
exact frame, **244**
expansion, **223**, **226**
field, **208**
Fourier coefficients, **232**
Fourier expansion, **232**
Fourier series, **232**
frame, **244**
frame bounds, **244**
fundamental, **225**
group, **207**
half-open interval, **297**, **298**
Hamel basis, **223**
hermitian, **266**
Hessian matrix, **205**
hypograph, **299**
inner product space, **271**
intervals on lattices, **297**
intervals on ordered sets, **297**
isometric, **339**
isosceles orthogonality, **294**
James orthogonality, **294**
linear basis, **223**
linear combination, **221**
linear manifold, **252**
linear space, **210**
linear subspace, **210**, **252**
minimum phase, **373**
multiplicative condition, **269**
negative part, **216**
norm, **287**
norm induced by the inner product, **277**
normal, **266**
normalized tight frame, **244**
normed algebra, **269**
normed linear space, **324**
normed space of linear operators, **324**
open interval, **297**, **298**
operator norm, **324**
optimal lower frame bound, **244**
optimal upper frame bound, **244**
ordered linear space, **214**
orthogonal basis, **232**
orthogonal complement, **257**
orthogonal in the Pythagorean sense, **296**
orthogonal in the sense of Birkhoff, **296**
orthogonal in the sense of James, **294**
orthonormal basis, **232**
outcome, **167**
Parseval frame, **244**

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

point spectrum, **192**
positive, **214**, **346**
positive cone, **214**
positive part, **216**
power mean, **308**
probability space, **167**, **172**
projection, **266**
Pythagorean orthogonality, **296**
real linear space, **210**
residual spectrum, **192**
resolvent, **265**
resolvent set, **192**
Riesz basis, **239**
Riesz space, **214**
ring, **207**
scalars, **210**
Schauder basis, **226**
self-adjoint, **266**
sequence, **369**
set projection operators, **159**
Smith-Barnwell filter, **377**
space of all absolutely square summable sequences, **369**
spans, **221**
spectral radius, **265**
spectrum, **192**, **265**
star-algebra, **266**
subspace intersection, **253**
subspace union, **253**
summation, **305**
tight frame, **244**
total, **225**
underlying set, **210**
unital, **265**
unitary, **342**
vector lattice, **214**
vector space, **210**
vectors, **210**
delay, 345
dense, 225, 226
Descartes, René, ix
detection, 65
deterministic, 71
difference, x, **212**
differentiable, 28
differentiable function, 28, 30, 31
differential operator, 397
dilation, 344
dilation equation, 410, 413
dimension, **223**
direction, 77
Dirichlet Lambda function, 403, 404
Discrete data whitening, 94
Discrete Time Fourier Series, xii
Discrete Time Fourier Transform, xii, 14, 359
discrete-time Fourier transform, **359**, 359–361, 365
discriminate, 274
disjoint, **218**
disjoint complement, **219**
distance, 289
distributes, 210
distributive, 12, 37–39, 72, 166–169, 212, 266–268, 271, 272, 333, 360, 371
distributive pseudocomplemented lattice, 166
domain, x
Double coin toss, **170**, 170
DTFT, 14–16, 111, 227, 360, 364, 367, 377–379, 411
DTFT periodicity, **359**
dyadic decomposition, 407

efficient, 88, 90–93
eigen-system, 22
Eigendecomposition, 246
eigenspace, **191**
eigenvalue, **191**, 246
eigenvector, **191**, 246
empty set, xi
energy, 373
energy spectral density, **46**
ensemble average, 46
epigraph, **299**
equal, **319**
equalities
Bessel's, 229
equality
triangle, 288
equality by definition, x
equality relation, x
equations
parallelogram law, 279
equivalent, **227**, 227, 239
ergodic in the mean, **46**, 46, 47
error vector, 72
estimate, **66**, 111, 143
Euclidean metric, 292
Euclidean space, 298
Euler formulas, 356, 398
even, 364
even/odd die, 172
even/odd die probability space, 171
even/odd die toss, **171**
event, **167**, 167, 168
events, 167
exact frame, **244**
examples
2 coefficient case, **412**
Double coin toss, **170**
even/odd die, 172
even/odd die toss, **171**
Mercedes Frame, **247**
Peace Frame, **247**
Rectangular pulse, 357
rectangular pulse, **356**, 398
single coin toss, **170**, 172
The usual norm, **289**
triangle, **357**
tuples in $\mathbb{F}^N$, **211**
excluded middle, 168, 169
exclusive OR, xi
existential quantifier, xi
exists, 46
expansion, **223**, **226**
expectation, 5, 9, 112
expectation operator, **3**, 3, 4, 6

false, x
FFT, 13
field, 207, **208**, 209, 211, 369
field of complex numbers, 333
field of probability, 166
filter banks, 377
finite modular lattice, 166
FontLab Studio, vi
for each, xi
Fourier analysis, 396
fourier analysis, 349
Fourier coefficients, **232**
Fourier expansion, **232**, 232, 233, **236**, 237, 238
Fourier kernel, **349**
Fourier Series, xi
Fourier series, **232**
Fourier series expansion, 234
Fourier Transform, xi, xii, 47, 243, 344, 345, 349, **350**, 353, 398
adjoint, **350**
Fourier transform, 227, 345, 354, 356, 357, 397
inverse, **350**
Fourier transform magnitude, 398, 399
Fourier transform scaling factor, 350
Fourier, Joseph, 349
frame, 242, **244**, 244, 248
frame bound, 245, 247
frame bounds, **244**, 248
frame operator, **244**, 244, 245, 247
frames, 222
Fredholm integral equation of the first kind, 18, **193**
Fredholm integral operators, 344
Fredholm operator, **193**, 193,

COPYRIGHT © 2019 DANIEL J. GREENHOE

194
Fredholm operators, 193
Free Software Foundation, vi
Frequency Response Function, 111
Frequency Response Identification, 111
Frequency Shift Keying
coherent, 147
FRF, 111
FSK
coherent, 147
function, 28, 175, 176, 319, 349, 384
even, 364
step, 383
function overload, 13
functional, 271, 300, 333
functions, xi
0 order B-spline, 401
0th order B-spline, 385
1st order B-spline, 386, 401
2nd order B-spline, 386, 401
3rd order B-spline, 388
4th order B-spline, 389
5th order B-spline, 389
6th order B-spline, 389
7th order B-spline, 389, 390
αth-order modified Bessel function of the first kind, **187**
nth derivative, 396
nth moment, 354
nth order B-spline, 387, 392, 394, 400, 401, 407, 409, 410
nth-order B-spline, 399
ϕ-mean, **306**
(ESD), 46
0 order B-spline, 398
1st order B-spline, 399
2nd order B-spline, 399
4th central moment, 101
absolute value, **208**, 217, 272, 287
adjoint, 268
anti-derivative, 384
arithmetic, 100
arithmetic mean, 97, 99, 114, **311**
arithmetic mean estimate, 67
Arithmetic Mean transfer function estimate, **114**
auto-correlation, **12**, 13, 19, 50
auto-correlation function, **18**
auto-power spectrum, 401
auto-spectral density, 16
average, **98**, 98–101, **311**
average power, **46**
B-spline, 36, 388, 390, 396, 397, 399, 400, 407, 410, 414–416
B-splines, 401
Bayesian estimate, **66**
biased sample variance, **100**, 100
binomial coefficient, 385, 386
Borel measure, 349
cdf, 27, **176**, 176
chi-square distribution, **184**
Chi-squared distribution with n degrees of freedom, **185**
coherence, 133, 134
complex coherence, **134**
complex exponential, 397
complex transmissibility, **116**, 116, 134
conditional cumulative distribution function, **176**, 176
conditional probability, 67, 83
conditional probability density function, **176**, 176
conjugate quadrature filter, **377**
continuous point spectrum, 397
cost function, 65, 72, 125
cross-correlation, **12**, 13, **18**
cummulative distribution function, **176**, 176
cumulative distribution function, 27, 170
dilation equation, 410, 413
Dirichlet Lambda function, 403, 404
Discrete Time Fourier Transform, 359
discrete-time Fourier transform, 359–361
DTFT, 364, 367, 378, 379
eigenvector, 246
energy spectral density, **46**
error vector, 72
estimate, **66**
Fourier kernel, **349**
Fourier transform, 354, 356, 357
Fredholm integral equation of the first kind, **193**
function, 176
Gamma distribution, **184**
gamma function, **183**
Gaussian distribution, **183**, **188**
generalized probability, **166**, 167
Geometric mean, 115
geometric mean, 115, **311**
geometric mean estimator, 115
Geometric mean transfer function estimate, **115**
Hanning window, 99
Harmonic mean, 116, 117
harmonic mean, **311**
Harmonic mean transfer function estimate, **116**
impulse response, 14, 15, 37, 39, 40, 58, 111
induced norm, 235, 276
inner product, 8, 9, 99, **271**, 276, 281, 349, 396
input, 14, 15, 37, 39
integer sequence, 390
Inverse Method transfer function estimate, **114**
Jensen's inequality, 5
joint cumulative distribution function, **176**, 176
Joint Gaussian distribution, 187
Joint Gaussian pdf, **188**
joint probability density function, **176**, 176
kernel, 397
Kronecker delta, 22
Kronecker delta function, **283**
kronecker delta function, 365
Least Squares Technique, 113
Least Squares transfer function estimate, **114**
linear functional, 329
magnitude, 115
MAP estimate, **66**, 81
maximum a-posteriori probability estimate, **66**
maximum likelihood estimate, **66**
mean, **3**, **12**, 12, **18**, 37
mean integrated square

