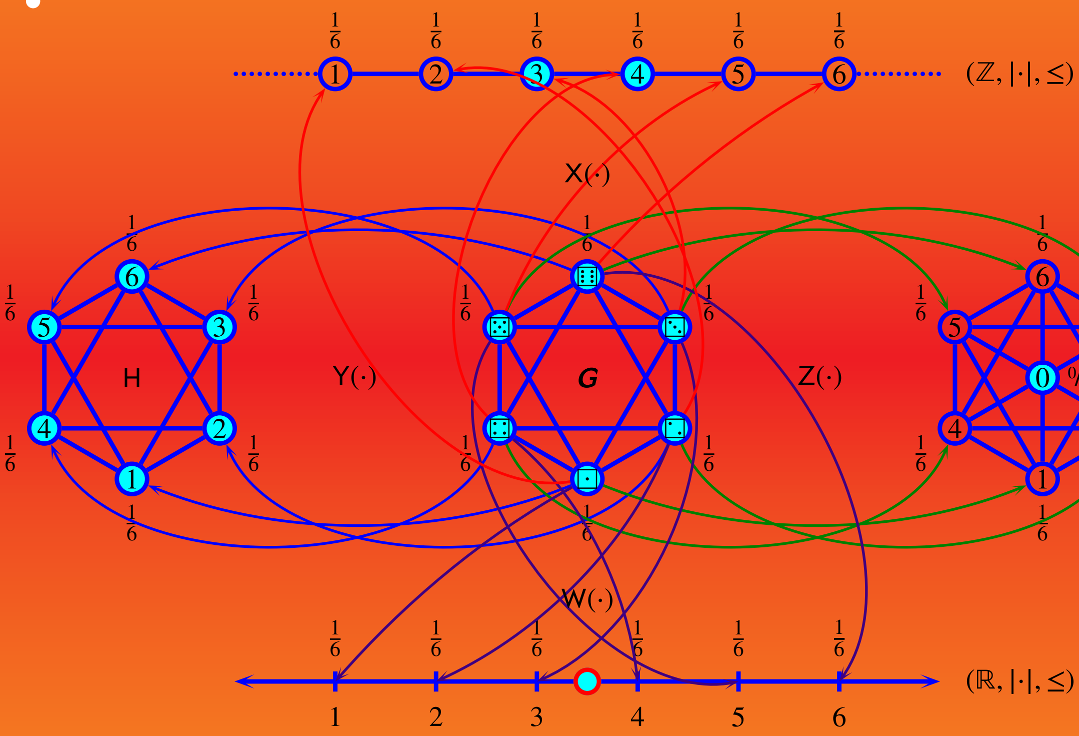
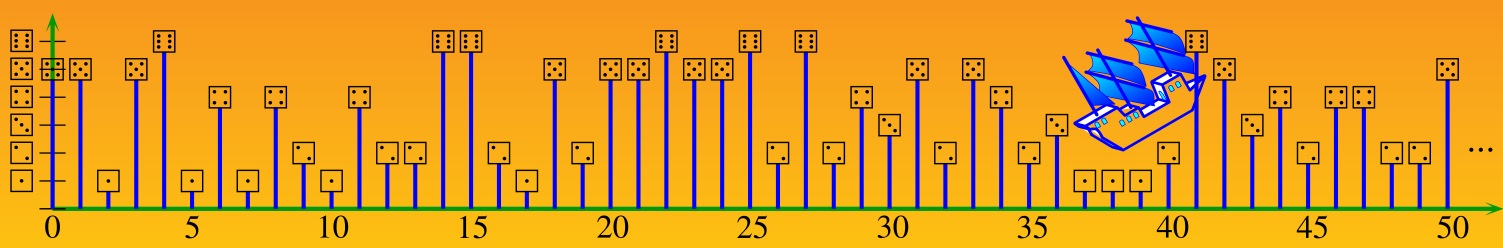
A Book Concerning Symbolic Sequence Processing

VERSION 0.52



Daniel J. Greenhoe



Signal Processing ABCs series

volume $\boxed{3}$

title:	*A Book Concerning Symbolic Sequence Processing*
document type:	`book`
series:	*Signal Processing ABCs*
volume:	$\boxed{3}$
author:	Daniel J. Greenhoe
version:	VERSION 0.52
time stamp:	2019 July 26 (Friday) 07:18pm UTC
copyright:	Copyright © 2019 Daniel J. Greenhoe
license:	Creative Commons license CC BY-NC-ND 4.0
typesetting engine:	XeLaTeX
document url:	https://www.researchgate.net/project/Signal-Processing-ABCs

This text was typeset using XeLaTeX, 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 version 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.

Abstract

A *real-valued random variable* X is a *measurable function* that maps *from* a *probability space* $(\Omega, \mathbb{E}, \mathsf{P})$ *to* $(\mathbb{R}, \le, \mathsf{d}, +, \cdot, \mathscr{B})$ where $\mathscr{B}$ is the *usual Borel σ-algebra* on the "real line" $(\mathbb{R}, \le, \mathsf{d}, +, \cdot)$ and where $\mathbb{R}$ is the *set of real numbers*, $\le$ is the standard linear order relation on $\mathbb{R}$, $\mathsf{d}(x, y) \triangleq |x - y|$ is the *usual metric* on $\mathbb{R}$, and $(\mathbb{R}, +, \cdot, 0, 1)$ is the standard *field* on $\mathbb{R}$. This text demonstrates that this definition of random variable is often a poor choice for computing statistics when the *probability space* that X maps *from* has structure that is dissimilar to that of the real line.

This text proposes two alternative statistical systems that, unlike the traditional method of a random variable X mapping exclusively to the real line, X instead maps more generally to (1) a *weighted graph* (2) an *ordered distance linear space* $\mathbb{R}^N$. In each mapping method, the structure that X maps to is preferably one that has order and metric geometry structures similar to that of the underlying stochastic process. And ideally the structure X maps *from* and the structure X maps *to* are, with respect to each other, both *isomorphic* and *isometric*.

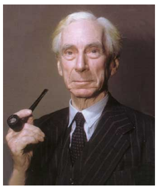
❝Here, on the level sand,
Between the sea and land,
What shall I build or write
Against the fall of night?❞

❝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) [1]

❝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 [2]



❝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. [3]

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

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

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

ACKNOWLEDGEMENTS

It is not far from the truth to say that much of the research presented in this text started with this email from Professor Po-Ning Chen of National Chiao-Tung University, Taiwan:[4]

```
                                          Sat, 25 Jan 2014 03:45:51 -0800 (PST)
Dear Dan: So far people are mostly (and are used to) dealing with
"frequency-transform" of a numerical sequences. However, sometimes, we
need to find out the "occurrence frequency" of a symbolic sequence. I am
wondering "Can we design a wavelet transform for a symbolic sequence like
DNA?" What do you think? See the attached paper as an example.
Po-Ning
```

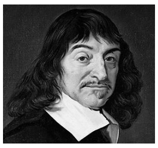
So I would like to say one more time to Professor Po-Ning Chen, “Thank you for introducing me to the topic”(!) which has to a large extent resulted in the research described in this text.[5]

[4] Po-Ning Chen:	陳伯寧	(pinyin: Chén Bó Nińg)
National Chiao-Tung University:	國立交通大學	(Gúo Lì Jiāo Tōng Dà Xúe).
Taiwan:	台灣	(pinyin: Tái Wān).

[5] The “attached paper” referred to by Professor Chen's email is Galleani and Garello (2010).

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 [6]

❝*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, [7]

symbol	description	reference
sets:		
$\varnothing$	*empty set*	Definition 1.1 (page 5)
$\mathbb{Z}$	*integers*	Definition 1.2 (page 5)
$\mathbb{W}$	*whole numbers*	Definition 1.2 (page 5)
$\mathbb{N}$	*natural numbers*	Definition 1.2 (page 5)
$\mathbb{Z}^*$	*extended set of integers*	Definition 1.2 (page 5)
$\mathbb{R}$	*real numbers*	Definition 1.2 (page 5)
$\mathbb{R}^{\vdash}$	*non-negative real numbers*	Definition 1.2 (page 5)
$\mathbb{R}^+$	*positive real numbers*	Definition 1.2 (page 5)
$\mathbb{R}^*$	*extended real numbers*	Definition 1.2 (page 5)
$\mathbb{C}$	*complex numbers*	Definition 1.10 (page 7)
values:		
π	*pi*	$3.14159265\ldots$
∞	*positive infinity*	
$-\infty$	*negative infinity*	
relations:		
®	*relation*	Definition 1.5 (page 6)
$(\triangle, \nabla)$	*ordered pair*	Definition 1.3 (page 5)
$X \times Y$	*Cartesian product of X and Y*	Definition 1.4 (page 6)
(a, b)	*ordered pair*	Definition 1.3 (page 5)
$®^{-1}$	*inverse* of relation ®	Definition 1.7 (page 6)
X^N		Definition 1.9 (page 6)
(a, b, c)	*ordered triple*	Definition 1.11 (page 7)
$\lvert z \rvert$	absolute value of a complex number z	Definition 1.24 (page 10)

...continued on next page...

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

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

symbol	description	reference
$=$	equality relation	
$\triangleq$	equality by definition	
$\to$	maps to	
$\in$	is an element of	
$\notin$	is not an element of	
$\mathcal{D}(®)$	*domain* of a relation ®	Definition 1.12 (page 7)
$\mathcal{I}(®)$	*image* of a relation ®	Definition 1.12 (page 7)
$\mathcal{R}(®)$	*range* of a relation ®	Definition 1.12 (page 7)
$\mathcal{N}(®)$	*null space* of a relation ®	Definition 1.12 (page 7)
relations on sets:		
$\subseteq$	subset	
$\subset$	proper subset	
$\supseteq$	super set	
$\supset$	proper superset	
$\not\subseteq$	is not a subset of	
$\not\subset$	is not a proper subset of	
sets of relations:		
2^{XY}	*set of all relations in* $X \times Y$	Definition 1.5 (page 6)
Y^X	*set of all functions in* $X \times Y$	Definition 1.6 (page 6)
operations on sets:		
$A \cup B$	set union	
$A \cap B$	set intersection	
$A \triangle B$	set symmetric difference	
$A \setminus B$	set difference	
A^c	set complement	
$\lvert \cdot \rvert$	set order	Definition 1.13 (page 7)
logic:		
$\neg$	logical NOT operation	
$\wedge$	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:		
$\vee$	join or least upper bound	Definition 1.26 (page 10)
$\wedge$	meet or greatest lower bound	Definition 1.27 (page 10)
$\le$	reflexive ordering relation	Definition 1.20 (page 9)
$\ge$	reflexive ordering relation	Definition 1.22 (page 9)
$<$	irreflexive ordering relation	Definition 1.22 (page 9)
$>$	irreflexive ordering relation	Definition 1.22 (page 9)
$(X, \le)$	*ordered set*	Definition 1.20 (page 9)
$[a : b]$	*closed interval* from a to b	Definition 1.23 (page 10)
$(a : b]$	*half-open interval* from a to b	Definition 1.23 (page 10)
$[a : b)$	*half-open interval* from a to b	Definition 1.23 (page 10)
$(a : b)$	*open interval* from a to b	Definition 1.23 (page 10)
$\lvert x \rvert$	*absolute value* of x	Definition 1.24 (page 10)

...continued on next page...

symbol	description	reference
$\lceil x\rceil$	*ceiling* of x	Definition 1.28 (page 11)
$\lfloor x\rfloor$	*floor* of x	Definition 1.28 (page 11)
ring operations:		
$\sum_{n\in\mathbb{D}} x_n$	*sum* over $\mathbb{D}$ of $(\!(x_n)\!)$	Definition D.12 (page 172)
$\mathsf{M}_\phi\big((\!(x_n)\!)\big)$	weighted *ϕ-mean*	Definition D.14 (page 174)
$\mathsf{M}_{\phi(x;p)}\big((\!(x_n)\!)\big)$	*power mean* with parameter p	Definition D.15 (page 174)
set structures:		
$\boldsymbol{T}$	a topology of sets	Definition D.16 (page 179)
$\boldsymbol{R}$	a ring of sets	
$\boldsymbol{A}$	an algebra of sets	
2^X	power set on a set X	Definition 1.8 (page 6)
topology:		
$(X,\boldsymbol{T})$	*topological space* on a set X	Definition D.16 (page 179)
A^-	*closure* of a set A	Definition D.18 (page 180)
A°	*interior* of a set A	Definition D.18 (page 180)
$\lim_{n\to\infty}(\!(x_n)\!)$	*limit* of $(\!(x_n)\!)$	Definition D.20 (page 181)
$(\!(x_n)\!)\to x$	*limit* of $(\!(x_n)\!)$	Definition D.20 (page 181)
distance spaces:		
$\mathsf{d}(p,q)$	*distance* function	Definition B.1 (page 133)
(X,d)	*distance space*	Definition B.1 (page 133)
$\mathsf{B}(x,r)$	*open ball* centered at x with radius r	Definition B.4 (page 134)
$\overline{\mathsf{B}}(x,r)$	*closed ball* centered at x with radius r	Definition B.4 (page 134)
$\operatorname{diam} A$	*diameter* of a set A	Definition B.2 (page 133)
τ	*power triangle function*	Definition C.1 (page 145)
$\triangle(p,\sigma;\mathsf{d})$	*power triangle inequality*	Definition C.2 (page 146)
probability:		
$(\Omega,\mathbb{E},\mathsf{P})$	*probability space*	Definition 1.38 (page 24)
Ω	*set of outcomes,*	Definition 1.38 (page 24)
$\mathbb{E}$	*σ-algebra* on Ω	Definition 1.38 (page 24)
P	*probability function* on $\mathbb{E}$	Definition 1.38 (page 24)
X	*random variable*	Definition 1.39 (page 24)
$\mathsf{E}(\mathsf{X})$	*traditional expected value* of X	Definition 1.40 (page 24)
$\mathsf{Var}(\mathsf{X})$	*traditional variance* of X	Definition 1.40 (page 24)
graphs:		
(X,E)	*graph*	Definition 1.17 (page 8)
$(X,E,\mathsf{d},\mathsf{w})$	*weighted graph*	Definition 1.18 (page 8)
$\mathsf{C}(\boldsymbol{G})$	*center* of a graph $\boldsymbol{G}$	Definition 1.19 (page 8)
outcome subspaces:		
$(\Omega,\le,\mathsf{d},\mathbb{E},\mathsf{P})$	*extended probability space*	Definition 2.1 (page 29)
$\mathsf{R}_{\mathsf{xy}}(n)$	*cross-correlation* of $(\!(x_n)\!)$ and $(\!(y_n)\!)$	Definition 3.2 (page 82)
$\mathsf{R}_{\mathsf{xx}}(n)$	*auto-correlation* of $(\!(x_n)\!)$	Definition 3.2 (page 82)
$\mathsf{m}_n(x,y)$	*nth-moment* from x to y	Definition 2.2 (page 29)
$\mathsf{m}(x,y)$	*moment* from x to y	Definition 2.2 (page 29)
$\dot{\mathsf{C}}(\boldsymbol{G})$	*outcome center* of $\boldsymbol{G}$	Definition 2.3 (page 30)
$\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G})$	*arithmetic center* of $\boldsymbol{G}$	Definition 2.4 (page 30)
$\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G})$	*geometric center* of $\boldsymbol{G}$	Definition 2.4 (page 30)
$\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G})$	*harmonic center* of $\boldsymbol{G}$	Definition 2.4 (page 30)
$\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G})$	*minimal center* of $\boldsymbol{G}$	Definition 2.4 (page 30)

...continued on next page...

symbol	description	reference
$\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G})$	*maxmin center* of $\boldsymbol{G}$	Definition 2.4 (page 30)
$\ddot{\mathsf{E}}(\mathsf{X})$	*outcome expected value*	Definition 2.14 (page 49)
$\ddot{\mathsf{Var}}(\mathsf{X}; \mathsf{E}_x)$	*outcome variance*	Definition 2.14 (page 49)
$\ddot{\mathsf{Var}}(\mathsf{X})$	*outcome variance*	Definition 2.14 (page 49)
$(X, \leq, \mathsf{d})$	*ordered distance space*	Definition 1.34 (page 22)
$(\mathbb{Z}, \leq, \lvert\cdot\rvert)$	*integer line*	Definition 1.36 (page 23)
$(\mathbb{R}, \lvert\cdot\rvert, \leq)$	*real line*	Definition 1.35 (page 22)
$\left(\{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}, \dot{\mathsf{d}}, \dot{\lesssim}, \dot{\mathsf{P}}\right)$		Definition 2.7 (page 30)
$\left(\{\text{①}, \text{②}, \text{③}, \text{④}, \text{⑤}, \text{⑥}\}, \dot{\mathsf{d}}, \varnothing, \dot{\mathsf{P}}\right)$		Definition 2.10 (page 31)
$\left(\{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}\}, \dot{\mathsf{d}}, \varnothing, \dot{\mathsf{P}}\right)$		Definition 2.11 (page 31)
$\left(\{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}, \boxed{N}\}, \dot{\mathsf{d}}, \varnothing, \dot{\mathsf{P}}\right)$		Definition 2.12 (page 31)
sequences:		
$(\!(x_n)\!)_{n\in\mathbb{D}}$	*sequence* over domain $\mathbb{D}$	Definition 1.41 (page 25)
$(\!(x_n)\!)$	*sequence*	Definition 1.41 (page 25)
$(\!(z_n)\!)_{\mathbb{D}} \triangleq (\!(x_n)\!)_{\mathbb{D}_1} \star (\!(y_n)\!)_{\mathbb{D}_2}$	*convolution* of $(\!(x_n)\!)$ and $(\!(y_n)\!)$	Definition 1.43 (page 25)
$\mathsf{DFT}(\!(x_n)\!)$	*Discrete Fourier Transform* on $(\!(x_n)\!)$	Definition 1.52 (page 27)
$\alpha + (\!(x_n)\!)$	*constant-sequence addition*	Definition 1.44 (page 26)
$(\!(x_n)\!) + \alpha$	*constant-sequence addition*	Definition 1.44 (page 26)
$\alpha\,(\!(x_n)\!)$	*constant-sequence multiplication*	Definition 1.44 (page 26)
$(\!(x_n)\!)\,\alpha$	*constant-sequence multiplication*	Definition 1.44 (page 26)
linear spaces:		
$\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$	*linear space* over X	Definition D.1 (page 157)
$\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d}\right)$	*metric linear space*	Definition D.3 (page 158)
$\left(\mathbb{R}^N, \leq, \mathsf{d}\right)$	$\mathbb{R}^N$ *ordered distance linear space*	
$\lVert\cdot\rVert$	vector norm	Definition D.4 (page 159)
ring operations:		
$\oplus$	*addition operator*	Definition 1.15 (page 7)
$\otimes$	*multiplication operator*	Definition 1.16 (page 8)

CONTENTS

Preface

❝*You know that I write slowly. This is chiefly because I am never satisfied until I have said as much as possible in a few words, and writing briefly takes far more time than writing at length.*❞
Karl Friedrich Gauss (1777–1855), German mathematician [8]

This book is largely based on four papers written by the same author as this text:

1. Greenhoe (2015): *Order and metric geometry compatible stochastic processing.*
   https://peerj.com/preprints/844/
   https://www.researchgate.net/publication/303057738
   This paper is the basis for CHAPTER 2.

2. Greenhoe (2016c): *Order and Metric Compatible Symbolic Sequence Processing.*
   https://peerj.com/preprints/2052/
   https://www.researchgate.net/publication/302947227
   This paper is the basis for CHAPTER 3.
   The C++ source code used to generate the 128 or so TeX files for the data plots in the paper can be downloaded from here:
   https://www.researchgate.net/publication/302953844

3. Greenhoe (2016a): *An extension to the spherical metric using polar linear interpolation.*
   https://www.researchgate.net/publication/303984226
   This paper is the basis for APPENDIX A.

4. Greenhoe (2016b): *Properties of distance spaces with power triangle inequalities.*
   http://www.journals.pu.if.ua/index.php/cmp/article/view/483
   https://peerj.com/preprints/2055
   https://www.researchgate.net/publication/281831459
   This paper is the basis for APPENDIX B and APPENDIX C. It has been published in the 2016 volume 8 number 1 edition of the journal *Carpathian Mathematical Publications.*

[8] quote: Simmons (2007), page 177
image: http://en.wikipedia.org/wiki/Karl_Friedrich_Gauss

CHAPTER 1

FOUNDATIONS

"Je me plaisois surtout aux mathématiques, à cause de la certitude et de l'évidence de leurs raisons: mais je ne remarquois point encore leur vrai usage; et, pensant qu'elles ne servoient qu'aux arts mécaniques, je m'étonnois de ce que leurs fondements étant si fermes et si solides, on n'avoit rien bâti dessus de plus relevé:"

"I was especially delighted with the mathematics, on account of the certitude and evidence of their reasonings; but I had not as yet a precise knowledge of their true use; and thinking that they but contributed to the advancement of the mechanical arts, I was astonished that foundations, so strong and solid, should have had no loftier superstructure reared on them."

René Descartes, philosopher and mathematician (1596–1650) [1]

1.1 Introduction

In *traditional stochastic processing*, a *real-valued random variable* X (or we might even say "a *traditional random variable* X) first maps the underlying *probability space* $(\Omega, \mathbb{E}, \mathsf{P})$ to $(\mathbb{R}, \le, \mathsf{d}, +, \cdot, \mathscr{B})$ where $\mathscr{B}$ is the *usual Borel σ-algebra* on the "real line" $(\mathbb{R}, \le, \mathsf{d}, +, \cdot)$ (see Figure 1.1 page 2), and then operations (such as the expected value operation $\mathsf{E}(X)$) is performed on X. Here, *real line* refers to the structure $(\mathbb{R}, |\cdot|, \le)$, where $\mathbb{R}$ is the *set of real numbers*, $\le$ is the standard linear order relation on $\mathbb{R}$, and $\mathsf{d}(x, y) \triangleq |x - y|$ is the *usual metric* on $\mathbb{R}$.

This is all well and good when the physical process being analyzed (in the case of statistical estimation or system analysis) or being processed (as in the case of signal processing, including digital signal processing) is also *linearly ordered* and has a *metric geometry* similar to the one induced by the *usual metric* on $\mathbb{R}$.

Be that as it may, in several real world applications, this is simply not the case. Take these processes for example:

[1] quote: Descartes (1637a)
translation: Descartes (1637b) ⟨part I, paragraph 10⟩
image: http://en.wikipedia.org/wiki/Image:Descartes_Discourse_on_Method.png

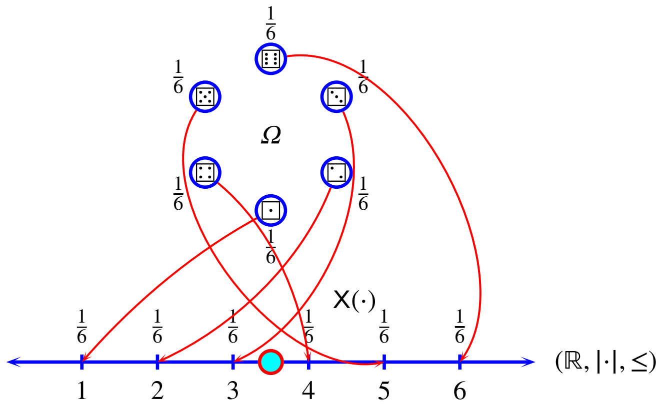
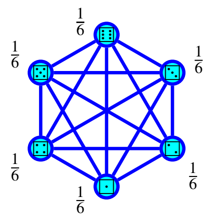
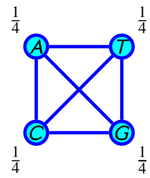


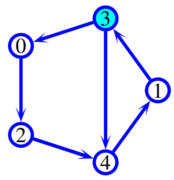
Figure 1.1: *traditional random variable* X mapping a *stochastic process* to the *real line*



The values of a *fair die* { ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ } have absolutely no order structure, and have no metric except the *discrete metric*. On a fair die, ⚁ is not greater or less than ⚀; rather ⚁ and ⚀ are simply symbols without order. Moreover, ⚀ is not "closer" to ⚁ than it is to ⚂; rather, ⚀, ⚁, and ⚂ are simply symbols without any inherit order or metric geometry.



Genomic Signal Processing (*GSP*) analyzes biological sequences called *genomes*. These sequences are constructed over a set of 4 symbols that are commonly referred to as A, T, C, and G, each of which corresponds to a nucleobase (adenine, thymine, cytosine, and guanine, respectively).[2]A typical genome sequence contains a large number of symbols (about 3 billion for humans, 29751 for the SARS virus).[3]



A *linear congruential pseudo-random number generator* induced by the equation $y_{n+1} = (y_n + 2) \mod 5$ with $y_0 = 1$. The sequential nature of the structure induces both a natural order and distance.

In all three processes, the symbols in general have an order structure and a *metric geometry* that is fundamentally dissimilar from that of the *real line*. Therefore, statistical inferences based on the *real line* will likely lead to results that arguably have little relationship with intuition or reality.

So we can observe the following:

1. A traditional random variable X maps to the real line.
2. The structure of the underlying stochastic process (the domain of X) may be very dissimilar to that of the real line.
3. To fix the problem, we need random variables that map to alternative structures that are more similar to the underlying stochastic processes.
4. Such a structure should have both an *order relation* and a *distance function* defined on it that are similar to the stochastic process. In particular, such a structure should be an

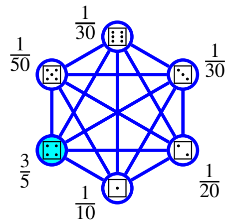
[1] Mendel (1853) ⟨Mendel (1853): gene coding uses discrete symbols⟩, Watson and Crick (1953a) page 737 ⟨Watson and Crick (1953): gene coding symbols are adenine, thymine, cytosine, and guanine⟩, Watson and Crick (1953b) page 965, Pommerville (2013) page 52

[2] GenBank (2014) ⟨http://www.ncbi.nlm.nih.gov/genome/guide/human/⟩⟨Homo sapiens, NC_000001–NC_000022 (22 chromosome pairs), NC_000023 (X chromosome), NC_000024 (Y chromosome), NC_012920 (mitochondria)⟩, GenBank (2014) ⟨http://www.ncbi.nlm.nih.gov/nuccore/30271926⟩⟨SARS coronavirus, NC_004718.3⟩ S. G. Gregory (2006) ⟨homo sapien chromosome 1⟩, Runtao He (2004) ⟨SARS coronavirus⟩

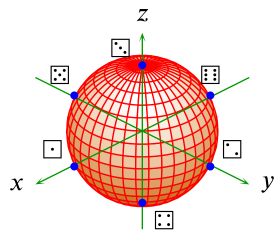
ordered distance space. And ideally, the stochastic process and the ordered distance space should be both **isomorphic** and **isometric** with respect to each other.

This text proposes two *ordered distance spaces* for stochastic processing:

CHAPTER 2 proposes mapping to *directed weighted graphs*. In such a structure, order is represented by direction of it's edges, and distance by the lengths of it's edges. Furthermore, probability can be represented by weights assigned to it's vertices.

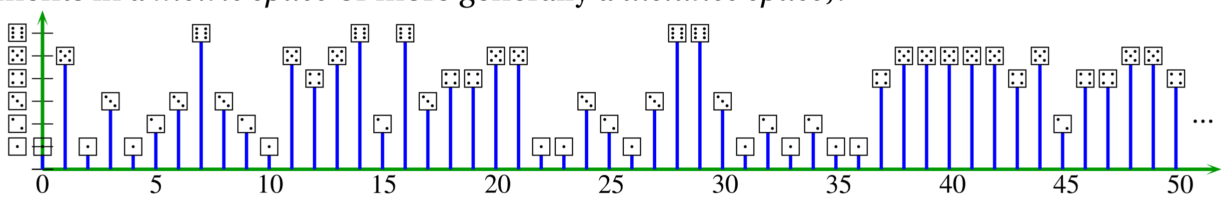


CHAPTER 3 proposes mapping to an *ordered distance linear space* $(\mathbb{R}^n, \le, \mathsf{d}, +, \cdot, \mathbb{R}, \dot{+}, \dot{\times})$, where $(\mathbb{R}, \dot{+}, \dot{\times}, 0, 1)$ is a *field*, $+$ is the vector addition operator on $\mathbb{R}^n \times \mathbb{R}^n$, and $\cdot$ is the scalar-vector multiplication operator on $\mathbb{R} \times \mathbb{R}^n$. Probability has no natural representation here, and must be assigned through a separate *probability function* P.



Mapping to a weighted graph as in CHAPTER 2 is useful for analysis of a single random variable X. But the traditional expectation value $\mathsf{E}(\mathsf{X})$ of X is often a poor choice of a statistic. For example, the traditional expected value of a fair die is $\mathsf{E}(X) = \frac{1}{6}(1+2+\cdots+6) = 3.5$. But this result has no relationship with reality or with intuition because the result implies that we expect the value of ⚂ or ⚃ more than we expect the outcome of say ⚀ or ⚁. The fact is, that for a fair die, we would expect any pair of values equally. The reason for this is that the values of the face of a fair die are merely symbols with no order, and with no metric geometry other than the *discrete metric geometry*. Weighted graphs on the other hand offer structures more similar to the underlying stochastic process. And the *expectation* EX of X can be defined simply as the *center* of the *weighted graph* (as illustrated above for a *weighted die* with EX shaded in blue).

However, the mapping has limitations with regards to a *sequence* of random variables in performing sequence analysis (using for example *Fourier analysis* or *wavelet analysis*), in performing sequence processing (using for example *FIR filtering* or *IIR filtering*), in making diagnostic measurements (using a post-transform metric space), or in making "optimal" decisions (based on "distance" measurements in a *metric space* or more generally a *distance space*).



Mapping to an *ordered distance linear space* $\boldsymbol{Y}$ as in CHAPTER 3 provides the *linear space* component of $\boldsymbol{Y}$, which in turn provides a much more convenient framework for *sequence* analysis and processing (as compared to the weighted graph). The *ordered set* and *distance space* components of $\boldsymbol{Y}$ allow one to preserve the order structure and distance geometry inherent in the underlying stochastic process, which can provide a less distorted (as compared to the *real line*) framework for analysis, diagnostics, and optimal decision making.

For any given stochastic process, there are an infinite number of possible random variable mappings, with some being "better"—with respect to some criteria (probably involving *isomorphic* and *isometric* properties)—than others. Multiple mappings using random variables W, X, Y, and Z for the "*real die*" stochastic process $\boldsymbol{G}$ are illustrated in Figure 1.2 (page 4).

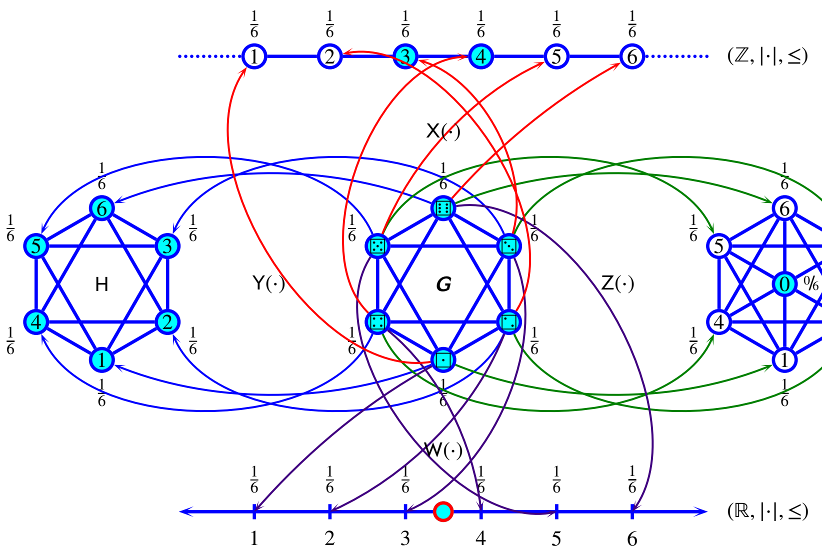
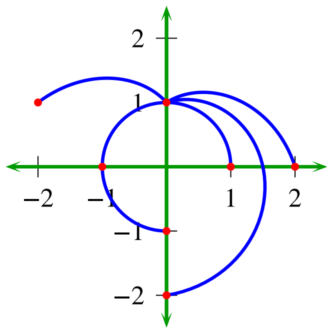


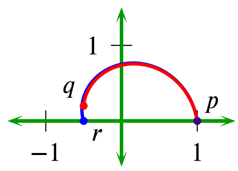
Figure 1.2: several random variable mappings for the *real die*

1.2 Further mathematical support

Further mathematical support for these two methods is provided in appendices. Here is partial summary of what may be found there:



APPENDIX A introduces what is herein called the *Lagrange arc distance* function. It is important in this text because it is used in CHAPTER 3 in the processing of *real die sequences* in $\mathbb{R}^3$ and *spinner sequences* in $\mathbb{R}^2$. The function is an extension to *all* of $\mathbb{R}^N$ of the *spherical metric*, which has as domain only a "spherical" *surface* in $\mathbb{R}^N$. The *Lagrange arc distance* $\mathsf{d}(p,q)$ draws arcs between certain pairs of points (p,q) using *Lagrange interpolation*.



APPENDIX B presents *distance spaces*. A *distance space* is a *metric space* in which the *triangle inequality* does not necessarily hold. *Distance spaces* are important in this text because the *Lagrange arc distance* is a *distance* function, and *not* a *metric* in that $\mathsf{d}(p,r) \le \mathsf{d}(p,q)+\mathsf{d}(q,r)$ does *not* necessarily hold for all *triples* (p,q,r) in $\mathbb{R}^N$.

APPENDIX C introduces what is herein called the *power distance space*. This space is a generalization of the *metric space*, and is a *distance space* that satisfies what is herein called the *power triangle inequality*:

$$\mathsf{d}(x,y) \;\le\; 2\sigma\left[\tfrac{1}{2}\mathsf{d}^p(x,z)+\tfrac{1}{2}\mathsf{d}^p(z,y)\right]^{\frac{1}{p}} \qquad \forall (p,\sigma)\in\mathbb{R}^*\times\mathbb{R}^{\vdash},\quad x,y,z\in X$$

It turns out that the inequality $2\sigma \le 2^{\frac{1}{p}}$ has special significance with regards to these spaces and appears repeatedly in APPENDIX C. It is plotted in Figure 1.3 (page 5). *Power distance spaces* are not explicitly used in this text. However, they may prove useful to future research in symbolic sequence processing space in which the *triangle inequality* fails to hold, but in which some of the key properties of a metric space are still required.

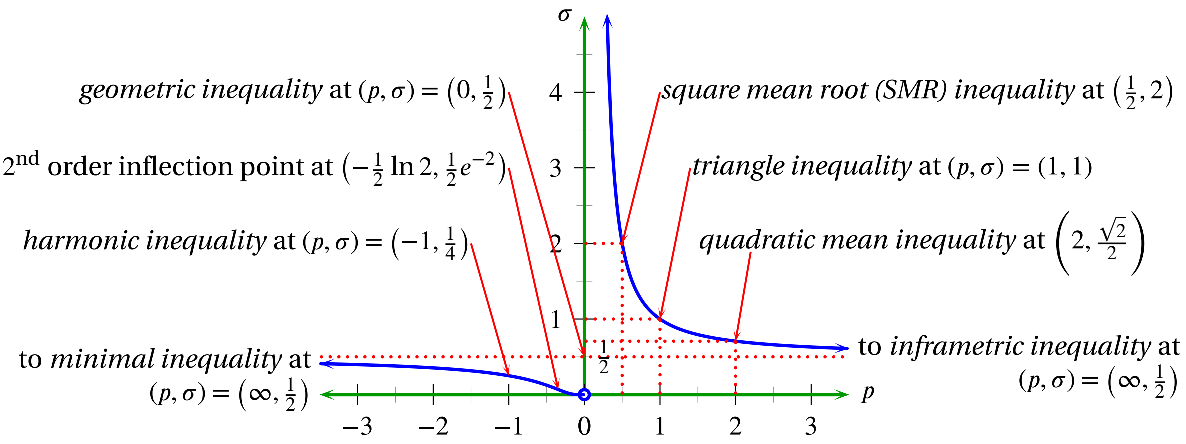


Figure 1.3: $\sigma = \frac{1}{2}(2^{\frac{1}{p}}) = 2^{\frac{1}{p}-1}$ or $p = \frac{\ln 2}{\ln(2\sigma)}$

1.3 Standard definitions

1.3.1 Sets

Definition 1.1. [3] *Let X be a* SET.[4]

DEF *The **empty set** $\varnothing$ is defined as* $\varnothing \triangleq \{x \in X | x \neq x\}$.

Definition 1.2. [5] *Let $\mathbb{R}$ be the **set of real numbers**. Let $\mathbb{R}^{\vdash} \triangleq \{x \in \mathbb{R} | x \geq 0\}$ be the **set of non-negative real numbers**. Let $\mathbb{R}^{+} \triangleq \{x \in \mathbb{R} | x > 0\}$ be the **set of postive real numbers**. Let $\mathbb{R}^{*} \triangleq \mathbb{R} \cup \{-\infty, \infty\}$ be the set of **extended real numbers**.* [6] *Let $\mathbb{Z}$ be the **set of integers**. Let $\mathbb{W} \triangleq \{n \in \mathbb{Z} | n \geq 0\}$ be the **set of whole numbers**. Let $\mathbb{N} \triangleq \{n \in \mathbb{Z} | n \geq 1\}$ be the **set of natural numbers**. Let $\mathbb{Z}^{*} \triangleq \mathbb{Z} \cup \{-\infty, \infty\}$ be the* EXTENDED SET OF INTEGERS.

1.3.2 Relations

One of the most fundamental structures in mathematics is the *ordered pair*, and one of the most common definitions of *ordered pair* is due to *Kuratowski* (1921) and is presented next:[7]

Definition 1.3. [8]

DEF *The **ordered pair** (a,b) is defined as* $(a,b) \triangleq \{\{a\},\{a,b\}\}$.