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

error, **67**
mean square error, **67**, 67, 99
Mean square estimate, **66**
measure, 3, 166, 167, 170, 176, 177
measure-theoretic, 167
mini-max estimate, **66**
ML estimate, **66**, 81, 83, 86, 88, 92
MM estimate, **66**
modulus, **208**
MS estimate, **66**
non-central chi-squared distribution, 186
noncentral chi-square distribution, **186**
noncentral chi-square with n-degrees of freedom, **187**
norm, 8, 9, 72, 99, 324
normal distribution, **183**, **188**
normalized rms error, **67**, 67
order n B-spline, 390
order n B-spline MRA coefficient sequence, 410
order n B-spline MRA scaling sequence, 410
order-k moment estimate, **103**
ordinary coherence, **134**
ordinary transmissibility, 116
output, 14, 15, 37, 39
pdf, **176**, 176
phase, 115, 116, 122
power mean, 308, 311
power spectral density, 16, 46
power sum, 390
probability, **167**
probability density function, 3, 27, 36, 170, **176**, 176, 184
probability function, **166**
quantization noise, 113
quantum probability, **166**, 167
random process, **17**, 19, 21, 46
random sequence, **11**, 11–14, 37, 41, 112
random variable, 3–8, **17**, 27, 28, 30, 31, 33, 35, 170, 175, **176**, 176, 177, 184
Rayleigh distribution, **187**
Rice distribution, **187**
Riesz sequence, **239**
sample variance, **100**
scalar product, **271**
scaling coefficient sequence, 412
scaling function, 409, 410
Scaling transfer function estimate, **117**, 117
sequence, 11, 36, 410
set indicator function, 357
signal-to-noise ratio, 113
sine sweep, 111
SNR, 113
spectral power, 16, **46**, 113
step function, **383**, 383–387
Total Least Squares transfer function estimate, **117**
Total least squares transfer function estimate, 117
traditional probability, **166**, 167
transfer function estimate $\hat{\mathsf{H}}_\kappa(\omega;\kappa)$, **117**
transfer function estimate $\hat{\mathsf{H}}_c(\omega)$, **118**
translation operator, 355
Transmissibility, 116
transmissibility, 113, 115
transmissibility $\tilde{\mathsf{T}}_{xy}(\omega)$, **113**
uniform distribution, **181**, 181
variance, **3**, **12**, 12, 97
wavelet, 407
windowed autocorrelation estimate, **105**
Z transform, 411
z transform, 373
Z-transform, 361
z-transform, 378
fundamental, **225**
Fundamental Theorem of Calculus, 384, 391
Fundamental theorem of calculus, 177, 179
Fundamental theorem of linear equations, 323

Gamma distribution, **184**, 184, 185
gamma function, **183**, 183
gap metric, 263
Gaussian, 35, 82, 84, 86
Gaussian distribution, 182, **183**, **188**
Gelfand-Mazur Theorem, **269**
General ML estimation, **86**
Generalized AM-GM inequality, 311
generalized arithmetic mean geometric mean inequality, 311
Generalized associative property, **305**
Generalized Parseval's Identity, 234
generalized probability, **166**, 167
Generalized triangle inequality, 9
geodesically between, **298**
Geometric mean, 115
geometric mean, 115, **311**
geometric mean estimator, 115
Geometric mean transfer function estimate, **115**
Geometric Series, **316**
geometry, 209
GNU Octave
for, 415
step, 415
Golden Hind, vi
gradient , 77
gradient of $\boldsymbol{y}$ with respect to $\boldsymbol{x}$, **195**
gradient of $\boldsymbol{y}^T$ with respect to $\boldsymbol{x}$, **195**
greatest lower bound, xi
group, **207**, 207
Gutenberg Press, vi

Hölder inequality, 315
Hölder's Inequality, **312**, 314
half-open interval, **297**, **298**
Hamel bases, 222
Hamel basis, **223**, 223, 225
Handbook of Algebras, 265
Hanning window, 99
Hardy, G.H., 305
harmonic analysis, 349
Harmonic mean, 116, 117
harmonic mean, **311**
Harmonic mean transfer function estimate, **116**
hermitian, **266**, 266, 267, **336**
hermitian components, 268
Hermitian representation, **268**
Hermitian symmetric, 73, 353
Hermitian Transpose, 73
Hessian, 78
Hessian matrix, **204**, **205**

COPYRIGHT © 2019 DANIEL J. GREENHOE

Heuristica, vi
high-pass filter, 377
Hilbert space, 232, 234, 236, 238, 239, 242, 244, 261, 332, 333, 336–338, 343, 349
homogeneous, 5, 208, 232, 233, 236, 240, 242, 259, 271, 272, 276, 280, 281, 283, 287, 293, 314, 320, 321, 323, 324
Housman, Alfred Edward, vii
hypograph, **299**
hypotenuse, 296

identities
Fourier expansion, 232, 233
Parseval frame, 232, 233
identity, 207, 210, 319
identity element, 319
identity operator, 142, 198, 319, **320**
if, xi
if and only if, xi
image, x
image set, 321, 323, 334–338, 343
imaginary part, xi, **267**
implied by, xi
implies, xi
implies and is implied by, xi
impossible, 400
impulse response, 14, 15, 37, 39, 40, 58, 111
inclusive OR, xi
increasing, 307, 406
independence, 83, 86, 171
independent, 7, 33, 36, 82, 86, 87, 144, **168**, 170, 184
indicator function, x
induced norm, 235, **276**, 276
induces, 289
induction, 390, 391, 400
inequalities
AM-GM, 311
Cauchy-Bunyakovsky-Schwarz, 272, 314
Cauchy-Bunyakovsky-Schwarz Inequality, 272
Cauchy-Schwartz inequality, 9, 99
Cauchy-Schwarz, 272, 314
Cauchy-Schwarz Inequality, 99, **272**, 272, 275, 276
Cauchy-Schwarz inequality, 315
Cauchy-Schwarz Inequality for inner-product spaces, 315
Cauchy-Schwarz inequality for inner-product spaces, 315
Cauchy-Schwarz Inequality for sequences, **314**, 315
Cauchy-Schwarz inequality for sequences, 315
Generalized triangle inequality, 9
Hölder, 312
Hölder inequality, 315
Hölder's Inequality, **312**, 314
Jensen's, 300
Jensen's Inequality, 307, 311
Minkowski (sequences), 313
Minkowski's Inequality, **276**, 307
Minkowski's Inequality for sequences, **313**
quadratic discriminate inequality, **274**
triangle inequality for vectors, 284
Young, 312
Young's Inequality, **312**
inequality
Bessel's, 230
triangle, 287, 324
infinite sum, 221
injective, xi, 322
inner product, 8, 9, 99, 209, **271**, 276, 277, 281, 349, 396
Polarization Identity, 278
uniqueness, 277
inner product space, 8, 9, 209, 228–231, 239, 257, 259, 260, **271**, 272, 275–277, 279, 283, 284, 294, 396
inner-product, xi
inner-product space, 294, 296
input, 14, 15, 37, 39
inside, 374
instantaneous error, 78
integer sequence, 390
integral, 396
integral domain, 207
integral test, 404
intersection, x, 167
interval, 297, 298
intervals on lattices, **297**
intervals on ordered sets, **297**
into, 237, 238
inverse, 207, 210, 245, 247, 412
inverse DTFT, **365**, 379
Inverse Fourier Transform, 344
Inverse Fourier transform, **350**
inverse Fourier Transform, 345
Inverse Method transfer function estimate, **114**
Inverse probability integral transform, **27**
invertible, 28, 170, 227, 244, 266
involutary, 333
involution, **266**, 266, 270
involutory, 15, 262, 266–268
irreflexive ordering relation, xi
isometric, 233, 241, 242, 330, **339**, 339, 343, 351
isometric in distance, 343
isometric in length, 343
isometric operator, 337, 340–342
isometry, 339
isomorphic, 237
isosceles, 294
Isosceles orthogonality, 294
isosceles orthogonality, **294**, 294–296

Jacobian matrix, **195**, 195
James orthogonality, **294**
Jensen's Inequality, 6, **300**, 307, 311
Jensen's inequality, 5, **6**, 6
join, xi
joint cumulative distribution function, **176**, 176
Joint Gaussian distribution, 187
Joint Gaussian pdf, **188**
joint probability density function, **176**, 176

Kaneyoshi, Urabe, 419
Karhunen-Loève Expansion, **21**
Kenko, Yoshida, 419
kernel, 18, 397
Kolmogorov's inequality, **7**
Kronecker delta, 22
Kronecker delta function, **283**
kronecker delta function, 198, 365
kronecker product, 196
kronecker products, 196

l'Hôpital's rule, 120, 123, 124, 310, 403, 405, 409
Laplace operator, 344
lattice, 214, 251, 256, 297, 298
lattice with negation, 166
Laurent series, 370

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

Law of the Unconscious Statistician, **6**
Least squares, **74**
least squares, 74, 75
Least squares estimation, 125
least squares estimation, 78
Least squares estimations, 125
Least Squares Technique, 113
Least Squares transfer function estimate, **114**
least upper bound, xi
least-squares, 125, 128
Lebesgue integral, 172
Lebesgue integration, 394, 395
Lebesgue square-integrable functions, 349
left distributive, 5, 132, 207, 208, 323
left inverse, 142
Leibnitz GPR, 380
Leibniz Integration Rule, 392
Leibniz, Gottfried, ix, 319
length, 289
linear, 4, 5, 37, 38, 56, 57, 59, 86, 112, 122, **320**, 320, 374
linear basis, **223**
linear bounded, xi
linear combination, **221**
linear combinations, 222
linear functional, 329
linear independence, 228
linear manifold, **252**
linear operator, 4, 5, 54, 56, 57, 240, 242
linear operators, 238, 320, 329
linear ops., 238, 242
Linear space, 251
linear space, 209, **210**, 210, 211, 221, 223, 252–254, 256, 257, 265, 271, 277, 287, 293, 297–300, 319
linear spaces, 319
linear span, **221**, 252, 253
linear subspace, **210**, 221, 251, **252**, 252–257, 259, 260
linear subspaces, 254
linear time invariant, 39, 43, 113, 125, 129, 130, 397
linear time invariant systems, 47
linear time-invariant, 59
linearity, 5, 19, 21, 22, 29, 32, 33, 47, 53, 57–59, 68, 73, 97, 101, 102, 126, 197, 321, 386, 387, 405, 408
Linearity of E, **4**
linearly dependent, **221**, 223–225
linearly independent, **221**, 223–225, 228
Lippmann, Bernard A., 182
Liquid Crystal, vi
low-pass filter, 377
lower bounded, 169
LTI, 14, 15, 37–40, 47, 58–61, 113, 114, 125, 126, 129–131, 137, 397
    operations on non-stationary random processes, 43
    operations on WSS random processes, 45

Machiavelli, Niccolò, 419
magnitude, 115, 116
MAP, 66, 81
MAP estimate, **66**, 81, 83
maps to, x
Markov's inequality, **6**, 7
matched filter, 94
MatLab, 410
matrix, 23
    rotation, 346
matrix algebra, 72
matrix calculus, 195
matrix:quadratic form, 200, 201
Maxima, 387, 389, 390
maximal likelihood (ML), 144
maximum a-posteriori, 66
maximum a-posteriori probability estimate, **66**
maximum a-posteriori probability estimation, 81
maximum likelihood, 66
maximum likelihood estimate, **66**
maximum likelihood estimation, 81, 86
    amplitude, 87
    general, 86
    phase, 91
Mazur-Ulam theorem, 330
mean, **3**, 3, **12**, 12, **18**, 37, 100
mean integrated square error, **67**
mean square error, **67**, 67, 99
Mean square estimate, **66**
measurable space, 167
measure, 3, 166, 167, 170, 176, 177
measure space, 167, 168
measure-theoretic, 167
measurement additive noise cross-correlation, 57
measurement additive noise model, **112**, 125
Measurement additive noise system, 57
measurement LTI additive noise model, **129**, **137**
measurement system, 128
meet, xi
memoryless, 144
Mercedes Frame, **247**
Mercer's Theorem, 21, **194**, 194
metric, xi, 209, 289, 290
    Euclidean, 292
    generated by norm, 290
    induced by norm, 290
    parabolic, 292
    sup, 292
    taxi-cab, 292
metric induced by the norm, 290
metric linear space, 209, 289, 290, 292
metric space, 209, 290, 297, 298
metrics
    gap, 263
    Schäffer, 263
mini-max estimate, **66**
Minimum mean square estimation, **72**
Minimum phase, 381
minimum phase, 40, 51, **373**, 373, 374
    energy, 373
minimum variance unbiased estimator, **68**
Minkowski addition, 159, 253
Minkowski's Inequality, **276**, 307
Minkowski's Inequality for sequences, **313**
ML, 66, 81
ML amplitude estimation, **87**
ML estimate, **66**, 81, 83, 86, 88, 92
ML estimation of a function of a parameter, **92**
ML phase estimation, **91**
MM estimate, **66**
modular, 166
modulus, **208**
monotone, 166, 168, 170
monotonically decreasing, 119
MRA, 410
MRA system, 410, 412
MS estimate, **66**
multiplicative condition, **269**
Multiresolution Analysis, 407
multiresolution analysis,