Proposition 1.1 (next) and Corollary 1.1 demonstrate that the the definition of *ordered pair* given by Definition 1.3 allows a and b to be unambiguously extracted from (a,b) and that (a,b) is well defined.

[3] Halmos (1960) page 8, Kelley (1955) page 3, Kuratowski (1961), page 26

[4] The mathematical structure called *set* is left undefined in this paper. For more information on *sets*, see for example Zermelo (1908a) pages 263–267 (7 axioms), Zermelo (1908b) ⟨(English translation of previous reference)⟩, Fraenkel (1922), Halmos (1960) pages 1–6 ⟨Naive set theory⟩, Wolf (1998), page 139

[5] Notation $\mathbb{R}, \mathbb{W}, \mathbb{N}$, etc.: *Bourbaki notation*. References: Davis (2005) page 9, Cohn (2012) page 3

[6] Rana (2002) pages 385–388 ⟨Appendix A⟩

[7] As an alternative to the Kuratowski definition, the *ordered pair* can also be taken as an *axiom*. References: Bourbaki (1968), page 72, Munkres (2000), page 13

[8] Suppes (1972) page 32, Halmos (1960) page 23, Kuratowski (1961), page 39, Kuratowski (1921) ⟨Def. V, page 171⟩, Wiener (1914)

Proposition 1.1.

THM
$$\begin{array}{rclclcl}\{a\} &=& \bigcap(a,b) &=& \bigcap\{\{a\},\{a,b\}\} &=& \{a\}\cap\{a,b\}\\ \{b\} &=& \triangle(a,b) &=& \triangle\{\{a\},\{a,b\}\} &=& \{a\}\triangle\{a,b\}\end{array}$$

Corollary 1.1. [9]

COR
$$(a,b)=(c,d) \iff \{a=c \text{ and } b=d\}$$

✎PROOF:
$$\begin{array}{rclclcll}\{a\} &=& \bigcap(a,b) &=& \bigcap(c,d) &=& \{c\} & \text{by Proposition 1.1 and left hypothesis}\\ \{b\} &=& \triangle(a,b) &=& \triangle(c,d) &=& \{d\} & \text{by Proposition 1.1 and left hypothesis}\\ (a,b) &=& (c,d) &&&&& \text{by right hypothesis}\end{array}$$
✏

Definition 1.4. [10] *Let X and Y be sets.*

DEF
*The **Cartesian product** $X\times Y$ is defined as*
$$X\times Y \triangleq \{(x,y)\,|\,(x\in X) \text{ and } (y\in Y)\}$$

Definition 1.5. [11] *Let X and Y be* SETS.

DEF
*A **relation** ® on X and Y is any subset of $X\times Y$ such that ® $\subseteq X\times Y$. The set 2^{XY} is the **set of all relations** in $X\times Y$.*

Definition 1.6. [12] *Let X and Y be* SETS.

DEF
A RELATION $\mathsf{f}\in 2^{XY}$ *is a **function** if*
$$\{(x,y_1)\in\mathsf{f} \quad\text{and}\quad (x,y_2)\in\mathsf{f}\} \implies \{y_1=y_2\}.$$
*The set Y^X is the **set of all functions** in 2^{XY}.*

A function does not always have an inverse that is also a function. But unlike functions, *every* relation has an inverse that is also a relation. Note that since all functions are relations, every function *does* have an inverse that is at least a relation, and in some cases this inverse is also a function.

Definition 1.7. [13] *Let ® be a relation in 2^{XY}.*

DEF
*®$^{-1}$ is the **inverse** of relation ® if*
$$®^{-1} \triangleq \{(y,x)\in Y\times X\,|\,(x,y)\in ®\}$$
*The inverse relation ®$^{-1}$ is also called the **converse** of ®.*

Definition 1.8. *Let X be a set.*

DEF
The quantity 2^X is the POWER SET OF X *such that*
$$2^X \triangleq \{A\subseteq X\} \quad \text{(the set of all subsets of } X\text{)}.$$

Definition 1.9. *Let Y be a* SET.

DEF
The structure Y^n for $n\in\mathbb{N}$ is a SET *defined as*
$$\begin{array}{rcll}Y^1 &\triangleq& Y & \text{and}\\ Y^n &\triangleq& Y\times Y^{n-1} & \text{for } n=2,3,4,\ldots\end{array}$$

[9] Apostol (1975) page 33, Hausdorff (1937) page 15

[10] Halmos (1960) page 24, G. Frege, 2007 August 25, http://groups.google.com/group/sci.logic/msg/3b3294f5ac3a76f0

[11] Maddux (2006) page 4, Halmos (1960) pages 26–30, Suppes (1972) page 86, Kelley (1955) page 10, Bourbaki (1939), Bottazzini (1986) page 7, Comtet (1974) page 4 ⟨$|Y^X|$⟩; The notation 2^{XY} is motivated by the fact that for finite X and Y, $|2^{XY}|=2^{|X|\cdot|Y|}$.

[12] Maddux (2006) page 4, Halmos (1960) pages 26–30, Suppes (1972) page 86, Kelley (1955) page 10, Bourbaki (1939), Bottazzini (1986) page 7, Comtet (1974) page 4 ⟨$|Y^X|$⟩; The notation Y^X is motivated by the fact that for finite X and Y, $|Y^X|=|Y|^{|X|}$.

[13] Suppes (1972) page 61 ⟨Defintion 6, inverse="converse"⟩
Kelley (1955) page 7
Peirce (1883) page 188 ⟨inverse="converse"⟩

Definition 1.10.

*The **set of complex numbers** $\mathbb{C}$ is defined as $\mathbb{C} \triangleq \mathbb{R}^2$.*

Definition 1.11. *Let $Y_1, Y_2, \ldots Y_N$ be* SETS.

*The structure $(x_1, x_2, \cdots, x_n)$ is an **n-tuple** on $Y_1 \times Y_2 \times \cdots \times Y_N$ if $(x_1, x_2, \cdots, x_n)$ is an element in the set $Y_1 \times Y_2 \times \cdots \times Y_N$. A 3-TUPLE is also called an **ordered triple** or simply a **triple**.*

Definition 1.12. [14] *Let $® \in 2^{XY}$ be a* RELATION (Definition 1.5 page 6).

*The **domain** of ® is*	$\mathcal{D}(®)$	$\triangleq$	$\{x \in X \mid \exists y$ such that $(x, y) \in ®\}$.
*The **image set** of ® is*	$\mathcal{I}(®)$	$\triangleq$	$\{y \in Y \mid \exists x$ such that $(x, y) \in ®\}$.
*The **null space** of ® is*	$\mathcal{N}(®)$	$\triangleq$	$\{x \in X \mid (x, 0) \in ®\}$.
*The **range** of ® is*	*any set $\mathcal{R}(®)$ such that $\mathcal{I}(®) \subseteq \mathcal{R}(®)$.*		

Definition 1.13.

The SET FUNCTION[15] $|A| \in \mathbb{Z}^{*2^X}$ *is the* CARDINALITY OF A *such that*

$$|A| \triangleq \begin{cases} \text{the number of elements in } A & \text{for FINITE } A \\ \infty & \text{otherwise} \end{cases} \qquad \forall A \in 2^X$$

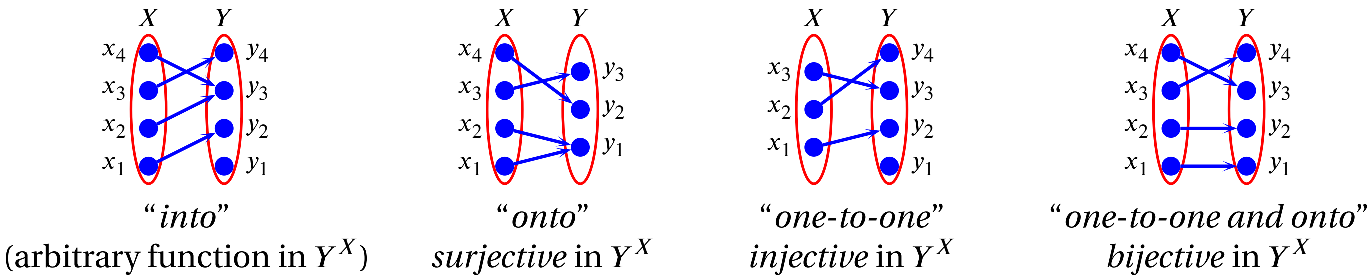


Figure 1.4: types of functions (Definition 1.14 page 7)

Definition 1.14. [16] *Let f be a* FUNCTION *in Y^X* (Definition 1.6 page 6).

f *is **surjective***	*(also called **onto**)*	*if*	$f(X) = Y$.
f *is **injective***	*(also called **one-to-one**)*	*if*	$f(x) = f(y) \implies x = y$.
f *is **bijective***	*(also called **one-to-one and onto**)*	*if*	f *is both* SURJECTIVE *and* INJECTIVE.

1.3.3 Field operator pairs

Definition 1.15. *Let X, Y, and Z be* SETS. *Let $(\mathbb{R}, +, \cdot, 0, 1)$ be the standard* FIELD OF REAL NUMBERS.

*The **addition operator** $\oplus : X \times Y \to Z$ is defined as shown in Table 1.1 (page 8).*

Moreover, for some sequence $(\!(x_n)\!)$,

$$\bigoplus_{n=1}^{N} x_n \triangleq x_1 \oplus x_2 \oplus \cdots \oplus x_N \qquad \text{and} \qquad \bigoplus_{n \in \mathbb{D}} x_n \triangleq x_{\alpha \in \mathbb{D}} \oplus x_{\beta \in \mathbb{D}} \oplus \cdots \oplus x_{\gamma \in \mathbb{D}}$$

---

[14] Munkres (2000), page 16, Kelley (1955) page 7

[15] *set function*: Pap (1995) page 8 ⟨Definition 2.3: extended real-valued set function⟩, Halmos (1950) page 30 ⟨§7. MEASURE ON RINGS⟩

[16] Michel and Herget (1993) pages 14–15, Fuhrmann (2012) page 2, Comtet (1974) page 5, Durbin (2000) pages 16–17

1.	If	$X \times Y \triangleq \mathbb{R} \times \mathbb{R}$	then	$x \oplus y$	$\triangleq$	$x + y$	$\in \mathbb{R}$.
2.	If	$X \times Y \triangleq \mathbb{C} \times \mathbb{C}$	then	$(a, b) \oplus (c, d)$	$\triangleq$	$(a + c, b + d)$	$\in \mathbb{C}$.
3.	If	$X \times Y \triangleq \mathbb{R}^n \times \mathbb{R}^n$	then	$\left(x_1, x_2, \cdots, x_n\right) \oplus \left(y_1, y_2, \cdots, y_n\right)$	$\triangleq$	$\left(x_1 + y_1, \cdots, x_n + y_n\right)$	$\in \mathbb{R}^n$.
4.	If	$X \times Y \triangleq \mathbb{C}^n \times \mathbb{C}^n$	then	$\left(x_1, x_2, \cdots, x_n\right) \oplus \left(y_1, y_2, \cdots, y_n\right)$	$\triangleq$	$\left(x_1 \oplus y_1, \cdots, x_n \oplus y_n\right)$	$\in \mathbb{C}^n$.
5.	If	$X \times Y \triangleq \mathbb{R} \times \mathbb{C}$	then	$x \oplus (a, b)$	$\triangleq$	$(x + a, x + b)$	$\in \mathbb{C}$.
6.	If	$X \times Y \triangleq \mathbb{C} \times \mathbb{R}$	then	$x \oplus y$	$\triangleq$	$y \oplus x$	$\in \mathbb{C}$.

Table 1.1: Definition of the *addition operator* $\oplus : X \times Y \to Z$ (see Definition 1.15 page 7)

Definition 1.16. *Let X, Y, and Z be* SETS. *Let* $(\mathbb{R}, +, \cdot, 0, 1)$ *be the standard* FIELD OF REAL NUMBERS. *Let the* JUXTAPOSITION *operator xy on x and y be equivalent to the real field operator $x \cdot y$ on x and y such that $xy \triangleq x \cdot y$.*

DEF *The **multiplication operator** $\otimes : X \times Y \to Z$ is defined as shown in Table 1.2 (page 8).*

1.	If	$X \times Y \triangleq \mathbb{R} \times \mathbb{R}$	then	$x \otimes y$	$\triangleq$	xy	$\in \mathbb{R}$.
2.	If	$X \times Y \triangleq \mathbb{R} \times \mathbb{C}$	then	$x \otimes (a, b)$	$\triangleq$	(xa, xb)	$\in \mathbb{C}$.
3.	If	$X \times Y \triangleq \mathbb{C} \times \mathbb{C}$	then	$(a, b) \otimes (c, d)$	$\triangleq$	$(ac - bd, ad + bc)$	$\in \mathbb{C}$.
4.	If	$X \times Y \triangleq \mathbb{R} \times \mathbb{R}^n$	then	$x \otimes \left(y_1, y_2, \cdots, y_n\right)$	$\triangleq$	$\left(xy_1, xy_2, \cdots, xy_n\right)$	$\in \mathbb{R}^n$.
5.	If	$X \times Y \triangleq \mathbb{C} \times \mathbb{R}^n$	then	$x \otimes \left(y_1, y_2, \cdots, y_n\right)$	$\triangleq$	$\left(y_1 \otimes x, y_2 \otimes x, \cdots, y_n \otimes x\right)$	$\in \mathbb{C}^n$.
6.	If	$X \times Y \triangleq \mathbb{C} \times \mathbb{C}^n$	then	$x \otimes \left(y_1, y_2, \cdots, y_n\right)$	$\triangleq$	$\left(x \otimes y_1, x \otimes y_2, \cdots, x \otimes y_n\right)$	$\in \mathbb{C}^n$.
7.	If	$X \times Y \triangleq \mathbb{C} \times \mathbb{R}$	then	$x \otimes y$	$\triangleq$	$y \otimes x$	$\in \mathbb{C}$.
8.	If	$X \times Y \triangleq \mathbb{R}^n \times \mathbb{R}$	then	$x \otimes y$	$\triangleq$	$y \otimes x$	$\in \mathbb{R}^n$.
9.	If	$X \times Y \triangleq \mathbb{R}^n \times \mathbb{C}$	then	$x \otimes y$	$\triangleq$	$y \otimes x$	$\in \mathbb{C}^n$.
10.	If	$X \times Y \triangleq \mathbb{C}^n \times \mathbb{C}$	then	$x \otimes y$	$\triangleq$	$y \otimes x$	$\in \mathbb{C}^n$.

Table 1.2: Definition of the *multiplication operation* $\otimes : X \times Y \to Z$ (see Definition 1.16 page 8)

1.3.4 Graph Theory

Definition 1.17. [17] *Let 2^{XX} be the set of all* RELATIONS (Definition 1.5 page 6) *on a set X.*

DEF *The pair (X, E) is a **graph** if $E \subseteq 2^{XX}$. A graph (X, E) is **undirected** if $(x, y) \in E \implies (y, x) \in E$. A graph (X, E) is **directed** if it is* NOT *undirected. A* GRAPH *that is* DIRECTED *is a **directed graph**. A* GRAPH *that is* UNDIRECTED *is an **undirected graph**. The elements of X are the **vertices** and the ordered pairs of E are the **arcs** of a* GRAPH (X, E). *The element x is the **tail** and y is the **head** of an arc (x, y).*

Definition 1.18.

DEF *The tuple $(X, E, \mathsf{d}, \mathsf{w})$ is a **weighted graph** if (X, E) is a* GRAPH (Definition 1.17 page 8), $\mathsf{d}(x, y)$ *is a* FUNCTION *in* $(\mathbb{R}^{\vdash})^{X \times X}$ (Definition 1.6 page 6), *and $\mathsf{w}(x)$ is a function in $(\mathbb{R}^{\vdash})^{X}$. The function d is called the **edge weight**, and the function w is called the **vertice weight**.*

Definition 1.19. *Let $\boldsymbol{G} \triangleq (X, E, \mathsf{d}, \mathsf{w})$ be a* WEIGHTED GRAPH (Definition 1.18 page 8).

DEF *The **center** $\mathbb{C}(\boldsymbol{G})$ of $\boldsymbol{G}$ is* $\quad \mathbb{C}(\boldsymbol{G}) \triangleq \arg\min_{x \in X} \max_{y \in X} \mathsf{d}(x, y)\, \mathsf{w}(y).$

[17] orgen Bang-Jensen and Gutin (2007), page 2 ⟨§1.2⟩, Harary (1969), page 9

1.4 Order space concepts

1.4.1 Order

Definition 1.20. [18] *Let X be a set.*

DEF

*A relation $\le$ is an **order relation** in 2^{XX} (Definition 1.5 page 6) if*

1. $x \le x$ $\quad \forall x \in X$ (REFLEXIVE) and — PREORDER
2. $x \le y$ and $y \le z \implies x \le z$ $\quad \forall x,y,z \in X$ (TRANSITIVE) and — PREORDER
3. $x \le y$ and $y \le x \implies x = y$ $\quad \forall x,y \in X$ (ANTI-SYMMETRIC).

*An **ordered set** is the pair $(X, \le)$. If $\le = \varnothing$ (Definition 1.1 page 5), then $(X, \le)$ is an **unordered set**. If $x \le y$ or $y \le x$, then elements x and y are said to be **comparable**; otherwise they are **incomparable**.*

Definition 1.21. [19]

DEF

*A relation $\le$ is a **linear order relation** on X if*

1. $\le$ is an ORDER RELATION (Definition 1.20 page 9) and
2. $x \le y$ or $y \le x$ $\quad \forall x,y \in X$ (COMPARABLE).

*A **linearly ordered set** is the pair $(X, \le)$.*
*A linearly ordered set is also called a **totally ordered set**, a **fully ordered set**, and a **chain**.*

The familiar relations $\ge$, $<$, and $>$ (next) can be defined in terms of the order relation $\le$ (Definition 1.20—previous).

Definition 1.22. [20] *Let $(X, \le)$ be an ordered set.*

DEF

The relations $\ge, <, > \in 2^{XX}$ are defined as follows:

$$x \ge y \stackrel{\text{def}}{\iff} y \le x \qquad \forall x,y \in X$$

$$x \lneq y \stackrel{\text{def}}{\iff} x \le y \text{ and } x \ne y \qquad \forall x,y \in X$$

$$x \gneq y \stackrel{\text{def}}{\iff} x \ge y \text{ and } x \ne y \qquad \forall x,y \in X$$

*The relation $\ge$ is called the **dual** of $\le$.*

Example 1.1 (Coordinatewise order relation). [21] Let $(X, \le)$ be an ordered set.
Let $\boldsymbol{x} \triangleq (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} \triangleq (y_1, y_2, \ldots, y_n)$.

EX

The **coordinatewise order relation** $\lesssim$ on the Cartesian product X^n is defined for all $\boldsymbol{x}, \boldsymbol{y} \in X^n$ as

$$\boldsymbol{x} \lesssim \boldsymbol{y} \stackrel{\text{def}}{\iff} \{x_1 \le y_1 \text{ and } x_2 \le y_2 \text{ and } \ldots \text{ and } x_n \le y_n\}$$

Example 1.2 (Lexicographical order relation). [22] Let $(X, \le)$ be an ordered set.
Let $\boldsymbol{x} \triangleq (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} \triangleq (y_1, y_2, \ldots, y_n)$.

---

[18] MacLane and Birkhoff (1999) page 470, Beran (1985) page 1, Korselt (1894) page 156 ⟨I, II, (1)⟩, Dedekind (1900) page 373 ⟨I–III⟩. An *order relation* is also called a **partial order relation**. An *ordered set* is also called a **partially ordered set** or **poset**.
[19] MacLane and Birkhoff (1999) page 470, Ore (1935) page 410
[20] Peirce (1880) page 2
[21] Shen and Vereshchagin (2002) page 43
[22] Shen and Vereshchagin (2002) page 44, Halmos (1960) page 58, Hausdorff (1937) page 54

EX

The **lexicographical order relation** $\precsim$ on the Cartesian product X^n is defined forall $\boldsymbol{x}, \boldsymbol{y} \in X^n$ as

$$\boldsymbol{x} \precsim \boldsymbol{y} \overset{\text{def}}{\iff} \left\{ \begin{array}{lllll} (& x_1 < y_1 & & &) \text{ or} \\ (& x_2 < y_2 & \text{and} & x_1 = y_1 &) \text{ or} \\ (& x_3 < y_3 & \text{and} & (x_1, x_2) = (y_1, y_2) &) \text{ or} \\ (& \dots & \dots & \dots &) \text{ or} \\ (& x_{n-1} < y_{n-1} & \text{and} & (x_1, x_2, \dots, x_{n-2}) = (y_1, y_2, \dots, y_{n-2}) &) \text{ or} \\ (& x_n \le y_n & \text{and} & (x_1, x_2, \dots, x_{n-1}) = (y_1, y_2, \dots, y_{n-1}) &) \end{array} \right\}$$

The lexicographical order relation is also called the **dictionary order relation** or **alphabetic order relation**.

Definition 1.23. [23]

DEF

In an ORDERED SET $(X, \le)$,

the SET	$[x : y]$	$\triangleq$	$\{z \in X \mid x \le z \le y\}$	*is a **closed interval** on* $(X, \le)$	*and*
the SET	$(x : y]$	$\triangleq$	$\{z \in X \mid x < z \le y\}$	*is a **half-open interval** on* $(X, \le)$	*and*
the SET	$[x : y)$	$\triangleq$	$\{z \in X \mid x \le z < y\}$	*is a **half-open interval** on* $(X, \le)$	*and*
the SET	$(x : y)$	$\triangleq$	$\{z \in X \mid x < z < y\}$	*is an **open interval**. on* $(X, \le)$	

Definition 1.24. *Let* $(\mathbb{R}, \le)$ *be the* ORDERED SET OF REAL NUMBERS *(Definition 1.20 page 9).*

DEF

*The **absolute value*** $|\cdot| \in \mathbb{R}^{\mathbb{R}}$ *is defined as*[24] $\quad |x| \triangleq \left\{ \begin{array}{ll} -x & \text{for } x \le 0 \\ x & \text{otherwise} \end{array} \right\}.$

Definition 1.25. [25] *Let* $(X, \le)$ *be an* ORDERED SET.

DEF

A subset $D \subseteq X$ *is **convex** in* X *if*

$$x, y \in D \implies (x : y) \subseteq D.$$

Example 1.3. *Convex* subsets of $\mathbb{Z}$ under the usual integer ordering relation include $\emptyset$, $\mathbb{Z}$, $\mathbb{W}$, $\mathbb{N}$, $\{0,1,2,3,4,5\}$, and $\{-2,-1,0,1,2,3\}$.

Definition 1.26. *Let* $(X, \le)$ *be an ordered set.*

DEF

For any set $A \in 2^X$, c *is an **upper bound** of* A *in* $(X, \le)$ *if* $x \le c \quad \forall x \in A$. *An element* b *is the **least upper bound**, or* **lub**, *of* A *in* $(X, \le)$ *if*

$$b \text{ and } c \text{ are UPPER BOUNDS of } A \implies b \le c.$$

The least upper bound of the set A *is denoted* $\bigvee A$.
*The **join*** $x \vee y$ *of* x *and* y *is defined as* $x \vee y \triangleq \bigvee \{x, y\}$.

Definition 1.27. *Let* $(X, \le)$ *be an ordered set.*

DEF

For any set $A \in 2^X$, p *is a **lower bound** of* A *in* $(X, \le)$ *if* $p \le x \quad \forall x \in A$. *An element* a *is the **greatest lower bound**, or* **glb**, *of* A *in* $(X, \le)$ *if*

$$a \text{ and } p \text{ are LOWER BOUNDS of } A \implies p \le a.$$

The greatest lower bound of the set A *is denoted* $\bigwedge A$.
*The **meet*** $x \wedge y$ *of* x *and* y *is defined as* $x \wedge y \triangleq \bigwedge \{x, y\}$.

---

[23] Apostol (1975) page 4, Ore (1935) page 409, Duthie (1942) page 2, Ore (1935) page 425 ⟨*quotient structures*⟩

[24] A more general definition for *absolute value* is available for any *commutative ring*: Let R be a *commutative ring*. A function $|\cdot|$ in R^R is an **absolute value**, or **modulus**, on R if

1.	$\lvert x\rvert$	$\ge$	0	$x \in \mathbb{R}$	(*non-negative*)	and
2.	$\lvert x\rvert$	$=$	$0 \iff x = 0$	$x \in \mathbb{R}$	(*nondegenerate*)	and
3.	$\lvert xy\rvert$	$=$	$\lvert x\rvert \cdot \lvert y\rvert$	$x, y \in \mathbb{R}$	(*homogeneous / submultiplicative*)	and
4.	$\lvert x+y\rvert$	$\le$	$\lvert x\rvert + \lvert y\rvert$	$x, y \in \mathbb{R}$	(*subadditive / triangle inequality*)	

Reference: Cohn (2002) page 312

[25] Barvinok (2002) page 5

Lemma 1.1. *Let* $(X, \leq)$ *be an* ORDERED SET (Definition 1.20 page 9). *Let* $\bigvee A$ *be the* LEAST UPPER BOUND (Definition 1.26 page 10) *of a set* $A \in 2^X$ (Definition 1.8 page 6). *Let* $\bigwedge A$ *be the* GREATEST LOWER BOUND (Definition 1.27 page 10) *of a set* $A \in 2^X$.

LEM
$$\{A = X\} \implies \left\{\begin{array}{lrcll} 1. & \bigvee A & = & \{a \in X | x \leq a,\ \forall x, a \in X\} & \text{and} \\ 2. & \bigwedge A & = & \{a \in X | a \leq x,\ \forall x, a \in X\} & . \end{array}\right\}$$

Definition 1.28. *Let* $(\mathbb{R}, \leq)$ *be the* STANDARD ORDERED SET OF REAL NUMBERS. *The* ***floor function*** $\lfloor x \rfloor \in \mathbb{Z}^{\mathbb{R}}$ *and the* ***ceiling function*** $\lceil x \rceil \in \mathbb{Z}^{\mathbb{R}}$ *are defined as*

$$\underbrace{\lfloor x \rfloor \triangleq \bigvee \{n \in \mathbb{Z} | n \leq x\}}_{\text{FLOOR FUNCTION}} \quad \text{and} \quad \underbrace{\lceil x \rceil \triangleq \bigwedge \{n \in \mathbb{Z} | n \geq x\}}_{\text{CEILING FUNCTION}}.$$

1.4.2 Lattices

The structure available in an *ordered set* (Definition 1.20 page 9) tends to be insufficient to ensure "well-behaved" mathematical systems. This situation is greatly remedied if every pair of elements in the ordered set has both a *least upper bound* and a *greatest lower bound* (Definition 1.27 page 10) in the set; in this case, that ordered set is a *lattice* (next definition).[26]

Definition 1.29. [27]

DEF
An algebraic structure $\mathbf{L} \triangleq (X, \vee, \wedge; \leq)$ *is a* ***lattice*** *if*

$$\begin{array}{llcll} 1. & (X, \leq) \text{ is an ORDERED SET} & & & \text{(Definition 1.20 page 9)} \quad \text{and} \\ 2. & x, y \in X & \implies & x \vee y \in X & \text{(Definition 1.26 page 10)} \quad \text{and} \\ 3. & x, y \in X & \implies & x \wedge y \in X & \text{(Definition 1.27 page 10)}. \end{array}$$

The LATTICE $\mathbf{L}$ *is* ***linear*** *if* $(X, \leq)$ *is a* LINEARLY ORDERED SET (Definition 1.21 page 9).

Theorem 1.1. [28]

THM
$(X, \vee, \wedge; \leq)$ *is a* LATTICE (Definition 1.29 page 11) $\iff$

$$\left\{\begin{array}{rcl|rcl|lll} x \vee x & = & x & x \wedge x & = & x & \forall x \in X & \text{(IDEMPOTENT)} & \text{and} \\ x \vee y & = & y \vee x & x \wedge y & = & y \wedge x & \forall x, y \in X & \text{(COMMUTATIVE)} & \text{and} \\ (x \vee y) \vee z & = & x \vee (y \vee z) & (x \wedge y) \wedge z & = & x \wedge (y \wedge z) & \forall x, y, z \in X & \text{(ASSOCIATIVE)} & \text{and} \\ x \vee (x \wedge y) & = & x & x \wedge (x \vee y) & = & x & \forall x, y \in X & \text{(ABSORPTIVE)}. & \end{array}\right\}$$

Minimax inequality. Suppose we arrange a finite sequence of values into m groups of n elements per group. This could be represented as an $m \times n$ matrix. Suppose now we find the minimum value in each row, and the maximum value in each column. We can call the maximum of all the minimum row values the *maximin*, and the minimum of all the maximum column values the *minimax*. Now, which is greater, the maximin or the minimax? The *minimax inequality* demonstrates that the maximin is always less than or equal to the minimax. The minimax inequality is illustrated below

[26] Gian-Carlo Rota (1932–1999) has illustrated the advantage of lattices over simple ordered sets by pointing out that the *ordered set of partitions of an integer* "is fraught with pathological properties", while the *lattice* of partitions of a set "remains to this day rich in pleasant surprises". Rota (1997) page 1440 ⟨(illustration)⟩, Rota (1964) page 498 ⟨partitions of a set⟩

[27] MacLane and Birkhoff (1999) page 473, Birkhoff (1948) page 16, Ore (1935), Birkhoff (1933) page 442, Maeda and Maeda (1970), page 1

[28] MacLane and Birkhoff (1999) pages 473–475 ⟨LEMMA 1, THEOREM 4⟩, Burris and Sankappanavar (1981) pages 4–7, Birkhoff (1938), pages 795–796, Ore (1935) page 409 ⟨(α)⟩, Birkhoff (1933) page 442, Dedekind (1900) pages 371–372 ⟨(1)–(4)⟩

and stated formerly in Theorem 1.2 (page 12).

$$\underbrace{\bigvee_1^m \left\{ \begin{array}{c} \bigwedge_1^n \left\{ x_{11} \quad x_{12} \quad \cdots \quad x_{1n} \right\} \\ \hline \bigwedge_1^n \left\{ x_{21} \quad x_{22} \quad \cdots \quad x_{2n} \right\} \\ \hline \bigwedge_1^n \left\{ \vdots \quad \ddots \quad \ddots \quad \vdots \right\} \\ \hline \bigwedge_1^n \left\{ x_{m1} \quad x_{m2} \quad \cdots \quad x_{mn} \right\} \end{array} \right\}}_{\text{maximin}} \le \underbrace{\bigwedge_1^n \left\{ \begin{array}{c|c|c|c} \bigvee_1^m & \bigvee_1^m & \bigvee_1^m & \bigvee_1^m \\ x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{array} \right\}}_{\text{minimax}}$$

Theorem 1.2 (minimax inequality). [29] *Let* $(X, \vee, \wedge; \le)$ *be a* LATTICE *(Definition 1.29 page 11).*

THM

$$\underbrace{\bigvee_{i=1}^m \bigwedge_{j=1}^n x_{ij}}_{\textit{maxmini: largest of the smallest}} \le \underbrace{\bigwedge_{j=1}^n \bigvee_{i=1}^m x_{ij}}_{\textit{minimax: smallest of the largest}} \qquad \forall x_{ij} \in X$$

✎PROOF:

$$\begin{aligned}
& \underbrace{\left(\bigwedge_{k=1}^n x_{ik}\right)}_{\text{smallest for any given } i} \le x_{ij} \le \underbrace{\left(\bigvee_{k=1}^n x_{kj}\right)}_{\text{largest for any given } j} && \forall i, j \\
\implies \quad & \underbrace{\bigvee_{i=1}^m \left(\bigwedge_{k=1}^n x_{ik}\right)}_{\text{largest amoung all } i\text{s of the smallest values}} \le \underbrace{\bigwedge_{j=1}^n \left(\bigvee_{k=1}^m x_{kj}\right)}_{\text{smallest amoung all } j\text{s of the largest values}} \\
\implies \quad & \underbrace{\bigvee_{i=1}^m \left(\bigwedge_{j=1}^n x_{ij}\right)}_{\text{maxmini}} \le \underbrace{\bigwedge_{j=1}^n \left(\bigvee_{i=1}^m x_{ij}\right)}_{\text{minimax}} && \text{(change of variables)}
\end{aligned}$$

✏

Special cases of the minimax inequality include three distributive *inequalities* (next theorem). If for some lattice any *one* of these inequalities is an *equality*, then *all three* are *equalities*; and in this case, the lattice is a called a **distributive** lattice.

Theorem 1.3 (distributive inequalities). [30]

THM

$(X, \vee, \wedge; \le)$ *is a* LATTICE *(Definition 1.29 page 11)* $\implies$

$$\left\{ \begin{array}{rcll} x \wedge (y \vee z) & \ge & (x \wedge y) \vee (x \wedge z) & \forall x,y,z \in X \quad \text{(JOIN SUPER-DISTRIBUTIVE)} \quad \textit{and} \\ x \vee (y \wedge z) & \le & (x \vee y) \wedge (x \vee z) & \forall x,y,z \in X \quad \text{(MEET SUB-DISTRIBUTIVE)} \quad \textit{and} \\ (x \wedge y) \vee (x \wedge z) \vee (y \wedge z) & \le & (x \vee y) \wedge (x \vee z) \wedge (y \vee z) & \forall x,y,z \in X \quad \text{(MEDIAN INEQUALITY).} \end{array} \right\}$$

✎PROOF:

[29] Birkhoff (1948) pages 19–20

[30] Davey and Priestley (2002) page 85, Grätzer (2003) page 38, Birkhoff (1933) page 444, Korselt (1894) page 157, Müller-Olm (1997) page 13 ⟨terminology⟩

1. Proof that $\wedge$ sub-distributes over $\vee$:

$$\begin{aligned}(x \wedge y) \vee (x \wedge z) &\leq (x \vee x) \wedge (y \vee z) && \text{by } \textit{minimax inequality} \text{ (Theorem 1.2 page 12)}\\ &= x \wedge (y \vee z) && \text{by } \textit{idempotent} \text{ property of lattices (Theorem 1.1 page 11)}\end{aligned}$$

$$\bigvee\left\{\begin{array}{c}\bigwedge\left\{\begin{array}{cc}x & y\end{array}\right\}\\ \hline \bigwedge\left\{\begin{array}{cc}x & z\end{array}\right\}\end{array}\right\} \quad\leq\quad \bigwedge\left\{\begin{array}{c|c}\bigvee & \bigvee\\ x & y\\ x & z\end{array}\right\}$$

2. Proof that $\vee$ super-distributes over $\wedge$:

$$\begin{aligned}x \vee (y \wedge z) &= (x \wedge x) \vee (y \wedge z) && \text{by } \textit{idempotent} \text{ property of lattices (Theorem 1.1 page 11)}\\ &\leq (x \vee y) \wedge (x \vee z) && \text{by } \textit{minimax inequality} \text{ (Theorem 1.2 page 12)}\end{aligned}$$

$$\bigvee\left\{\begin{array}{c}\bigwedge\left\{\begin{array}{cc}x & x\end{array}\right\}\\ \hline \bigwedge\left\{\begin{array}{cc}y & z\end{array}\right\}\end{array}\right\} \quad\leq\quad \bigwedge\left\{\begin{array}{c|c}\bigvee & \bigvee\\ x & x\\ y & z\end{array}\right\}$$

3. Proof that of median inequality: by *minimax inequality* (Theorem 1.2 page 12)

✏

Besides the distributive property, another consequence of the minimax inequality is the *modularity inequality* (next theorem). A lattice in which this inequality becomes equality is said to be **modular**.

Theorem 1.4 (Modular inequality). [31] *Let* $(X, \vee, \wedge; \leq)$ *be a* LATTICE *(Definition 1.29 page 11)*.

THM $\quad x \leq y \quad\Longrightarrow\quad x \vee (y \wedge z) \leq y \wedge (x \vee z)$

✎PROOF:

$$\begin{aligned}x \vee (y \wedge z) &= (x \wedge x) \vee (y \wedge z) && \text{by } \textit{absorptive} \text{ property (Theorem 1.1 page 11)}\\ &\leq (x \vee y) \wedge (x \vee z) && \text{by the } \textit{minimax inequality} \text{ (Theorem 1.2 page 12)}\\ &= y \wedge (x \vee z) && \text{by left hypothesis}\end{aligned}$$

$$\bigvee\left\{\begin{array}{c}\bigwedge\left\{\begin{array}{cc}x & x\end{array}\right\}\\ \hline \bigwedge\left\{\begin{array}{cc}y & z\end{array}\right\}\end{array}\right\} \quad\leq\quad \bigwedge\left\{\begin{array}{c|c}\bigvee & \bigvee\\ x & x\\ y & z\end{array}\right\}$$

✏

1.4.3 Isomorphic spaces

Definition 1.30. [32] *Let* $\boldsymbol{X} \triangleq (X, \leq)$ *and* $\boldsymbol{Y} \triangleq (Y, \precsim)$ *be ordered sets.*

DEF *A function* $\theta \in Y^X$ *is* ***order preserving*** *in* $(\boldsymbol{X}, \boldsymbol{Y})$ *if* $\quad (x \leq y) \implies (\theta(x) \precsim \theta(y)) \quad \forall x,y \in X.$

[31] Birkhoff (1948) page 19, Burris and Sankappanavar (1981) page 11, Dedekind (1900) page 374
[32] Burris and Sankappanavar (2000), page 10

Definition 1.31. *Let* $\mathbf{L}_1 \triangleq (X, \vee, \wedge; \le)$ *and* $\mathbf{L}_2 \triangleq (Y, \ovee, \owedge; \precsim)$ *be* LATTICES.