COPYRIGHT © 2019 DANIEL J. GREENHOE

398, 409
mutually exclusive, 169, 172
MVUE, **68**

negative part, **216**, 217–219
Neumann Expansion Theorem, 331
Newton's method, 77
no input noise, 131
no output noise, 131
noise
colored, 94
non-central chi-squared distribution, 186
non-contradiction, 168
non-isometric, 276
non-isotropic, 260, 271, 272, 280, 281, 284
non-linear, 56, 57, 112, 128
non-negative, 19–21, 208, 271, 276, 279, 281, 284, 308, 324, 394, 400
non-positive, 274
non-stationary, 46
non-structured, 143
non-zero, 412
Noncentral chi-square distribution, 186
noncentral chi-square distribution, **186**
noncentral chi-square with n-degrees of freedom, **187**
nondegenerate, 166, 168, 171, 172, 208, 287, 293, 324
nonnegative, 166
nonparametric, 65
norm, 8, 9, 72, 99, 209, 273, 275–277, **287**, 287, 289, 290, 293, 324
induced by inner product, 276, 277
Polarization Identity, 278
usual, 289
norm induced by the inner product, **277**, 277
normal, 193, **266**, 336, **337**, 337, 338, 343, 344
normal distribution, 182, **183**, **188**
normal operator, 337, 342
normalized, 21, 22, 166, 168, 171
normalized rms error, **67**, 67
normalized tight frame, **244**
normed algebra, **269**, 269, 270
normed linear space, 209, 277, 284, 287–290, 292, 294, 296, **324**
normed linear spaces, 329, 339
normed space of linear operators, **324**
NOT, x
not ergodic in the mean, 46
not included
metric, 290
not invertible, 28
not necessarily linear, 39, 53
not symmetric, 395
not total, 235
not unique, 256
null space, x, 191, 321–323, 332, 334–338, 343

O_6 lattice, 166
Octave, 410
one sided shift operator, 341
only if, xi
open ball, 290
open interval, **297**, **298**
opening, 263
operations
0 order B-spline dyadic decomposition, 408
1st order B-spline dyadic decomposition, 409
2nd order B-spline dyadic decomposition, 409
adjoint, **329**, 329, **332**
auto-correlation, 19
auto-correlation operator, **19**, 21
auto-spectral density, **14**
Bartlett Method, 107
Blackman-Tukey Method, 107
closure, 394
communications additive noise model, **112**
communications LTI additive noise model, **129**, **137**
Continuous data whitening, 94
convolution, 33, 34, 37, 38, 40, 59, 184, **369**, 370, 383
convolution operation, **352**
cross spectral density, **14**
CSD, **14**
detection, 65
differential operator, 397
Discrete data whitening, 94
Discrete Time Fourier Series, xii
Discrete Time Fourier Transform, xii, 14
discrete-time Fourier transform, **359**
DTFT, 14–16, 111, 227, 360, 377, 379, 411
Eigendecomposition, 246
ensemble average, 46
estimate, 111
expectation, 5, 9, 112
expectation operator, **3**, 3, 4, 6
FFT, 13
Fourier analysis, 396
Fourier Series, xi
Fourier Transform, xi, xii, 47, 243, 344, 345, **350**, 353
Fourier transform, 227, 397
Fourier transform magnitude, 398, 399
frame operator, **244**, 244, 245, 247
Fredholm integral equation of the first kind, 18
Fredholm operator, **193**
Frequency Response Function, 111
Frequency Response Identification, 111
FRF, 111
gradient of $\boldsymbol{y}$ with respect to $\boldsymbol{x}$, **195**
gradient of $\boldsymbol{y}^T$ with respect to $\boldsymbol{x}$, **195**
Hermitian Transpose, 73
identity operator, **320**
imaginary part, **267**
integral, 396
inverse, 245, 247, 412
Inverse Fourier Transform, 344
inverse Fourier Transform, 345
involution, **266**
Jacobian matrix, **195**
kernel, 18
Laplace operator, 344
Least squares estimation, 125
Least squares estimations, 125
Lebesgue integral, 172
Lebesgue integration, 394, 395
linear operator, 4, 5, 54, 56, 57
linear operators, 329
linear span, **221**
matrix, 23
measure, 167
measurement additive noise model, **112**

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

measurement LTI additive noise model, **129**, **137**
operator, 244, **319**
operator adjoint, 333
Periodogram, 107
power spectral density, **14**
project, 396
projection, **334**
PSD, **14**
rationalizing the denominator, 124
real part, **267**
sampling, 4
Sum, 34
summation operator, 300
support, 394
system identication, 111
time average, 46
unitary Fourier Transform, **350**
variance, 46
vector addition, 197
Welch Method 0% overlap, 107
Welch Method 50% overlap, 107
z-domain cross spectral density, **14**
z-domain power spectral density, **14**
Z-Transform, xii
z-transform, **370**, 370
Z-transform operator, 14
operator, 209, 244, **319**
adjoint, 267
autocorrelation, 336
bounded, 327
definition, 319
delay, 345
dilation, 344
identity, 319
isometric, 337, 340–342
linear, 320
norm, 324
normal, 337, 338, 342
null space, 334
positive, 346, 347
projection, 334
range, 334
self-adjoint, 336
shift, 341
translation, 344
unbounded, 327
unitary, 337, 342, 343, 351
operator adjoint, 332, 333
operator norm, xi, **324**
operator star-algebra, 333
optimal lower frame bound, **244**
optimal receiver, 81
optimal upper frame bound, **244**
order, x, xi
order n B-spline, 390
order n B-spline MRA coefficient sequence, 410
order n B-spline MRA scaling sequence, 410
order relation, 251, 297
order-k moment estimate, **103**
ordered linear space, **214**
ordered pair, x
ordered set, 262, 297, 298
ordinary coherence, **134**
ordinary transmissibility, 116
orthocomplemented lattice, 168, 263
orthogonal, 19, 20, 22, 228, 231, 232, **283**, 283, 284, 336, 396
orthogonal basis, **232**
orthogonal complement, **257**, 257, 261
orthogonal in the Pythagorean sense, **296**
orthogonal in the sense of Birkhoff, **296**
orthogonal in the sense of James, **294**
orthogonality, 228, 294, 296
Birkhoff, 296
inner product space, 283
isosceles, 294
James, 294
Pythagorean, 296
orthomodular, 166, 262
orthomodular identity, 262
orthomodular lattice, 166, 262
orthonomal, **283**
orthonormal, 229–235, 248, **283**
orthonormal bases, 222
orthonormal basis, 146, **232**, 236–241, 243, 406, 407
Orthonormal decomposition, 94
orthonormal expansion, 232, 236
orthonormal quadrature conditions, **365**
orthonormality, 229, 240, 248
orthornormal basis, 233
outcome, **167**
outcomes, 167
output, 14, 15, 37, 39
over estimate, 114
over-estimated, 130

PAM, 154
parabolic metric, 292
parallelogram, 279, 294
Parallelogram Law, 9
Parallelogram law, **279**, 282
parallelogram law, 261, 277, 279, 282
parametric, 65
Parseval frame, 232, 233, **244**, 244
Parseval frames, 222
Parseval's equation, 351
Parseval's Identity, 234, 236, 238, 244
Parseval's Theorem, 16
partition, 169, 170
partition of unity, 36, 168, 400, 401, 412, 413
Pascal's identity, 388
Pascal, Blaise, 165
pdf, **176**, 176
Peace Frame, **247**
Per Enflo, 226
Perfect reconstruction, **232**
periodic, 316, 359, 402
periodicity, 402
Periodogram, 107
phase, 115, 116, 122
Phase Shift Keying, 151
Plancheral's formula, 47, 87
Plancherel's formula, 351
Poincaré, Henri, 182
point spectrum, **192**
pointwise addition, 211
Polar Identity, **275**, 276, 295
polar identity, 229
Polarization Identities, 278
Polarization Identity, 277, 279, 280, 282
polarization identity, 285, 295
pole, 381
poles, 51, 375
polynomial approximation, 75
Popper, Karl, 111
positive, 20, 21, **214**, **346**
positive cone, **214**
positive definite, 20, 21
positive part, **216**, 217–219
power mean, **308**, 308, 311
power rule, 405
power set, xi
power spectral density, **14**, 16, 46
power sum, 390
Power Sums, **316**
probability, **167**
probability density function,

COPYRIGHT © 2019 DANIEL J. GREENHOE

3, 27, 36, 170, **176**, 176, 184
probability function, **166**
Probability integral transform, **27**
probability measure, 167
probability space, 3–5, 8, 9, 11, 17, 35, **167**, 167–171, **172**, 172, 176, 177, 184
probabilty space, 172
product rule, 197, 200, 393
project, 396
projection, **266**, **334**
projection operator, 334, 336
projection statistics
Additive *Gaussian* noise channel, 84
Additive noise channel, 83
Additive white *Gaussian* noise channel, 85
Additive white noise channel, 85
Projection Theorem, **262**, 262
projections, 86
proper subset, x
proper superset, x
properties
σ-additive, 166, 172
σ-additivity, 166
absolute value, x
absolutely summable, 15, 370, 371
additive, 20, 166–169, 171, 259, 271, 272, 281, 283, 284, 320, 323
additive identity, 12, 68, 212, 321
additive inverse, 212, 321
additivity, 20, 166, 167, 248, 275, 332
admissibility condition, 412
admissibility equation, 412
affine, 192, **299**, 330
algebra of sets, xi
AND, xi
anti-symmetric, 13, 14, 50, 262, 353
anti-symmetric , 50
antiautomorphic, 15, 266, 267, 272, 333
antitone, 168, 171, 257, 259, 262
associates, 210
associative, 5, 207, 208, 210, 212, 305, 306, 322, 323, 342
asymptotically unbiased, 105
AWGN, 81
Bayesian, 65
biased, 100, 101, 114
bijection, 237
bijective, 237, 330
biorthogonal, 241, 242
Birkhoff orthogonality, 296
Boole's inequality, 169
Boolean, 167, 169
Boolean lattice, 169
bounded, 168, 227, 336, 347
bounded bijective, 240
Cartesian product, x
Cauchy, 261
causal, **374**
characteristic function, x
closed, 259, 261
closed support, 394
closure, 225
colored, 19
commutative, 12, 208, 210, 217, 322, 369, 370
compact, 193
compact support, 407
comparable, 297
complement, x
complemented, 262
complex, 122
complex-valued, 116
concave, **298**, 298, **299**, 299, **300**, 307, 308
conjugate linear, 266, 333
conjugate quadrature filter condition, **377**
conjugate symmetric, 14, 15, 18, 50, 194, 271, 272, 280, 281, 284
conjugate symmetry, 20
conjugate-symmetric, 15
consistent, 98, 99, 101, 105
constant, 12, 45, 67
Continuity, 290
continuity, 177, 178
continuous, 27, 170, 177, 194, 275, 290, 306, 308, 322, 396, 402, 409
convergence, 403
convergence in probability, 21
convergent, 46, 227
converges, 403–405
converges absolutely, 194
convex, 5, 6, 252, 253, 261, 290, 292, 298, **299**, 299, **300**, 300, 307, 308
convexity, 261
correlated, 54, 112, 113
countably infinite, 193, 226
counting measure, xi
CQF condition, **377**, 379
critical point, **28**, 28, 30, 31
de Morgan, 263
decreasing, 307, 406
dense, 225, 226
difference, x
differentiable, 28
differentiable function, 30, 31
distributes, 210
distributive, 12, 37–39, 72, 166–169, 212, 266–268, 271, 272, 333, 360, 371
domain, x
dyadic decomposition, 407
efficient, 88, 90–93
empty set, xi
equality by definition, x
equality relation, x
equivalent, 227, 239
ergodic in the mean, **46**, 46, 47
even, 364
excluded middle, 168, 169
exclusive OR, xi
existential quantifier, xi
exists, 46
false, x
for each, xi
Fredholm operator, 193, 194
Fredholm operators, 193
Gamma distribution, 184
Gaussian, 35, 82, 84, 86
geodesically between, **298**
greatest lower bound, xi
hermitian, 266, 267, **336**
Hermitian symmetric, 73, 353
homogeneous, 5, 208, 232, 233, 236, 240, 242, 259, 271, 272, 276, 280, 281, 283, 287, 293, 314, 320, 321, 323, 324
hypotenuse, 296
identity, 207, 210
identity operator, 198
if, xi
if and only if, xi

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

image, x
imaginary part, xi
implied by, xi
implies, xi
implies and is implied by, xi
impossible, 400
inclusive OR, xi
increasing, 307, 406
independence, 83, 86, 171
independent, 7, 33, 36, 82, 86, 87, **168**, 170, 184
indicator function, x
injective, 322
inner-product, xi
inside, 374
intersection, x
into, 237, 238
inverse, 207
invertible, 28, 170, 227, 244, 266
involutary, 333
involution, 266, 270
involutory, 15, 262, 266–268
irreflexive ordering relation, xi
isometric, 233, 241, 242, 330, 339, 343, 351
isometric in distance, 343
isometric in length, 343
isomorphic, 237
Isosceles orthogonality, 294
isosceles orthogonality, 294–296
join, xi
kronecker delta function, 198
least upper bound, xi
least-squares, 125, 128
left distributive, 5, 132, 207, 208, 323
linear, 4, 5, 37, 38, 56, 57, 59, 86, 112, 122, **320**, 320
linear independence, 228
linear space, 211
linear time invariant, 39, 113, 125, 129, 130, 397
linear time-invariant, 59
linearity, 5, 19, 21, 22, 29, 32, 33, 47, 53, 57–59, 68, 73, 97, 101, 102, 126, 197, 321, 386, 387, 405, 408
linearly dependent, **221**, 223–225
linearly independent, **221**, 223–225, 228
lower bounded, 169
LTI, 14, 15, 37–40, 58–61, 113, 114, 125, 126, 129–131, 137, 397
maps to, x
meet, xi
metric, xi
Minimum phase, 381
minimum phase, 40, 373, 374
minimum variance unbiased estimator, **68**
modular, 166
monotone, 166, 168, 170
monotonically decreasing, 119
mutually exclusive, 169, 172
MVUE, **68**
no input noise, 131
no output noise, 131
non-contradiction, 168
non-isometric, 276
non-isotropic, 260, 271, 272, 280, 281, 284
non-linear, 56, 57, 112, 128
non-negative, 19–21, 208, 271, 276, 279, 281, 284, 308, 324, 394, 400
non-positive, 274
non-stationary, 46
non-zero, 412
nondegenerate, 166, 168, 171, 172, 208, 287, 324
nonnegative, 166
nonparametric, 65
normal, 193, 336, **337**, 343, 344
normalized, 21, 22, 166, 168, 171
NOT, x
not ergodic in the mean, 46
not invertible, 28
not necessarily linear, 39, 53
not symmetric, 395
not total, 235
not unique, 256
null space, x
only if, xi
operator norm, xi
order, x, xi
ordered pair, x
orthogonal, 19, 20, 22, 228, 232, **283**, 283, 284, 336, 396
orthogonality, 228, 294, 296
orthomodular, 166, 262
orthomodular identity, 262
orthonomal, **283**
orthonormal, 229, 230, 232–235, 248, **283**
orthonormality, 229, 240, 248
over estimate, 114
over-estimated, 130
parallelogram, 279, 294
parallelogram law, 282
parametric, 65
Parseval's Theorem, 16
partition of unity, 36, 168, 400, 401, 412, 413
periodic, 316, 359, 402
periodicity, 402
pointwise addition, 211
Polarization Identity, 282
positive, 20, 21
positive definite, 20, 21
power set, xi
proper subset, x
proper superset, x
pseudo-distributes, 210
Pythagorean orthogonality, 296
Pythagorean Theorem, 296
randomness, 175
range, x
real, 20, 290, 295, 362, 375
real part, xi
real-valued, 15, 16, 20, 72, 116, 336, 353, 360, 361, 364
reality condition, 352
reflexive, 262
reflexive ordering relation, xi
regular, 193
relation, x
relational and, x
reverse triangle inequality, 288
right distributive, 5, 207, 208, 323
ring, 73
ring of sets, xi
scalar commutative, 208
self adjoint, 20, 336
self-adjoint, 19, **336**, 336
separable, 226, 227, 238
set of algebras of sets, xi
set of rings of sets, xi
set of topologies, xi
space of linear transforms, 321
span, xi