DEF

$\mathbf{L}_1$ *and* $\mathbf{L}_2$ *are* ***isomorphic*** *on* $(\boldsymbol{X}, \boldsymbol{Y})$ *if there exists a function* $\theta \in Y^X$ *such that*

1. $\theta(x \vee y) = \theta(x) \ovee \theta(y) \quad \forall x,y \in X$ (PRESERVES JOINS) *and*
2. $\theta(x \wedge y) = \theta(x) \owedge \theta(y) \quad \forall x,y \in X$ (PRESERVES MEETS).

In this case, the function θ *is said to be an* ***isomorphism*** *from* $\mathbf{L}_1$ *to* $\mathbf{L}_2$*, and the isomorphic relationship between* $\mathbf{L}_1$ *and* $\mathbf{L}_2$ *is denoted as* $\mathbf{L}_1 \equiv \mathbf{L}_2$.

Theorem 1.5. [33] *Let* $(X, \vee, \wedge; \le)$ *and* $(Y, \ovee, \owedge; \precsim)$ *be lattices and* $\theta \in Y^X$ *be a* BIJECTIVE *function with inverse* $\theta^{-1} \in X^Y$.

THM

$$\underbrace{\left.\begin{array}{lllll} x_1 \le x_2 & \implies & \theta(x_1) \precsim \theta(x_2) & \forall x_1,x_2 \in X & \text{and} \\ y_1 \precsim y_2 & \implies & \theta^{-1}(y_1) \le \theta^{-1}(y_2) & \forall y_1,y_2 \in Y & \end{array}\right\}}_{\theta \text{ and } \theta^{-1} \text{ are ORDER PRESERVING}} \iff \underbrace{(X, \vee, \wedge; \le) \equiv (Y, \ovee, \owedge; \precsim)}_{\text{ISOMORPHIC}}$$

✎PROOF: Let $\theta \in Y^X$ be the isomorphism between lattices $(X, \vee, \wedge; \le)$ and $(Y, \ovee, \owedge; \precsim)$.

1. Proof that *order preserving* $\implies$ *preserves joins*:
   (a) Proof that $\theta(x_1 \vee x_2) \ogreaterthan \theta(x_1) \ovee \theta(x_2)$:
      i. Note that
      $$x_1 \le x_1 \vee x_2$$
      $$x_2 \le x_1 \vee x_2.$$
      ii. Because θ is *order preserving*
      $$\theta(x_1) \precsim \theta(x_1 \vee x_2)$$
      $$\theta(x_2) \precsim \theta(x_1 \vee x_2).$$
      iii. We can then finish the proof of item (1a):
      $$\begin{aligned} \theta(x_1) \ovee \theta(x_2) &\precsim \underbrace{\theta(x_1 \vee x_2)}_{x_1 \le x_1 \vee x_2} \ovee \underbrace{\theta(x_1 \vee x_2)}_{x_2 \le x_1 \vee x_2} && \text{by } \textit{order preserving} \text{ hypothesis} \\ &= \theta(x_1 \vee x_2) && \text{by } \textit{idempotent} \text{ property page 11} \end{aligned}$$
   (b) Proof that $\theta(x_1 \vee x_2) \precsim \theta(x_1) \ovee \theta(x_2)$:
      i. Just as in item (1a), note that $\theta^{-1}(y_1) \vee \theta^{-1}(y_2) \le \theta^{-1}(y_1 \ovee y_2)$:
      $$\begin{aligned} \theta^{-1}(y_1) \vee \theta^{-1}(y_2) &\le \underbrace{\theta^{-1}(y_1 \ovee y_2)}_{y_1 \precsim y_1 \ovee y_2} \vee \underbrace{\theta^{-1}(y_1 \ovee y_2)}_{y_2 \precsim y_1 \ovee y_2} && \text{by } \textit{order preserving} \text{ hypothesis} \\ &= \theta^{-1}(y_1 \ovee y_2) && \text{by } \textit{idempotent} \text{ property page 11} \end{aligned}$$
      ii. Because θ is *order preserving*
      $$\begin{aligned} \theta\left[\theta^{-1}(y_1) \vee \theta^{-1}(y_2)\right] &\precsim \theta\theta^{-1}(y_1 \ovee y_2) && \text{by item (1(b)i) page 14} \\ &= y_1 \ovee y_2 && \text{by definition of inverse function } \theta^{-1} \end{aligned}$$
      iii. Let $u_1 \triangleq \theta(x_1)$ and $u_2 \triangleq \theta(x_2)$.
      iv. We can then finish the proof of item (1b):
      $$\begin{aligned} \theta(x_1 \vee x_2) &= \theta\left[\theta^{-1}\theta(x_1) \vee \theta^{-1}\theta(x_2)\right] && \text{by definition of inverse function } \theta^{-1} \\ &= \theta\left[\theta^{-1}(u_1) \vee \theta^{-1}(u_2)\right] && \text{by definition of } u_1, u_2 \text{, item (1(b)iii)} \\ &\precsim u_1 \ovee u_2 && \text{by item (1(b)ii)} \\ &= \theta(x_1) \ovee \theta(x_2) && \text{by definition of } u_1, u_2 \text{, item (1(b)iii)} \end{aligned}$$

[33] Burris and Sankappanavar (2000), page 10

(c) And so, combining item (1a) and item (1b), we have

$$\left.\begin{array}{lcll} \theta(x_1 \vee x_2) & \ogreaterthan & \theta(x_1) \ovee \theta(x_2) & \text{(item (1a) page 14)} \quad \text{and} \\ \theta(x_1 \vee x_2) & \lesssim & \theta(x_1) \ovee \theta(x_2) & \text{(item (1b) page 14)} \end{array}\right\} \implies \theta(x_1 \vee x_2) = \theta(x_1) \ovee \theta(x_2)$$

2. Proof that *order preserving* $\implies$ *preserves meets*:

(a) Proof that $\theta(x_1 \wedge x_2) \lesssim \theta(x_1) \owedge \theta(x_2)$:

$$\begin{aligned} \theta(x_1) \owedge \theta(x_2) &\ogreaterthan \underbrace{\theta(x_1 \wedge x_2)}_{x_1 \geq x_1 \wedge x_2} \owedge \underbrace{\theta(x_1 \wedge x_2)}_{x_2 \geq x_1 \wedge x_2} && \text{by } \textit{order preserving} \text{ hypothesis} \\ &= \theta(x_1 \wedge x_2) && \text{by } \textit{idempotent} \text{ property page 11} \end{aligned}$$

(b) Proof that $\theta(x_1 \wedge x_2) \ogreaterthan \theta(x_1) \owedge \theta(x_2)$:

i. Just as in item (2a), note that $\theta^{-1}(y_1) \wedge \theta^{-1}(y_2) \geq \theta^{-1}(y_1 \owedge y_2)$:

$$\begin{aligned} \theta^{-1}(y_1) \wedge \theta^{-1}(y_2) &\geq \underbrace{\theta^{-1}(y_1 \owedge y_2)}_{y_1 \ogreaterthan y_1 \owedge y_2} \owedge \underbrace{\theta^{-1}(y_1 \owedge y_2)}_{y_2 \ogreaterthan y_1 \owedge y_2} && \text{by } \textit{order preserving} \text{ hypothesis} \\ &= \theta^{-1}(y_1 \owedge y_2) && \text{by } \textit{idempotent} \text{ property page 11} \end{aligned}$$

ii. Because θ is *order preserving*

$$\begin{aligned} \theta\left[\theta^{-1}(y_1) \wedge \theta^{-1}(y_2)\right] &\ogreaterthan \theta\theta^{-1}(y_1 \owedge y_2) && \text{by item (2(b)i)} \\ &= y_1 \owedge y_2 \end{aligned}$$

iii. Let $v_1 \triangleq \theta(x_1)$ and $v_2 \triangleq \theta(x_2)$.

iv. We can then finish the proof of item (2a):

$$\begin{aligned} \theta(x_1 \wedge x_2) &= \theta\left[\theta^{-1}\theta(x_1) \wedge \theta^{-1}\theta(x_2)\right] \\ &= \theta\left[\theta^{-1}(v_1) \wedge \theta^{-1}(v_2)\right] && \text{by item (2(b)iii)} \\ &\ogreaterthan v_1 \owedge v_2 && \text{by item (2(b)ii)} \\ &= \theta(x_1) \owedge \theta(x_2) && \text{by item (2(b)iii)} \end{aligned}$$

(c) And so, combining item (2a) and item (2b), we have

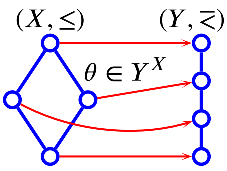
$$\left.\begin{array}{lcll} \theta(x_1 \wedge x_2) & \lesssim & \theta(x_1) \owedge \theta(x_2) & \text{(item (2a) page 15)} \quad \text{and} \\ \theta(x_1 \wedge x_2) & \ogreaterthan & \theta(x_1) \owedge \theta(x_2) & \text{(item (2b) page 15)} \end{array}\right\} \implies \theta(x_1 \wedge x_2) = \theta(x_1) \owedge \theta(x_2)$$

3. Proof that *order preserving* $\impliedby$ *isomorphic*:

$$\begin{aligned} x \leq y &\implies \theta(y) = \theta(x \vee y) = \theta(x) \ovee \theta(y) && \text{by right hypothesis} \\ &\implies \theta(x) \lesssim \theta(y) \\ x \leq y &\implies \theta(x) = \theta(x \wedge y) = \theta(x) \owedge \theta(y) && \text{by right hypothesis} \\ &\implies \theta(x) \lesssim \theta(y) \end{aligned}$$

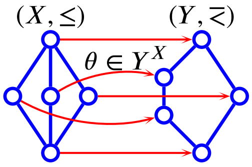
⇰

Example 1.4. [34] In the diagram to the right, the function $\theta \in Y^X$ is *order preserving* with respect to $\leq$ and $\lesssim$. Note that θ^{-1} is *not order preserving* and that the ordered sets $(X, \leq)$ and $(Y, \lesssim)$ are *not isomorphic*—as already demonstrated by Theorem 1.5 that they cannot be.

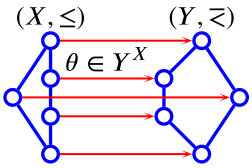


[34] Burris and Sankappanavar (2000), page 10

Example 1.5. In the diagram to the right, the function $\theta \in Y^X$ is *order preserving* with respect to $\le$ and $\lessapprox$. Note that θ^{-1} is *not* order preserving. Like Example 1.4 (page 15), this example also illustrates the fact that that *order preserving* does not imply *isomorphic*.



Example 1.6. In the diagram to the right, the function $\theta \in Y^X$ is *order preserving* with respect to $\le$ and $\lessapprox$. Note that θ^{-1} is also *order preserving* and that the ordered sets $(X, \le)$ and $(Y, \lessapprox)$ are *isomorphic*—as already demonstrated by Theorem 1.5 that they must be.



1.4.4 Monotone functions on ordered sets

Definition 1.32. [35] *Let* $(X, \le)$ *and* $(Y, \sqsubseteq)$ *be* ORDERED SETS (Definition 1.20 page 9). *Let* ϕ *be a function in* Y^X (Definition 1.6 page 6).

DEF

$$\begin{array}{llll}
\phi \text{ is } \textbf{isotone} & \text{in } (Y, \sqsubseteq)^{(X,\le)} & \text{if} & x \le y \implies \phi(x) \sqsubseteq \phi(y) \quad \forall x,y \in X. \\
\phi \text{ is } \textbf{strictly isotone} & \text{in } (Y, \sqsubseteq)^{(X,\le)} & \text{if} & x < y \implies \phi(x) \sqsubset \phi(y) \quad \forall x,y \in X. \\
\psi \text{ is } \textbf{antitone} & \text{in } (Y, \sqsubseteq)^{(X,\le)} & \text{if} & x \le y \implies \psi(y) \sqsubseteq \psi(x) \quad \forall x,y \in X. \\
\psi \text{ is } \textbf{strictly antitone} & \text{in } (Y, \sqsubseteq)^{(X,\le)} & \text{if} & x < y \implies \psi(y) \sqsubset \psi(x) \quad \forall x,y \in X.
\end{array}$$

A FUNCTION *is* ***monotone*** *if it is* ISOTONE *or* ANTITONE *and* ***strictly monotone*** *if it is* STRICTLY ISOTONE *or* STRICTLY ANTITONE. *An* ISOTONE *function in* $(Y, \sqsubseteq)^{(X,\le)}$ *is also said to be* ***order preserving*** *in* $(Y, \sqsubseteq)^{(X,\le)}$.

Lemma 1.2. *Let* $(X, \vee, \wedge; \le)$ *and* $(Y, \sqcup, \sqcap; \sqsubseteq)$ *be* LATTICES (Definition 1.29 page 11). *Let* f *be a* FUNCTION *in* X^X. *Let* ϕ *be a* FUNCTION *in* Y^X (Definition 1.6 page 6).

LEM

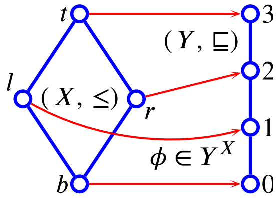
$$\left\{ \begin{array}{l} \phi \text{ is ISOTONE} \\ \text{(Definition 1.32 page 16)} \end{array} \right\} \implies \left\{ \begin{array}{lllll} 1. & \arg \bigvee_{x\in X} \mathsf{f}(x) & \subseteq & \arg \bigsqcup_{x\in X} \phi[\mathsf{f}(x)] & \text{and} \\ 2. & \arg \bigwedge_{x\in X} \mathsf{f}(x) & \subseteq & \arg \bigsqcap_{x\in X} \phi[\mathsf{f}(x)] & . \end{array} \right\}$$

✎PROOF:

$$\begin{array}{rll}
\arg \bigvee_{x\in X} \mathsf{f}(x) &= \arg_x \{\mathsf{f}(x) | \mathsf{f}(y) \le \mathsf{f}(x) \quad \forall x,y \in X\} & \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11} \\
&\subseteq \arg_x \{\mathsf{f}(x) | \phi[\mathsf{f}(y)] \le \phi[\mathsf{f}(x)] \quad \forall x,y \in X\} & \text{by } \textit{isotone} \text{ hypothesis (Definition 1.32 page 16)} \\
&= \arg_x \{\phi[\mathsf{f}(x)] | \phi[\mathsf{f}(y)] \le \phi[\mathsf{f}(x)] \quad \forall x,y \in X\} & \text{because } \arg_x \{\mathsf{f}(x) | P(x)\} = \arg_x \{g[\mathsf{f}(x)] | P(x)\} \\
&\triangleq \arg \bigsqcup_{x\in X} \phi[\mathsf{f}(x)] & \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11} \\
\arg \bigwedge_{x\in X} \mathsf{f}(x) &= \arg_x \{\mathsf{f}(x) | \mathsf{f}(x) \le \mathsf{f}(y) \quad \forall x,y \in X\} & \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11} \\
&\subseteq \arg_x \{\mathsf{f}(x) | \phi[\mathsf{f}(x)] \le \phi[\mathsf{f}(y)] \quad \forall x,y \in X\} & \text{by } \textit{isotone} \text{ hypothesis (Definition 1.32 page 16)} \\
&= \arg_x \{\mathsf{f}(x) | \phi[\mathsf{f}(x)] \le \phi[\mathsf{f}(y)] \quad \forall x,y \in X\} & \text{by } \textit{isotone} \text{ hypothesis (Definition 1.32 page 16)} \\
&= \arg_x \{\phi[\mathsf{f}(x)] | \mathsf{f}(x) \le \mathsf{f}(y) \quad \forall x,y \in X\} & \text{because } \arg_x \{\mathsf{f}(x) | P(x)\} = \arg_x \{g[\mathsf{f}(x)] | P(x)\} \\
&\triangleq \arg \bigsqcap_{x\in X} \phi[\mathsf{f}(x)] & \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11}
\end{array}$$

✏

[35] Rudeanu (2001) page 182 ⟨Definition 2.1⟩, Burris and Sankappanavar (2000), page 10, Topkis (1978) page 308



Remark 1.1. [36] Let $(X, \leq)$ and $(Y, \sqsubseteq)$ be *ordered sets* (Definition 1.20 page 9). Let ϕ be a *function* in Y^X (Definition 1.6 page 6). Note that even if ϕ is *bijective* (Definition 1.14 page 7) and *strictly isotone* (Definition 1.32 page 16),

$$x < y \quad \not\Longleftarrow \quad \phi(x) \sqsubset \phi(y) \quad \forall x,y \in X \; .$$

An example is illustrated to the right where $\phi(l) \sqsubset \phi(r)$, but $l \not< r$.

Lemma 1.3. *Let* $\boldsymbol{X} \triangleq (X, \leq)$ *and* $\boldsymbol{Y} \triangleq (Y, \sqsubseteq)$ *be* ORDERED SETS. *Let* ϕ *be a* FUNCTION *in* Y^X.

$$\text{LEM}\quad \left.\begin{array}{ll} \text{A.} & \phi \text{ is STRICTLY ISOTONE and} \\ \text{B.} & \boldsymbol{X} \text{ and } \boldsymbol{Y} \text{ are LINEARLY ORDERED} \end{array}\right\} \Longrightarrow \left\{\begin{array}{llllll} 1. & x \leq y & \iff & \phi(x) \sqsubseteq \phi(y) & \forall x,y \in X & \text{and} \\ 2. & x < y & \iff & \phi(x) \sqsubset \phi(y) & \forall x,y \in X & . \end{array}\right.$$

✎PROOF:

$$\begin{array}{llll} \boxed{\phi(x) \sqsubseteq \phi(y)} & \Longrightarrow & y \not< x & \text{by contrapositive of } \textit{strictly isotone} \text{ hypothesis (A)} \\ & \Longrightarrow & \boxed{x \leq y} & \text{by } \textit{linear} \text{ hypothesis (B)} \\ & \Longrightarrow & \boxed{\phi(x) \sqsubseteq \phi(y)} & \text{by } \textit{strictly isotone} \text{ hypothesis (A)} \\ \boxed{\phi(x) \sqsubset \phi(y)} & \Longrightarrow & y \not\leq x & \text{by contrapositive of } \textit{strictly isotone} \text{ hypothesis (A)} \\ & \Longrightarrow & \boxed{x < y} & \text{by } \textit{linear} \text{ hypothesis (B)} \\ & \Longrightarrow & \boxed{\phi(x) \sqsubset \phi(y)} & \text{by } \textit{strictly isotone} \text{ hypothesis (A)} \end{array}$$

⇒

Lemma 1.4. *Let* $\boldsymbol{X} \triangleq (X, \vee, \wedge; \leq)$ *and* $\boldsymbol{Y} \triangleq (Y, \sqcup, \sqcap; \sqsubseteq)$ *be* LATTICES *(Definition 1.29 page 11).*[37] *Let* f *be a* FUNCTION *in* X^X. *Let* ϕ *be a* FUNCTION *in* Y^X *(Definition 1.6 page 6).*

$$\text{LEM}\quad \left.\begin{array}{lll} \text{A.} & \phi \text{ is STRICTLY ISOTONE} & \text{and} \\ \text{B.} & \boldsymbol{X} \text{ is LINEARLY ORDERED} & \text{and} \\ \text{C.} & \boldsymbol{Y} \text{ is LINEARLY ORDERED} & \end{array}\right\} \Longrightarrow \left\{\begin{array}{lllll} 1. & \displaystyle\bigsqcup_{x\in X} \phi[\mathsf{f}(x)] & = & \phi\Big[\displaystyle\bigvee_{x\in X} \mathsf{f}(x)\Big] & \text{and} \\ 2. & \displaystyle\bigsqcap_{x\in X} \phi[\mathsf{f}(x)] & = & \phi\Big[\displaystyle\bigwedge_{x\in X} \mathsf{f}(x)\Big]. & \end{array}\right.$$

✎PROOF:

$$\begin{aligned} \phi\Big[\bigvee_{x\in X} \mathsf{f}(x)\Big] &= \phi[\{\mathsf{f}(a) | \mathsf{f}(x) \leq \mathsf{f}(a) \quad \forall x, a \in X\}] && \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11} \\ &= \{\phi[\mathsf{f}(a)] | \mathsf{f}(x) \leq \mathsf{f}(a) \quad \forall x \in X\} \\ &= \{\phi[\mathsf{f}(a)] | \phi[\mathsf{f}(x)] \sqsubseteq \phi[\mathsf{f}(a)] \quad \forall x \in X\} && \text{by Lemma 1.3 (page 17)} \\ &\triangleq \bigsqcup_{x\in X} \phi[\mathsf{f}(x)] && \text{by definition of } (Y, \sqcup, \sqcap; \sqsubseteq) \\ \phi\Big[\bigwedge_{x\in X} \mathsf{f}(x)\Big] &= \phi[\{\mathsf{f}(a) | \mathsf{f}(a) \leq \mathsf{f}(x) \quad \forall x \in X\}] && \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11} \\ &= \{\phi[\mathsf{f}(a)] | \mathsf{f}(a) \leq \mathsf{f}(x) \quad \forall x \in X\} \\ &= \{\phi[\mathsf{f}(a)] | \phi[\mathsf{f}(a)] \sqsubseteq \phi[\mathsf{f}(x)] \quad \forall x \in X\} && \text{by Lemma 1.3 (page 17)} \\ &\triangleq \bigsqcap_{x\in X} \phi[\mathsf{f}(x)] && \text{by definition of } (Y, \sqcup, \sqcap; \sqsubseteq) \end{aligned}$$

⇒

[36] Burris and Sankappanavar (2000), page 10

[37] *strictly isotone* Definition 1.32 page 16
linearly ordered Definition 1.21 page 9
linearly ordered Definition 1.21 page 9

Lemma 1.5. *Let* $\boldsymbol{X} \triangleq (X, \vee, \wedge; \leq)$ *and* $\boldsymbol{Y} \triangleq (Y, \sqcup, \sqcap; \sqsubseteq)$ *be* LATTICES.[38]

LEM

$$\left.\begin{array}{lll} \text{A.} & \psi \text{ is STRICTLY ANTITONE} & \text{and} \\ \text{B.} & \boldsymbol{X} \text{ is LINEARLY ORDERED} & \text{and} \\ \text{C.} & \boldsymbol{Y} \text{ is LINEARLY ORDERED} & \end{array}\right\} \implies \left\{\begin{array}{lllll} \text{1.} & \bigsqcup_{x\in X} \psi[\mathsf{f}(x)] & = & \psi\left[\bigwedge_{x\in X} \mathsf{f}(x)\right] & \text{and} \\ \text{2.} & \bigsqcap_{x\in X} \psi[\mathsf{f}(x)] & = & \psi\left[\bigvee_{x\in X} \mathsf{f}(x)\right]. & \end{array}\right.$$

✎PROOF:

$$\begin{aligned}
\psi\left[\bigvee_{x\in X} \mathsf{f}(x)\right] &= \psi[\{\mathsf{f}(a)|\mathsf{f}(x) \leq \mathsf{f}(a) \quad \forall x \in X\}] && \text{by definition of } (X, \vee, \wedge; \leq) \\
&= \{\psi[\mathsf{f}(a)]|\mathsf{f}(x) \leq \mathsf{f}(a) \quad \forall x \in X\} \\
&= \{\psi[\mathsf{f}(a)]|\psi[\mathsf{f}(a)] \sqsubseteq \psi[\mathsf{f}(x)] \quad \forall x \in X\} && \text{by definition of } \textit{strictly antitone} \text{ (Definition 1.32 page 16)} \\
&\triangleq \bigsqcap_{x\in X} \psi[\mathsf{f}(x)] && \text{by definition of } (Y, \sqcup, \sqcap; \sqsubseteq) \\
\psi\left[\bigwedge_{x\in X} \mathsf{f}(x)\right] &= \psi[\{\mathsf{f}(a)|\mathsf{f}(a) \leq \mathsf{f}(x) \quad \forall x \in X\}] && \text{by definition of } (X, \vee, \wedge; \leq) \\
&= \{\psi[\mathsf{f}(a)]|\mathsf{f}(a) \leq \mathsf{f}(x) \quad \forall x \in X\} \\
&= \{\psi[\mathsf{f}(a)]|\psi[\mathsf{f}(x)] \sqsubseteq \psi[\mathsf{f}(a)] \quad \forall x \in X\} && \text{by definition of } \textit{strictly antitone} \text{ (Definition 1.32 page 16)} \\
&\triangleq \bigsqcup_{x\in X} \psi[\mathsf{f}(x)] && \text{by definition of } (Y, \sqcup, \sqcap; \sqsubseteq)
\end{aligned}$$

✏

Lemma 1.6. *Let* $\boldsymbol{X} \triangleq (X, \vee, \wedge; \leq)$ *and* $\boldsymbol{Y} \triangleq (Y, \sqcup, \sqcap; \sqsubseteq)$ *be* LATTICES *(Definition 1.29 page 11). Let* f *be a function in* X^X*. Let* ϕ *and* ψ *be functions in* Y^X.[39]

LEM

$$\left\{\begin{array}{lll} \text{A.} & \phi \text{ is STRICTLY ISOTONE} & \text{and} \\ \text{B.} & \boldsymbol{X} \text{ is LINEARLY ORDERED} & \end{array}\right\} \implies \left\{\begin{array}{lllll} \text{1.} & \arg \bigvee_{x\in X} \mathsf{f}(x) & = & \arg \bigsqcup_{x\in X} \phi[\mathsf{f}(x)] & \text{and} \\ \text{2.} & \arg \bigwedge_{x\in X} \mathsf{f}(x) & = & \arg \bigsqcap_{x\in X} \phi[\mathsf{f}(x)] & . \end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
\arg \bigvee_{x\in X} \mathsf{f}(x) &\triangleq \arg \bigvee \{\mathsf{f}(x)|x \in X\} \\
&= \arg \{\mathsf{f}(a)|\mathsf{f}(x) \leq \mathsf{f}(a) \quad \forall x, a \in X\} && \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11} \\
&= \arg \{\mathsf{f}(a)|\phi[\mathsf{f}(x)] \leq \phi[\mathsf{f}(a)] \quad \forall x, a \in X\} && \text{by hypothesis (A) and Lemma 1.5 page 18} \\
&= \arg \{\phi[\mathsf{f}(a)]|\mathsf{f}(x) \leq \mathsf{f}(a) \quad \forall x, a \in X\} && \text{because } \arg_x \{\mathsf{f}(x)|P(x)\} = \arg_x \{\mathsf{g}[\mathsf{f}(x)]|P(x)\} \\
&\triangleq \bigsqcup_{x\in X} \phi[\mathsf{f}(x)] && \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11} \\
\arg \bigwedge_{x\in X} \mathsf{f}(x) &\triangleq \arg \bigwedge \{\mathsf{f}(x)|x \in X\} \\
&= \arg \{\mathsf{f}(a)|\mathsf{f}(a) \leq \mathsf{f}(x) \quad \forall x, a \in X\} && \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11} \\
&= \arg \{\mathsf{f}(a)|\phi[\mathsf{f}(a)] \leq \phi[\mathsf{f}(x)] \quad \forall x, a \in X\} && \text{by hypothesis (A) and Lemma 1.5 page 18} \\
&= \arg \{\phi[\mathsf{f}(a)]|\mathsf{f}(a) \leq \mathsf{f}(x) \quad \forall x, a \in X\} && \text{because } \arg_x \{\mathsf{f}(x)|P(x)\} = \arg_x \{\mathsf{g}[\mathsf{f}(x)]|P(x)\}
\end{aligned}$$

---

[38] Let f be a *function* in X^X.
Let ψ be a *function* in Y^X. *strictly antitone* Definition 1.32 page 16
linearly ordered Definition 1.21 page 9
linearly ordered Definition 1.21 page 9

[39] *strictly isotone* Definition 1.32 page 16
linearly ordered Definition 1.21 page 9

$$\triangleq \bigsqcap_{x \in X} \phi[\mathsf{f}(x)] \qquad \text{because } \mathsf{f} \in X^X \text{ and by Lemma 1.1 page 11}$$

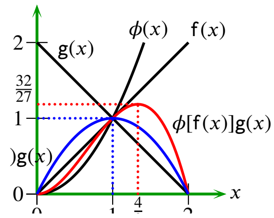
✏

Remark 1.2. Using the definitions of Lemma 1.6 (page 18), and letting g be a function in X^X, and despite the results of Lemma 1.6, note that

REM
$$\left.\begin{array}{ll} \text{A.} & \phi \text{ is } \textit{strictly isotone} \quad \text{and} \\ \text{B.} & \boldsymbol{X} \text{ is } \textit{linearly ordered} \end{array}\right\} \not\Longrightarrow \left\{ \arg \bigwedge_{x \in X} [\mathsf{f}(x)\mathsf{g}(x)] = \arg \bigsqcap_{x \in X} \phi[\mathsf{f}(x)]\mathsf{g}(x) \right.$$

For example, let $\mathsf{f}(x) \triangleq x$, $\mathsf{g}(x) \triangleq -x+2$, and $\phi(x) \triangleq x^2$. Then

$$\begin{aligned} \arg \bigwedge_{x \in X} [\mathsf{f}(x)\mathsf{g}(x)] &\triangleq \arg \bigwedge_{x \in X} [-x^2 + 2x] \\ &= 1 \neq \frac{4}{3} \\ &= \arg \bigsqcap_{x \in X} [-x^3 + 2x^2] \\ &\triangleq \arg \bigsqcap_{x \in X} \phi[\mathsf{f}(x)]\mathsf{g}(x) \end{aligned}$$



Lemma 1.7. *Let* $\boldsymbol{X} \triangleq (X, \vee, \wedge; \leq)$ *and* $\boldsymbol{Y} \triangleq (Y, \sqcup, \sqcap; \sqsubseteq)$ *be* LATTICES (Definition 1.29 page 11). *Let* f *be a function in* $(X \times X)^X$. *Let* ϕ *and* ψ *be functions in* Y^X.

LEM
$$\left.\begin{array}{ll} \text{A.} & \phi \text{ is STRICTLY ISOTONE} \quad \textit{and} \\ \text{B.} & \boldsymbol{X} \text{ is LINEARLY ORDERED} \end{array}\right\}$$
$$\Longrightarrow \left\{\begin{array}{llll} 1. & \arg \bigwedge_{x \in X} \bigvee_{y \in X} \mathsf{f}(x,y) = \arg \bigsqcap_{x \in X} \bigsqcup_{y \in X} \phi[\mathsf{f}(x,y)] & \forall x,y \in X & \textit{and} \\ 2. & \arg \bigwedge_{x \in X} \bigvee_{y \in X} \mathsf{f}(x,y) = \arg \bigsqcap_{x \in X} \bigsqcap_{y \in X} \phi[\mathsf{f}(x,y)] & \forall x,y \in X & \textit{and} \\ 3. & \arg \bigvee_{x \in X} \bigwedge_{y \in X} \mathsf{f}(x,y) = \arg \bigsqcup_{x \in X} \bigsqcap_{y \in X} \phi[\mathsf{f}(x,y)] & \forall x,y \in X & \end{array}\right.$$

✎PROOF:

$$\begin{aligned} \arg \bigsqcap_{x \in X} \bigsqcup_{y \in X} \phi[\mathsf{f}(x,y)] &= \arg \bigsqcap_{x \in X} \phi\left[\bigvee_{y \in X} \mathsf{f}(x,y)\right] && \text{by Lemma 1.4 (page 17)} \\ &= \arg \bigwedge_{x \in X} \bigvee_{y \in X} \mathsf{f}(x,y) && \text{by Lemma 1.2 (page 16)} \end{aligned}$$

$$\begin{aligned} \arg \bigsqcap_{x \in X} \bigsqcap_{y \in X} \phi[\mathsf{f}(x,y)] &= \arg \bigsqcap_{x \in X} \phi\left[\bigwedge_{y \in X} \mathsf{f}(x,y)\right] && \text{by Lemma 1.4 (page 17)} \\ &= \arg \bigwedge_{x \in X} \bigwedge_{y \in X} \mathsf{f}(x,y) && \text{by Lemma 1.2 (page 16)} \end{aligned}$$

$$\begin{aligned} \arg \bigsqcup_{x \in X} \bigsqcap_{y \in X} \phi[\mathsf{f}(x,y)] &= \arg \bigsqcap_{x \in X} \phi\left[\bigwedge_{y \in X} \mathsf{f}(x,y)\right] && \text{by Lemma 1.4 (page 17)} \\ &= \arg \bigvee_{x \in X} \bigwedge_{y \in X} \mathsf{f}(x,y) && \text{by Lemma 1.2 (page 16)} \end{aligned}$$

✏

Remark 1.3. Let $(X, \leq)$ be an *ordered set* (Definition 1.20 page 9). Let ϕ be a function in X^X.

REM

If ϕ is *strictly isotone* then

$$\sum_{x\in X} \mathsf{f}(x) < \sum_{x\in X} \mathsf{g}(x) \implies \phi\left[\sum_{x\in X} \mathsf{f}(x)\right] < \phi\left[\sum_{x\in X} \mathsf{g}(x)\right] \quad \textit{but}$$

$$\sum_{x\in X} \phi[\mathsf{f}(x)] < \sum_{x\in X} \phi[\mathsf{g}(x)] \;\not\Longrightarrow\; \phi\left[\sum_{x\in X} \mathsf{f}(x)\right] < \phi\left[\sum_{x\in X} \mathsf{g}(x)\right].$$

✎PROOF:

1. Proof for (1): this follows directly from the definition of *isotone* (Definition 1.32 page 16).

2. Proof for (2): Let $\mathsf{f}(x) = \left(\!\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)\!\right)$, $\mathsf{g}(x) = (\!(0,\ 0,\ 1)\!)$, and $\phi(x) = x^2$.

$$\text{Then} \quad \sum_{x\in X} \phi[\mathsf{f}(x)] = \left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{2}\right)^2 = \tfrac{3}{4} < 1 = 0^2 + 0^2 + 1^2 = \sum_{x\in X} \phi[\mathsf{g}(x)]$$

$$\text{but} \quad \phi\left[\sum_{x\in X} \mathsf{f}(x)\right] = \left(\tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2}\right)^2 = \tfrac{9}{4} > 1 = (0+0+1)^2 = \phi\left[\sum_{x\in X} \mathsf{g}(x)\right]$$

✏

1.5 Metric space concepts

1.5.1 Motivation

Why might we care about *metric*s, or more generally *distance functions*, in symbolic sequence processing? *Metric balls* in a *metric space* induce a *topology*. Topologies are necessary for the concept of *convergence*. Some topologies are also *algebra of sets*;[40] an algebra of sets (or in particular a *sigma-algebra*) is used for the concept of measure. Loosely speaking then, we care about *distance* and *distance spaces* for two reasons:

1. In analysis, *metric spaces* allow us to define the concepts of convergence and limit of a sequence as in $\sum_{n=0}^{\infty} x_n \triangleq \lim_{N\to\infty} \sum_{n=0}^{n=N} x_n$. That is, without the implicit or explicit definition of convergence and limit, the expression $\sum_{n=0}^{\infty} x_n$ is meaningless.[41]
2. In signal processing, "optimal" decisions may be made with respect to a *distance space*. For example, a point may be selected (identified as "optimal") based on it being measured as having the smallest distance to some reference point.

1.5.2 Isometric spaces

Definition 1.33. [42] *Let* (X, d) *and* (Y, p) *be* DISTANCE SPACES *(Definition B.1 page 133).*

DEF

The function $\mathsf{f} \in Y^X$ *is an* ***isometry*** *on* $(Y, \mathsf{p})^{(X,\mathsf{d})}$ *if* $\mathsf{d}(x, y) = \mathsf{p}(\mathsf{f}(x), \mathsf{f}(y)) \quad \forall x,y\in X.$
The spaces (X, d) *and* (Y, p) *are* ***isometric*** *if there exists an isometry on* $(Y, \mathsf{p})^{(X,\mathsf{d})}$.

[40] For example on the three element set $\{x, y, z\}$, there are a total of 29 topologies; on of these 29, five are algebras of sets. References: Isham (1999), page 44, Isham (1989), page 1516, Steiner (1966), page 386, Sloane (2014) ⟨http://oeis.org/A000798⟩, Brown and Watson (1996), page 31, Comtet (1974) page 229, Comtet (1966), Chatterji (1967), page 7, Evans et al. (1967), Krishnamurthy (1966), page 157

[41] Klauder (2010) page 4

[42] Thron (1966), page 153 ⟨definition 19.4⟩, Giles (1987) page 124 ⟨Definition 6.22⟩, Khamsi and Kirk (2001) page 15 ⟨Definition 2.4⟩, Kubrusly (2001) page 110

Theorem 1.6. [43] *Let* (X, d) *and* (Y, p) *be* DISTANCE SPACES. *Let* f *be a function in* Y^X *and* f^{-1} *its inverse in* X^Y.

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

If a function p is a *metric* and a function g is *injective*, then the function $\mathsf{d}(x, y) \triangleq \mathsf{p}(\mathsf{g}(x), \mathsf{g}(y))$ is also a *metric* (next theorem). For an example of this with $\mathsf{p}(x, y) \triangleq |x - y|$ and $\mathsf{g} \triangleq \arctan(x)$, see Example 1.7 (page 21).