COPYRIGHT © 2019 DANIEL J. GREENHOE

spans, 223, 224
stability condition, 242, 244
stable, 374
stationary in correlation, 45
stationary in the 1st moment, 12
stationary in the 2nd moment, 12
stationary in the mean, 45, 46
Strang-Fix condition, 355
strictly concave, **300**
strictly convex, **299, 300**
strictly increasing, 308
strictly monotone, 170
strictly monotonic, 306, 308
strictly monotonic increasing, 121
strictly positive, 287
strong convergence, 226, 236
subadditive, 166, 208, 276, 287, 314, 324
submultiplicative, 208
subset, x
sufficient, 94
sufficient statistic, 83
super set, x
surjective, 343
symmetric, 13, 14, 50, 290, 353, 362, 394, 395, 402, 406
symmetric , 50
symmetric difference, x
symmetry, 283, 373
there exists, xi
tight frame, 245
time-invariant, **374**
Toeplitz, 23
topology of sets, xi
total, 225, 234, 235, 239, 240
transitive, 262
translation invariant, 293
triangle inequality, 208, 297
triangle inquality, 287
true, x
unbiased, 68, 88, 89, 97–101, 131
uncorrelated, 19, 53–61, 82, 85, 86, 97, 98, 101, 112–114, 126, 130, 132, 135
uncorrelatedness, 86
under estimate, 114
under estimates, 114
under-estimated, 130
uniformly, 194
uniformly distributed, 27, 36
union, x
unique, 28, 212, 223, 226, 256, 260, 261, 399
unit area, 394
unit length, 341, 343
unitary, 342, 343, 351
universal quantifier, xi
upper bounded, 166
vanishing 0th moment, 412
vector norm, xi
von Neumann-Jordan condition, 279
white, **19**, 19, 85
wide sense stationary, **12**, 12–14, 39, **45**, 45, 46, 54, 97, 98, 113
wide-sense stationary, 130
WSS, **12**, 13–15, 38–40, **45**, 45, 53, 55–61, 97, 98, 101, 113, 114, 126, 130
zero at −1, 412
zero measurement error, 113
zero measurement noise, 113
zero-mean, 7, 53–61, 126
PSD, **14**
pseudo-distributes, 210
PSF, 355, 356
PSK, 151
pstricks, vi
Pulse Amplitude Modulation, 154
Pythagorean orthogonality, 294, **296**, 296
Pythagorean Theorem, 228, 229, 231, 233, 240, **284**, 284, 296
Pythagorean theorem, 228

QAM, 149
quadratic discriminate inequality, **274**
Quadratic Equation, 133
quadratic equation, 274
Quadratic form, **200**
quadratic form, 73, 200, 201
Quadrature Amplitude Modulation, 149
quantization noise, 113
quantum probability, **166**, 167
quotes
Abel, 314
Abel, Niels Henrik, 419
Aristotle, 11, 17
Bak, Per, 125
Bohr, Harald, 305
Cardano, Gerolamo, 207
Descartes, René, ix
Fourier, Joseph, 349
Hardy, G.H., 305
Housman, Alfred Edward, vii
Kaneyoshi, Urabe, 419
Kenko, Yoshida, 419
Leibniz, Gottfried, ix, 319
Lippmann, Bernard A., 182
Machiavelli, Niccolò, 419
Pascal, Blaise, 165
Poincaré, Henri, 182
Popper, Karl, 111
Russull, Bertrand, vii
Stravinsky, Igor, vii
Ulam, Stanislaus M., 330
Quotient Rule, 132, 133
quotient rule, 405
quotient structures, 297

random process, **17**, 17, 19, 21, 46
random sequence, **11**, 11–14, 37, 41, 112
random variable, 3–8, **17**, 17, 27, 28, 30, 31, 33, 35, 170, 175, **176**, 176, 177, 184
randomness, 175
range, x
range space, 332
rational expression, 40, 51
rationalizing factor, 124
rationalizing the denominator, 124
Rayleigh distribution, **187**, 187
real, 20, 290, 295, 362, 375
real interval, 177
real linear space, **210**
real numbers, 289
real part, xi, **267**
real-valued, 15, 16, 20, 72, 116, 336, 353, 360, 361, 364
reality condition, 352
Rectangular pulse, 357
rectangular pulse, **356**, 398
recursively, 77
reflection, 330
reflexive, 262
reflexive ordering relation, xi
regular, 193
relation, x, 28, 319
relational and, x
relations, xi
order relation, 297
residual spectrum, **192**

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

resolvent, **265**
resolvent set, **192**, 193
Reverse Triangle Inequality, **288**, 290
reverse triangle inequality, 288
Rice distribution, **187**, 187
Riesz bases, 222
Riesz basis, **239**, 239–243, 398, 406, 407
Riesz sequence, **239**, 241, 409
Riesz space, **214**, 214, 216–219
Riesz-Fischer Theorem, **236**
Riesz-Fischer Thm., 237, 238
right distributive, 5, 207, 208, 323
right inverse, 142
ring, 73, **207**, 207, 208
    absolute value, 208
    commutative, 207
    modulus, 208
ring of complex square $n \times n$ matrices, 333
ring of sets, xi
Robinson's Energy Delay Theorem, **373**
rotation matrix, 346
Russull, Bertrand, vii

sample variance, **100**
sampling, 4
scalar commutative, 208
scalar product, **271**
scalars, **210**
scaling coefficient sequence, 412
scaling function, **117**, 409, 410
scaling parameter, **117**
Scaling transfer function estimate, **117**, 117
Schauder bases, 222, 227
Schauder basis, **226**, 226, 227, 232, 238
Schäffer's metric, 263
self adjoint, 20, 336
self-adjoint, 19, **266**, **336**, 336
semilinear, 267
separable, 226, 227, 238
separable Hilbert space, 237, 239, 241, 369
separable Hilbert spaces, 237
sequence, 11, 36, 289, **369**, 410
sequences, 377
set, 167, 209, 298
set indicator, 383
set indicator function, 357
set of algebras of sets, xi
set of rings of sets, xi
set of topologies, xi
set projection operators, **159**
sets
    open ball, 290
    real numbers, 289
shift operator, 341
shift relation, 356, 357
sigma-algebra, 167
Signal matching, 94
signal to noise ratio, 94
signal-to-noise ratio, 113
sinc, 356, 357
sine sweep, 111
single coin toss, **170**, 172
Smith-Barnwell filter, **377**
SNR, 113
source code, 415
space, 167
    inner product, 271
    linear, 209
    linear subspace, 252
    metric, 290
    Minkowski addition, 253
    normed vector, 287
    orthogonal, 257
    vector, 209
space of all absolutely square Lebesgue integrable functions, 369
space of all absolutely square summable sequences, **369**
space of linear transforms, 321
span, xi, 225, 227
spans, **221**, 223, 224
spectral power, 16, **46**, 113
spectral radius, **265**
Spectral Theorem, **193**
spectrum, **192**, 193, **265**
stability, 374
stability condition, 242, 244
stable, 374
standard othornormal basis, 233
star-algebra, **266**, 266, 333
star-algebras, 332, 333
state , 77
stationary in correlation, 45
stationary in the 1st moment, 12
stationary in the 2nd moment, 12
stationary in the mean, 45, 46
statistical average, 78
statistics, 144
steepest descent, 77, 78
step function, **383**, 383–387
step size, 77
Stifel formual, 408
Stifel formula, 388
stochastic, 71
Strang-Fix condition, **355**, 355
Stravinsky, Igor, vii
strictly concave, **300**
strictly convex, **299**, 299, **300**
strictly increasing, 308
strictly monotone, 170
strictly monotonic, 306, 308
strictly monotonic increasing, 121
strictly positive, 287, 293
strong convergence, 226, 236
structured, 143
structures
    C^* algebra, **269**
    C^*-algebra, 270
    $*$-algebra, **12**, **15**, **37–39**, **266**, 266–268, 333, **360**
    N-tuple, 211, 287
    *-algebra, 72
    *-algebras, 333
    σ-algebra, 167, 172, 174
    absolute value, 218, 219
    additive identity element, 211, 212
    additive inverse, 212
    adjoint, 333
    algebra, **208**, 209, **265**, 265, 266
    algebraic, 209
    ball, 298
    Banach space, 226, 227
    basis, 21, 22, 223–225, 232, 400
    basis sequence, 399
    bijection, 237
    bijective, 237
    bilinear functionals, 278
    Boolean algebra, 172
    Boolean lattice, 166, 169
    Borel algebra on $\mathbb{R}$, 172
    Borel sets, 349
    bounded lattice, 169
    bounded linear operator, 343
    bounded linear operators, 329, 331, 332, 334, 335, 337, 338, 340–343
    C star algebra, **269**
    calculus, 209
    closed interval, **297**, **298**, 298
    coefficient functionals, 226
    coefficients, 375
    communication system, 128
    communications additive noise model, 125
    Communications additive noise system, 57
    commutative ring, 207,