Theorem 1.7 (Pullback metric/g-transform metric). [44] *Let* X *and* Y *be sets.*

THM
$$\left\{\begin{array}{llll} (1). & \mathsf{p} \text{ is a METRIC on } Y & \text{(Definition D.7 page 163)} & \text{and} \\ (2). & \mathsf{g} \text{ is a FUNCTION in } Y^X & \text{(Definition 1.6 page 6)} & \text{and} \\ (3). & \mathsf{g} \text{ is INJECTIVE} & \text{(Definition 1.14 page 7)} & \end{array}\right\} \implies \left\{\begin{array}{l} \mathsf{d}(x, y) = \mathsf{p}(\mathsf{g}(x), \mathsf{g}(y)) \quad \forall x,y \in X \\ \text{is a METRIC on } X \end{array}\right\}$$

✎PROOF:

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

$$\begin{aligned} \mathsf{d}(x, y) &\triangleq \mathsf{p}(\phi(x), \phi(y)) && \text{by definition of } \mathsf{d} \\ &= \mathsf{p}(\phi(x), \phi(x)) && \text{by } x = y \text{ hypothesis} \\ &= 0 && \text{by \textit{nondegenerate} property of metric } \mathsf{p} \text{ (Definition D.7 page 163)} \\ &= 0 \end{aligned}$$

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

$$\begin{aligned} 0 &= \mathsf{d}(x, y) && \text{by right hypothesis} \\ &\triangleq \mathsf{p}(\phi(x), \phi(y)) && \text{by definition of } \mathsf{d} \\ \implies \mathsf{p}(\phi(x), \phi(y)) &= 0 \text{ for } n = 1, 2, \ldots, N && \text{because } \mathsf{p} \text{ is non-negative} \\ \implies x &= y && \text{by \textit{injective} hypothesis (3)} \end{aligned}$$

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

$$\begin{aligned} \mathsf{d}(x, y) &\triangleq \mathsf{p}(\phi(x), \phi(y)) && \text{by definition of } \mathsf{d} \\ &\le \big(\mathsf{p}(\phi(x), \phi(z)) + \mathsf{d}(\phi(z), \phi(y))\big) && \text{by \textit{subadditive} property of } \mathsf{p} \text{ (Definition D.7 page 163)} \\ &= \mathsf{p}(\phi(z), \phi(x)) + \mathsf{p}(\phi(z), \phi(y)) && \text{by \textit{symmetry} property of metric } \mathsf{p} \text{ (Definition D.7 page 163)} \\ &\triangleq \mathsf{d}(z, x) + \mathsf{d}(z, y) && \text{by definition of } \mathsf{d} \end{aligned}$$

✏

Example 1.7 (Inverse tangent metric). [45] Let $(x_1, x_2, \ldots, x_N)$ and $(y_1, y_2, \ldots, y_N)$ be points in $\mathbb{R}^N$.

EX
$$\left\{\mathsf{d}\big((x_1, x_2, \ldots, x_N), (y_1, y_2, \ldots, y_N)\big) \triangleq \sum_{n=1}^{N} \left|\arctan x_n - \arctan y_n\right|\right\} \quad \text{is a \textit{metric}.}$$

✎PROOF:

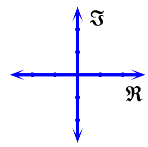
1. The function $\mathsf{d}(x, y) \triangleq |x - y|$ is a *metric* (the *usual metric*, Definition D.9 page 166).
2. The function $\mathsf{g}(x) \triangleq \arctan(x)$ is *injective* in $\mathbb{R}^{\mathbb{R}}$.
3. Therefore, d is a *Pullback metric* (or g-*transform metric*), and by Theorem 1.7 (page 21), d is a *metric*.

✏

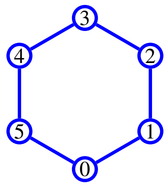
[43] Thron (1966), page 153 ⟨theorem 19.5⟩
[44] Deza and Deza (2009) page 81
[45] Copson (1968), page 25, Khamsi and Kirk (2001) page 14

(A) *integer line* (Example 1.8)
linear ordered metric space

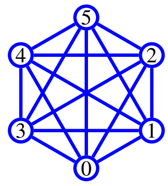
(B) *real line* (Example 1.9)
linear ordered metric space



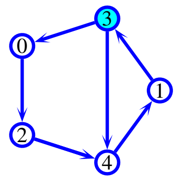
(C) *complex plane* (Example 1.10)
ordered/unordered metric space



(D) *6 element ring* (Example 1.11)
unordered metric space



(E) *6 element discrete metric* (Example 1.12)
unordered metric space



(F) *directed PRNG state machine* (Example 1.13)
unordered quasi-metric space

Figure 1.5: Examples of *ordered quasi-metric spaces* (Definition 1.34 page 22)

1.6 Ordered metric spaces

1.6.1 Definitions

Definition 1.34.

DEF

A triple $\mathsf{G} \triangleq (X, \leq, \mathsf{d})$ *is an* ***ordered quasi-metric space*** *if* (X, d) *is a* QUASI-METRIC SPACE *(Definition D.6 page 163) and* $(X, \leq)$ *is an* ORDERED SET *(Definition 1.20 page 9).*

G *is an* ***ordered metric space***	*if* d *is a* METRIC *(Definition D.7 page 163).*
G *is an* ***unordered quasi-metric space***	*if* $\leq = \emptyset$.
G *is an* ***unordered metric space***	*if* d *is a* METRIC *and* $\leq = \emptyset$.

Remark 1.4. Note that the four structures defined in Definition 1.34 are not mutually exclusive. For example, by Definition 1.34,

$$\begin{array}{lclcl}\{\textit{unordered metric space}\} & \subseteq & \{\textit{unordered quasi-metric space}\} & \subseteq & \{\textit{ordered quasi-metric space}\}\\ \{\textit{unordered metric space}\} & \subseteq & \{\textit{ordered metric space}\} & \subseteq & \{\textit{ordered quasi-metric space}\}.\end{array}$$

Remark 1.5. The use of the *quasi-metric* rather than exclusive use of the more restrictive *metric* in Definition 1.34 is motivated by state machines, where metrics measuring distances between states are in some cases by nature *non-symmetric*. One such example is the *linear congruential pseudo-random number generator* (Example 2.19 page 64).

Remark 1.6. This text makes extensive reference to the *real line* (next definition). There are several ways to define the real line. In particular, there are many possible ordering relations on $\mathbb{R}$ and several possible topologies on $\mathbb{R}$.[46] In fact, order and topology are closely related in that an order relation $\leq$ (Definition 1.20 page 9) on a set always induces a topology (called the *order topology* / *interval topology*[47]); and in the case of the real line, a topology induces an order structure up to the order relation's *dual* (Definition 1.22 page 9).[48] This text uses a fairly standard structure, as defined next.

Definition 1.35. *The triple* $(\mathbb{R}, |\cdot|, \leq)$ *is the* ***real line*** *if* $\mathbb{R}$ *is the* SET OF REAL NUMBERS *(Definition 1.2 page 5),* $\mathsf{d}(x,y) \triangleq |x-y|$ *is the* USUAL METRIC *on* $\mathbb{R}$ *(Definition D.9 page 166), and* $\leq$ *is the standard* LINEAR ORDER RELATION *(Definition 1.21 page 9) on* $\mathbb{R}$.

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

[47] Salzmann et al. (2007) page 23 ⟨3.2 Topology induced by an ordering⟩ Willard (1970) page 43 ⟨6D. Ordered Spaces⟩, Steen and Seebach (1978) page 66 ⟨39. Order Topology⟩

[48] Hocking and Young (1961) page 52 ⟨2–5 The interval and the circle⟩, Salzmann et al. (2007) pages 69–70 ⟨5.75 Note: Ordering and topology on $\mathbb{R}$, see also 5.10 Theorem page 36⟩

Definition 1.36. *The triple* $(\mathbb{Z}, \le, |\cdot|)$ *is the* ***integer line*** *if* $\mathbb{Z}$ *is the* SET OF INTEGERS *(Definition 1.2 page 5)*, $\mathsf{d}(m, n) \triangleq |m - n|$ *is the* USUAL METRIC *(Definition D.9 page 166) on* $\mathbb{R}$ *restricted to* $\mathbb{Z}$*, and* $\le$ *is the standard* LINEAR ORDER RELATION *on* $\mathbb{Z}$ *as induced by* PEANO'S AXIOMS.[49]

1.6.2 Examples

Example 1.8. The *integer line* (Definition 1.36 page 23) is an *ordered metric space* (Definition 1.34 page 22), and is illustrated in Figure 1.5 page 22 (A).

Example 1.9. The *real line* (Definition 1.35 page 22) is an *ordered metric space* (Definition 1.34 page 22), and is illustrated in Figure 1.5 page 22 (B).

Example 1.10. The **complex plane** $(\mathbb{C}, |\cdot|, \le)$ is an *ordered metric space* (Definition 1.34 page 22) where $\mathbb{C} \triangleq \mathbb{R}^2$ is the *set of complex numbers*, $\mathsf{d}(x, y) \triangleq |x - y| \triangleq \sqrt{(\Re x - \Re y)^2 + (\Im x - \Im y)^2}$, $\Re x \triangleq \Re\,(a, b) \triangleq a$ $\forall_{(a,b)\in\mathbb{C}}$ ($\Re x$ is the **real part** of x), $\Im x \triangleq \Im\,(a, b) \triangleq b$ $\forall_{(a,b)\in\mathbb{C}}$ ($\Im x$ is the **imaginary part** of x), and $\le$ is any *order relation* defined on $\mathbb{C}$. Possible order relations include the *coordinatewise order relation* (Example 1.1 page 9), the *lexicographical order relation* (Example 1.2 page 9), and $\le = \emptyset$ (in which case the *complex plane* is *unordered*). The *complex plane* is illustrated in Figure 1.5 page 22 (C).

Example 1.11. A **6 element ring** $(\{0, 1, 2, 3, 4, 5\}, \mathsf{d}, \emptyset)$ is an *unordered metric space* (Definition 1.34 page 22) where the metric d is defined on a ring as illustrated in Figure 1.5 page 22 (D), with each line segment representing a distance of 1.

Example 1.12. A **6 element discrete metric** $(\{0, 1, 2, 3, 4, 5\}, \mathsf{d}, \emptyset)$ is an *unordered metric space* (Definition 1.34 page 22) where the metric d is the *discrete metric* (Definition D.8 page 166). This structure is illustrated in Figure 1.5 page 22 (E).

Example 1.13. Figure 1.5 page 22 (F) illustrates a *linear congruential pseudo-random number generator* induced by the equation $y_{n+1} = (y_n + 2) \mod 5$ with $y_0 = 1$. The structure is an *unordered quasi-metric space*. See Example 2.17 (page 61)–Example 2.19 (page 64) for further demonstration.

1.7 Traditional probability

❝ *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[50]

Definition 1.37. [51] *Let* X *be a* SET.

DEF

A function $\mathsf{P} \in \mathbb{R}^{\vdash X}$ *is a* ***probability function*** *if*

(1).	$\mathsf{P}(1) = 1$		(NORMALIZED)	and
(2).	$\mathsf{P}(x) \ge 0$	$\forall x \in X$	(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)	.

[49] Landau (1966) page 2, Halmos (1960) page 46, Thurston (1956) page 51, Peano (1889a), Peano (1889b) page 94, Dedekind (1888a), Dedekind (1888b) page 67, Cori and Lascar (2001) pages 8–15 ⟨recursion theory⟩

[50] quote: Snell (1997) page 307, Snell (2005) page 251. image: http://www.dartmouth.edu/~chance/Doob/conversation.html

[51] Papoulis (1991) pages 21–22, Kolmogorov (1933), page 2 ⟨§1. Axioms I–V⟩

Definition 1.38. [52]

DEF: *The triple* $(\Omega, \mathbb{E}, \mathsf{P})$ *is a* ***probability space*** *if* Ω *is a* SET, $\mathbb{E}$ *is a* σ-ALGEBRA *on* Ω, *and* P *is a* PROBABILITY FUNCTION *in* $[0:1]^{\mathbb{E}}$. *In this case,* Ω *is called the* ***set of outcomes***.

Before defining a random variable formally, note two things that a random variable is *not*:[53]

- 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 variable itself) is very deterministic and well-defined.

Definition 1.39. [54]

DEF: *A* ***real-valued random variable***, *or* ***traditional random variable***, X *on a* PROBABILITY SPACE $(\Omega, \mathbb{E}, \mathsf{P})$ *is any* MEASURABLE FUNCTION *that maps from* $(\Omega, \mathbb{E}, \mathsf{P})$ *to* $(\mathbb{R}, \le, \mathsf{d}, +, \cdot, \mathscr{B})$ *where* $\mathscr{B}$ *is the* USUAL BOREL σ-ALGEBRA *on the "real line"* $(\mathbb{R}, \le, \mathsf{d}, +, \cdot)$ (Definition 1.35 page 22).

Definition 1.40. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE (Definition 1.38 page 24) *and* $\mathsf{X} \in \mathbb{R}^{\Omega}$ *a* RANDOM VARIABLE (Definition 1.39 page 24).

DEF:

The ***traditional expected value*** $\mathsf{E}(\mathsf{X})$ *of* X *is* $\quad \mathsf{E}(\mathsf{X}) \triangleq \int_{\mathbb{R}} x\mathsf{p}(x)\,\mathsf{d}x.$

The ***traditional variance*** $\mathsf{Var}(\mathsf{X})$ *of* X *is* $\quad \mathsf{Var}(\mathsf{X}) \triangleq \int_{\mathbb{R}} [x - \mathsf{E}(\mathsf{X})]^2\mathsf{p}(x)\,\mathsf{d}x.$

Proposition 1.2. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$ *be a* PROBABILITY SPACE (Definition 1.38 page 24), $\mathsf{X} \in \mathbb{R}^{\Omega}$ *a* TRADITIONAL RANDOM VARIABLE (Definition 1.39 page 24), $\mathsf{E}(\mathsf{X})$ *the* TRADITIONAL EXPECTED VALUE *of* X, *and* $\mathsf{Var}(\mathsf{X})$ *the* TRADITIONAL VARIANCE *of* X (Definition 1.40 page 24).

PRP:

$$\left\{ \mathsf{P}(x) = 0 \quad \forall x \notin \mathbb{Z} \right\} \implies \left\{ \begin{array}{lrcll} 1. & \mathsf{E}(\mathsf{X}) &=& \displaystyle\sum_{x\in\mathbb{Z}} x\mathsf{P}(x) & \textit{and} \\ 2. & \mathsf{Var}(\mathsf{X}) &=& \displaystyle\sum_{x\in\mathbb{Z}} [x - \mathsf{E}(\mathsf{X})]^2\mathsf{P}(x) & \end{array} \right\}$$

Proposition 1.3. *Let* $(\Omega, \mathbb{E}, \mathsf{P})$, X, *and* E *be defined as in Proposition 1.2 (page 24).*

PRP:

$$\underbrace{\mathsf{P}(\gamma - x) = \mathsf{P}(\gamma + x) \qquad \forall x\in\mathbb{R}}_{(\mathsf{P}\text{ is SYMMETRIC about a point }\gamma)} \implies \{\mathsf{E}(\mathsf{X}) = \gamma\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{E}(\mathsf{X}) &\triangleq \int_{-\infty}^{\infty} x\mathsf{p}(x)\,\mathsf{d}x && \text{by definition of } \mathsf{E} \text{ (Definition 1.40 page 24)} \\
&= \int_{x=-\infty}^{x=\gamma} x\mathsf{p}(x)\,\mathsf{d}x + \int_{x=\gamma}^{x=\infty} x\mathsf{p}(x)\,\mathsf{d}x && \text{by \textit{additive property} of Lebesgue integration op.}^{55} \\
&= \int_{u+\gamma=-\infty}^{u+\gamma=\gamma} (u+\gamma)\mathsf{p}(u+\gamma)\,\mathsf{d}u + \int_{u+\gamma=\gamma}^{u+\gamma=\infty} (u+\gamma)\mathsf{p}(u+\gamma)\,\mathsf{d}u && \text{where } u \triangleq x - \gamma \implies x = u + \gamma \\
&= -\int_{0}^{-\infty} (u+\gamma)\mathsf{p}(u+\gamma)\,\mathsf{d}u + \int_{0}^{\infty} (u+\gamma)\mathsf{p}(u+\gamma)\,\mathsf{d}u \\
&= \int_{0}^{\infty} (-v+\gamma)\mathsf{p}(-v+\gamma)\,\mathsf{d}v + \int_{0}^{\infty} (u+\gamma)\mathsf{p}(u+\gamma)\,\mathsf{d}u && \text{where } v \triangleq -u
\end{aligned}$$

[52] Greenhoe (2015)
[53] Miller (2006) page 130, Feldman and Valdez-Flores (2010) page 4, Curry and Feldman (2010) page 4
[54] Bryc (2012) page 73, Papoulis (1991), page 63

$$
\begin{aligned}
&= -\int_0^\infty v\mathsf{p}(-v+\gamma)\,\mathrm{d}v + \int_0^\infty u\mathsf{p}(u+\gamma)\,\mathrm{d}u + \gamma\int_0^\infty \mathsf{p}(-v+\gamma)\,\mathrm{d}v + \gamma\int_0^\infty \mathsf{p}(u+\gamma)\,\mathrm{d}u \\
&= \underbrace{-\int_0^\infty v\mathsf{p}(v+\gamma)\,\mathrm{d}v + \int_0^\infty u\mathsf{p}(u+\gamma)\,\mathrm{d}u}_{\text{by symmetry hypothesis; cancels to 0}} + \gamma\left(\int_0^\infty \mathsf{p}(-v+\gamma)\,\mathrm{d}v + \int_0^\infty \mathsf{p}(u+\gamma)\,\mathrm{d}u\right) \\
&= \gamma\left(-\int_0^{-\infty} \mathsf{p}(w+\gamma)\,\mathrm{d}\omega + \int_0^\infty \mathsf{p}(u+\gamma)\,\mathrm{d}u\right) \qquad \text{where } w \triangleq -v \\
&= \gamma\left(\int_{-\infty}^0 \mathsf{p}(w+\gamma)\,\mathrm{d}\omega + \int_0^\infty \mathsf{p}(u+\gamma)\,\mathrm{d}u\right) = \gamma\int_{-\infty}^\infty \mathsf{p}(u+\gamma)\,\mathrm{d}u = \gamma\int_{-\infty}^\infty \mathsf{p}(u)\,\mathrm{d}u = \gamma
\end{aligned}
$$

✑

1.8 Sequences

1.8.1 Sequences

Definition 1.41. [56]

DEF *A* FUNCTION *in* $X^{\mathbb{D}}$ (Definition 1.6 page 6) *is an* X*-valued* ***sequence*** *if* $\mathbb{D} \neq \varnothing$ *and* $\mathbb{D}$ *is a* CONVEX (Definition 1.25 page 10) *subset of* $\mathbb{Z}$. *A sequence may be denoted in the form* $(\!(x_n)\!)_{n\in\mathbb{D}}$, *or simply as* $(\!(x_n)\!)$.

Definition 1.42. *The sequence* $(\!(y_n)\!)_{\mathbb{D}_2}$ *is the sequence* $(\!(x_n)\!)_{\mathbb{D}_1}$ ***down sampled by a factor of*** M, *where* $M \in \mathbb{N}$, *if* $n \in \mathbb{D}_2 \iff Mn \in \mathbb{D}_1$ *and* $y_n = x_{Mn} \quad \forall n\in\mathbb{D}_2$.

Definition 1.43. *Let* $\oplus$ *be the* ADDITION OPERATOR (Definition 1.15 page 7) *and* $\otimes$ *the* MULTIPLICATION OPERATOR (Definition 1.16 page 8). *Let* $\mathbb{D}_1$ *and* $\mathbb{D}_2$ *be* CONVEX SUBSETS *of* $\mathbb{Z}$.
Let $\mathbb{D} \triangleq \left(\bigwedge \mathbb{D}_1 + \bigwedge \mathbb{D}_2 - 1 : \bigvee \mathbb{D}_1 + \bigvee \mathbb{D}_2 + 1\right)$.
Let $(\!(x_n)\!)_{\mathbb{D}_1}$ *be a* SEQUENCE *over a* FIELD $\mathbb{F}_1$ *and* $(\!(y_n)\!)_{\mathbb{D}_2}$ *a* SEQUENCE *over a* FIELD $\mathbb{F}_2$.

DEF *The* ***convolution*** $(\!(z_n)\!)_{\mathbb{D}} \triangleq (\!(x_n)\!)_{\mathbb{D}_1} \star (\!(y_n)\!)_{\mathbb{D}_2}$ *of* $(\!(x_n)\!)$ *and* $(\!(y_n)\!)$ *is defined as*

$$z_n \triangleq \bigoplus_{m\in\mathbb{D}_1} \mathsf{f}(n,m) \quad \text{where } \mathsf{f} \text{ is defined as } \mathsf{f}(n,m) \triangleq \left\{\begin{array}{ll} x_m \otimes y_{n-m} & \text{if } m \in \mathbb{D}_1 \text{ and } (n-m) \in \mathbb{D}_2 \\ 0 & \text{otherwise} \end{array}\right\} \quad \forall n,m\in\mathbb{D}$$

Proposition 1.4. *Let* $(\!(x_n)\!)$ *and* $(\!(y_n)\!)$ *be finite* SEQUENCES *with lengths* N *and* M, *respectively. Then the length of* $(\!(x_n)\!) \star (\!(y_n)\!)$ *is* $N + M - 1$.

Example 1.14. [57] Let $(\!(x_n)\!)_{n\in\mathbb{Z}}$ and $(\!(y_n)\!)_{n\in\mathbb{Z}}$ be *sequences* over a *field* $\mathbb{F}$.
Then the *domain* $\mathbb{D}$ of the *convolution* $(\!(z_n)\!)_{n\in\mathbb{D}} \triangleq (\!(x_n)\!) \star (\!(y_n)\!)$ is
$\mathbb{D} \triangleq \left(\bigwedge \mathbb{Z} + \bigwedge \mathbb{Z} - 1 : \bigvee \mathbb{Z} + \bigvee \mathbb{Z} + 1\right) = \mathbb{Z}$ and $z_n \triangleq \sum_{m\in\mathbb{Z}} x_m y_{n-m}$.

Example 1.15. Let $(\!(x_n)\!)_{[0:1]} \triangleq (\!(1,2)\!)$ and $(\!(y_n)\!)_{[0:2]} \triangleq (\!(10,20,50)\!)$ be *sequences* over the *field* $(\mathbb{R},+,\cdot,0,1)$.
Then the *domain* $\mathbb{D}$ of the *convolution* $(\!(z_n)\!)_{n\in\mathbb{D}} \triangleq (\!(x_n)\!) \star (\!(y_n)\!)$ is

[55] Burkill (2004) page 35
[56] Simmons (2016) ⟨"Formal definition"⟩
[57] historical references: Cauchy (1821) ⟨Chapter IV⟩, Apostol (1975) page 204 ⟨note that convolution is a single element in a series that is the "Cauchy product"⟩, Dominguez-Torres (2010) page 20 ⟨section 4.2: connection to the work of Cauchy⟩, Dominguez-Torres (2015) ⟨history of the continuous convolution operation⟩

$\mathbb{D} \triangleq (0+0-1 : 1+2+1) = [0 : 3] = \{0,1,2,3\}$ and

$$\begin{aligned}
(\!(z_n)\!)_{n\in\mathbb{D}} &\triangleq \left(\!\!\left(\sum_{m\in\{0,1\}} \mathsf{f}(0,m),\ \sum_{m\in\{0,1\}} \mathsf{f}(1,m),\ \sum_{m\in\{0,1\}} \mathsf{f}(2,m),\ \sum_{m\in\{0,1\}} \mathsf{f}(3,m)\right)\!\!\right)_{\{0,1,2,3\}} \\
&\triangleq (\!((1\times 10+0),\ (1\times 20+2\times 10),\ (1\times 50+2\times 20),\ (0+2\times 50))\!)_{\{0,1,2,3\}} \\
&= \left(\!\!\left(\underbrace{10}_{z_0},\ \underbrace{40}_{z_1},\ \underbrace{90}_{z_2},\ \underbrace{100}_{z_3}\right)\!\!\right)_{\{0,1,2,3\}}
\end{aligned}$$

Example 1.16. Let $(\!(x_n)\!)_{[0:1]} \triangleq (\!(1,2)\!)$ and $(\!(y_n)\!)_{[3:5]} \triangleq (\!(10,20,50)\!)$ be *sequence*s over the field $(\mathbb{R},+,\cdot,0,1)$. Then the *domain* $\mathbb{D}$ of the *convolution* $(\!(z_n)\!)_{n\in\mathbb{D}} \triangleq (\!(x_n)\!) \star (\!(y_n)\!)$ is

$\mathbb{D} \triangleq (0+3-1 : 1+5+1) = [3 : 6] = \{3,4,5,6\}$ and

$$\begin{aligned}
(\!(z_n)\!)_{n\in\mathbb{D}} &\triangleq \left(\!\!\left(\sum_{m\in\{0,1\}} \mathsf{f}(3,m),\ \sum_{m\in\{0,1\}} \mathsf{f}(4,m),\ \sum_{m\in\{0,1\}} \mathsf{f}(5,m),\ \sum_{m\in\{0,1\}} \mathsf{f}(6,m)\right)\!\!\right)_{\{3,4,5,6\}} \\
&\triangleq (\!((1\times 10+0),\ (1\times 20+2\times 10),\ (1\times 50+2\times 20),\ (0+2\times 50))\!)_{\{3,4,5,6\}} \\
&= \left(\!\!\left(\underbrace{10}_{z_3},\ \underbrace{40}_{z_4},\ \underbrace{90}_{z_5},\ \underbrace{100}_{z_6}\right)\!\!\right)_{\{3,4,5,6\}}
\end{aligned}$$

Example 1.17. Let $(\!(x_n)\!)_{[0:1]} \triangleq (\!(1,2)\!)$ and $(\!(y_n)\!)_{[0:2]} \triangleq (\!((3,4),(5,6),(7,8))\!)$ be *sequence*s. Then the *domain* $\mathbb{D}$ of the *convolution* $(\!(z_n)\!)_{n\in\mathbb{D}} \triangleq (\!(x_n)\!) \star (\!(y_n)\!)$ is $\mathbb{D} = \{0,1,2,3\}$ and

$$\begin{aligned}
(\!(z_n)\!)_{n\in\mathbb{D}} &\triangleq \left(\!\!\left(\bigoplus_{m\in\{0,1\}} \mathsf{f}(0,m),\ \bigoplus_{m\in\{0,1\}} \mathsf{f}(1,m),\ \bigoplus_{m\in\{0,1\}} \mathsf{f}(2,m),\ \bigoplus_{m\in\{0,1\}} \mathsf{f}(3,m)\right)\!\!\right)_{\{0,1,2,3\}} \\
&\triangleq (\!([1\otimes(3,4)\oplus 0],\ [1\otimes(5,6)\oplus 2\otimes(3,4)],\ [1\otimes(7,8)\oplus 2\otimes(5,6)],\ [0\oplus 2\otimes(7,8)])\!)_{\{0,1,2,3\}} \\
&\triangleq (\!([(3,4)],\ [(5,6)\oplus(6,8)],\ [(7,8)\oplus(10,12)],\ [(14,16)])\!)_{\{0,1,2,3\}} \\
&= \left(\!\!\left(\underbrace{(3,4)}_{z_0},\ \underbrace{(11,14)}_{z_1},\ \underbrace{(17,20)}_{z_2},\ \underbrace{(14,16)}_{z_3}\right)\!\!\right)_{\{0,1,2,3\}}
\end{aligned}$$

Definition 1.44. *Let $(\!(x_n)\!)_{n\in\mathbb{D}}$ and $(\!(y_n)\!)_{n\in\mathbb{D}}$ be sequences over a* FIELD $\mathbb{F} \triangleq (X,+,\cdot,0,1)$*, and α an element in $\mathbb{F}$. The operations $\alpha + (\!(x_n)\!)$, $(\!(x_n)\!) + \alpha$, $\alpha\,(\!(x_n)\!)$, and $(\!(x_n)\!)\,\alpha$ are defined as*

DEF

$$\begin{array}{rclcll}
\alpha + (\!(x_n)\!)_{n\in\mathbb{D}} & \triangleq & (\!(x_n)\!)_{n\in\mathbb{D}} + \alpha & \triangleq & (\!(x_n+\alpha)\!)_{n\in\mathbb{D}} & \forall\alpha\in\mathbb{F} \quad \text{and} \\
\alpha\,(\!(x_n)\!)_{n\in\mathbb{D}} & \triangleq & (\!(x_n)\!)_{n\in\mathbb{D}}\,\alpha & \triangleq & (\!(\alpha x_n)\!)_{n\in\mathbb{D}} & \forall\alpha\in\mathbb{F} \quad .
\end{array}$$

1.8.2 Filtering

Definition 1.45. *Let $(\!(x_n)\!)_{\mathbb{D}_1}$ and $(\!(y_n)\!)_{\mathbb{D}_2}$ be* SEQUENCES *(Definition 1.41 page 25).*
The SEQUENCE $(\!(z_n)\!)_{n\in\mathbb{D}}$ *is said to be $(\!(x_n)\!)$* ***filtered*** *by $(\!(y_n)\!)$ if $(\!(z_n)\!) \triangleq (\!(x_n)\!) \star (\!(y_n)\!)$ (Definition 1.43 page 25).*
Moreover, in this case, the operation $\star\,(\!(y_n)\!)$ is a ***filter*** *on the* SEQUENCE $(\!(x_n)\!)$.

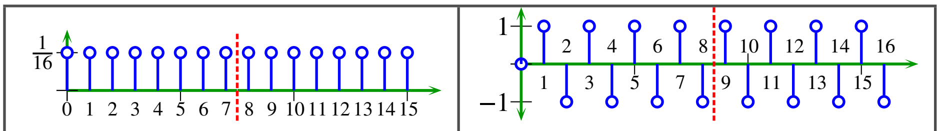
Definition 1.46. *A* ***length*** M ***low pass rectangular sequence*** $(\!(h_n)\!)_{n\in[0:M-1]}$ *is here defined as*
$h_n = {}^1\!/\!_M$ *for* $n \in [0 : M-1]$.

Definition 1.47. *A* ***length*** M ***high pass rectangular sequence*** $(\!(h_n)\!)_{n\in[0:M]}$ *is here defined as*

$$h_n \triangleq \left\{\begin{array}{ll} 0 & \text{for } n = 0 \\ (-1)^{n+1} & \text{for } n = 1,2,\ldots,M \end{array}\right\}$$

Note that in this definition, the sequence has been offset by 1 on the x-axis from what might normally be expected. This is for the purpose of computational convenience used in Section 3.3.2 (page 95).

Example 1.18. A *length 16 low pass rectangular sequence* and *length 16 high pass rectangular sequence* are illustrated below:

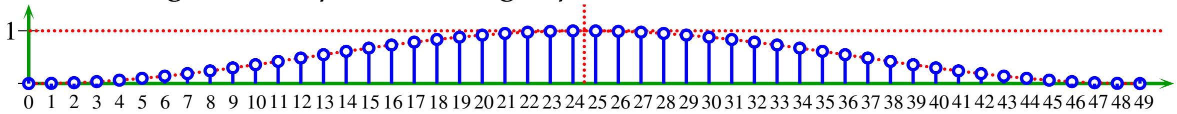


length 16 low pass rectangular sequence | *length 16 high pass rectangular sequence*

Definition 1.48. [58] *A **length M low pass Hanning sequence** $(\!(h_n)\!)_{n\in[0:M-1]}$ is here defined as*

$$h_n \triangleq \frac{1}{2}\left[1-\cos\left(\frac{2\pi n}{M-1}\right)\right] \text{ for } n = 0,1,2,\ldots,M-1.$$

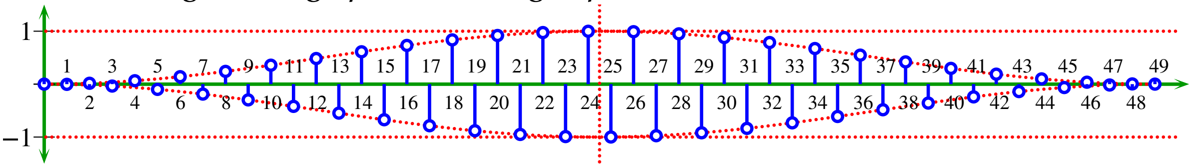
Example 1.19. A *length 50 low pass Hanning sequence* is illustrated below:



Definition 1.49. *A **length M high pass Hanning sequence** $(\!(h_n)\!)_{n\in[0:M-1]}$ is here defined as*

$$h_n \triangleq (-1)^n \frac{1}{2}\left[1-\cos\left(\frac{2\pi n}{M-1}\right)\right] \text{ for } n = 0,1,2,\ldots,M-1$$

Example 1.20. A *length 50 high pass Hanning sequence* is illustrated below:



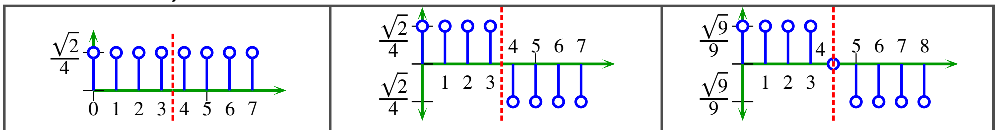
Definition 1.50. *A **length M Haar scaling sequence** $(\!(h_n)\!)_{n\in[0:M-1]}$ is here defined as*

$$h_n = \sqrt{1/M} \quad \text{for } n \in [0:M-1].$$

Definition 1.51. *A **length M Haar wavelet sequence** $(\!(h_n)\!)_{n\in[0:M-1]}$ is here defined as*

$$h_n \triangleq \left\{\begin{array}{rl} +\sqrt{1/M} & \text{for } n = 0,\ 1,\ \ldots,\ \lfloor M/2\rfloor - 1 \\ -\sqrt{1/M} & \text{for } n = \lceil M/2\rceil,\ \lceil M/2\rceil + 1,\ \ldots,\ M-1 \\ 0 & \text{otherwise} \end{array}\right\}$$

Example 1.21. A *length 8 Haar scaling sequence*, a *length 8 Haar wavelet sequence*, and *length 9 Haar wavelet sequence* are illustrated below:



length 8 Haar scaling sequence | *length 8 Haar wavelet sequence* | *length 9 Haar wavelet sequence*

1.8.3 Discrete Fourier Transform

Definition 1.52. *Let ⊕ be the* ADDITION OPERATOR (Definition 1.15 page 7) *and ⊗ the* MULTIPLICATION OPERATOR (Definition 1.16 page 8). *Let $(\!(x_n)\!)_{n\in\mathbb{D}}$ be a length N* SEQUENCE (Definition 1.41 page 25).

[58] Blackman and Tukey (1958) page 502 ⟨B.5 Particular Pairs of Windows⟩, Blackman and Tukey (1959), page 98 ⟨B.5 Particular Pairs of Windows⟩, Oppenheim and Schafer (1999) page 763, Prabhu (2013) page 148

DEF *The* ***discrete Fourier transform*** $\mathsf{DFT}⦅x_n⦆$ *of* $⦅x_n⦆$ *is a sequence* $⦅y_k⦆_{k\in\mathbb{D}}$ *over* $\mathbb{C}$, *where the element* y_k *is defined as*

$$y_k \triangleq \sqrt{\frac{1}{N}} \bigoplus_{n\in\mathbb{D}} \left[x_n \otimes \exp\left(\frac{-i2\pi nk}{N}\right)\right]$$

Example 1.22. Suppose $⦅x_n⦆_{n\in[1:N]}$ is a length N *sequence* over $\mathbb{R}$. Then the *discrete Fourier transform* $\mathsf{DFT}⦅x_n⦆$ of $⦅x_n⦆$ is the sequence $⦅y_k⦆_{k\in[0:N-1]}$ over $\mathbb{C}$, where the element y_k is defined as

$$y_k \triangleq \sqrt{\frac{1}{N}} \sum_{n=0}^{n=N-1} x_n \exp\left(\frac{-i2\pi nk}{N}\right)$$

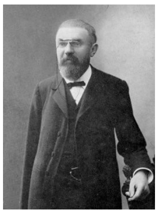
Example 1.23. Suppose $\boldsymbol{x} \triangleq \left(\left(\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}, \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}\right)\right)$ is a length $N = 3$ *sequence* over $\mathbb{R}^3$. Then the *discrete Fourier transform* $\mathsf{DFT}\boldsymbol{x}$ of $\boldsymbol{x}$ is the sequence $⦅\theta_0, \theta_1, \theta_2⦆$ where

$$\begin{aligned}
\theta_k &= \sqrt{\frac{1}{N}} \bigoplus_{n=0}^{n=2} \left[\begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} \otimes \exp\left(\frac{-i2\pi nk}{N}\right)\right] && \text{by definition of DFT (Definition 1.52 page 27)} \\
&= \sqrt{\frac{1}{3}} \left(\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \oplus \begin{bmatrix} x_1 \exp\left(\frac{-i2\pi k}{3}\right) \\ y_1 \exp\left(\frac{-i2\pi k}{3}\right) \\ z_1 \exp\left(\frac{-i2\pi k}{3}\right) \end{bmatrix} \oplus \begin{bmatrix} x_2 \exp\left(\frac{-i2\pi 2k}{3}\right) \\ y_2 \exp\left(\frac{-i2\pi 2k}{3}\right) \\ z_2 \exp\left(\frac{-i2\pi 2k}{3}\right) \end{bmatrix}\right) && \text{by definition of } \otimes \text{ (Definition 1.16 page 8)} \\
&= \sqrt{\frac{1}{3}} \begin{bmatrix} x_0 + x_1 \exp\left(\frac{-i2\pi k}{3}\right) + x_2 \exp\left(\frac{-i4\pi k}{3}\right) \\ y_0 + y_1 \exp\left(\frac{-i2\pi k}{3}\right) + y_2 \exp\left(\frac{-i4\pi k}{3}\right) \\ z_0 + z_1 \exp\left(\frac{-i2\pi k}{3}\right) + z_2 \exp\left(\frac{-i4\pi k}{3}\right) \end{bmatrix} && \text{by definition of } \oplus \text{ (Definition 1.15 page 7)}
\end{aligned}$$

CHAPTER 2

STOCHASTIC PROCESSING ON WEIGHTED GRAPHS

❝*Les mathématiciens n'étudient pas des objets, mais des relations entre les objets; il leur est donc indifférent de remplacer ces objets par d'autres, pourvu que les relations ne changent pas. La matière ne leur importe pas, la forme seule les intéresse.*❞



❝*Mathematicians do not study objects, but the relations between objects; to them it is a matter of indifference if these objects are replaced by others, provided that the relations do not change. Matter does not engage their attention, they are interested in form alone.*❞

Jules Henri Poincaré (1854-1912), physicist and mathematician [1]

2.1 Outcome subspaces

2.1.1 Definitions

Traditional probability theory is performed in a *probability space* $(\Omega, \mathbb{E}, \mathsf{P})$. This section extends[2] the probability space structure to include what herein is called an *outcome subspace* (next definition).

Definition 2.1.

DEF *An **extended probability space** is the tuple* $(\Omega, \le, \mathsf{d}, \mathbb{E}, \mathsf{P})$ *where* $(\Omega, \mathbb{E}, \mathsf{P})$ *is a* PROBABILITY SPACE (Definition 1.38 page 24) *and* $(\Omega, \mathsf{d}, \le)$ *is an* ORDERED QUASI-METRIC SPACE (Definition 1.34 page 22). *The 4-tuple* $\left(\Omega, \dot{\le}, \dot{\mathsf{d}}, \dot{\mathsf{P}}\right)$ *is an **outcome subspace** of the* EXTENDED PROBABILITY SPACE $(\Omega, \le, \mathsf{d}, \mathbb{E}, \mathsf{P})$.

Definition 2.2. *Let* $\boldsymbol{G} \triangleq \left(\Omega, \dot{\le}, \dot{\mathsf{d}}, \dot{\mathsf{P}}\right)$ *be an* OUTCOME SUBSPACE (Definition 2.1 page 29).

DEF

*The **nth-moment** $\mathsf{m}_n(x,y)$ from x to y in $\boldsymbol{G}$ is defined as*	$\mathsf{m}(x,y) \triangleq [\mathsf{d}(x,y)]^n \mathsf{P}(y)$	$\forall x,y \in \Omega,\ n \in \mathbb{N}$.
*The **moment** $\mathsf{m}(x,y)$ from x to y in $\boldsymbol{G}$ is defined as*	$\mathsf{m}(x,y) \triangleq \mathsf{m}_1(x,y)$	$\forall x,y \in \Omega$.

This paper introduces a quantity called the *outcome center* of an *outcome subspace* (next definition) which is in essence the same as the *center* of a *graph* (Definition 1.19 page 8).

[1] quote: Poincaré (1902a) ⟨Chapter 2⟩
translation: Poincaré (1902b), page 20
image: http://www-groups.dcs.st-and.ac.uk/~history/PictDisplay/Poincare.html

[2] 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..."⟩

Definition 2.3. *Let* $\boldsymbol{G} \triangleq (\Omega, \dot{\leq}, \dot{\mathsf{d}}, \dot{\mathsf{P}})$ *be an* OUTCOME SUBSPACE *(Definition 2.1 page 29).*

DEF

$$\dot{\complement}(\boldsymbol{G}) \triangleq \arg\min_{x\in\Omega} \max_{y\in\Omega} \underbrace{\mathsf{d}(x,y)\,\mathsf{P}(y)}_{\mathsf{m}(x,y)}$$ *is the* ***outcome center*** *of* $\boldsymbol{G}$.

The following additional definitions are of interest due in part to Corollary D.2 (page 177) and the *minimax inequality* (Theorem 1.2 page 12). They are illustrated in several examples in this section. However, most of them are not used outside this section.

Definition 2.4. *Let* $\boldsymbol{G} \triangleq (\Omega, \dot{\leq}, \dot{\mathsf{d}}, \dot{\mathsf{P}})$ *be an* OUTCOME SUBSPACE *(Definition 2.1 page 29).*

DEF

$\dot{\complement}_{\mathsf{a}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\Omega} \sum_{y\in\Omega} \mathsf{d}(x,y)\,\mathsf{P}(y)$	*is the* ***arithmetic center***	*of* $\boldsymbol{G}$.
$\dot{\complement}_{\mathsf{g}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\Omega} \prod_{y\in\Omega\setminus\{x\}} \left[\mathsf{d}(x,y)^{\mathsf{P}(y)}\right]$	*is the* ***geometric center***	*of* $\boldsymbol{G}$.
$\dot{\complement}_{\mathsf{h}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\Omega} \left(\sum_{y\in\Omega\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y)\right)^{-1}$	*is the* ***harmonic center***	*of* $\boldsymbol{G}$.
$\dot{\complement}_{\mathsf{m}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\Omega} \min_{y\in\Omega\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y)$	*is the* ***minimal center***	*of* $\boldsymbol{G}$.
$\dot{\complement}_{\mathsf{M}}(\boldsymbol{G}) \triangleq \arg\max_{x\in\Omega} \min_{y\in\Omega\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y)$	*is the* ***maxmin center***	*of* $\boldsymbol{G}$.

In a manner similar to the traditional *variance* function (Definition 1.40 page 24), the *outcome variance* (next) is a kind of measure of the quality of the outcome center as a representative estimate of all the values of the *outcome subspace*. Said another way, it is in essence the expected error of the center measure.

Definition 2.5. *Let* $\boldsymbol{G} \triangleq (\Omega, \dot{\leq}, \dot{\mathsf{d}}, \dot{\mathsf{P}})$ *be an* OUTCOME SUBSPACE *(Definition 2.1 page 29).*

DEF

The ***outcome variance*** $\dot{\mathrm{Var}}(\boldsymbol{G}; \dot{\complement}_x)$ *of* $\boldsymbol{G}$ *with respect to* $\dot{\complement}_x$ *is*

$$\dot{\mathrm{Var}}(\boldsymbol{G}; \dot{\complement}_x) \triangleq \sum_{x\in\Omega} \underbrace{\mathsf{d}^2\left(\dot{\complement}_x(\boldsymbol{G}), x\right)\mathsf{P}(x)}_{\mathsf{m}_2(\dot{\complement}(\boldsymbol{G}),x)}.$$

where $\dot{\complement}_x$ *is any of the operators defined in Definition 2.3 or Definition 2.4 (page 30). Moreover,* $\dot{\mathrm{Var}}(\boldsymbol{G}) \triangleq \dot{\mathrm{Var}}(\boldsymbol{G}; \dot{\complement})$, *where* $\dot{\complement}$ *is the* OUTCOME CENTER *(Definition 2.3 page 30).*

Remark 2.1. The quantity $\mathsf{p}(x,\Omega) \triangleq \sum_{y\in\Omega} \mathsf{d}(x,y)\,\mathsf{P}(y)$ in the *arithmetic center* $\dot{\complement}_{\mathsf{a}}(\boldsymbol{G})$ (Definition 2.4 page 30) is itself a *metric* (Definition D.7 page 163). Thus, $\dot{\complement}_{\mathsf{a}}(\boldsymbol{G})$ is the x that produces the minimum of all the metrics with center x.

✎PROOF: This follows directly from *power mean metrics* theorem with $r = 1$ (Theorem D.10 page 170). ⇨

2.1.2 Specific outcome subspaces

Definition 2.6.

DEF

The structure $\boldsymbol{G} \triangleq (\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}, \dot{\mathsf{d}}, \dot{\leq}, \dot{\mathsf{P}})$ *is the* ***weighted die outcome subspace*** *if* $\boldsymbol{G}$ *is an* OUTCOME SUBSPACE, $\dot{\leq} = \varnothing$ (UNORDERED *Definition 1.20 page 9*), *and* $\dot{\mathsf{d}}$ *is the* DISCRETE METRIC *(Definition D.8 page 166).*

Definition 2.7.

DEF

The structure $\boldsymbol{G} \triangleq (\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}, \dot{\mathsf{d}}, \dot{\leq}, \dot{\mathsf{P}})$ *is the* ***fair die outcome subspace*** *if* $\boldsymbol{G}$ *is a* WEIGHTED DIE OUTCOME SUBSPACE *(Definition 2.6), and*

$$\dot{\mathsf{P}}(⚀) = \dot{\mathsf{P}}(⚁) = \dot{\mathsf{P}}(⚂) = \dot{\mathsf{P}}(⚃) = \dot{\mathsf{P}}(⚄) = \dot{\mathsf{P}}(⚅) = \tfrac{1}{6}.$$

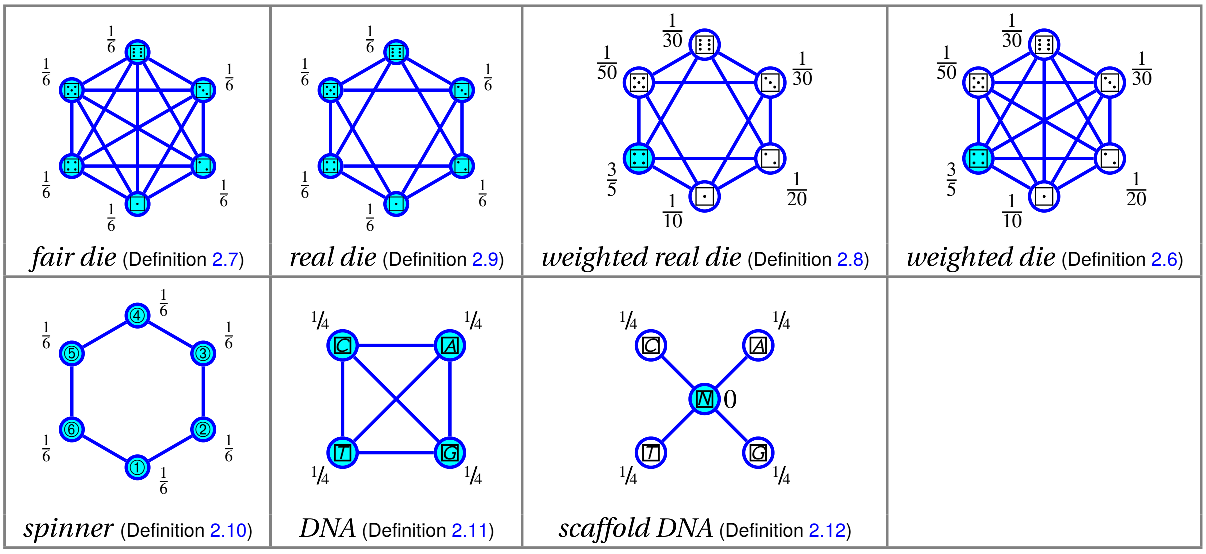


Figure 2.1: example *outcome subspaces* (Definition 2.1 page 29) illustrated by *weighted graphs* with shaded *expected values*.

$\dot{\mathsf{d}}(x,y)$	⚀	⚁	⚂	⚃	⚄	⚅
⚀	0	1	1	1	1	2
⚁	1	0	1	1	2	1
⚂	1	1	0	2	1	1
⚃	1	1	2	0	1	1
⚄	1	2	1	1	0	1
⚅	2	1	1	1	1	0

Definition 2.8. *The structure* $\boldsymbol{G} \triangleq (\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}, \dot{\mathsf{d}}, \varnothing, \dot{\mathsf{P}})$ *is the* ***weighted real die outcome subspace*** *if* $\boldsymbol{G}$ *is an* OUTCOME SUBSPACE, *and* METRIC $\dot{\mathsf{d}}$ *is defined as in the table to the right.*

Definition 2.9. *The structure* $\boldsymbol{G} \triangleq (\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}, \dot{\mathsf{d}}, \varnothing, \dot{\mathsf{P}})$ *is the* ***real die outcome subspace*** *if* $\boldsymbol{G}$ *is a* WEIGHTED REAL DIE OUTCOME SUBSPACE *(Definition 2.8 page 31) with*

$$\dot{\mathsf{P}}(⚀) = \dot{\mathsf{P}}(⚁) = \dot{\mathsf{P}}(⚂) = \dot{\mathsf{P}}(⚃) = \dot{\mathsf{P}}(⚄) = \dot{\mathsf{P}}(⚅) = \tfrac{1}{6}.$$

$\dot{\mathsf{d}}(x,y)$	①	②	③	④	⑤	⑥
①	0	1	2	3	2	1
②	1	0	1	2	3	2
③	2	1	0	1	2	3
④	3	2	1	0	1	2
⑤	2	3	2	1	0	1
⑥	1	2	3	2	1	0

Definition 2.10. *The structure* $\boldsymbol{G} \triangleq (\{①, ②, ③, ④, ⑤, ⑥\}, \dot{\mathsf{d}}, \varnothing, \dot{\mathsf{P}})$ *is the* ***spinner outcome subspace*** *if* $\boldsymbol{G}$ *is an* OUTCOME SUBSPACE,

$$\dot{\mathsf{P}}(①) = \dot{\mathsf{P}}(②) = \dot{\mathsf{P}}(③) = \dot{\mathsf{P}}(④) = \dot{\mathsf{P}}(⑤) = \dot{\mathsf{P}}(⑥) = \tfrac{1}{6},$$

and METRIC $\dot{\mathsf{d}}$ *is defined as in the table to the right.*

Definition 2.11. *The structure* $\boldsymbol{H} \triangleq (\{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}\}, \dot{\mathsf{d}}, \varnothing, \dot{\mathsf{P}})$ *is the* ***DNA outcome subspace***, *or* ***genome outcome subspace***, *if* $\boldsymbol{H}$ *is an* OUTCOME SUBSPACE, *and* $\dot{\mathsf{d}}$ *is the* DISCRETE METRIC *(Definition D.8 page 166).*

Definition 2.12. *The structure* $\boldsymbol{H} \triangleq (\{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}, \boxed{N}\}, \dot{\mathsf{d}}, \dot{\le}, \dot{\mathsf{P}})$ *is the* ***DNA scaffold outcome subspace***, *or* ***genome scaffold outcome subspace***, *if* $\boldsymbol{H}$ *is an* OUTCOME SUBSPACE,

$$\dot{\le} = \{(\boxed{N}, \boxed{A}), (\boxed{N}, \boxed{T}), (\boxed{N}, \boxed{C}), (\boxed{N}, \boxed{G}), \}$$

($\boxed{N} < \boxed{A}$, $\boxed{N} < \boxed{T}$, $\boxed{N} < \boxed{C}$, *and* $\boxed{N} < \boxed{G}$, *but otherwise* UNORDERED*),* $\dot{\mathsf{P}}$ *is a* PROBABILITY FUNCTION, *and* METRIC $\dot{\mathsf{d}}$ *is defined as in the table to the right.*

$\dot{\mathsf{d}}(x,y)$	$\boxed{N}$	$\boxed{A}$	$\boxed{T}$	$\boxed{C}$	$\boxed{G}$
$\boxed{N}$	0	$\sqrt{2}/2$	$\sqrt{2}/2$	$\sqrt{2}/2$	$\sqrt{2}/2$
$\boxed{A}$	$\sqrt{2}/2$	0	1	1	1
$\boxed{T}$	$\sqrt{2}/2$	1	0	1	1
$\boxed{C}$	$\sqrt{2}/2$	1	1	0	1
$\boxed{G}$	$\sqrt{2}/2$	1	1	1	0

2.1.3 Example calculations

Example 2.1. Let $\boldsymbol{G} \triangleq (\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}, \dot{\mathsf{d}}, \dot{\le}, \dot{\mathsf{P}})$ be the *fair die outcome subspace* (Definition 2.7 page 30). This structure is illustrated by the *weighted graph* (Definition 1.18 page 8) in Figure 2.1 (page 31)

(A), where each line segment represents a distance of 1. This structure has the following geometric values:

$$\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}$$
$$\dot{\mathsf{Var}}(\boldsymbol{G}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{a}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{g}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{h}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{m}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{M}}) = 0$$

✎PROOF:

$$\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\frac{1}{6} && \text{by definition of } \boldsymbol{G}\\
&= \arg\min_{x\in\boldsymbol{G}} \frac{1}{6}\{1,\ 1,\ 1,\ 1,\ 1,\ 1\} && \text{by definition of \textit{discrete metric} (Definition D.8 page 166)}\\
&= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} && \text{by definition of } \boldsymbol{G}\\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\frac{1}{6} && \text{by definition of } \boldsymbol{G}\\
&= \arg\min_{x\in\boldsymbol{G}} \frac{1}{6}\{5,\ 5,\ 5,\ 5,\ 5,\ 5\} && \text{by definition of \textit{discrete metric} (Definition D.8 page 166)}\\
&= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} && \text{by definition of } \boldsymbol{G}\\
\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{\mathsf{P}(y)} && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{g}} \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} 1^{\frac{1}{6}} && \text{by definition of } \boldsymbol{G} \text{ and \textit{discrete metric} (Definition D.8 page 166)}\\
&= \arg\min_{x\in\boldsymbol{G}} \{1, 1, 1, 1, 1, 1\}\\
&= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} && \text{by definition of } \boldsymbol{G}\\
\dot{\mathsf{C}}_{\mathsf{h}}(\mathsf{X}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \left(\sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y)\right)^{-1} && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{h}} \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \left(\sum_{y\in\boldsymbol{G}\setminus\{x\}} 1\times\frac{1}{6}\right)^{-1} && \text{by definition of } \boldsymbol{G}\\
&= \arg\min_{x\in\boldsymbol{G}} \left\{\frac{6}{5}, \frac{6}{5}, \frac{6}{5}, \frac{6}{5}, \frac{6}{5}, \frac{6}{5}\right\}\\
&= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} && \text{by definition of } \boldsymbol{G}\\
\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{m}} \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} 1\times\frac{1}{6} && \text{by definition of } \boldsymbol{G} \text{ and \textit{discrete metric} (Definition D.8 page 166)}\\
&= \arg\min_{x\in\boldsymbol{G}} \frac{1}{6}\{1, 1, 1, 1, 1, 1\}\\
&= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} && \text{by definition of } \boldsymbol{G}\\
\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &\triangleq \arg\max_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{M}} \text{ (Definition 2.4 page 30)}\\
&= \arg\max_{x\in\boldsymbol{G}} \frac{1}{6}\{1, 1, 1, 1, 1, 1\} && \text{by } \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \text{ result}\\
&= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} && \text{by definition of } \boldsymbol{G}
\end{aligned}$$

$$\dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{a}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{g}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{h}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{m}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{M}}) = \dot{\mathsf{Var}}(\boldsymbol{G})$$

$$
\begin{aligned}
&\triangleq \sum_{x\in\boldsymbol{G}} \left[\mathsf{d}\left(\dot{\mathsf{C}}(\boldsymbol{G}), x\right)\right]^2 \mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} (0^2)\frac{1}{6} && \text{because } \dot{\mathsf{C}}(\boldsymbol{G}) = \boldsymbol{G}\\
&= 0 && \text{by field property of } \textit{additive identity element } 0
\end{aligned}
$$

✏

Example 2.2. Let $\boldsymbol{G} \triangleq (\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}, \varnothing, \le, \mathsf{P})$ be the *real die outcome subspace* (Definition 2.9 page 31). This structure is illustrated by the *weighted graph* (Definition 1.18 page 8) in Figure 2.1 (page 31) (B), where each line segment represents a distance of 1. The structure has the following geometric values:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}\\
\dot{\mathsf{Var}}(\boldsymbol{G}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{a}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{g}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{h}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{M}}) &= 0
\end{aligned}
$$

✎PROOF:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\frac{1}{6} && \text{by definition of } \boldsymbol{G}\\
&= \arg\min_{x\in\boldsymbol{G}} \frac{1}{6}\{2,\ 2,\ 2,\ 2,\ 2,\ 2\} && \text{because for each } x\text{, there is a } y \text{ such that } \mathsf{d}(x,y)=2\\
&= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} && \text{by definition of } \boldsymbol{G}\\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\frac{1}{6} && \text{by definition of } \boldsymbol{G}\\
&= \arg\min_{x\in\boldsymbol{G}} \frac{1}{6}\left\{\begin{matrix}0+1+1+1+1+2\\1+0+1+1+2+1\\1+1+0+2+1+1\\1+1+2+0+1+1\\1+2+1+1+0+1\\2+1+1+1+1+0\end{matrix}\right\} = \arg\min_{x\in\boldsymbol{G}} \frac{1}{6}\left\{\begin{matrix}6\\6\\6\\6\\6\\6\end{matrix}\right\} = \left\{\begin{matrix}⚀\\⚁\\⚂\\⚃\\⚄\\⚅\end{matrix}\right\}\\
\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{\mathsf{P}(y)} && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{g}} \text{ (Definition 2.4 page 30)}\\
&\triangleq \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{\frac{1}{6}} && \text{by definition of } \boldsymbol{G}\\
&= \arg\min_{x\in\boldsymbol{G}} \left(\prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\right)^{\frac{1}{6}}\\
&= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y) && \text{because } \mathsf{f}(x) = x^{\frac{1}{6}} \text{ is } \textit{strictly isotone} \text{ and by Lemma 1.2 (page 16)}\\
&= \arg\min_{x\in\boldsymbol{G}} \{2,2,2,2,2,2\}\\
&= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}\\
\dot{\mathsf{C}}_{\mathsf{h}}(\mathsf{X}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \left(\sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y)\right)^{-1} && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{h}} \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \left(\sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)}\frac{1}{6}\right)^{-1} && \text{by definition of } \boldsymbol{G}
\end{aligned}
$$

$$
\begin{aligned}
&= \arg\min_{x\in\boldsymbol{G}} 6\left(\sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)}\right)^{-1} \\
&= \arg\max_{x\in\boldsymbol{G}} \frac{1}{2}6 \sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{2}{\mathsf{d}(x,y)} && \text{because } \phi(x) \triangleq x^{-1} \text{ is \textit{strictly antitone} and by Lemma 1.5 page 18} \\
&= \arg\min_{x\in\boldsymbol{G}} 3\left\{\begin{matrix} 0+2+2+2+2+1 \\ 2+0+2+2+1+2 \\ 2+2+0+1+2+2 \\ 2+2+1+0+2+2 \\ 2+1+2+2+0+2 \\ 1+2+2+2+2+0 \end{matrix}\right\} = \arg\min_{x\in\boldsymbol{G}} \frac{1}{6}\left\{\begin{matrix}9\\9\\9\\9\\9\\9\end{matrix}\right\} = \left\{\begin{matrix}\text{⚀}\\\text{⚁}\\\text{⚂}\\\text{⚃}\\\text{⚄}\\\text{⚅}\end{matrix}\right\} \\
\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{m}} \text{ (Definition 2.4 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} 1\times\frac{1}{6} && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{m}} \text{ (Definition 2.3 page 30)} \\
&= \frac{1}{6}\{2,2,2,2,2,2\} \\
&= \{\text{⚀},\text{⚁},\text{⚂},\text{⚃},\text{⚄},\text{⚅}\} && \text{by definition of } \boldsymbol{G} \\
\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &\triangleq \arg\max_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{M}} \text{ (Definition 2.4 page 30)} \\
&= \frac{1}{6}\{2,2,2,2,2,2\} && \text{by } \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \text{ result} \\
&= \{\text{⚀},\text{⚁},\text{⚂},\text{⚃},\text{⚄},\text{⚅}\} && \text{by definition of } \boldsymbol{G} \\
\dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{a}}) &= \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{g}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{h}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{M}}) = \dot{\mathsf{Var}}(\boldsymbol{G}) \\
&\triangleq \sum_{x\in\boldsymbol{G}} \left[\mathsf{d}\left(\dot{\mathsf{C}}(\boldsymbol{G}),x\right)\right]^2 \mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} (0)\frac{1}{6} && \text{because } \dot{\mathsf{C}}(\boldsymbol{G}) = \boldsymbol{G} \\
&= 0 && \text{by field property of \textit{additive identity element} } 0
\end{aligned}
$$

✏

Remark 2.2. Let $\boldsymbol{G} \triangleq (\Omega, \mathsf{d}, \leq, \mathsf{P})$ be the *fair die outcome subspace* (Example 2.1 page 31). Let $\boldsymbol{H} \triangleq (\Omega, \mathsf{p}, \leq, \mathsf{P})$ be the *real die outcome subspace* (Example 2.2 page 33). These two subspaces are identical except for their metrics d and p. So we can say that $\boldsymbol{G}$ and $\boldsymbol{H}$ *are* distinguished by their metrics. However, note that they are *indistinguishable* by the topologies induced by their metrics, because they both induce the same topology—the *discrete topology* 2^{Ω} (Definition 1.8 page 6). That is, the geometric distinction provided in metric spaces is in general lost in topological spaces. Thus, topological spaces are arguably too general for the type of stochastic processing presented in this paper; rather, the stochastic processing discussed in this paper calls for metric space structure. And in this paper, this type of metric space structure is referred to as **metric geometry**.

✎PROOF:

1. Every *metric space* (Ω, d) (Definition D.7 page 163) induces a **topological space** $(\Omega, \boldsymbol{T})$.

2. In particular, a *metric* d induces an **open ball** $\mathsf{B}(x,r) \in \left(2^{\Omega}\right)^{(\Omega\times\mathbb{R}^+)}$ centered at x with radius r such that $\mathsf{B}(x,r) \triangleq \{y\in\Omega | \mathsf{d}(x,y) < r\}$.

3. At each outcome x in $\boldsymbol{G}$, only two *open balls* are possible: $\mathsf{B}(x,r) = \left\{\begin{array}{ll} \{x\} & \text{for } 0 < r \leq 1 \\ \Omega & \text{for } r > 1 \end{array}\right\}$.

4. Let x' represent the die face which, when its numeric value is summed "in the usual way" with the numeric value of the die face x, equals 7. Then at each point x in $\boldsymbol{H}$, three *open balls* are possible: $\mathsf{B}(x,r) = \left\{\begin{array}{ll} \{x\} & \text{for } 0 < r \leq 1 \\ \{x, x'\} & \text{for } 1 < r \leq 2 \\ \Omega & \text{for } r > 2 \end{array}\right\}$.

5. The *open balls* of (Ω, d) or (Ω, p) in turn induce a **base** for a **topology** $\boldsymbol{T}$, such that $\boldsymbol{T} = \left\{U \in 2^{\Omega} | U \text{ is a union of } \textit{open balls}\right\}$. The topology induced by $\boldsymbol{G}$ is the *discrete topology* 2^{Ω} (Definition 1.8 page 6). The topology induced by $\boldsymbol{H}$ is also the *discrete topology* 2^{Ω}.

6. So the metrics of $\boldsymbol{G}$ and $\boldsymbol{H}$ are different. And the balls induced by $\boldsymbol{G}$ and those induced by $\boldsymbol{H}$ are different. However, the topologies induced by $\boldsymbol{G}$ and $\boldsymbol{H}$ are the same.

✎

Example 2.3. The *weighted real die outcome subspace* (Definition 2.6 page 30) illustrated in Figure 2.1 page 31 (C) has the following geometric characteristics:

$$\begin{array}{rcl@{\qquad}rcccl}
\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &=& \{\text{⚃}\} & \dot{\mathsf{Var}}(\boldsymbol{G}) &=& \frac{101}{300} &\approx& 0.33 \\
\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) &=& \{\text{⚁}\} & \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{g}}) &=& \frac{127}{150} &\approx& 0.847 \\
\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) &=& \{\text{⚄}\} & \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{h}}) &=& \frac{145}{150} &\approx& 0.967\ . \\
\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) &=& \{\text{⚀}, \text{⚂}, \text{⚃}, \text{⚅}\} & \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{m}}) &=& \frac{11}{50} &=& 0.22 \\
\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &=& \{\text{⚁}, \text{⚄}\} & \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{M}}) &=& \frac{11}{50} &\approx& 0.767
\end{array}$$

Note that the *outcome center* $\dot{\mathsf{C}}(\boldsymbol{G})$ and *arithmetic center* $\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G})$ again yield identical results. Also note that of the four center measures of cardinality 1 ($\left|\dot{\mathsf{C}}(\boldsymbol{G})\right| = \left|\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G})\right| = \left|\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G})\right| = \left|\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G})\right| = 1$ Definition 1.13 page 7), $\dot{\mathsf{C}}$ and $\dot{\mathsf{C}}_{\mathsf{a}}$ yield by far the lowest variance measures.

✎PROOF:

$$\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \arg\min_{x \in \boldsymbol{G}} \max_{y \in \boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)} \\
&\triangleq \arg\min_{x \in \boldsymbol{G}} \max_{y \in \boldsymbol{G}} \frac{1}{300}\mathsf{d}(x,y)\,\mathsf{P}(y)300 \\
&\triangleq \arg\min_{x \in \boldsymbol{G}} \max_{y \in \boldsymbol{G}} \mathsf{d}(x,y)\,300\mathsf{P}(y) && \text{by Lemma 1.7 (page 19)} \\
&= \arg\min_{x \in \boldsymbol{G}} \max_{y \in \boldsymbol{G}} 300 \left\{\begin{array}{ccccc}
\mathsf{d}(\text{⚀},\text{⚀})\mathsf{P}(\text{⚀}) & \mathsf{d}(\text{⚀},\text{⚁})\mathsf{P}(\text{⚁}) & \mathsf{d}(\text{⚀},\text{⚂})\mathsf{P}(\text{⚂}) & \cdots & \mathsf{d}(\text{⚀},\text{⚅})\mathsf{P}(\text{⚅}) \\
\mathsf{d}(\text{⚁},\text{⚀})\mathsf{P}(\text{⚀}) & \mathsf{d}(\text{⚁},\text{⚁})\mathsf{P}(\text{⚁}) & \mathsf{d}(\text{⚁},\text{⚂})\mathsf{P}(\text{⚂}) & \cdots & \mathsf{d}(\text{⚁},\text{⚅})\mathsf{P}(\text{⚅}) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\mathsf{d}(\text{⚅},\text{⚀})\mathsf{P}(\text{⚀}) & \mathsf{d}(\text{⚅},\text{⚁})\mathsf{P}(\text{⚁}) & \mathsf{d}(\text{⚅},\text{⚂})\mathsf{P}(\text{⚂}) & \cdots & \mathsf{d}(\text{⚅},\text{⚅})\mathsf{P}(\text{⚅})
\end{array}\right\} \\
&= \arg\min_{x \in \boldsymbol{G}} \max_{y \in \boldsymbol{G}} \left\{\begin{array}{cccccc}
0\times30 & 1\times15 & 1\times10 & 1\times180 & 1\times6 & 2\times10 \\
1\times30 & 0\times15 & 1\times10 & 1\times180 & 2\times6 & 1\times10 \\
1\times30 & 1\times15 & 0\times10 & 2\times180 & 1\times6 & 1\times10 \\
1\times30 & 1\times15 & 2\times10 & 0\times180 & 1\times6 & 1\times10 \\
1\times30 & 2\times15 & 1\times10 & 1\times180 & 0\times6 & 1\times10 \\
2\times30 & 1\times15 & 1\times10 & 1\times180 & 1\times6 & 0\times10
\end{array}\right\} = \arg\min_{x \in \boldsymbol{G}} \left\{\begin{array}{c} 180 \\ 180 \\ 360 \\ 30 \\ 180 \\ 180 \end{array}\right\} = \left\{\text{⚃}\right\} \\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &\triangleq \arg\min_{x \in \boldsymbol{G}} \sum_{y \in \boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)} \\
&= \arg\min_{x \in \boldsymbol{G}} \frac{1}{300} \sum_{y \in \boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y)300 && \text{by definition of } \boldsymbol{G} \\
&= \arg\min_{x \in \boldsymbol{G}} \sum_{y \in \boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y)300 && \text{by Lemma 1.2 (page 16)} \\
&= \arg\min_{x \in \boldsymbol{G}} \left\{\begin{array}{ccccccccccc}
0 & + & 15 & + & 10 & + & 180 & + & 6 & + & 20 \\
30 & + & 0 & + & 10 & + & 180 & + & 12 & + & 10 \\
30 & + & 15 & + & 0 & + & 360 & + & 6 & + & 10 \\
30 & + & 15 & + & 20 & + & 0 & + & 6 & + & 10 \\
30 & + & 30 & + & 10 & + & 180 & + & 0 & + & 10 \\
60 & + & 15 & + & 10 & + & 180 & + & 6 & + & 0
\end{array}\right\} = \arg\min_{x \in \boldsymbol{G}} \left\{\begin{array}{c} 231 \\ 242 \\ 421 \\ 81 \\ 260 \\ 271 \end{array}\right\} = \left\{\text{⚃}\right\} \\
\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) &\triangleq \arg\min_{x \in \boldsymbol{G}} \prod_{y \in \boldsymbol{G} \setminus \{x\}} \left[\mathsf{d}(x,y)^{\mathsf{P}(y)}\right] && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{g}} \text{ (Definition 2.4 page 30)} \\
&= \arg\min_{x \in \boldsymbol{G}} \prod_{y \in \boldsymbol{G} \setminus \{x\}} \left[\mathsf{d}(x,y)^{300\mathsf{P}(y)\frac{1}{300}}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \arg\min_{x\in\boldsymbol{G}} \left(\prod_{y\in\boldsymbol{G}\setminus\{x\}} \left[\mathsf{d}(x,y)^{300\mathsf{P}(y)}\right] \right)^{\frac{1}{300}} \\
&= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \left[\mathsf{d}(x,y)^{300\mathsf{P}(y)}\right] && \text{by Lemma 1.2 (page 16)} \\
&= \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccccc}
 & \times & 1^{15} & \times & 1^{10} & \times & 1^{180} & \times & 1^{6} & \times & 2^{10} \\
1^{30} & \times & & \times & 1^{10} & \times & 1^{180} & \times & 2^{6} & \times & 1^{10} \\
1^{30} & \times & 1^{15} & \times & & \times & 2^{180} & \times & 1^{6} & \times & 1^{10} \\
1^{30} & \times & 1^{15} & \times & 2^{10} & \times & & \times & 1^{6} & \times & 1^{10} \\
1^{30} & \times & 2^{15} & \times & 1^{10} & \times & 1^{180} & \times & & \times & 1^{10} \\
2^{30} & \times & 1^{15} & \times & 1^{10} & \times & 1^{180} & \times & 1^{6} & &
\end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 2^{10} \\ 2^{6} \\ 2^{180} \\ 2^{10} \\ 2^{15} \\ 2^{30} \end{array} \right\} = \left\{ \begin{array}{c} \text{⚁} \\ \\ \\ \\ \\ \\ \end{array} \right\} \\
\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \left(\sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y) \right)^{-1} && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{h}} \text{ (Definition 2.4 page 30)} \\
&= \arg\max_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y) && \text{by Lemma 1.4 page 17} \\
&= \arg\max_{x\in\boldsymbol{G}} \frac{1}{300} \sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)} 300\mathsf{P}(y) \\
&= \arg\max_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{300\mathsf{P}(y)}{\mathsf{d}(x,y)} && \text{by Lemma 1.2 (page 16)} \\
&= \arg\max_{x\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccccc}
 & + & \frac{15}{1} & + & \frac{10}{1} & + & \frac{180}{1} & + & \frac{6}{1} & + & \frac{10}{2} \\
\frac{30}{1} & + & & + & \frac{10}{1} & + & \frac{180}{1} & + & \frac{6}{2} & + & \frac{10}{1} \\
\frac{30}{1} & + & \frac{15}{1} & + & & + & \frac{180}{2} & + & \frac{6}{1} & + & \frac{10}{1} \\
\frac{30}{1} & + & \frac{15}{1} & + & \frac{10}{2} & + & & + & \frac{6}{1} & + & \frac{10}{1} \\
\frac{30}{1} & + & \frac{15}{2} & + & \frac{10}{1} & + & \frac{180}{1} & + & & + & \frac{10}{1} \\
\frac{30}{2} & + & \frac{15}{1} & + & \frac{10}{1} & + & \frac{180}{1} & + & \frac{6}{1} & + &
\end{array} \right\} = \arg\max_{x\in\boldsymbol{G}} \left\{ \begin{array}{r} 216.0 \\ 233.0 \\ 151.0 \\ 66.0 \\ 237.5 \\ 226.0 \end{array} \right\} = \left\{ \begin{array}{c} \\ \\ \\ \\ \text{⚄} \\ \\ \end{array} \right\} \\
\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{m}} \text{ (Definition 2.4 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \left\{ \begin{array}{rrrrrr}
0 & 15 & 10 & 180 & 6 & 20 \\
30 & 0 & 10 & 180 & 12 & 10 \\
30 & 15 & 0 & 360 & 6 & 10 \\
30 & 15 & 20 & 0 & 6 & 10 \\
30 & 30 & 10 & 180 & 0 & 10 \\
60 & 15 & 10 & 180 & 6 & 0
\end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 6 \\ 10 \\ 6 \\ 6 \\ 10 \\ 6 \end{array} \right\} = \left\{ \begin{array}{c} \text{⚀} \\ \\ \text{⚂} \\ \text{⚃} \\ \\ \text{⚅} \end{array} \right\} \\
\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &\triangleq \arg\max_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{M}} \text{ (Definition 2.4 page 30)} \\
&= \arg\max_{x\in\boldsymbol{G}} \{6, 10, 6, 6, 10, 6\} && \text{by } \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \text{ result} \\
&= \{\text{⚁}, \text{⚄}\} && \text{by definition of } \boldsymbol{G} \\
\dot{\mathsf{Var}}(\boldsymbol{G}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}(\boldsymbol{G}), x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\text{⚃}, x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}(\boldsymbol{G}) \text{ result} \\
&= 1^2 \times \frac{1}{10} + 1^2 \times \frac{1}{20} + 2^2 \times \frac{1}{30} + 0^2 \times \frac{3}{5} + 1^2 \times \frac{1}{50} + 1^2 \times \frac{1}{30} \\
&= \frac{1}{300}(30 + 15 + 40 + 0 + 6 + 10) = \frac{101}{300} \approx 0.337 \\
\dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{g}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}), x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\text{⚁}, x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) \text{ result}
\end{aligned}
$$

$$\begin{aligned}
&= 1^2 \times \frac{1}{10} + 0^2 \times \frac{1}{20} + 1^2 \times \frac{1}{30} + 1^2 \times \frac{3}{5} + 2^2 \times \frac{1}{50} + 1^2 \times \frac{1}{30} \\
&= \frac{1}{300}(30 + 0 + 10 + 180 + 24 + 10) = \frac{254}{300} = \frac{127}{150} \approx 0.847 \\
\dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{h}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\big(\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}), x\big)\,\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\text{⚄}, x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) \text{ result} \\
&= 1^2 \times \frac{1}{10} + 2^2 \times \frac{1}{20} + 1^2 \times \frac{1}{30} + 1^2 \times \frac{3}{5} + 0^2 \times \frac{1}{50} + 1^2 \times \frac{1}{30} \\
&= \frac{1}{300}(30 + 60 + 10 + 180 + 0 + 10) = \frac{290}{300} = \frac{145}{150} \approx 0.967 \\
\dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{m}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\big(\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}), x\big)\,\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\text{⚀}, \text{⚂}, \text{⚃}, \text{⚅}, x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \text{ result} \\
&= 0^2 \times \frac{1}{10} + 1^2 \times \frac{1}{20} + 0^2 \times \frac{1}{30} + 0^2 \times \frac{3}{5} + 1^2 \times \frac{1}{50} + 0^2 \times \frac{1}{30} \\
&= \frac{1}{300}(0 + 60 + 0 + 0 + 6 + 0) = \frac{66}{300} = \frac{11}{50} = 0.22 \\
\dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{M}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\big(\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}), x\big)\,\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\text{⚁}, \text{⚄}, x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) \text{ result} \\
&= 1^2 \times \frac{1}{10} + 0^2 \times \frac{1}{20} + 1^2 \times \frac{1}{30} + 1^2 \times \frac{3}{5} + 0^2 \times \frac{1}{50} + 1^2 \times \frac{1}{30} \\
&= \frac{1}{300}(30 + 0 + 10 + 180 + 0 + 10) = \frac{230}{300} = \frac{11}{50} \approx 0.767
\end{aligned}$$