COPYRIGHT © 2019 DANIEL J. GREENHOE

208, 208
complete metric space, 225
complete set, 225
complex linear space, **210**
conditional probability, 169
conditional probability of x given y, **168**
conjugate pairs, 375
conjugate quadrature filter, 377
Conjugate quadrature filters, 377
continuous spectrum, **192**
convex function, 299, 300
convex set, **298**, 298–300
convolution operator, 369, 370
coordinates, **223**
CQF, 377, 378
critical point, 28
differentiable function, 28
discrete-time Fourier transform, 365
distributive pseudo-complemented lattice, 166
Double coin toss, 170
eigen-system, 22
eigenspace, **191**
epigraph, **299**
Euclidean space, 298
even/odd die probability space, 171
even/odd die toss, **171**
event, **167**, 167, 168
expansion, **223**
field, 207, **208**, 209, 211, 369
field of complex numbers, 333
field of probability, 166
filter banks, 377
finite modular lattice, 166
Fourier Transform, 349
frame, 242, **244**, 244, 248
frames, 222
function, 28, 175, 349, 384
functional, 271, 300, 333
geometry, 209
group, **207**, 207
half-open interval, **297**, **298**
Hamel bases, 222
Hamel basis, **223**, 223, 225
high-pass filter, 377
Hilbert space, 232, 234, 236, 238, 239, 242, 244, 261, 332, 333, 336–338, 343, 349
hypograph, **299**
identity, 210
identity element, 319
image set, 321, 323, 334–338, 343
infinite sum, 221
inner product, 209, 277
inner product space, 8, 9, 209, 228–231, 239, 257, 259, 260, **271**, 272, 275–277, 279, 283, 284, 294, 396
inner-product space, 294, 296
integral domain, 207
interval, 297, 298
intervals on lattices, **297**
intervals on ordered sets, **297**
inverse, 210
isometry, 339
lattice, 251, 256, 297, 298
lattice with negation, 166
Laurent series, 370
Lebesgue square-integrable functions, 349
linear basis, **223**
linear combination, **221**
linear combinations, 222
linear operator, 240, 242
linear operators, 238
linear ops., 238, 242
Linear space, 251
linear space, 209, **210**, 210, 221, 223, 252–254, 256, 257, 265, 271, 277, 287, 293, 297–300, 319
linear spaces, 319
linear span, 252, 253
linear subspace, **210**, 221, 251, **252**, 252–257, 259, 260
low-pass filter, 377
matrix algebra, 72
measurable space, 167
measure, 167
measure space, 167, 168
measurement additive noise model, 125
Measurement additive noise system, 57
measurement system, 128
metric, 209, 290
metric linear space, 209, 289, 290, 292
metric space, 209, 290, 297, 298
MRA, 410
MRA system, 410, 412
Multiresolution Analysis, 407
multiresolution analysis, 398, 409
negative part, 217–219
norm, 209, 273, 275–277, 287, 289, 290, 293
normalized tight frame, **244**
normed algebra, 269, 270
normed linear space, 209, 277, 284, 287–290, 292, 294, 296
normed linear spaces, 329, 339
normed space of linear operators, **324**
null space, 191, 322, 336–338, 343
O_6 lattice, 166
open ball, 290
open interval, **297**, **298**
operator, 209
order relation, 251
ordered set, 262, 297, 298
orthocomplemented lattice, 168, 263
orthogonal basis, **232**
orthogonal complement, 257, 261
orthomodular lattice, 166, 262
orthonormal bases, 222
orthonormal basis, **232**, 236–241, 243, 406, 407
orthornormal basis, 233
outcome, **167**
Parseval frame, **244**
Parseval frames, 222
Parseval's equation, 351
partition, 169, 170
Plancherel's formula, 351
point spectrum, **192**
pole, 381
poles, 375
positive part, 217–219
probability space, 3–5, 8, 9, 11, 17, 35, **167**, 167–171, **172**, 172, 176, 177, 184
probabilty space, 172
projection operator, 336
quotient structures, 297
random variable, 176

real interval, 177
real linear space, **210**
relation, 28
residual spectrum, **192**
resolvent, **265**
resolvent set, **192**, 193
Riesz bases, 222
Riesz basis, 239–243, 398, 406, 407
Riesz sequence, 241, 409
Riesz space, 214, 216–219
ring, **207**, 207, 208
ring of complex square $n \times n$ matrices, 333
scalars, **210**
Schauder bases, 222, 227
Schauder basis, **226**, 226, 227, 232, 238
separable Hilbert space, 237, 239, 241, 369
separable Hilbert spaces, 237
sequence, 289
sequences, 377
set, 167, 209, 298
sigma-algebra, 167
space, 167
space of all absolutely square Lebesgue integrable functions, 369
space of all absolutely square summable sequences, **369**
span, 225, 227
spectral radius, **265**
spectrum, **192**, 193, **265**
standard othornormal basis, 233
star-algebra, **266**, 333
star-algebras, 332
subset, 253, 258, 261
subspace, 253, 258–262
support, 394
system, 14, 39, 53–61, 111, 113–117, 121, 125, 126, 128–132, 134
tight frame, **244**, 245, 247
tight frames, 222
topological dual space, 329
topological linear space, 209, 225
topology, 167, 209, 221
total set, 225
trivial linear space, 252
underlying set, **210**
unit vector, 241
unital $*$-algebra, 266
unital algebra, 265
vector, 223
vector additive identity element, 257
vector space, 8, 209, **210**
vector space of random variables, 4
vectors, **210**, 297
weights, 300
zero, 381
zeros, 373, 375
subadditive, 166, 208, 276, 287, 314, 324
submultiplicative, 208
subset, x, 253, 258, 261
subspace, 253, 258–262
subspace intersection, **253**
subspace union, **253**
sufficient, 94
sufficient statistic, 81, 83
Sufficient statistic theorem, **81**
Sum, 34
sum of products, 29, 33, **169**, 177
summation, **305**
Summation around unit circle, 246
summation operator, 300
sup metric, 292
super set, x
support, 394
surjective, xi, 343
symmetric, 13, 14, 50, 290, 353, 362, 394, 395, 402, 406
conjugate, 18
symmetric , 50
symmetric difference, x
symmetry, 283, 373
system, 14, 39, 53–61, 111, 113–117, 121, 125, 126, 128–132, 134
system identication, 111

taxi-cab metric, 292
The basis problem, 226
The Book Worm, 419
The Fourier Series Theorem, **234**
The usual norm, **289**
Theorem of Reversibility, **143**
theorems
Affine equations, **199**
affine equations, 73
Alternating Series Test, 406
B-spline dyadic decomposition, 407, 413
B-spline general form, 394
B-spline partition of unity, 400
B-spline recursion, 392
B-spline scaling coefficients, **410**
Battle-Lemarié orthogonalization, **243**
Bessel's Equality, **229**, 230, 231, 235
Bessel's Inequality, 237
Bessel's inequality, **230**
Best Approximation Theorem, **230**, 230
Binomial Theorem, 5, 68, 102
binomial theorem, 411
Cauchy Schwartz inequality, 120, 122
Cauchy-Schwarz Inequality, 95
Cauchy-Schwarz inequality, 245
chain rule, 29
Chebyshev's inequality, 7
Closed Set Theorem, 259
communications additive noise cross-correlation, 57
comparison test, 403
Convolution Theorem, 40, 398
convolution theorem, **352**, 357, **372**
CQF theorem, **377**, 380
CS Inequality, 248
dilation equation, 410
DTFT periodicity, **359**
Euler formulas, 356, 398
Fourier expansion, 233, **236**, 237, 238
Fourier series expansion, 234
Fundamental Theorem of Calculus, 384, 391
Fundamental theorem of calculus, 177, 179
Fundamental theorem of linear equations, 323
Gelfand-Mazur Theorem, **269**
General ML estimation, **86**
Generalized AM-GM inequality, 311
Generalized associative property, **305**
Generalized Parseval's Identity, 234
Geometric Series, **316**
Hermitian representation, **268**