⇰

Example 2.4 (board game spinner outcome subspace). The six value *spinner outcome subspace* (Definition 2.10 page 31) has the following geometric values:

$$\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &= \{1, 2, 3, 4, 5, 6\} \\
\dot{\mathsf{Var}}(\boldsymbol{G}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{a}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{g}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{h}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{m}}) = \dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{M}}) &= 0
\end{aligned}$$

✎PROOF:

$$\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x, y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x, y)\,\frac{1}{6} && \text{by definition of } \boldsymbol{G} \\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x, y) && \text{because } \mathsf{f}(x) = \frac{1}{6}x \text{ is } \textit{strictly isotone} \text{ and by Lemma 1.7 (page 19)} \\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \left\{\begin{matrix} \mathsf{d}(1,1) & \cdots & \mathsf{d}(1,6) \\ \mathsf{d}(2,1) & \cdots & \mathsf{d}(2,6) \\ \mathsf{d}(3,1) & \cdots & \mathsf{d}(3,6) \\ \mathsf{d}(4,1) & \cdots & \mathsf{d}(4,6) \\ \mathsf{d}(5,1) & \cdots & \mathsf{d}(5,6) \\ \mathsf{d}(6,1) & \cdots & \mathsf{d}(6,6) \end{matrix}\right\} = \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \left\{\begin{matrix} 0 & 1 & 2 & 3 & 2 & 1 \\ 1 & 0 & 1 & 2 & 3 & 2 \\ 2 & 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 1 & 0 & 1 & 2 \\ 2 & 3 & 2 & 1 & 0 & 1 \\ 1 & 2 & 3 & 2 & 1 & 0 \end{matrix}\right\} = \arg\min_{x\in\boldsymbol{G}} \left\{\begin{matrix} 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{matrix}\right\} \\
&= \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\} && \text{by definition of } \boldsymbol{G} \\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x, y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x, y)\,\frac{1}{6} && \text{by definition of } \boldsymbol{G}
\end{aligned}$$

$$= \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y) \qquad \text{because } \mathsf{f}(x) = \frac{1}{6}x \text{ is \textit{strictly isotone} and by Lemma 1.2 (page 16)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccccc} 0 & + & 1 & + & 2 & + & 3 & + & 2 & + & 1 \\ 1 & + & 0 & + & 1 & + & 2 & + & 3 & + & 2 \\ 2 & + & 1 & + & 0 & + & 1 & + & 2 & + & 3 \\ 3 & + & 2 & + & 1 & + & 0 & + & 1 & + & 2 \\ 2 & + & 3 & + & 2 & + & 1 & + & 0 & + & 1 \\ 1 & + & 2 & + & 3 & + & 2 & + & 1 & + & 0 \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 9 \\ 9 \\ 9 \\ 9 \\ 9 \\ 9 \end{array} \right\} = \left\{ \begin{array}{c} \text{⚀} \\ \text{⚁} \\ \text{⚂} \\ \text{⚃} \\ \text{⚄} \\ \text{⚅} \end{array} \right\}$$

$$\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \left[\mathsf{d}(x,y)^{\mathsf{P}(y)}\right] \qquad \text{by definition of } \dot{\mathsf{C}}_{\mathsf{g}} \text{ (Definition 2.4 page 30)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \left[\mathsf{d}(x,y)^{\frac{1}{6}}\right] \qquad \text{by definition of } \boldsymbol{G}$$

$$= \arg\min_{x\in\boldsymbol{G}} \left(\prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y) \right)^{\frac{1}{6}}$$

$$= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y) \qquad \text{by Lemma 1.2 (page 16)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccccc} & \times & 1 & \times & 2 & \times & 3 & \times & 2 & \times & 1 \\ 1 & \times & & \times & 1 & \times & 2 & \times & 3 & \times & 2 \\ 2 & \times & 1 & \times & & \times & 1 & \times & 2 & \times & 3 \\ 3 & \times & 2 & \times & 1 & \times & & \times & 1 & \times & 2 \\ 2 & \times & 3 & \times & 2 & \times & 1 & \times & & \times & 1 \\ 1 & \times & 2 & \times & 3 & \times & 2 & \times & 1 & \times & \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 12 \\ 12 \\ 12 \\ 12 \\ 12 \\ 12 \end{array} \right\} = \left\{ \begin{array}{c} \text{⚀} \\ \text{⚁} \\ \text{⚂} \\ \text{⚃} \\ \text{⚄} \\ \text{⚅} \end{array} \right\}$$

$$\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\boldsymbol{G}} \left(\sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)} \mathsf{P}(y) \right)^{-1} \qquad \text{by definition of } \dot{\mathsf{C}}_{\mathsf{h}} \text{ (Definition 2.4 page 30)}$$

$$= \arg\max_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)} \mathsf{P}(y) \qquad \text{because } \phi(x) \triangleq x^{-1} \text{ is \textit{strictly antitone} and by Lemma 1.5 page 18}$$

$$= \arg\max_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)} \frac{1}{6}$$

$$= \arg\max_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)} \qquad \text{by Lemma 1.2 (page 16)}$$

$$= \arg\max_{x\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccccc} & + & \frac{1}{1} & + & \frac{1}{2} & + & \frac{1}{3} & + & \frac{1}{2} & + & \frac{1}{1} \\ \frac{1}{1} & + & & + & \frac{1}{1} & + & \frac{1}{2} & + & \frac{1}{3} & + & \frac{1}{2} \\ \frac{1}{2} & + & \frac{1}{1} & + & & + & \frac{1}{1} & + & \frac{1}{2} & + & \frac{1}{3} \\ \frac{1}{3} & + & \frac{1}{2} & + & \frac{1}{1} & + & & + & \frac{1}{1} & + & \frac{1}{2} \\ \frac{1}{2} & + & \frac{1}{3} & + & \frac{1}{2} & + & \frac{1}{1} & + & & + & \frac{1}{1} \\ \frac{1}{1} & + & \frac{1}{2} & + & \frac{1}{3} & + & \frac{1}{2} & + & \frac{1}{1} & + & \end{array} \right\} = \arg\max_{x\in\boldsymbol{G}} \frac{1}{6} \left\{ \begin{array}{c} 20 \\ 20 \\ 20 \\ 20 \\ 20 \\ 20 \end{array} \right\} = \left\{ \begin{array}{c} \text{⚀} \\ \text{⚁} \\ \text{⚂} \\ \text{⚃} \\ \text{⚄} \\ \text{⚅} \end{array} \right\}$$

$$\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \dot{\mathsf{C}}_{\mathsf{m}} \text{ (Definition 2.4 page 30)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\frac{1}{6}$$

$$= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y) \qquad \text{by Lemma 1.7 (page 19)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \{1, 1, 1, 1, 1, 1\}$$

$$= \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\} \qquad \text{by definition of } \boldsymbol{G}$$

$$\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) \triangleq \arg\max_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \dot{\mathsf{C}}_{\mathsf{M}} \text{ (Definition 2.4 page 30)}$$

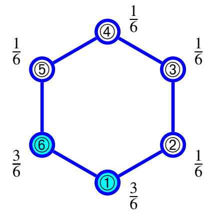
$$
\begin{aligned}
&= \arg\max_{x\in \boldsymbol{G}} \{1,1,1,1,1,1\} && \text{by } \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \text{ result}\\
&= \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} && \text{by definition of } \boldsymbol{G}
\end{aligned}
$$

$$
\begin{aligned}
\dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{a}}) &= \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{g}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{h}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) = \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{M}}) = \dot{\mathsf{Var}}(\boldsymbol{G})\\
&\triangleq \sum_{x\in\boldsymbol{G}} \left[\mathsf{d}\left(\dot{\mathsf{C}}(\boldsymbol{G}), x\right)\right]^2 \mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} (0^2)\frac{1}{6} && \text{because } \dot{\mathsf{C}}(\boldsymbol{G}) = \boldsymbol{G}\\
&= 0 && \text{by field property of } \textit{additive identity element } 0
\end{aligned}
$$

✏

Example 2.5 (weighted spinner outcome subspace).

The six value *weighted spinner outcome subspace* $\boldsymbol{G}$ (Definition 2.1 page 29) illustrated to the right has the following geometric values:



$$
\begin{array}{rcl@{\qquad}lcll}
\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) &=& \{1,6\} & \dot{\mathsf{Var}}(\boldsymbol{G}) &=& \frac{5}{3} \approx 1.667\\
\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) &=& \{2,5\} & \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{h}}) &=& \frac{4}{3} \approx 1.333\\
\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &=& \{1,2,3,4,5,6\} & \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) &=& 0 = 0
\end{array}
$$

The *outcome center* result is used later in Example 2.16 (page 60). Note that, unlike the *weighted real die outcome subspace* (Example 2.3 page 35), of the center measures of cardinality 2 or less, the *harmonic center* $\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G})$ yields the lowest *outcome variance* (Definition 2.5 page 30). This is surprising since it suggests that $\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G})$ is superior to all the other *center measures* (Definition 2.3 page 30, Definition 2.4 page 30), but yet unlike the other center measures, it yields center values that are not maximally likely.

✎PROOF:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \frac{1}{10}\left\{\begin{array}{cccccc}
0\times3 & 1\times1 & 2\times1 & 3\times1 & 2\times1 & 1\times3\\
1\times3 & 0\times1 & 1\times1 & 2\times1 & 3\times1 & 2\times3\\
2\times3 & 1\times1 & 0\times1 & 1\times1 & 2\times1 & 3\times3\\
3\times3 & 2\times1 & 1\times1 & 0\times1 & 1\times1 & 2\times3\\
2\times3 & 3\times1 & 2\times1 & 1\times1 & 0\times1 & 1\times3\\
1\times3 & 2\times1 & 3\times1 & 2\times1 & 1\times1 & 0\times3
\end{array}\right\}
= \arg\min_{x\in\boldsymbol{G}} \frac{1}{10}\left\{\begin{array}{c}3\\6\\9\\9\\6\\3\end{array}\right\}
= \left\{\begin{array}{c}1\\ \\ \\ \\ \\6\end{array}\right\}\\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \frac{1}{10}\left\{\begin{array}{c}
0\times3+1\times1+2\times1+3\times1+2\times1+1\times3\\
1\times3+0\times1+1\times1+2\times1+3\times1+2\times3\\
2\times3+1\times1+0\times1+1\times1+2\times1+3\times3\\
3\times3+2\times1+1\times1+0\times1+1\times1+2\times3\\
2\times3+3\times1+2\times1+1\times1+0\times1+1\times3\\
1\times3+2\times1+3\times1+2\times1+1\times1+0\times3
\end{array}\right\}
= \arg\min_{x\in\boldsymbol{G}} \frac{1}{10}\left\{\begin{array}{c}11\\15\\19\\19\\15\\11\end{array}\right\}
= \left\{\begin{array}{c}1\\ \\ \\ \\ \\6\end{array}\right\}\\
\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \left[\mathsf{d}(x,y)^{\mathsf{P}(y)}\right] && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{g}} \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \left[\mathsf{d}(x,y)^{6\mathsf{P}(y)\frac{1}{6}}\right]\\
&= \arg\min_{x\in\boldsymbol{G}} \left(\prod_{y\in\boldsymbol{G}\setminus\{x\}} \left[\mathsf{d}(x,y)^{6\mathsf{P}(y)}\right]\right)^{\frac{1}{6}}\\
&= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \left[\mathsf{d}(x,y)^{6\mathsf{P}(y)}\right] && \text{by Lemma 1.2 (page 16)}
\end{aligned}
$$

$$= \arg\min_{x\in\mathbf{G}} \left\{ \begin{array}{ccccccccccc} & \times & 1^1 & \times & 2^1 & \times & 3^1 & \times & 2^1 & \times & 1^3 \\ 1^3 & \times & & \times & 1^1 & \times & 2^1 & \times & 3^1 & \times & 2^3 \\ 2^3 & \times & 1^1 & \times & & \times & 1^1 & \times & 2^1 & \times & 3^3 \\ 3^3 & \times & 2^1 & \times & 1^1 & \times & & \times & 1^1 & \times & 2^3 \\ 2^3 & \times & 3^1 & \times & 2^1 & \times & 1^1 & \times & & \times & 1^3 \\ 1^3 & \times & 2^1 & \times & 3^1 & \times & 2^1 & \times & 1^1 & & \end{array} \right\} = \arg\min_{x\in\mathbf{G}} \left\{ \begin{array}{c} 2^2\times 3^1 \\ 2^4\times 3^1 \\ 2^4\times 3^3 \\ 2^4\times 3^3 \\ 2^4\times 3^1 \\ 2^2\times 3^1 \end{array} \right\} = \left\{ \begin{array}{c} 1 \\ \\ \\ \\ 6 \end{array} \right\}$$

$$\dot{\mathsf{C}}_{\mathsf{h}}(\mathbf{G}) \triangleq \arg\min_{x\in\mathbf{G}} \left(\sum_{y\in\mathbf{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)} \mathsf{P}(y) \right)^{-1}$$ by definition of $\dot{\mathsf{C}}_{\mathsf{h}}$ (Definition 2.4 page 30)

$$= \arg\max_{x\in\mathbf{G}} \sum_{y\in\mathbf{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)} \mathsf{P}(y)$$ because $\phi(x) \triangleq x^{-1}$ is *strictly antitone* and by Lemma 1.5 page 18

$$= \arg\max_{x\in\mathbf{G}} \sum_{y\in\mathbf{G}\setminus\{x\}} \frac{1}{\mathsf{d}(x,y)} 6\mathsf{P}(y)\frac{1}{6}$$

$$= \arg\max_{x\in\mathbf{G}} \sum_{y\in\mathbf{G}\setminus\{x\}} \frac{6\mathsf{P}(y)}{\mathsf{d}(x,y)}$$ by Lemma 1.2 (page 16)

$$= \arg\max_{x\in\mathbf{G}} \left\{ \begin{array}{ccccccccccc} & + & \frac{1}{1} & + & \frac{1}{2} & + & \frac{1}{3} & + & \frac{1}{2} & + & \frac{3}{1} \\ \frac{3}{1} & + & & + & \frac{1}{1} & + & \frac{1}{2} & + & \frac{1}{3} & + & \frac{3}{2} \\ \frac{3}{2} & + & \frac{1}{1} & + & & + & \frac{1}{1} & + & \frac{1}{2} & + & \frac{3}{3} \\ \frac{3}{3} & + & \frac{1}{2} & + & \frac{1}{1} & + & & + & \frac{1}{1} & + & \frac{3}{2} \\ \frac{3}{2} & + & \frac{1}{3} & + & \frac{1}{2} & + & \frac{1}{1} & + & & + & \frac{3}{1} \\ \frac{3}{1} & + & \frac{1}{2} & + & \frac{1}{3} & + & \frac{1}{2} & + & \frac{1}{1} & + & \end{array} \right\} = \arg\max_{x\in\mathbf{G}} \frac{1}{6} \left\{ \begin{array}{c} 32 \\ 38 \\ 30 \\ 30 \\ 38 \\ 32 \end{array} \right\} = \left\{ \begin{array}{c} 2 \\ \\ 5 \end{array} \right\}$$

$$\dot{\mathsf{C}}_{\mathsf{m}}(\mathbf{G}) \triangleq \arg\min_{x\in\mathbf{G}} \min_{y\in\mathbf{G}} \mathsf{d}(x,y)\,\mathsf{P}(y)$$ by definition of $\dot{\mathsf{C}}_{\mathsf{m}}$ (Definition 2.4 page 30)

$$= \arg\min_{x\in\mathbf{G}} \min_{y\in\mathbf{G}\setminus\{x\}} \left\{ \begin{array}{cccccc} & 1 & 2 & 3 & 2 & 1 \\ 3 & & 1 & 2 & 3 & 2 \\ 6 & 1 & & 1 & 2 & 3 \\ 9 & 2 & 1 & & 1 & 2 \\ 6 & 3 & 2 & 1 & & 1 \\ 3 & 2 & 3 & 2 & 1 & \end{array} \right\} = \arg\min_{x\in\mathbf{G}} \left\{ \begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{array} \right\} = \left\{ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \right\}$$

$$\dot{\mathsf{C}}_{\mathsf{M}}(\mathbf{G}) \triangleq \arg\max_{x\in\mathbf{G}} \min_{y\in\mathbf{G}} \mathsf{d}(x,y)\,\mathsf{P}(y)$$ by definition of $\dot{\mathsf{C}}_{\mathsf{M}}$ (Definition 2.4 page 30)

$$= \arg\max_{x\in\mathbf{G}} \{1,1,1,1,1,1\}$$ by $\dot{\mathsf{C}}_{\mathsf{m}}(\mathbf{G})$ result

$$= \{1,2,3,4,5,6\}$$

$$\dot{\mathsf{Var}}(\mathbf{G};\dot{\mathsf{C}}_{\mathsf{a}}) = \dot{\mathsf{Var}}(\mathbf{G};\dot{\mathsf{C}}_{\mathsf{g}}) = \dot{\mathsf{Var}}(\mathbf{G})$$

$$\triangleq \sum_{x\in\mathbf{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}(\mathbf{G}),x\right)\mathsf{P}(x)$$ by definition of $\dot{\mathsf{Var}}$ (Definition 2.5 page 30)

$$= \sum_{x\in\mathbf{G}} \mathsf{d}^2(\{1,6\},x)\,\mathsf{P}(x)$$ by $\dot{\mathsf{C}}(\mathbf{G})$ result

$$= (0)^2\frac{3}{6} + (1)^2\frac{1}{6} + (2)^2\frac{1}{6} + (2)^2\frac{1}{6} + (1)^2\frac{1}{6} + (0)^2\frac{3}{6}$$

$$= \frac{10}{6} = \frac{5}{3} = 1\frac{2}{3} \approx 1.667$$

$$\dot{\mathsf{Var}}(\mathbf{G};\dot{\mathsf{C}}_{\mathsf{h}}) \triangleq \sum_{x\in\mathbf{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{h}}(\mathbf{G}),x\right)\mathsf{P}(x)$$ by definition of $\dot{\mathsf{Var}}$ (Definition 2.5 page 30)

$$= \sum_{x\in\mathbf{G}} \mathsf{d}^2(\{2,5\},x)\,\mathsf{P}(x)$$ by $\dot{\mathsf{C}}_{\mathsf{h}}(\mathbf{G})$ result

$$= (1)^2\frac{3}{6} + (0)^2\frac{1}{6} + (1)^2\frac{1}{6} + (1)^2\frac{1}{6} + (0)^2\frac{1}{6} + (1)^2\frac{3}{6}$$

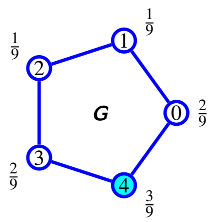
$$= \frac{8}{6} = \frac{4}{3} \approx 1.333$$

$$
\begin{aligned}
\dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{M}}) &= \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) \\
&\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}), x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{1,2,3,4,5,6\}, x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) \text{ result} \\
&= \sum_{x\in\boldsymbol{G}} 0^2\mathsf{P}(x) = 0
\end{aligned}
$$

✏

Example 2.6 (weighted ring). The weighted five element ring illustrated to the right has the geometric values below. The *outcome center* result is used later in Example 2.17 (page 61).

$$
\begin{array}{rclrcll}
\dot{\mathsf{C}}(\boldsymbol{G}) &=& \{4\} & \dot{\mathsf{Var}}(\boldsymbol{G}) &=& \frac{11}{9} &\approx 1.222 \\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &=& \{3,4\} & \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{a}}) &=& \frac{7}{9} &\approx 0.778 \\
\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) &=& \{3\} & \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{g}}) &=& \frac{16}{9} &\approx 1.778 \\
\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) &=& \{1,2,3\} & \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{h}}) &=& \frac{5}{9} &\approx 0.556 \\
\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) &=& \{0,3,4\} & \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) &=& \frac{2}{9} &\approx 0.222 \\
\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &=& \{1,2\} & \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{M}}) &=& \frac{16}{9} &\approx 1.778
\end{array}.
$$



✎PROOF:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)} \\
&= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \frac{1}{9}\mathsf{d}(x,y)\,\mathsf{P}(y)9 \\
&= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y)9 && \text{because } \mathsf{f}(x) = \frac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.7 (page 19)} \\
&= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \left\{\begin{array}{ccccc}
\mathsf{d}(0,0)\mathsf{P}(0)9 & \mathsf{d}(0,1)\mathsf{P}(1)9 & \mathsf{d}(0,2)\mathsf{P}(2)9 & \mathsf{d}(0,3)\mathsf{P}(3)9 & \mathsf{d}(0,4)\mathsf{P}(4)9 \\
\mathsf{d}(1,0)\mathsf{P}(0)9 & \mathsf{d}(1,1)\mathsf{P}(1)9 & \mathsf{d}(1,2)\mathsf{P}(2)9 & \mathsf{d}(1,3)\mathsf{P}(3)9 & \mathsf{d}(1,4)\mathsf{P}(4)9 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\mathsf{d}(4,0)\mathsf{P}(0)9 & \mathsf{d}(4,1)\mathsf{P}(1)9 & \mathsf{d}(4,2)\mathsf{P}(2)9 & \mathsf{d}(4,3)\mathsf{P}(3)9 & \mathsf{d}(4,4)\mathsf{P}(4)9
\end{array}\right\} \\
&= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \left\{\begin{array}{ccccc}
0\times2 & 1\times1 & 2\times1 & 2\times2 & 1\times3 \\
1\times2 & 0\times1 & 1\times1 & 2\times2 & 2\times3 \\
2\times2 & 1\times1 & 0\times1 & 1\times2 & 2\times3 \\
2\times2 & 2\times1 & 1\times1 & 0\times2 & 1\times3 \\
1\times2 & 2\times1 & 2\times1 & 1\times2 & 0\times3
\end{array}\right\}
= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \left\{\begin{array}{c} 4\\6\\6\\4\\2 \end{array}\right\}
= \left\{\begin{array}{c} \\ \\ \\ \\ 4 \end{array}\right\} \\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &\triangleq \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)} \\
&= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \frac{1}{9}\mathsf{d}(x,y)\,\mathsf{P}(y)9 \\
&= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y)9 && \text{because } \mathsf{f}(x) = \frac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.2 (page 16)} \\
&= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \left\{\begin{array}{ccccccccc}
0\times2 & + & 2\times1 & + & 1\times1 & + & 1\times2 & + & 2\times3 \\
2\times2 & + & 0\times1 & + & 2\times1 & + & 1\times2 & + & 1\times3 \\
1\times2 & + & 2\times1 & + & 0\times1 & + & 2\times2 & + & 1\times3 \\
1\times2 & + & 1\times1 & + & 2\times1 & + & 0\times2 & + & 1\times3 \\
2\times2 & + & 1\times1 & + & 1\times1 & + & 1\times2 & + & 0\times3
\end{array}\right\}
= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \left\{\begin{array}{c} 11\\11\\11\\8\\8 \end{array}\right\}
= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \left\{\begin{array}{c} \\ \\ \\ 3\\4 \end{array}\right\} \\
\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) &\triangleq \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{\mathsf{P}(y)} && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{g}} \text{ (Definition 2.4 page 30)} \\
&= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{9\mathsf{P}(y)\frac{1}{9}} \\
&= \operatorname*{arg\,min}_{x\in\boldsymbol{G}} \left[\prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{9\mathsf{P}(y)}\right]^{\frac{1}{9}}
\end{aligned}
$$

$$= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{9\mathsf{P}(y)} \qquad \text{because } \mathsf{f}(x) \triangleq x^{\frac{1}{9}} \text{ is \textit{strictly isotone} and by Lemma 1.2 (page 16)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccc} & & 2^1 & \times & 1^1 & \times & 1^2 & \times & 2^3 \\ 2^2 & \times & & & 2^1 & \times & 1^2 & \times & 1^3 \\ 1^2 & \times & 2^1 & \times & & & 2^2 & \times & 1^3 \\ 1^2 & \times & 1^1 & \times & 2^1 & \times & & & 1^3 \\ 2^2 & \times & 1^1 & \times & 1^1 & \times & 1^2 & & \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 16 \\ 8 \\ 8 \\ 2 \\ 4 \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} \\ \\ 3 \\ \\ \\ \end{array} \right\}$$

$$\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\Omega} \left(\sum_{y\in\Omega} \frac{1}{\mathsf{d}(x,y)} \mathsf{P}(y) \right)^{-1} \qquad \text{by definition of } \dot{\mathsf{C}}_{\mathsf{h}} \text{ (Definition 2.4 page 30)}$$

$$= \arg\max_{x\in\Omega} \left(\sum_{y\in\Omega} \frac{1}{\mathsf{d}(x,y)} \mathsf{P}(y) \right) \qquad \text{because } \phi(x) \triangleq x^{-1} \text{ is \textit{strictly antitone} and by Lemma 1.5 page 18}$$

$$= \arg\max_{x\in\Omega} \left(\frac{1}{9} \sum_{y\in\Omega} \frac{1}{\mathsf{d}(x,y)} \mathsf{P}(y) 9 \right)$$

$$= \arg\max_{x\in\Omega} \sum_{y\in\Omega} \frac{9\mathsf{P}(y)}{\mathsf{d}(x,y)} \qquad \text{because } \mathsf{f}(x) = \frac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.2 (page 16)}$$

$$= \arg\max_{x\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccc} 0 & + & \frac{1}{2} & + & \frac{1}{1} & + & \frac{2}{1} & + & \frac{3}{2} \\ \frac{2}{2} & + & 0 & + & \frac{1}{2} & + & \frac{2}{1} & + & \frac{3}{1} \\ \frac{2}{1} & + & \frac{1}{2} & + & 0 & + & \frac{2}{2} & + & \frac{3}{1} \\ \frac{2}{1} & + & \frac{1}{1} & + & \frac{1}{2} & + & 0 & + & \frac{3}{1} \\ \frac{2}{2} & + & \frac{1}{1} & + & \frac{1}{1} & + & \frac{2}{1} & + & 0 \end{array} \right\} = \arg\max_{x\in\boldsymbol{G}} \frac{1}{2} \left\{ \begin{array}{c} 10 \\ 13 \\ 13 \\ 13 \\ 11 \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \\ \end{array} \right\}$$

$$\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \dot{\mathsf{C}}_{\mathsf{m}} \text{ (Definition 2.4 page 30)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{9} \mathsf{d}(x,y)\,\mathsf{P}(y) 9$$

$$= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) 9 \qquad \text{because } \mathsf{f}(x) = \frac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.7 (page 19)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \left\{ \begin{array}{ccccc} & 2\times 1 & 1\times 1 & 1\times 2 & 2\times 3 \\ 2\times 2 & & 2\times 1 & 1\times 2 & 1\times 3 \\ 1\times 2 & 2\times 1 & & 2\times 2 & 1\times 3 \\ 1\times 2 & 1\times 1 & 2\times 1 & & 1\times 3 \\ 2\times 2 & 1\times 1 & 1\times 1 & 1\times 2 & \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 1 \\ 2 \\ 2 \\ 1 \\ 1 \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 0 \\ \\ \\ 3 \\ 4 \end{array} \right\}$$

$$\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) \triangleq \arg\max_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \dot{\mathsf{C}}_{\mathsf{M}} \text{ (Definition 2.4 page 30)}$$

$$= \arg\max_{x\in\boldsymbol{G}} \{1,2,2,1,1\} \qquad \text{by } \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \text{ result}$$

$$= \{1,2\}$$

$$\dot{\mathsf{Var}}(\boldsymbol{G}) \triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}(\boldsymbol{G}), x\right) \mathsf{P}(x) \qquad \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}$$

$$= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{4\}, x)\,\mathsf{P}(x) \qquad \text{by } \dot{\mathsf{C}}(\boldsymbol{G}) \text{ result}$$

$$= (1)^2\frac{2}{9} + (2)^2\frac{1}{9} + (2)^2\frac{1}{9} + (1)^2\frac{2}{9} + (0)^2\frac{3}{9} = \frac{11}{9} \approx 1.222$$

$$\dot{\mathsf{Var}}(\boldsymbol{G}; \dot{\mathsf{C}}_{\mathsf{a}}) \triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}), x\right) \mathsf{P}(x) \qquad \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}$$

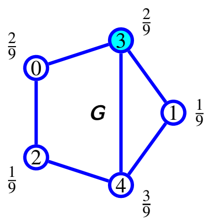
$$= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{3,4\}, x)\,\mathsf{P}(x) \qquad \text{by } \dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) \text{ result}$$

$$= (1)^2\frac{2}{9} + (2)^2\frac{1}{9} + (1)^2\frac{1}{9} + (0)^2\frac{2}{9} + (0)^2\frac{3}{9} = \frac{7}{9} \approx 0.778$$

$$
\begin{aligned}
\dot{\mathsf{V}}\mathsf{ar}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{g}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}),x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{V}}\mathsf{ar} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{3\},x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) \text{ result}\\
&= (2)^2\frac{2}{9} + (2)^2\frac{1}{9} + (1)^2\frac{1}{9} + (0)^2\frac{2}{9} + (1)^2\frac{3}{9} = \frac{16}{9} \approx 1.778\\
\dot{\mathsf{V}}\mathsf{ar}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{h}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}),x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{V}}\mathsf{ar} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{1,2,3\},x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) \text{ result}\\
&= (1)^2\frac{2}{9} + (0)^2\frac{1}{9} + (0)^2\frac{1}{9} + (0)^2\frac{2}{9} + (1)^2\frac{3}{9} = \frac{5}{9} \approx 0.556\\
\dot{\mathsf{V}}\mathsf{ar}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}),x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{V}}\mathsf{ar} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{0,3,4\},x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \text{ result}\\
&= (0)^2\frac{2}{9} + (1)^2\frac{1}{9} + (1)^2\frac{1}{9} + (0)^2\frac{2}{9} + (0)^2\frac{3}{9} = \frac{2}{9} \approx 0.222\\
\dot{\mathsf{V}}\mathsf{ar}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{M}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}),x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{V}}\mathsf{ar} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{1,2\},x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \text{ result}\\
&= (1)^2\frac{2}{9} + (0)^2\frac{1}{9} + (0)^2\frac{1}{9} + (1)^2\frac{2}{9} + (2)^2\frac{3}{9} = \frac{16}{9} \approx 1.778
\end{aligned}
$$

✏

Example 2.7. The weighted five element structure illustrated to the right has the following geometric values:



$$
\begin{array}{rclcrcll}
\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) &=& \{3\} && \dot{\mathsf{V}}\mathsf{ar}(\boldsymbol{G}) &=& \frac{10}{9} & \approx 1.111\\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &=& \{3,4\} && \dot{\mathsf{V}}\mathsf{ar}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{a}}) &=& \frac{4}{9} & \approx 0.444\\
\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) &=& \{1,2,3\} && \dot{\mathsf{V}}\mathsf{ar}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{h}}) &=& \frac{5}{9} & \approx 0.555\ .\\
\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) &=& \{0,3,4\} && \dot{\mathsf{V}}\mathsf{ar}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) &=& \frac{2}{9} & \approx 0.222\\
\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &=& \{1,2\} && \dot{\mathsf{V}}\mathsf{ar}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{M}}) &=& \frac{7}{9} & \approx 0.778
\end{array}
$$

The *outcome center* result is used later in Example 2.18 (page 63). Note that only the operators $\dot{\mathsf{C}}$ and $\dot{\mathsf{C}}_{\mathsf{g}}$ were able to successfully isolate a single center point ($\left|\dot{\mathsf{C}}(\boldsymbol{G})\right| = \left|\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G})\right| = |\{3\}| = 1$).

✎PROOF:

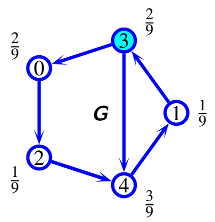
$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \frac{1}{9}\mathsf{d}(x,y)\,\mathsf{P}(y)9\\
&= \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y)9 && \text{because } \mathsf{f}(x) = \frac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.7 (page 19)}\\
&= \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \left\{\begin{array}{ccccc}
\mathsf{d}(0,0)\mathsf{P}(0)9 & \mathsf{d}(0,1)\mathsf{P}(1)9 & \mathsf{d}(0,2)\mathsf{P}(2)9 & \mathsf{d}(0,3)\mathsf{P}(3)9 & \mathsf{d}(0,4)\mathsf{P}(4)9\\
\mathsf{d}(1,0)\mathsf{P}(0)9 & \mathsf{d}(1,1)\mathsf{P}(1)9 & \mathsf{d}(1,2)\mathsf{P}(2)9 & \mathsf{d}(1,3)\mathsf{P}(3)9 & \mathsf{d}(1,4)\mathsf{P}(4)9\\
\vdots & \vdots & \vdots & \vdots & \vdots\\
\mathsf{d}(4,0)\mathsf{P}(0)9 & \mathsf{d}(4,1)\mathsf{P}(1)9 & \mathsf{d}(4,2)\mathsf{P}(2)9 & \mathsf{d}(4,3)\mathsf{P}(3)9 & \mathsf{d}(4,4)\mathsf{P}(4)9
\end{array}\right\}\\
&= \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \left\{\begin{array}{ccccc}
0\times2 & 2\times1 & 1\times1 & 1\times2 & 2\times3\\
2\times2 & 0\times1 & 2\times1 & 1\times2 & 1\times3\\
1\times2 & 2\times1 & 0\times1 & 2\times2 & 1\times3\\
1\times2 & 1\times1 & 2\times1 & 0\times2 & 1\times3\\
2\times2 & 1\times1 & 1\times1 & 1\times2 & 0\times3
\end{array}\right\} = \arg\min_{x\in\boldsymbol{G}}\left\{\begin{array}{c}6\\4\\4\\3\\4\end{array}\right\} = \arg\min_{x\in\boldsymbol{G}}\left\{\begin{array}{c}\\ \\ 3\\ \\ \end{array}\right\}\\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}}\sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)}
\end{aligned}
$$

$$
\begin{aligned}
&= \arg\min_{x\in\boldsymbol{G}} \left\{\begin{array}{ccccccccc}
0\times 2 & + & 2\times 1 & + & 1\times 1 & + & 1\times 2 & + & 2\times 3\\
2\times 2 & + & 0\times 1 & + & 2\times 1 & + & 1\times 2 & + & 1\times 3\\
1\times 2 & + & 2\times 1 & + & 0\times 1 & + & 2\times 2 & + & 1\times 3\\
1\times 2 & + & 1\times 1 & + & 2\times 1 & + & 0\times 2 & + & 1\times 3\\
2\times 2 & + & 1\times 1 & + & 1\times 1 & + & 1\times 2 & + & 0\times 3
\end{array}\right\}
= \arg\min_{x\in\boldsymbol{G}} \left\{\begin{array}{c}11\\11\\11\\8\\8\end{array}\right\}
= \arg\min_{x\in\boldsymbol{G}} \left\{\begin{array}{c}\\ \\ \\3\\4\end{array}\right\}
\\
\dot{C}_g(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{\mathsf{P}(y)} && \text{by definition of } \dot{C}_g \text{ (Definition 2.4 page 30)}
\\
&= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{9\mathsf{P}(y)\frac{1}{9}}
\\
&= \arg\min_{x\in\boldsymbol{G}} \left[\prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{9\mathsf{P}(y)}\right]^{\frac{1}{9}}
\\
&= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{9\mathsf{P}(y)} && \text{because } \mathsf{f}(x)\triangleq x^{\frac{1}{9}} \text{ is \textit{strictly isotone} and by Lemma 1.2 (page 16)}
\\
&= \arg\min_{x\in\boldsymbol{G}} \left\{\begin{array}{ccccccccc}
 & & 2^1 & \times & 1^1 & \times & 1^2 & \times & 2^3\\
2^2 & \times & & & 2^1 & \times & 1^2 & \times & 1^3\\
1^2 & \times & 2^1 & \times & & & 2^2 & \times & 1^3\\
1^2 & \times & 1^1 & \times & 2^1 & \times & & & 1^3\\
2^2 & \times & 1^1 & \times & 1^1 & \times & 1^2 & &
\end{array}\right\}
= \arg\min_{x\in\boldsymbol{G}} \left\{\begin{array}{c}2^4\\2^3\\2^3\\2^1\\2^2\end{array}\right\}
= \arg\min_{x\in\boldsymbol{G}} \left\{\begin{array}{c}\\ \\ \\3\\ \\ \end{array}\right\}
\\
\dot{C}_h(\boldsymbol{G}) &\triangleq \arg\min_{x\in\Omega} \left(\sum_{y\in\Omega} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y)\right)^{-1} && \text{by definition of } \dot{C}_h \text{ (Definition 2.4 page 30)}
\\
&= \arg\max_{x\in\Omega} \left(\sum_{y\in\Omega} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y)\right) && \text{because } \phi(x)\triangleq x^{-1} \text{ is \textit{strictly antitone} and by Lemma 1.5 page 18}
\\
&= \arg\max_{x\in\Omega} \left(\frac{1}{9}\sum_{y\in\Omega} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y)9\right)
\\
&= \arg\max_{x\in\Omega} \sum_{y\in\Omega} \frac{9\mathsf{P}(y)}{\mathsf{d}(x,y)} && \text{because } \mathsf{f}(x) = \frac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.2 (page 16)}
\\
&= \arg\max_{x\in\boldsymbol{G}} \left\{\begin{array}{ccccccccc}
0 & + & \frac{1}{2} & + & \frac{1}{1} & + & \frac{2}{1} & + & \frac{3}{2}\\
\frac{2}{2} & + & 0 & + & \frac{1}{2} & + & \frac{2}{1} & + & \frac{3}{1}\\
\frac{2}{1} & + & \frac{1}{2} & + & 0 & + & \frac{2}{2} & + & \frac{3}{1}\\
\frac{2}{1} & + & \frac{1}{1} & + & \frac{1}{2} & + & 0 & + & \frac{3}{1}\\
\frac{2}{2} & + & \frac{1}{1} & + & \frac{1}{1} & + & \frac{2}{1} & + & 0
\end{array}\right\}
= \arg\max_{x\in\boldsymbol{G}} \frac{1}{2}\left\{\begin{array}{c}10\\13\\13\\13\\10\end{array}\right\}
= \left\{\begin{array}{c}1\\2\\3\end{array}\right\}
\\
\dot{C}_m(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{C}_m \text{ (Definition 2.4 page 30)}
\\
&= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{9}\mathsf{d}(x,y)\,\mathsf{P}(y)9
\\
&= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y)9 && \text{because } \mathsf{f}(x) = \frac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.7 (page 19)}
\\
&= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \left\{\begin{array}{ccccc}
 & 2\times 1 & 1\times 1 & 1\times 2 & 2\times 3\\
2\times 2 & & 2\times 1 & 1\times 2 & 1\times 3\\
1\times 2 & 2\times 1 & & 2\times 2 & 1\times 3\\
1\times 2 & 1\times 1 & 2\times 1 & & 1\times 3\\
2\times 2 & 1\times 1 & 1\times 1 & 1\times 2 &
\end{array}\right\}
= \arg\min_{x\in\boldsymbol{G}} \left\{\begin{array}{c}1\\2\\2\\1\\1\end{array}\right\}
= \left\{\begin{array}{c}0\\ \\ \\3\\4\end{array}\right\}
\\
\dot{C}_M(\boldsymbol{G}) &\triangleq \arg\max_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{C}_M \text{ (Definition 2.4 page 30)}
\\
&= \arg\max_{x\in\boldsymbol{G}} \{1,2,2,1,1\} && \text{by } \dot{C}_m(\boldsymbol{G}) \text{ result}
\\
&= \{1,2\}
\\
\dot{\mathrm{Var}}(\boldsymbol{G}) &= \dot{\mathrm{Var}}(\boldsymbol{G};\dot{C}_g) && \text{by } \dot{C}(\boldsymbol{G}) \text{ and } \dot{C}_g(\boldsymbol{G}) \text{ results}
\end{aligned}
$$

$$
\begin{aligned}
&\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}),x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{3\},x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) \text{ result}\\
&= (1)^2\frac{2}{9}+(1)^2\frac{1}{9}+(2)^2\frac{1}{9}+(0)^2\frac{2}{9}+(1)^2\frac{3}{9} = \frac{10}{9} \approx 1.111\\
\dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{a}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}),x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{3,4\},x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) \text{ result}\\
&= (1)^2\frac{2}{9}+(1)^2\frac{1}{9}+(1)^2\frac{1}{9}+(0)^2\frac{2}{9}+(0)^2\frac{3}{9} = \frac{4}{9} \approx 0.444\\
\dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{h}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}),x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{1,2,3\},x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) \text{ result}\\
&= (1)^2\frac{2}{9}+(0)^2\frac{1}{9}+(0)^2\frac{1}{9}+(0)^2\frac{2}{9}+(1)^2\frac{3}{9} = \frac{5}{9} \approx 0.555\\
\dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}),x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{0,3,4\},x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) \text{ result}\\
&= (0)^2\frac{2}{9}+(1)^2\frac{1}{9}+(1)^2\frac{1}{9}+(0)^2\frac{2}{9}+(0)^2\frac{3}{9} = \frac{2}{9} \approx 0.222\\
\dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{M}}) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}),x\right)\mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{1,2\},x)\,\mathsf{P}(x) && \text{by } \dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) \text{ result}\\
&= (1)^2\frac{2}{9}+(0)^2\frac{1}{9}+(0)^2\frac{1}{9}+(1)^2\frac{2}{9}+(1)^2\frac{3}{9} = \frac{7}{9} \approx 0.778
\end{aligned}
$$

✏

Example 2.8. The *outcome subspace* (Definition 2.1 page 29) illustrated to the right, with a *quasi-metric* (Definition D.6 page 163) has the following geometric values:



$$
\begin{array}{rcl@{\qquad}rcll}
\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{g}}(\boldsymbol{G}) = \dot{\mathsf{C}}_{\mathsf{h}}(\boldsymbol{G}) &=& \{3\} & \dot{\mathsf{Var}}(\boldsymbol{G}) &=& \frac{12}{9} &\approx 1.333\\
\dot{\mathsf{C}}_{\mathsf{m}}(\boldsymbol{G}) &=& \{0,4\} & \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{m}}) &=& \frac{10}{9} &\approx 1.111\\
\dot{\mathsf{C}}_{\mathsf{M}}(\boldsymbol{G}) &=& \{1,2,3\} & \dot{\mathsf{Var}}(\boldsymbol{G};\dot{\mathsf{C}}_{\mathsf{M}}) &=& \frac{5}{9} &\approx 0.555
\end{array}
$$

This is the first example in this section to use a *directed graph* (rather than an *undirected graph* Definition 1.17 page 8) and to require the use of a *quasi-metric* (Definition D.6 page 163) that is not a *metric*. Unlike Example 2.7 (page 43), which had neither of these restrictions, twice as many center operators (4 rather than 2) were able to successfully isolate a single center point. The *outcome center* result is used later in Example 2.19 (page 64).

✎PROOF:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}\\
&\triangleq \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \frac{1}{9}\mathsf{d}(x,y)\,\mathsf{P}(y)9\\
&\triangleq \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y)9 && \text{because } \mathsf{f}(x)=\frac{1}{9}x \text{ is } \textit{strictly isotone} \text{ and by Lemma 1.7 (page 19)}
\end{aligned}
$$

$$= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \left\{ \begin{array}{ccccc} \mathsf{d}(0,0)\mathsf{P}(0)9 & \mathsf{d}(0,1)\mathsf{P}(1)9 & \mathsf{d}(0,2)\mathsf{P}(2)9 & \mathsf{d}(0,3)\mathsf{P}(3)9 & \mathsf{d}(0,4)\mathsf{P}(4)9 \\ \mathsf{d}(1,0)\mathsf{P}(0)9 & \mathsf{d}(1,1)\mathsf{P}(1)9 & \mathsf{d}(1,2)\mathsf{P}(2)9 & \mathsf{d}(1,3)\mathsf{P}(3)9 & \mathsf{d}(1,4)\mathsf{P}(4)9 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathsf{d}(4,0)\mathsf{P}(0)9 & \mathsf{d}(4,1)\mathsf{P}(1)9 & \mathsf{d}(4,2)\mathsf{P}(2)9 & \mathsf{d}(4,3)\mathsf{P}(3)9 & \mathsf{d}(4,4)\mathsf{P}(4)9 \end{array} \right\}$$

$$= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \left\{ \begin{array}{ccccc} 0\times 2 & 3\times 1 & 1\times 1 & 4\times 2 & 2\times 3 \\ 2\times 2 & 0\times 1 & 3\times 1 & 1\times 2 & 2\times 3 \\ 4\times 2 & 2\times 1 & 0\times 1 & 3\times 2 & 1\times 3 \\ 1\times 2 & 2\times 1 & 2\times 1 & 0\times 2 & 1\times 3 \\ 3\times 2 & 1\times 1 & 4\times 1 & 2\times 2 & 0\times 3 \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 8 \\ 6 \\ 8 \\ 3 \\ 6 \end{array} \right\} = \left\{ \begin{array}{c} \\ \\ \\ 3 \\ \\ \end{array} \right\}$$

$$\begin{aligned}
\dot{\mathfrak{C}}_{\mathsf{a}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathfrak{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \frac{1}{9}\mathsf{d}(x,y)\,\mathsf{P}(y)9 \\
&= \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y)9 && \text{because } \mathsf{f}(x)=\tfrac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.2 (page 16)}
\end{aligned}$$

$$= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccc} 0\times 2 & + & 3\times 1 & + & 1\times 1 & + & 4\times 2 & + & 2\times 3 \\ 2\times 2 & + & 0\times 1 & + & 3\times 1 & + & 1\times 2 & + & 2\times 3 \\ 4\times 2 & + & 2\times 1 & + & 0\times 1 & + & 3\times 2 & + & 1\times 3 \\ 1\times 2 & + & 2\times 1 & + & 2\times 1 & + & 0\times 2 & + & 1\times 3 \\ 3\times 2 & + & 1\times 1 & + & 4\times 1 & + & 2\times 2 & + & 0\times 3 \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 18 \\ 15 \\ 19 \\ 9 \\ 15 \end{array} \right\} = \left\{ \begin{array}{c} \\ \\ \\ 3 \\ \\ \end{array} \right\}$$

$$\begin{aligned}
\dot{\mathfrak{C}}_{\mathsf{g}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{\mathsf{P}(y)} && \text{by definition of } \dot{\mathfrak{C}}_{\mathsf{g}} \text{ (Definition 2.4 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{9\mathsf{P}(y)\frac{1}{9}} \\
&= \arg\min_{x\in\boldsymbol{G}} \left[\prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{9\mathsf{P}(y)} \right]^{\frac{1}{9}} \\
&= \arg\min_{x\in\boldsymbol{G}} \prod_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)^{9\mathsf{P}(y)} && \text{because } \mathsf{f}(x)\triangleq x^{\frac{1}{9}} \text{ is \textit{strictly isotone} and by Lemma 1.2 (page 16)}
\end{aligned}$$

$$= \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccc} & & 3^1 & \times & 1^1 & \times & 4^2 & \times & 2^3 \\ 2^2 & \times & & & 3^1 & \times & 1^2 & \times & 2^3 \\ 4^2 & \times & 2^1 & \times & & & 3^2 & \times & 1^3 \\ 1^2 & \times & 2^1 & \times & 2^1 & \times & & & 1^3 \\ 3^2 & \times & 1^1 & \times & 4^1 & \times & 2^2 & & \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 384 \\ 192 \\ 432 \\ 24 \\ 144 \end{array} \right\} = \left\{ \begin{array}{c} \\ \\ \\ 3 \\ \\ \end{array} \right\}$$

$$\begin{aligned}
\dot{\mathfrak{C}}_{\mathsf{h}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\Omega} \left(\sum_{y\in\Omega} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y) \right)^{-1} && \text{by definition of } \dot{\mathfrak{C}}_{\mathsf{h}} \text{ (Definition 2.4 page 30)} \\
&= \arg\max_{x\in\Omega} \left(\sum_{y\in\Omega} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y) \right) && \text{because } \phi(x)\triangleq x^{-1} \text{ is \textit{strictly antitone} and by Lemma 1.5 page 18} \\
&= \arg\max_{x\in\Omega} \left(\frac{1}{9} \sum_{y\in\Omega} \frac{1}{\mathsf{d}(x,y)}\mathsf{P}(y)9 \right) \\
&= \arg\max_{x\in\Omega} \sum_{y\in\Omega} \frac{9\mathsf{P}(y)}{\mathsf{d}(x,y)} && \text{because } \mathsf{f}(x)=\tfrac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.2 (page 16)}
\end{aligned}$$

$$= \arg\max_{x\in\boldsymbol{G}} \left\{ \begin{array}{ccccccccc} 0 & + & \frac{1}{3} & + & \frac{1}{1} & + & \frac{2}{4} & + & \frac{3}{2} \\ \frac{2}{2} & + & 0 & + & \frac{1}{3} & + & \frac{2}{1} & + & \frac{3}{2} \\ \frac{2}{4} & + & \frac{1}{2} & + & 0 & + & \frac{2}{3} & + & \frac{3}{1} \\ \frac{2}{1} & + & \frac{1}{2} & + & \frac{1}{2} & + & 0 & + & \frac{3}{1} \\ \frac{2}{3} & + & \frac{1}{1} & + & \frac{1}{4} & + & \frac{2}{2} & + & 0 \end{array} \right\} = \arg\max_{x\in\boldsymbol{G}} \frac{1}{2} \left\{ \begin{array}{c} 20 \\ 27 \\ 28 \\ 36 \\ 35 \end{array} \right\} = \left\{ \begin{array}{c} \\ \\ 3 \\ \\ \\ \end{array} \right\}$$

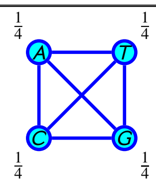
$$\begin{aligned}
\dot{C}_m(\boldsymbol{G}) & \\
&\triangleq \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{C}_m \text{ (Definition 2.4 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \frac{1}{9}\mathsf{d}(x,y)\,\mathsf{P}(y)9 \\
&= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y)9 && \text{because } \mathsf{f}(x) = \frac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.7 (page 19)} \\
&= \arg\min_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \left\{ \begin{array}{ccccc} & 3\times1 & 1\times1 & 4\times2 & 2\times3 \\ 2\times2 & & 3\times1 & 1\times2 & 2\times3 \\ 4\times2 & 2\times1 & & 3\times2 & 1\times3 \\ 1\times2 & 2\times1 & 2\times1 & & 1\times3 \\ 3\times2 & 1\times1 & 4\times1 & 2\times2 & \end{array} \right\} = \arg\min_{x\in\boldsymbol{G}} \left\{ \begin{array}{c} 1 \\ 2 \\ 2 \\ 2 \\ 1 \end{array} \right\} = \left\{ \begin{array}{c} 0 \\ \\ \\ \\ 4 \end{array} \right\} \\
\dot{C}_M(\boldsymbol{G}) &\triangleq \arg\max_{x\in\boldsymbol{G}} \min_{y\in\boldsymbol{G}\setminus\{x\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{C}_M \text{ (Definition 2.4 page 30)} \\
&\triangleq \arg\max_{x\in\boldsymbol{G}} \{1,2,2,2,1\} && \text{by } \dot{C}_m(\boldsymbol{G}) \text{ result} \\
&= \{1,2,3\} \\
\dot{\mathrm{Var}}(\boldsymbol{G}) &= \dot{\mathrm{Var}}(\boldsymbol{G};\dot{C}_a) = \dot{\mathrm{Var}}(\boldsymbol{G};\dot{C}_g) = \dot{\mathrm{Var}}(\boldsymbol{G};\dot{C}_h) && \text{by } \dot{C}(\boldsymbol{G}),\ \dot{C}_a(\boldsymbol{G}),\ \dot{C}_g(\boldsymbol{G}), \text{ and } \dot{C}_h(\boldsymbol{G}) \text{ results} \\
&\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{C}_h(\boldsymbol{G}), x\right) \mathsf{P}(x) && \text{by definition of } \dot{\mathrm{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{3\}, x)\,\mathsf{P}(x) && \text{by } \dot{C}_h(\boldsymbol{G}) \text{ result} \\
&= (1)^2\frac{2}{9} + (2)^2\frac{1}{9} + (2)^2\frac{1}{9} + (0)^2\frac{2}{9} + (1)^2\frac{3}{9} = \frac{12}{9} = \frac{4}{3} \approx 1.333 \\
\dot{\mathrm{Var}}(\boldsymbol{G};\dot{C}_m) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{C}_m(\boldsymbol{G}), x\right) \mathsf{P}(x) && \text{by definition of } \dot{\mathrm{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{0,4\}, x)\,\mathsf{P}(x) && \text{by } \dot{C}_m(\boldsymbol{G}) \text{ result} \\
&= (0)^2\frac{2}{9} + (1)^2\frac{1}{9} + (1)^2\frac{1}{9} + (2)^2\frac{2}{9} + (0)^2\frac{3}{9} = \frac{10}{9} \approx 1.111 \\
\dot{\mathrm{Var}}(\boldsymbol{G};\dot{C}_M) &\triangleq \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\dot{C}_M(\boldsymbol{G}), x\right) \mathsf{P}(x) && \text{by definition of } \dot{\mathrm{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2(\{1,2,3\}, x)\,\mathsf{P}(x) && \text{by } \dot{C}_m(\boldsymbol{G}) \text{ result} \\
&= (1)^2\frac{2}{9} + (0)^2\frac{1}{9} + (0)^2\frac{1}{9} + (0)^2\frac{2}{9} + (1)^2\frac{3}{9} = \frac{5}{9} \approx 0.555
\end{aligned}$$

✏

Example 2.9 (DNA). *Genomic Signal Processing* (*GSP*) analyzes biological sequences called *genomes*. These sequences are constructed over a set of 4 symbols that are commonly referred to as $\boxed{A}$, $\boxed{T}$, $\boxed{C}$, and $\boxed{G}$, each of which corresponds to a nucleobase (adenine, thymine, cytosine, and guanine, respectively).[3] A typical genome sequence contains a large number of symbols (about 3 billion for humans, 29751 for the SARS virus).[4]

[3] Mendel (1853) ⟨Mendel (1853): gene coding uses discrete symbols⟩, Watson and Crick (1953a) page 737 ⟨Watson and Crick (1953): gene coding symbols are adenine, thymine, cytosine, and guanine⟩, Watson and Crick (1953b) page 965, Pommerville (2013) page 52

[4] GenBank (2014) ⟨http://www.ncbi.nlm.nih.gov/genome/guide/human/⟩⟨Homo sapiens, NC_000001–



Let $\boldsymbol{G} \triangleq (\{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}\}, \mathsf{d}, \leq, \mathsf{P})$ be the *outcome subspace* (Definition 2.1 page 29) generated by a *genome* where d is the *discrete metric* (Definition D.8 page 166), $\leq \triangleq \varnothing$ (completely unordered set), and $\mathsf{P}(\boxed{A}) = \mathsf{P}(\boxed{T}) = \mathsf{P}(\boxed{C}) = \mathsf{P}(\boxed{G}) = \frac{1}{4}$. This space is illustrated by the *graph* (Definition 1.17 page 8) to the right with shaded *center* (Definition 2.3 page 30).

The graph has the following geometric values:

$$
\begin{array}{rcll}
\dot{\mathsf{C}}(\boldsymbol{G}) &=& \{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}\} & \text{(Definition 2.3 page 30)} \quad \text{(shaded in illustration)} \\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &=& \{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}\} & \text{(Definition 2.4 page 30)} \quad \text{(shaded in illustration)} \\
\dot{\mathsf{Var}}(\boldsymbol{G}) &=& 0 & \text{(Definition 2.5 page 30)}
\end{array}
$$

✎PROOF:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\frac{1}{4} && \text{by definition of } \boldsymbol{G} \\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y) && \text{because } \mathsf{f}(x) = \tfrac{1}{4}x \text{ is } \textit{strictly isotone} \text{ and by Lemma 1.7 (page 19)} \\
&= \arg\min_{x\in\boldsymbol{G}} \{1,\ 1,\ 1,\ 1\} && \text{because for the } \textit{discrete metric} \text{ (Definition D.8 page 166), } \max \mathsf{d} = 1 \\
&= \{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}\} && \text{by definition of } \boldsymbol{G} \\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\frac{1}{6} && \text{by definition of } \boldsymbol{G} \\
&= \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y) && \text{because } \mathsf{f}(x) = \tfrac{1}{4}x \text{ is } \textit{strictly isotone} \text{ and by Lemma 1.7 (page 19)} \\
&= \arg\min_{x\in\boldsymbol{G}} \left\{\begin{array}{c} 0+1+1+1 \\ 1+0+1+1 \\ 1+1+0+1 \\ 1+1+1+0 \end{array}\right\} = \arg\min_{x\in\boldsymbol{G}} \left\{\begin{array}{c} 3 \\ 3 \\ 3 \\ 3 \end{array}\right\} = \{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}\} \\
\dot{\mathsf{Var}}(\boldsymbol{G}) &\triangleq \sum_{x\in\boldsymbol{G}} \left[\mathsf{d}\left(\dot{\mathsf{C}}(\boldsymbol{G}), x\right)\right]^2 \mathsf{P}(x) && \text{by definition of } \dot{\mathsf{Var}} \text{ (Definition 2.5 page 30)} \\
&= \sum_{x\in\boldsymbol{G}} (0)\frac{1}{6} && \text{because } \dot{\mathsf{C}}(\boldsymbol{G}) = \boldsymbol{G} \\
&= 0 && \text{by field property of } \textit{additive identity element } 0
\end{aligned}
$$

✏

2.2 Random variables on outcome subspaces

2.2.1 Definitions

The traditional *random variable* (Definition 1.39 page 24) is a mapping from a *probability space* (Definition 1.38 page 24) to the *real line* (Definition 1.35 page 22). This paper extends this definition to include functions with additional structure in the domain and expanded structure in the range (next definition).

NC_000022 (22 chromosome pairs), NC_000023 (X chromosome), NC_000024 (Y chromosome), NC_012920 (mitochondria)⟩, GenBank (2014) ⟨http://www.ncbi.nlm.nih.gov/nuccore/30271926⟩⟨SARS coronavirus, NC_004718.3⟩ S. G. Gregory (2006) ⟨homo sapien chromosome 1⟩, Runtao He (2004) ⟨SARS coronavirus⟩

Definition 2.13.

DEF *A function* $\mathsf{X} \in \mathsf{H}^{\boldsymbol{G}}$ (Definition 1.6 page 6) *is an* ***outcome random variable*** *if* $\boldsymbol{G}$ *is an* OUTCOME SUBSPACE (Definition 2.1 page 29) *and* H *is an* ORDERED QUASI-METRIC SPACE (Definition 1.34 page 22).

The definitions of *outcome expected value* and *outcome variance* (next definition) of an *outcome random variable* are, in essence, identical to the *outcome center* (Definition 2.14 page 49) and *outcome variance* (Definition 2.14 page 49) of *outcome subspaces* (Definition 2.1 page 29) that *outcome random variables* map from and by induction, to.

Definition 2.14. *Let* $\boldsymbol{G}$ *be an* OUTCOME SUBSPACE (Definition 2.1 page 29), H *an* ORDERED QUASI-METRIC SPACE (Definition 1.34 page 22), *and* X *be an* OUTCOME RANDOM VARIABLE (Definition 2.13 page 49) *in* $\in \mathsf{H}^{\boldsymbol{G}}$. *Let* $\boldsymbol{H} \triangleq \left(\Omega, \dot{\leq}, \dot{\mathsf{d}}, \dot{\mathsf{P}}\right)$ *be the* OUTCOME SUBSPACE *induced by* H, $\boldsymbol{G}$, *and* X. *Let* $\ddot{\mathsf{E}}_x$ *be a function from* Ω *to the power set* 2^{Ω}.

DEF
The ***outcome expected value*** $\ddot{\mathsf{E}}(\mathsf{X})$ *of* X *is* $\ddot{\mathsf{E}}(\mathsf{X}) \triangleq \arg\min_{x\in\Omega} \max_{y\in\Omega} \mathsf{d}(x,y)\,\mathsf{P}(y)$.

The ***outcome variance*** $\ddot{\mathsf{Var}}(\mathsf{X}; \mathsf{E}_x)$ *of* X *is* $\ddot{\mathsf{Var}}(\mathsf{X}) \triangleq \sum_{x\in\Omega} \mathsf{d}^2\left(\mathsf{E}_x(\mathsf{X}), x\right)\mathsf{P}(x)$.

Moreover, $\ddot{\mathsf{Var}}(\mathsf{X}) \triangleq \ddot{\mathsf{Var}}(\mathsf{X}; \ddot{\mathsf{E}})$, *where* $\ddot{\mathsf{E}}$ *is the* OUTCOME EXPECTED VALUE *function.*

2.2.2 Properties

Theorem 2.1. *Let* $\mathsf{X} \in \mathsf{H}^{\boldsymbol{G}}$ *be a* RANDOM VARIABLE (Definition 2.13 page 49) *on an* ORDERED QUASI-METRIC SPACE (Definition 1.34 page 22) H. *Let* $\boldsymbol{H} \triangleq \left(\Omega, \dot{\leq}, \dot{\mathsf{d}}, \dot{\mathsf{P}}\right)$ *be the* OUTCOME SUBSPACE (Definition 2.1 page 29) *induced by* H, $\boldsymbol{G}$, *and* X. *Let* $\mathsf{Var}(\mathsf{X})$ *be the* TRADITIONAL VARIANCE (Definition 1.40 page 24) *of* X. *Let* $\ddot{\mathsf{Var}}(\mathsf{X})$ *be the* OUTCOME SUBSPACE VARIANCE *of* X (Definition 2.14 page 49).

THM $\left\{ \mathsf{H} \triangleq (\mathbb{R}, |\cdot|, \leq) \text{ is the REAL LINE (Definition 1.35 page 22)} \right\} \implies \left\{ \ddot{\mathsf{Var}}(\mathsf{X}; \mathsf{E}) = \mathsf{Var}(\mathsf{X}) \right\}$

✎PROOF:

$$\begin{aligned}
\ddot{\mathsf{Var}}(\mathsf{X}; \mathsf{E}) &\triangleq \sum_{x\in\mathsf{H}} \mathsf{d}^2(\mathsf{E}(\mathsf{X}), x)\,\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\mathbb{R}} |\mathsf{E}(\mathsf{X}) - x|^2 \mathsf{P}(x) && \text{by definition of real line } \mathsf{H} \text{ (Definition 1.35 page 22)}\\
&= \int_{\mathbb{R}} (x - \mathsf{E}(\mathsf{X}))^2 \mathsf{P}(x)\,\mathsf{d}x && \text{by definition of Lebesgue integration on } \mathbb{R}\\
&= \mathsf{Var}(\mathsf{X}) && \text{by definition of } \mathsf{Var} \text{ (Definition 1.40 page 24)}
\end{aligned}$$

✏

Remark 2.3. Despite the correspondence of traditional variance and outcome variance on the *real line* as demonstrated in Theorem 2.1 (page 49), the situation is different for expected values. Even when both are calculated on the same *real line*, the *traditional expected value* $\mathsf{E}(\mathsf{X})$ (Definition 1.40 page 24) and the *outcome expected value* $\ddot{\mathsf{E}}(\mathsf{X})$ (Definition 2.14 page 49) don't always yield the same value. Demonstrations of this include Example 2.14 (page 57) and Example 2.17 (page 61). However, there is one common situation in which the two statistics do correspond (next theorem).

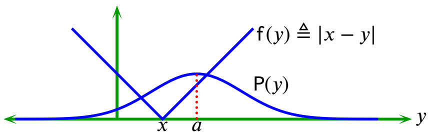
Theorem 2.2. *Let* X, H, $\boldsymbol{G}$ *be defined as in Theorem 2.1 (page 49). Let* $\mathsf{E}(\mathsf{X})$ *be the* TRADITIONAL EXPECTED VALUE (Definition 1.40 page 24) *and* $\ddot{\mathsf{E}}(\mathsf{X})$ *the* OUTCOME EXPECTED VALUE *of* X (Definition 2.14 page 49).

THM $\left\{ \begin{array}{lll} 1. & \mathsf{H} \triangleq (\mathbb{R}, |\cdot|, \leq) & \text{(REAL LINE Definition 1.35 page 22) and}\\ 2. & \mathsf{P}(a - x) = \mathsf{P}(a + x) \quad \forall x \in \mathbb{R} & \text{(SYMMETRIC about } a\text{)} \end{array} \right\} \implies \left\{ \ddot{\mathsf{E}}(\mathsf{X}) = a = \mathsf{E}(\mathsf{X}) \right\}$

✎PROOF:

$$\boxed{\ddot{\mathsf{E}}(\mathsf{X})} \triangleq \arg\min_{x\in\mathbb{R}}\max_{y\in\mathbb{R}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}$$

$$= \arg\min_{x\in\mathbb{R}}\max_{y\in\mathbb{R}} |x-y|\mathsf{P}(y) \qquad \text{by definition of } \textit{real line} \text{ (Definition 1.35 page 22)}$$



$$= \boxed{a} \qquad \text{because } \mathsf{h}(x) \triangleq \max_{y\in\mathbb{R}} |x-y|\mathsf{P}(y) \text{ is minimized when } x=a$$

$$= \boxed{\mathsf{E}(\mathsf{X})} \qquad \text{by Proposition 1.3 (page 24)}$$

✏

Theorem 2.3.

Let $\boldsymbol{G} \triangleq (\Omega_{\boldsymbol{G}}, \mathsf{d}_{\boldsymbol{G}}, \leq_{\boldsymbol{G}}, \mathsf{P}_{\boldsymbol{G}})$	*be an* OUTCOME SUBSPACE	*(Definition 2.1 page 29).*
Let $\mathsf{H} \triangleq (\Omega_{\mathsf{H}}, \mathsf{d}_{\mathsf{H}}, \leq_{\mathsf{H}})$	*be an* ORDERED QUASI-METRIC SPACE	*(Definition 1.34 page 22).*
Let $\mathsf{K} \triangleq (\Omega_{\mathsf{K}}, \mathsf{d}_{\mathsf{K}}, \leq_{\mathsf{K}})$	*be an* ORDERED QUASI-METRIC SPACE	*(Definition 1.34 page 22).*
Let $\mathsf{X} \in \mathsf{H}^{\boldsymbol{G}}$	*be a* RANDOM VARIABLE *from* $\boldsymbol{G}$ *onto* H	*(Definition 2.13 page 49).*
Let $\mathsf{f} \in {\Omega_{\mathsf{K}}}^{\Omega_{\mathsf{H}}}$	*be a function from* Ω_{H} *onto* Ω_{K} *(*PULLBACK*)*	*(Theorem 1.7 page 21).*
Let $\phi \in \mathbb{R}^{\mathbb{R}}$	*be a function from* $\mathbb{R}$ *into* $\mathbb{R}$ *(*PUSHFORWARD*)*	*(Definition D.11 page 166).*
Let $\boldsymbol{H} \triangleq (\Omega_{\mathsf{H}}, \mathsf{d}_{\mathsf{H}}, \leq_{\mathsf{H}}, \mathsf{P}_{\mathsf{H}})$	*be an* OUTCOME SUBSPACE *induced by* $\boldsymbol{G}$, H, *and* X.	
Let $\boldsymbol{K} \triangleq (\Omega_{\mathsf{K}}, \mathsf{d}_{\mathsf{K}}, \leq_{\mathsf{K}}, \mathsf{P}_{\mathsf{K}})$	*be an* OUTCOME SUBSPACE *induced by* K, $\boldsymbol{H}$ *and* f.	

THM

$$\left\{\begin{array}{lll} 1. & \mathsf{f} \text{ is INJECTIVE} & \textit{and} \\ 2. & \phi \text{ is STRICTLY ISOTONE} & \textit{and} \\ 3. & \mathsf{d}_{\mathsf{H}}(\mathsf{f}(x),\mathsf{f}(y))\mathsf{P}(y) = \phi\left[\mathsf{d}_{\mathsf{H}}(x,y)\mathsf{P}(y)\right] & \end{array}\right\} \implies \left\{\ddot{\mathsf{E}}[\mathsf{f}(\mathsf{X})] = \mathsf{f}\left[\ddot{\mathsf{E}}(\mathsf{X})\right]\right\}$$

✎PROOF:

$$\ddot{\mathsf{E}}[\mathsf{f}(\mathsf{X})] = \arg\min_{x\in\Omega_{\mathsf{K}}}\max_{y\in\Omega_{\mathsf{K}}} \mathsf{d}_{\mathsf{K}}(x,y)\mathsf{P}_{\mathsf{K}}(y) \qquad \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49) and } \mathsf{K}$$

$$= \mathsf{f}\left[\arg\min_{x\in\Omega_{\mathsf{H}}}\max_{y\in\Omega_{\mathsf{H}}} \mathsf{d}_{\mathsf{K}}(\mathsf{f}(x),\mathsf{f}(y))\mathsf{P}_{\mathsf{K}}(\mathsf{f}(y))\right] \qquad \text{by } \mathsf{f} \textit{ bijection} \text{ hypothesis}$$

$$= \mathsf{f}\left[\arg\min_{x\in\Omega_{\mathsf{H}}}\max_{y\in\Omega_{\mathsf{H}}} \mathsf{d}_{\mathsf{H}}(\mathsf{f}(x),\mathsf{f}(y))\mathsf{P}_{\mathsf{H}}(y)\right] \qquad \text{by } \mathsf{f} \textit{ bijection} \text{ hypothesis}$$

$$= \mathsf{f}\left[\arg\min_{x\in\Omega_{\mathsf{H}}}\max_{y\in\Omega_{\mathsf{H}}} \phi\left[\mathsf{d}_{\mathsf{H}}(x,y)\mathsf{P}_{\mathsf{H}}(y)\right]\right] \qquad \text{by } \mathsf{d}_{\mathsf{H}} \text{ hypothesis}$$

$$= \mathsf{f}\left[\arg\min_{x\in\Omega_{\mathsf{H}}}\max_{y\in\Omega_{\mathsf{H}}} \mathsf{d}_{\mathsf{H}}(x,y)\mathsf{P}_{\mathsf{H}}(y)\right] \qquad \text{by } \phi \text{ is } \textit{strictly isotone} \text{ hypothesis and Lemma 1.7 page 19}$$

$$= \mathsf{f}\ddot{\mathsf{E}}(\mathsf{X}) \qquad \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49) and } \mathsf{X}$$

✏

Corollary 2.1. *Let* H *be an* ORDERED METRIC SPACE *(Definition 1.34 page 22) and* $\mathsf{X} \in \mathsf{H}^{\boldsymbol{G}}$ *a* RANDOM VARIABLE *(Definition 2.13 page 49) onto* H. *Let* $(\mathbb{R}, |\cdot|, \leq)$ *be the* REAL LINE ORDERED METRIC SPACE *(Definition 1.35 page 22).*

THM

$$\underbrace{\mathsf{H} = (\mathbb{R}, |\cdot|, \leq)}_{\text{(real line)}} \implies \left\{ \ddot{\mathsf{E}}(a\mathsf{X}) = a\ddot{\mathsf{E}}(\mathsf{X}) \quad \forall a\in\mathbb{R}^{\vdash} \right\}$$

✎PROOF:

1. Proof for $a = 0$ case:

$$\begin{aligned}
\ddot{\mathsf{E}}(0\cdot\mathsf{X}) &= \arg\min_{x\in 0\cdot\mathsf{H}}\max_{y\in 0\cdot\mathsf{H}} \mathsf{d}(x,y)\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\{0\}}\max_{y\in\{0\}} \mathsf{d}(x,y)\mathsf{P}(y)\\
&= \arg\min_{x\in\{0\}}\max_{y\in\{0\}} \mathsf{d}(0,0)\mathsf{P}(y)\\
&= \arg\min_{x\in\{0\}}\max_{y\in\{0\}} 0\mathsf{P}(y) && \text{by \textit{nondegenerate} property of } \mathsf{d} \text{ (Definition D.7 page 163)}\\
&= 0\\
&= 0\cdot\ddot{\mathsf{E}}(\mathsf{X})
\end{aligned}$$

2. Proof for $a > 0$ case: $\mathsf{d}(\mathsf{f}(x),\mathsf{f}(y))\mathsf{P}(y) \triangleq |ax-ay|\mathsf{P}(y) = |a||x-y|\mathsf{P}(y) \triangleq |a||\mathsf{d}(x,y)\mathsf{P}(y)|$

$$\ddot{\mathsf{E}}(a\mathsf{X}) = a\ddot{\mathsf{E}}(\mathsf{X}) \quad \text{because } \mathsf{f}(x) = ax \text{ is \textit{strictly isotone} on the real line and by Theorem 2.3 (page 50)}$$

⇰

2.2.3 Problem statement

The *traditional random variable* X (Definition 1.39 page 24) is a function that maps from a *stochastic process* to the *real line* (Definition 1.35 page 22). The traditional expectation value E(X) of X is then often a poor choice of a statistic when the stochastic process that X maps from is a structure other than the real line or some substructure of the real line. There are two fundamental problems:

1. A traditional random variable X maps **to** the *linearly ordered* real line. However, X often maps **from** a random process that is *non-linearly ordered* (or even *unordered* Definition 1.20 page 9, Definition 1.21 page 9).

2. A traditional random variable X maps **to** the real line with a *metric geometry* (Remark 2.2 page 34) induced by the *usual metric* (Definition D.9 page 166). But many random processes have a fundamentally different *metric geometry*, a common one being that induced by the *discrete metric* (Definition D.8 page 166).

Thus, the order structure of the domain and range of X are often fundamentally dissimilar, leading to statistics, such as E(X), that are of poor quality with regards to qualitative intuition and quantitative variance (expected error) measurements, and of dubious suitability for tasks such as decision making, prediction, and hypothesis testing.