COPYRIGHT © 2019 DANIEL J. GREENHOE

induced norm, **276**
induction, 390, 391, 400
integral test, 404
inverse DTFT, **365**, 379
Inverse Fourier transform, **350**
Inverse probability integral transform, **27**
Jensen's Inequality, 6, **300**
Jensen's inequality, **6**, 6
Karhunen-Loève Expansion, **21**
Kolmogorov's inequality, **7**
l'Hôpital's rule, 120, 123, 124, 310, 403, 405, 409
Law of the Unconscious Statistician, **6**
Least squares, **74**
Leibnitz GPR, 380
Leibniz Integration Rule, 392
Linearity of E, **4**
Markov's inequality, **6**, 7
Mazur-Ulam theorem, 330
measurement additive noise cross-correlation, 57
Mercer's Theorem, 21, **194**, 194
Minimum mean square estimation, **72**
ML amplitude estimation, **87**
ML estimation of a function of a parameter, **92**
ML phase estimation, **91**
Neumann Expansion Theorem, 331
operator star-algebra, 333
orthonormal quadrature conditions, **365**
Parallelogram Law, 9
Parallelogram law, **279**
parallelogram law, 261, 279
Parseval's Identity, 234, 236, 238, 244
Pascal's identity, 388
Perfect reconstruction, **232**
Plancheral's formula, 47, 87
Polar Identity, **275**, 276, 295
polar identity, 229
Polarization Identities, 278
Polarization Identity, 279, 280, 282
polarization identity, 285, 295
power rule, 405
Power Sums, **316**
Probability integral transform, **27**
Projection Theorem, **262**, 262
PSF, 355, 356
Pythagorean Theorem, 228, 229, 231, 233, 240, **284**, 284
Pythagorean theorem, 228
Quadratic Equation, 133
Quadratic form, **200**
quadratic form, 73
Quotient Rule, 132, 133
quotient rule, 405
Reverse Triangle Inequality, **288**, 290
Riesz-Fischer Theorem, **236**
Riesz-Fischer Thm., 237, 238
Robinson's Energy Delay Theorem, **373**
shift relation, 356, 357
Spectral Theorem, **193**
Stifel formual, 408
Stifel formula, 388
Strang-Fix condition, **355**
Sufficient statistic theorem, **81**
sum of products, 29, 33, **169**, 177
Summation around unit circle, 246
The Fourier Series Theorem, **234**
Theorem of Reversibility, **143**
triangle inequality, **287**
Wiener-Hopf equations, **73**, 73
Wold's Theorem, 181
Young's Inequality, 313
there exists, xi
tight frame, **244**, 245, 247
tight frames, 222
time average, 46
time-invariant, **374**, 374
Toeplitz, 23
topological dual space, 329
topological linear space, 209, 225
topological space, 167, 289
topology, 167, 209, 221
topology of sets, xi
total, **225**, 225, 234, 235, 239, 240
Total Least Squares transfer function estimate, **117**
Total least squares transfer function estimate, 117
total set, 225
traditional probability, **166**, 167
transfer function estimate $\hat{H}_\kappa(\omega;\kappa)$, **117**
transfer function estimate $\hat{H}_c(\omega)$, **118**
transform
inverse Fourier, **350**
transitive, 262
translation, 344
translation invariant, 293
translation operator, 355
Transmissibility, 116
transmissibility, 113, 115
transmissibility $\tilde{T}_{xy}(\omega)$, **113**
triangle, **357**
triangle inequality, 208, **287**, 287, 297, 324
triangle inequality for vectors, 284
triangle inquality, 287
triangle-inequality, 294
trivial linear space, 252
true, x
tuples in $\mathbb{F}^N$, **211**

Ulam, Stanislaus M., 330
unbiased, 68, 88, 89, 97–101, 131
uncorrelated, 19, 53–61, 82, 85, 86, 97, 98, 101, 112–114, 126, 130, 132, 135
uncorrelatedness, 86
under estimate, 114
under estimates, 114
under-estimated, 130
underlying set, **210**
uniform distribution, **181**, 181
uniformly, 194
uniformly distributed, 27, 36
union, x, 167
unique, 28, 212, 223, 226, 256, 260, 261, 399
unit area, 394
unit length, 341, 343
unit vector, 241
unital, **265**
unital $*$-algebra, 266
unital algebra, 265
unitary, **342**, 342, 343, 345, 350, 351
unitary Fourier Transform, **350**
unitary operator, 337, 342

COPYRIGHT © 2019 DANIEL J. GREENHOE
BY-NC-ND

universal quantifier, xi
upper bounded, 166
usual norm, 289
Utopia, vi

values
  nth moment, **354**
  Bessel's correction, 100
  Cramér-Rao bound, 92
  Cramér-Rao lower bound, 91
  dimension, **223**
  eigenvalue, **191**, 246
  eigenvector, **191**
  frame bound, 245, 247
  frame bounds, 248
  MAP estimate, 83
  mean, 100
  rationalizing factor, 124
  scaling function, **117**
  scaling parameter, **117**
  sufficient statistic, 81
Vandermonde matrix, 75
vanishing 0th moment, 412
vanishing moments, 354, 367
variance, **3**, 3, **12**, 12, 46, 97
vector, 223
vector addition, 197
vector additive identity element, 257
vector lattice, **214**
vector norm, xi
vector space, 8, 209, **210**
vector space of random variables, 4
vectors, **210**, 297
von Neumann-Jordan condition, 279

wavelet, 407
weighted, 307
weights, 300
Welch Method 0% overlap, 107
Welch Method 50% overlap, 107
white, **19**, 19, 85
whitening filter, 40, 50
wide sense stationary, **12**, 12–14, 39, **45**, 45, 46, 54, 97, 98, 113
wide-sense stationary, 130
Wiener-Hopf equations, **73**, 73
windowed auto-correlation estimate, **105**
Wold's Theorem, 181
Wronskian, 205
WSS, **12**, 12–15, 38–40, **45**, 45, 53, 55–61, 97, 98, 101, 113, 114, 126, 130

Young's Inequality, **312**, 313

Z transform, 411
z transform, 373
z-domain cross spectral density, **14**
z-domain power spectral density, **14**
Z-Transform, xii
Z-transform, 361
z-transform, **370**, 370, 378
Z-transform operator, 14
zero, 381
zero at −1, 362, 412
zero measurement error, 113
zero measurement noise, 113
zero-mean, 7, 53–61, 126
zeros, 51, 373, 375

License

This document is provided under the terms of the Creative Commons license CC BY-NC-ND 4.0. For an exact statement of the license, see

https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode

The icon appearing throughout this document is based on one that was once at

https://creativecommons.org/

where it was stated, “Except where otherwise noted, content on this site is licensed under a Creative Commons Attribution 4.0 International license.”

…last page …please stop reading …

Copyright © 2019 Daniel J. Greenhoe
BY-NC-ND