Remark 2.4. Unlike in traditional statistical processing, it in general **not true** that $\ddot{\mathsf{E}}(\mathsf{X}+\mathsf{Y}) = \ddot{\mathsf{E}}(\mathsf{X}) + \ddot{\mathsf{E}}(\mathsf{Y})$. See Example 2.33 (page 78) for a counter example.

Remark 2.5. A possible solution to the traditional random variable order and metric geometry problem is to allow the random variable to map into the *complex plane* (Example 1.10 page 23) with the usual metric, rather than into the real line only. However, this is a poor solution, as demonstrated in Example 2.21 (page 67).

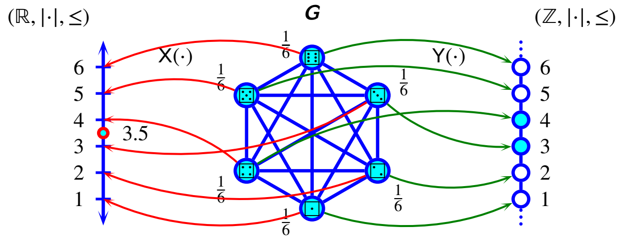


Figure 2.2: random variable mappings from the fair die to the real line and integer line

2.2.4 Examples

Fair die examples

Example 2.10 (fair die mappings to real line and integer line). Let $\boldsymbol{G}$ be the *fair die outcome subspace* (Example 2.1 page 31). Let $\mathsf{X} \in (\mathbb{R}, |\cdot|, \le)^{\boldsymbol{G}}$ be a *random variable* (Definition 2.13 page 49) mapping from $\boldsymbol{G}$ to the *real line* (Definition 1.35 page 22), and $\mathsf{Y} \in (\mathbb{Z}, \le, |\cdot|)^{\boldsymbol{G}}$ be a *random variable* (Definition 2.13 page 49) mapping from $\boldsymbol{G}$ to the *integer line* (Definition 1.36 page 23), as illustrated in Figure 2.2 (page 52). Let E be the *traditional expected value* function (Definition 1.40 page 24), Var the *traditional variance* function (Definition 1.40 page 24), $\ddot{\mathsf{E}}$ the *outcome expected value* function (Definition 2.14 page 49), and $\ddot{\mathsf{Var}}$ the *outcome variance* function (Definition 2.14 page 49). This yields the following statistics:

geometry of $\boldsymbol{G}$:	$\dot{\mathsf{C}}(\boldsymbol{G})$	$=$	$\{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$					
traditional statistics on *real line*:	$\mathsf{E}(\mathsf{X})$	$=$	3.5	$\ddot{\mathsf{Var}}(\mathsf{X}; \mathsf{E})$	$=$	$\frac{35}{12}$	$\approx$	2.917
outcome subspace statistics on *real line*:	$\ddot{\mathsf{E}}(\mathsf{X})$	$=$	$\{3.5\}$	$\ddot{\mathsf{Var}}(\mathsf{X}; \ddot{\mathsf{E}})$	$=$	$\frac{35}{12}$	$\approx$	2.917
outcome subspace statistics on *integer line*:	$\ddot{\mathsf{E}}(\mathsf{Y})$	$=$	$\{3, 4\}$	$\ddot{\mathsf{Var}}(\mathsf{Y}; \ddot{\mathsf{E}})$	$=$	$\frac{20}{12}$	$\approx$	1.667

✎PROOF:

$\dot{\mathsf{C}}(\boldsymbol{G}) = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$ by Example 2.1 page 31

$$\mathsf{E}(\mathsf{X}) \triangleq \sum_{x\in\mathbb{R}} x\mathsf{P}(x) \qquad \text{by definition of } \mathsf{E} \text{ (Definition 1.40 page 24)}$$

$$= \sum_{x\in\mathbb{R}} x\frac{1}{6} = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = \frac{7}{2} = 3.5$$

$$\ddot{\mathsf{Var}}(\mathsf{X}; \mathsf{E}) = \mathsf{Var}(\mathsf{X}) \qquad \text{by Theorem 2.1 page 49}$$

$$\triangleq \sum_{x\in\mathbb{R}} [x - \mathsf{E}(\mathsf{X})]^2 \mathsf{P}(x) \qquad \text{by definition of } \mathsf{Var} \text{ (Definition 1.40 page 24)}$$

$$= \sum_{x\in\mathbb{R}} \left(x - \frac{7}{2}\right)^2 \frac{1}{6} \qquad \text{by } \mathsf{E}(\mathsf{X}) \text{ result}$$

$$= \left[\left(1-\frac{7}{2}\right)^2 + \left(2-\frac{7}{2}\right)^2 + \left(3-\frac{7}{2}\right)^2 + \left(4-\frac{7}{2}\right)^2 + \left(5-\frac{7}{2}\right)^2 + \left(6-\frac{7}{2}\right)^2\right]\frac{1}{6}$$

$$= \left[(2-7)^2 + (4-7)^2 + (6-7)^2 + (8-7)^2 + (10-7)^2 + (12-7)^2\right]\frac{1}{2^2\times 6}$$

$$= \frac{25+9+1+1+9+25}{24} = \frac{70}{24} = \frac{35}{12} \approx 2.917$$

$$\ddot{\mathsf{E}}(\mathsf{X}) = \mathsf{E}(\mathsf{X}) \qquad \text{by Theorem 2.2 (page 49)}$$

$$= \left\{\frac{7}{2}\right\} = \{3.5\} \qquad \text{by } \mathsf{E}(\mathsf{X}) \text{ result}$$

$$\ddot{\mathsf{Var}}(\mathsf{X}; \ddot{\mathsf{E}}) \triangleq \sum_{x\in\mathbb{R}} \mathsf{d}^2\left(\ddot{\mathsf{E}}(\mathsf{X}), x\right) \mathsf{P}(x) \qquad \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}$$

$$= \sum_{x\in\mathbb{R}} \mathsf{d}^2(\mathsf{E}(\mathsf{X}), x)\, \mathsf{P}(x) \qquad \text{by } \mathsf{E}(\mathsf{X}) \text{ and } \ddot{\mathsf{E}}(\mathsf{X}) \text{ results}$$

$$
\begin{aligned}
&\ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \frac{35}{12} \approx 2.917 && \text{by } \mathsf{Var}(\mathsf{X}) \text{ result}\\
\ddot{\mathsf{E}}(\mathsf{Y}) &\triangleq \arg\min_{x\in\mathbb{Z}} \max_{y\in\mathsf{H}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\mathbb{Z}} \max_{y\in\mathsf{H}} |x-y|\frac{1}{6} && \text{by definition of \textit{integer line} (Definition 1.36 page 23) and } \boldsymbol{G}\\
&= \arg\min_{x\in\mathbb{Z}} \max_{y\in\mathsf{H}} |x-y| && \text{because } \phi(x) = \tfrac{1}{6}x \text{ is \textit{strictly isotone} and by Lemma 1.7 page 19}\\
&= \arg\min_{x\in\mathbb{Z}} \max_{y\in\mathsf{H}} \left\{\begin{matrix} 0&1&2&3&4&5\\ 1&0&1&2&3&4\\ 2&1&0&1&2&3\\ 3&2&1&0&1&2\\ 4&3&2&1&0&1\\ 5&4&3&2&1&0 \end{matrix}\right\} = \arg\min_{x\in\mathbb{Z}} \left\{\begin{matrix}5\\4\\3\\3\\4\\5\end{matrix}\right\} = \arg\min_{x\in\mathbb{Z}} \left\{\begin{matrix}3\\4\end{matrix}\right\}\\
\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) &\triangleq \sum_{x\in\mathbb{Z}} \mathsf{d}^2\big(\ddot{\mathsf{E}}(\mathsf{Y}),x\big)\,\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\mathbb{Z}} \mathsf{d}^2(\{3,4\},x)\,\frac{1}{6} && \text{by } \ddot{\mathsf{E}}(\mathsf{Y}) \text{ result and definition of } \boldsymbol{G}\\
&= \frac{1}{6}\big(|3-1|^2 + |3-2|^2 + |3-3|^2 + |4-4|^2 + |4-5|^2 + |4-6|^2\big)\\
&= \frac{1}{6}(4+1+0+0+1+4) = \frac{10}{6} = \frac{5}{3}
\end{aligned}
$$

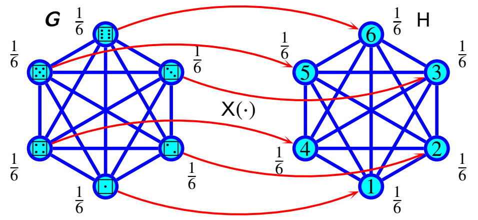
The random variable mappings in Example 2.10 (page 52) have two fundamental problems:

1. The order structure of the *fair die* and the order structure of *real line* are inherently dissimilar in that while the *bijective* (Definition 1.14 page 7) mapping X is trivially *order preserving* (Definition 1.32 page 16), its inverse is *not order preserving*. And this is a problem. In the *linearly ordered* (Definition 1.21 page 9) range of X, it is true that X(⚀) = 1 < 2 = X(⚁). But in the unordered domain of X {⚀, ⚁, …, ⚅}, it is *not* true that ⚀ < ⚁; rather ⚀ and ⚁ are simply symbols without order. This causes problems when we attempt to use the random variable to make statistical inferences involving moments (Definition 2.2 page 29). The *traditional expected value* (Definition 1.40 page 24) of a *fair die* (Example 2.1 page 31) is $\mathsf{E}(X) = \frac{1}{6}(1+2+\cdots+6) = 3.5$. This implies that we expect the outcome of ⚂ or ⚃ more than we expect the outcome of say ⚀ or ⚁. But these results have no relationship with reality or with intuition because the values of a fair die are merely symbols. For a fair die, we would expect any pair of values equally. We would not expect the outcome [⚂ or ⚃] more than we would expect the outcome [⚀ or ⚁], or more than we would expect any other outcome pair.

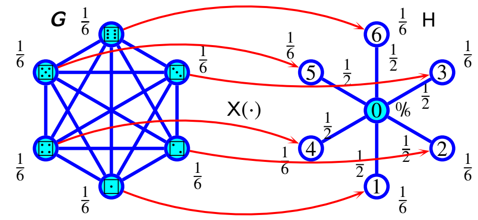
2. The metric geometry (Remark 2.2 page 34) of the *fair die outcome subspace* is very dissimilar to the metric geometry of the *real line* (Definition 1.35 page 22) that it is mapped to by the random variable X. And this is a problem. In the metric geometry of the fair die induced by the *discrete metric* (Definition D.8 page 166), ⚀ is no closer to ⚁ than it is to ⚂ (d(⚀, ⚁) = 1 = d(⚀, ⚂)). However in the metric geometry of the real line induced by the *usual metric* $\mathsf{d}(x,y) \triangleq |x-y|$ (Definition D.9 page 166), X(⚀) = 1 is closer to X(⚁) = 2 than it is to X(⚂) = 3 ($|1-2| = 1 \neq 2 = |1-3|$).

Example 2.11 (fair die mapping to isomorphic structure). Let $\boldsymbol{G} \triangleq$ ({⚀, ⚁, ⚂, ⚃, ⚄, ⚅}, d, ∅, P) be a *fair die outcome subspace* (Example 2.1 page 31), and $\mathsf{H} \triangleq (\{1,2,3,4,5,6\},\mathsf{d},\varnothing)$ be an *unordered metric space* (Definition 1.34 page 22). Example 2.10 (page 52) presented mappings from $\boldsymbol{G}$ to structures with structures dissimilar to $\boldsymbol{G}$. Figure 2.3 page 54 (A) illustrates a mapping to the isomorphic structure $\boldsymbol{H} \triangleq (\{1,2,3,4,5,6\},\mathsf{d},\varnothing,\mathsf{X}(\mathsf{P}))$, yielding the following statistics:

[h]



(A) mapping to isomorphic structure



(B) mapping to extended structure

Figure 2.3: order preserving random variable mappings from *fair die*

$\ddot{\mathsf{E}}(\mathsf{X}) = \{1, 2, 3, 4, 5, 6\} \qquad \ddot{\mathsf{Var}}(\mathsf{X}) = 0$

Here, $\ddot{\mathsf{E}}(\mathsf{X})$ equals the entire base set of H, indicating a statistic carrying no information about an expected outcome. That is, there is no best guess concerning outcome. This is much different than the traditional probability of 3.5 (Example 2.10 page 52) which deceptively suggests a likely outcome of ⚂ or ⚃. And one could easily argue that no information is much better than misleading information.

✎PROOF:

$$\begin{aligned}
\ddot{\mathsf{E}}(\mathsf{X}) &\triangleq \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{because } \boldsymbol{G} \text{ and H are } \textit{isomorphic}\\
&= \mathsf{X}\left[\dot{\mathsf{C}}(\boldsymbol{G})\right] && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}\\
&= \mathsf{X}[\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}] && \text{by Example 2.10 (page 52)}\\
&= \{1,2,3,4,5,6\} && \text{by definition of X}\\
\ddot{\mathsf{Var}}(\mathsf{X}) &\triangleq \sum_{x\in\mathsf{H}} \mathsf{d}^2\left(\ddot{\mathsf{E}}(\mathsf{X}), x\right)\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\boldsymbol{G}} \mathsf{d}^2\left(\mathsf{X}(\dot{\mathsf{C}}(\boldsymbol{G}), x\right)\mathsf{P}(x) && \text{because } \boldsymbol{G} \text{ and H are } \textit{isomorphic}\\
&\triangleq \ddot{\mathsf{Var}}(\boldsymbol{G}) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.5 page 30)}\\
&= 0 && \text{by Example 2.10 (page 52)}
\end{aligned}$$

✏

Although all the coefficients of the polynomial equation $x^2-2x+2=0$ are in the set of real numbers $\mathbb{R}$, the solutions of the equation ($x = 1+i$ and $x = 1-i$) are not. Rather, the two solutions are in the *complex plane* $\mathbb{R}^2$ (Example 1.10 page 23), of which $\mathbb{R}$ is a substructure. This is an example of extending a structure (from $\mathbb{R}$ to $\mathbb{R}^2$) to achieve more useful results. The same idea can be applied to a random variable $\mathsf{X} \in \mathsf{H}^{\boldsymbol{G}}$. The definition of an *outcome random variable* (Definition 2.13 page 49) does not require a bijection between $\boldsymbol{G}$ and H; rather, it only requires that the mapping be "into" the base set of H (Definition 1.14 page 7). In Example 2.11 (page 53) in which $\boldsymbol{G}$ is isomorphic to $\boldsymbol{H}$, the expected value of X is a set with six values. However, we could extend H, while still preserving the order and metric geometry of $\boldsymbol{G}$, to produce a random variable with a simpler expected value (next example).

Example 2.12 (fair die mapping with extended range). Let $\boldsymbol{G} \triangleq (\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}, \mathsf{d}, \varnothing, \mathsf{P})$ be a *fair die outcome subspace* (Example 2.1 page 31), and $\mathsf{H} \triangleq (\{1,2,3,4,5,6,0\}, \mathsf{p}, \varnothing)$ be an *unordered metric space* (Definition 1.34 page 22). Figure 2.3 page 54 (B) illustrates a random variable mapping X from $\boldsymbol{G}$ to the extended structure H, yielding the following statistics:

$\ddot{\mathsf{E}}(\mathsf{X}) = \{0\} \qquad \ddot{\mathsf{Var}}(\mathsf{X}) = \frac{1}{4}$

As in Example 2.11, order and metric geometry are still preserved. Here, an expected value of $\{0\}$ simply means that no real physical value is expected more or less than any other real physical value. Note also that the variance (expected error) is more than 11 times smaller than that of the corresponding statistical estimates on the real line ($^3/_{12}$ versus $^{35}/_{12}$ Example 2.10 page 52).

✎PROOF:

$$\begin{aligned}
\ddot{\mathsf{E}}(\mathsf{X}) &\triangleq \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}\setminus\{0\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{because } \mathsf{P}(0)=0\\
&= \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}\setminus\{0\}} \mathsf{d}(x,y)\,\frac{1}{6}\\
&= \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}\setminus\{0\}} \mathsf{d}(x,y) && \text{because } \mathsf{f}(x)=\frac{1}{6}x \text{ is \textit{strictly isotone} and by Lemma 1.7 (page 19)}\\
&= \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}} \left\{\begin{array}{cccc}
\mathsf{d}(1,1) & \mathsf{d}(1,2) & \cdots & \mathsf{d}(1,6)\\
\mathsf{d}(2,1) & \mathsf{d}(2,2) & \cdots & \mathsf{d}(2,6)\\
\mathsf{d}(3,1) & \mathsf{d}(3,2) & \cdots & \mathsf{d}(3,6)\\
\mathsf{d}(4,1) & \mathsf{d}(4,2) & \cdots & \mathsf{d}(4,6)\\
\mathsf{d}(5,1) & \mathsf{d}(5,2) & \cdots & \mathsf{d}(5,6)\\
\mathsf{d}(6,1) & \mathsf{d}(6,2) & \cdots & \mathsf{d}(6,6)\\
\mathsf{d}(0,1) & \mathsf{d}(0,2) & \cdots & \mathsf{d}(0,6)
\end{array}\right\}
= \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}} \left\{\begin{array}{cccccc}
0&1&1&1&1&2\\
1&0&1&1&2&1\\
1&1&0&2&1&1\\
1&1&2&0&1&1\\
1&2&1&1&0&1\\
2&1&1&1&1&0\\
1&1&1&1&1&1
\end{array}\right\}
= \arg\min_{x\in\mathsf{H}} \left\{\begin{array}{c}2\\2\\2\\2\\2\\2\\1\end{array}\right\}\\
&= \{0\}\\
\ddot{\mathsf{Var}}(\mathsf{X}) &\triangleq \sum_{x\in\mathsf{H}} \mathsf{d}^2\left(\ddot{\mathsf{E}}(\mathsf{X}),x\right)\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\mathsf{H}} \mathsf{d}^2(\{0\},x)\,\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{X}) \text{ result}\\
&= \sum_{x\in\mathsf{H}\setminus\{0\}} \mathsf{d}^2(\{0\},x)\,\frac{1}{6} && \text{by definition of } \boldsymbol{G}\\
&= 6\left(\frac{1}{2}\right)^2\frac{1}{6} = \frac{1}{4} && \text{by definition of } \mathsf{H}
\end{aligned}$$

✏

Real die examples

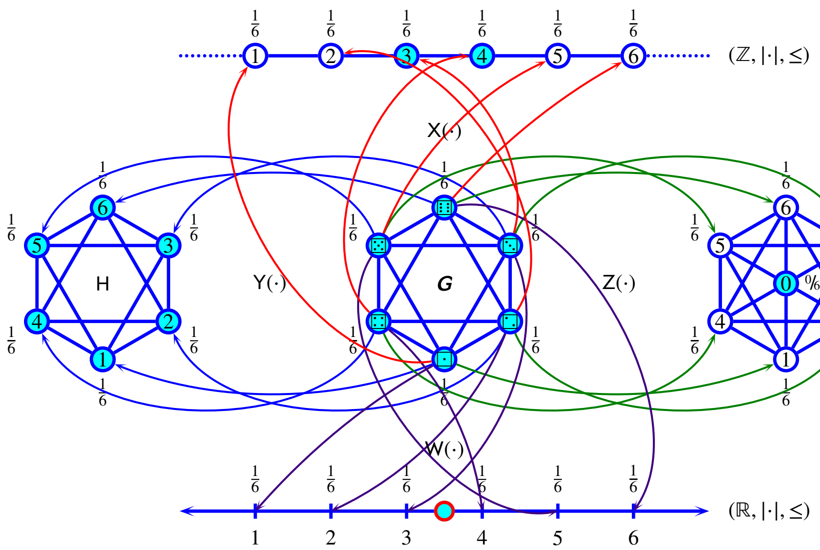


Figure 2.4: random variable mappings from the *real die outcome subspace* to several *ordered metric spaces* (Example 2.13 page 55)

Example 2.13 (real die mappings). Let $\boldsymbol{G}$ be the *real die outcome subspace* (Example 2.2 page 33). Let W, X, Y and Z be *random variable* (Definition 2.13 page 49) mappings as illustrated in Figure 2.4 (page 55). Let E, Var, Ë, and V̈ar be defined as in Example 2.10 (page 52). This yields the following statistics:

geometry of $\boldsymbol{G}$:	$\dot{\mathsf{C}}(\boldsymbol{G})$	$=$	$\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}$		
traditional statistics on real line:	$\mathsf{E}(\mathsf{W})$	$=$	3.5	$\ddot{\mathsf{Var}}(\mathsf{W};\mathsf{E}) = \frac{35}{12}$	≈ 2.917
outcome subspace statistics on real line:	$\ddot{\mathsf{E}}(\mathsf{W})$	$=$	$\{3.5\}$	$\ddot{\mathsf{Var}}(\mathsf{W};\ddot{\mathsf{E}}) = \frac{35}{12}$	≈ 2.917
outcome subspace statistics on integer line:	$\ddot{\mathsf{E}}(\mathsf{X})$	$=$	$\{3, 4\}$	$\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) = \frac{20}{12}$	≈ 1.667
outcome subspace statistics on isomorphic structure:	$\ddot{\mathsf{E}}(\mathsf{Y})$	$=$	$\{1, 2, \ldots, 6\}$	$\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) = 0$	
outcome subspace statistics on extended structure:	$\ddot{\mathsf{E}}(\mathsf{Z})$	$=$	$\{0\}$	$\ddot{\mathsf{Var}}(\mathsf{Z};\ddot{\mathsf{E}}) = 1$	

Similar to Example 2.11 (page 53), the statistic $\ddot{\mathsf{E}}(\mathsf{Z}) = \{0\}$ indicates a statistic carrying no information about an expected outcome. Again, one could easily argue that no information is much better than misleading information.

✎PROOF:

$$\dot{\mathsf{C}}(\boldsymbol{G}) = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} \qquad \text{by Example 2.2 (page 33)}$$

$$\mathsf{E}(\mathsf{W}) = \frac{7}{2} = 3.5 \qquad \text{by } \mathsf{E}(\mathsf{X}) \text{ result of Example 2.10 page 52}$$

$$\ddot{\mathsf{Var}}(\mathsf{W};\mathsf{E}) = \frac{35}{12} \approx 2.917 \qquad \text{by } \mathsf{Var}(\mathsf{X}) \text{ result of Example 2.10 page 52}$$

$$\ddot{\mathsf{E}}(\mathsf{W}) = \{3.5\} \qquad \text{by } \ddot{\mathsf{E}}(\mathsf{X}) \text{ result of Example 2.10 page 52}$$

$$\ddot{\mathsf{Var}}(\mathsf{W};\ddot{\mathsf{E}}) = \frac{35}{12} \approx 2.917 \qquad \text{by } \ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) \text{ result of Example 2.10 page 52}$$

$$\ddot{\mathsf{E}}(\mathsf{X}) = \{3, 4\} \qquad \text{by } \ddot{\mathsf{E}}(\mathsf{Y}) \text{ result of Example 2.10 page 52}$$

$$\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) = \frac{5}{3} \approx 1.667 \qquad \text{by } \ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) \text{ result of Example 2.10 page 52}$$

$$\begin{aligned}
\ddot{\mathsf{E}}(\mathsf{Y}) &= \mathsf{Y}(\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}) && \text{by } \dot{\mathsf{C}}(\boldsymbol{G}) \text{ result of Example 2.2 page 33}\\
&= \{1, 2, 3, 4, 5, 6\} && \text{by definition of } \mathsf{Y}
\end{aligned}$$

$$\begin{aligned}
\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) &\triangleq \sum_{x\in\mathsf{H}} \mathsf{d}^2\big(\ddot{\mathsf{E}}(\mathsf{Y}), x\big)\,\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\mathsf{H}} \mathsf{d}^2(\mathsf{H}, x)\,\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{Y}) \text{ result}\\
&= \sum_{x\in\mathsf{H}} 0^2 x\mathsf{P}(x) && \text{by \textit{nondegenerate} property of \textit{quasi-metrics} (Definition D.6 page 163)}\\
&= 0
\end{aligned}$$

$$\begin{aligned}
\ddot{\mathsf{E}}(\mathsf{Z}) &\triangleq \arg\min_{x\in\mathsf{K}} \max_{y\in\mathsf{K}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\mathsf{K}} \max_{y\in\mathsf{K}\setminus\{0\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{because } \mathsf{P}(0) = 0\\
&= \arg\min_{x\in\mathsf{K}} \max_{y\in\mathsf{K}\setminus\{0\}} \mathsf{d}(x,y)\,\frac{1}{6} && \text{by definition of } \boldsymbol{G} \text{ and } \mathsf{Z}\\
&= \arg\min_{x\in\mathsf{K}} \max_{y\in\mathsf{K}\setminus\{0\}} \mathsf{d}(x,y) && \text{because } \mathsf{f}(x) = \frac{1}{6}x \text{ is \textit{strictly isotone} and by Lemma 1.7 (page 19)}\\
&= \arg\min_{x\in\mathsf{K}} \max_{y\in\mathsf{K}} \left\{\begin{array}{cccc}
\mathsf{d}(1,1) & \mathsf{d}(1,2) & \cdots & \mathsf{d}(1,6)\\
\mathsf{d}(2,1) & \mathsf{d}(2,2) & \cdots & \mathsf{d}(2,6)\\
\mathsf{d}(3,1) & \mathsf{d}(3,2) & \cdots & \mathsf{d}(3,6)\\
\mathsf{d}(4,1) & \mathsf{d}(4,2) & \cdots & \mathsf{d}(4,6)\\
\mathsf{d}(5,1) & \mathsf{d}(5,2) & \cdots & \mathsf{d}(5,6)\\
\mathsf{d}(6,1) & \mathsf{d}(6,2) & \cdots & \mathsf{d}(6,6)\\
\mathsf{d}(0,1) & \mathsf{d}(0,2) & \cdots & \mathsf{d}(0,6)
\end{array}\right\}
= \arg\min_{x\in\mathsf{K}} \max_{y\in\mathsf{K}} \left\{\begin{array}{cccccc}
0&1&1&1&1&2\\
1&0&1&1&2&1\\
1&1&0&2&1&1\\
1&1&2&0&1&1\\
1&2&1&1&0&1\\
2&1&1&1&1&0\\
1&1&1&1&1&1
\end{array}\right\}
= \arg\min_{x\in\mathsf{K}} \left\{\begin{array}{c}2\\2\\2\\2\\2\\2\\1\end{array}\right\}
= \left\{\begin{array}{c}\\ \\ \\ \\ \\ \\0\end{array}\right\}
\end{aligned}$$

$$\begin{aligned}
\ddot{\mathsf{Var}}(\mathsf{Z}) &\triangleq \sum_{x\in\mathsf{K}} \mathsf{d}^2\big(\ddot{\mathsf{E}}(\mathsf{Z}), x\big)\,\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\mathsf{K}\setminus\{0\}} \mathsf{d}^2(\{0\}, x)\,\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{Z}) \text{ result}\\
&= 6 \times 1^2 \times \frac{1}{6} = 1 && \text{by } \ddot{\mathsf{E}}(\mathsf{Z}) \text{ result}
\end{aligned}$$

✏

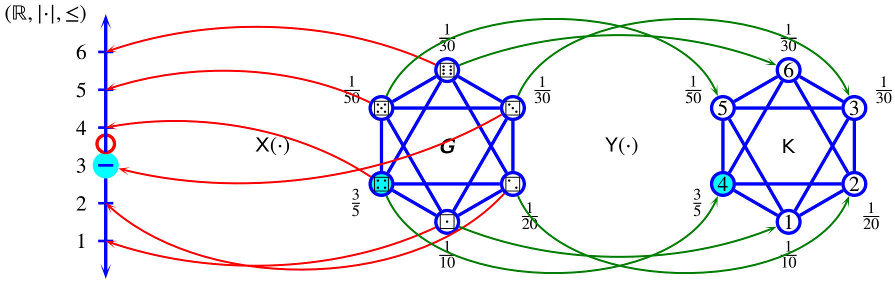


Figure 2.5: *weighted die* mappings (Example 2.14 page 57)

Example 2.14 (weighted die mappings). Let $\boldsymbol{G}$ be *weighted die outcome subspace* (Example 2.3 page 35), and X and Y be *random variables*, as illustrated in Figure 2.5 (page 57). Let E, Var, $\ddot{\mathsf{E}}$, and $\ddot{\mathsf{Var}}$ be defined as in Example 2.10 (page 52). This yields the following statistics:

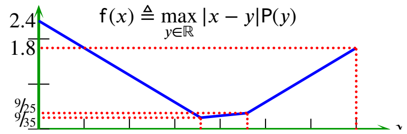
$$\begin{array}{llcl lcl}
\text{geometry of } \boldsymbol{G}: & \dot{\mathsf{C}}(\boldsymbol{G}) & = & \{⚃\} & & & \\
\text{traditional statistics on real line:} & \mathsf{E}(\mathsf{X}) & = & 3 & \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) & = & \frac{143}{100} = 1.43 \\
\text{outcome subspace statistics on real line:} & \ddot{\mathsf{E}}(\mathsf{X}) & = & \left\{\frac{25}{7}\right\} \approx \{3.57\} & \ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) & = & \frac{3361}{2940} \approx 1.143 \\
\text{outcome subspace stats. on isomorphic structure} & \ddot{\mathsf{E}}(\mathsf{Y}) & = & \{4\} & \ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) & = & \frac{101}{300} \approx 0.337
\end{array}$$

The statistic $\mathsf{E}(\mathsf{X}) = 3$ evaluated on the *real line* is arguably very poor because it suggests that we "expect" the event ⚂ rather than ⚃, even though $\mathsf{P}(⚃)$ is very large, $\mathsf{P}(⚂)$ is very small, and the physical distance $\mathsf{d}(⚃, ⚂) = 2$ on the die from ⚃ to ⚂ is twice as much as it is to any of the other four die faces. If we retain use of the real line but replace the *traditional expected value* $\mathsf{E}(\mathsf{X})$ with the *outcome expected value* $\ddot{\mathsf{E}}(\mathsf{X})$, a small but significant improvement is made ($\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) \approx 1.143 < 1.43 = \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E})$). Arguably a better choice still is to abandon the real line altogether in favor of the isomorphic structure K and the statistic $\ddot{\mathsf{E}}(\mathsf{Y}) = \{4\}$ evaluated on K, yielding not only an intuitively better result but also a variance $\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}})$ that is more than 4 times smaller than that of $\mathsf{E}(\mathsf{X})$ ($\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) \approx 0.337 < 1.43 = \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E})$).

✎PROOF:

$$\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &= \{⚃\} && \text{by Example 2.3 (page 35)} \\
\mathsf{E}(\mathsf{X}) &\triangleq \int_{\mathbb{R}} x\mathsf{P}(x)\,\mathrm{d}x && \text{by definition of } \mathsf{E} \text{ (Definition 1.40 page 24)} \\
&= \sum_{x\in\mathbb{Z}} x\mathsf{P}(x) && \text{by definition of } \mathsf{P} \\
&= 1\times\frac{1}{10} + 2\times\frac{1}{20} + 3\times\frac{1}{30} + 4\times\frac{3}{5} + 5\times\frac{1}{50} + 6\times\frac{1}{30} \\
&= \frac{1}{300}(1\times30 + 2\times15 + 3\times10 + 4\times180 + 5\times6 + 6\times10) = \frac{900}{300} = 3 \\
\ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) &= \mathsf{Var}(\mathsf{X}) && \text{by Theorem 2.1 page 49} \\
&\triangleq \int_{\mathbb{R}} [x-\mathsf{E}(\mathsf{X})]^2\mathsf{P}(x) && \text{by definition of } \mathsf{Var} \text{ (Definition 1.40 page 24)} \\
&= \sum_{x\in\mathbb{Z}} [x-\mathsf{E}(\mathsf{X})]^2\mathsf{P}(x) && \text{by definition of } \mathsf{P} \\
&= \frac{1}{10}(1-3)^2 + \frac{1}{20}(2-3)^2 + \frac{1}{30}(3-3)^2 + \frac{3}{5}(4-3)^2 + \frac{1}{50}(5-3)^2 + \frac{1}{30}(6-3)^2 \\
&= \frac{1}{300}(120 + 15 + 0 + 180 + 24 + 90) = \frac{429}{300} = \frac{143}{100} = 1.43 \\
\ddot{\mathsf{E}}(\mathsf{X}) &\triangleq \arg\min_{x\in\mathbb{R}} \max_{y\in\mathbb{R}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}
\end{aligned}$$

$$\begin{aligned}
&\triangleq \arg\min_{x\in\mathbb{R}} \max_{y\in\mathbb{R}} |x-y|\mathsf{P}(y) && \text{by definition standard metric on real line (Definition 1.35 page 22)}\\
&= \arg\min_{x\in\mathbb{R}} \begin{cases} |x-1|\frac{1}{10} & \text{for } \frac{25}{7} \le x \le \frac{23}{5} \\ |x-4|\frac{3}{5} & \text{otherwise} \end{cases} \\
&= \left\{\frac{25}{7}\right\} \approx \{3.5714\} && \text{because } \mathsf{f}(x) \text{ is minimized at argument } x = \frac{25}{7}\\
\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) &\triangleq \sum_{x\in\mathbb{R}} \mathsf{d}^2\big(\ddot{\mathsf{E}}(\mathsf{X}), x\big)\,\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\mathbb{R}} \mathsf{d}^2\left(\frac{25}{7}, x\right)\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{X}) \text{ result}\\
&= \left(\frac{25}{7}-1\right)^2\frac{1}{10} + \left(\frac{25}{7}-2\right)^2\frac{1}{20} + \left(\frac{25}{7}-3\right)^2\frac{1}{30} + \left(\frac{25}{7}-4\right)^2\frac{3}{5} + \left(\frac{25}{7}-5\right)^2\frac{1}{50} + \left(\frac{25}{7}-6\right)^2\frac{1}{30}\\
&= \frac{16805}{49*300} = \frac{3361}{49*60} = \frac{3361}{2940} \approx 1.143\\
\ddot{\mathsf{E}}(\mathsf{Y}) &= \mathsf{Y}\big(\dot{\mathsf{C}}(\boldsymbol{G})\big) && \text{because } \boldsymbol{G} \text{ and } \boldsymbol{H} \text{ are } \textit{isomorphic}\\
&= \mathsf{Y}(\{⚃\}) && \text{by Example 2.3 (page 35)}\\
&= \{4\} && \text{by definition of } \mathsf{Y}\\
\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) &= \dot{\mathsf{Var}}(\boldsymbol{G}) = \frac{101}{300} \approx 0.337 && \text{by Example 2.3 (page 35)}
\end{aligned}$$



✑

Spinner examples

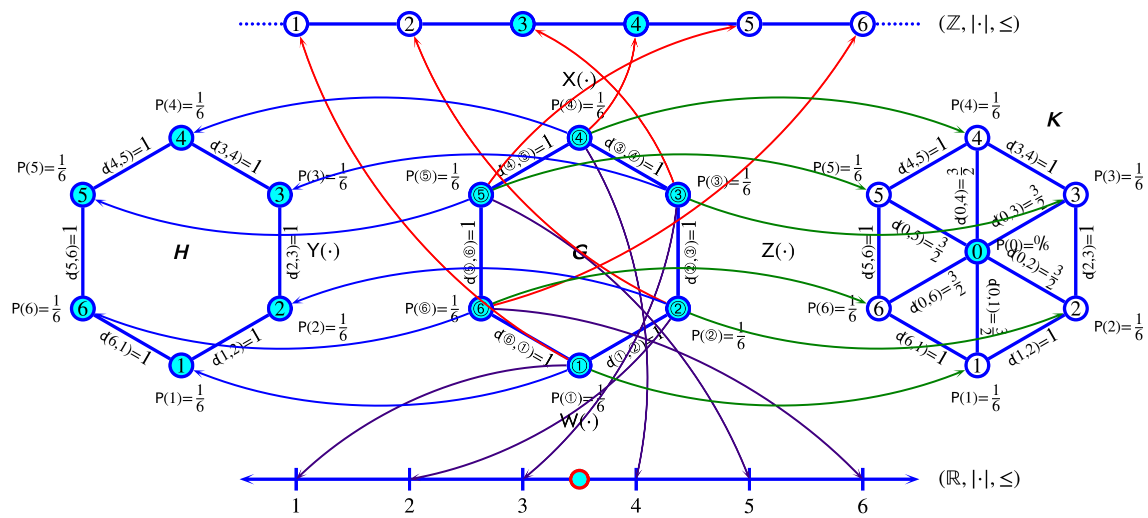


Figure 2.6: Six value fair spinner with assorted random variable mappings (Example 2.15 page 58)

Example 2.15 (spinner mappings). A six value board game spinner has a cyclic structure as illustrated in Figure 2.6 (page 58). Again, the order and metric geometry of the real line mapped to by the random variable X is very dissimilar to that of the *outcome subspace* that it is supposed to represent. Therefore, statistical inferences based on X will likely result in values that are arguably unacceptable. Both random variables Y and Z map to structures in which order and metric geometry are preserved. The mappings yield the following statistics:

geometry of $\boldsymbol{G}$:	$\dot{\mathsf{C}}(\boldsymbol{G})$	$=$	$\{①, ②, ③, ④, ⑤, ⑥\}$				
traditional statistics on real line:	$\mathsf{E}(\mathsf{W})$	$=$	3.5	$\ddot{\mathsf{Var}}(\mathsf{W};\mathsf{E})$	$=$	$\frac{35}{12}$	≈ 2.917
outcome subspace statistics on real line:	$\ddot{\mathsf{E}}(\mathsf{W})$	$=$	$\{3.5\}$	$\ddot{\mathsf{Var}}(\mathsf{W};\ddot{\mathsf{E}})$	$=$	$\frac{35}{12}$	≈ 2.917
outcome subspace statistics on integer line:	$\ddot{\mathsf{E}}(\mathsf{X})$	$=$	$\{3,\ 4\}$	$\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}})$	$=$	$\frac{20}{12}$	≈ 1.667
outcome subspace stats. on isomorphic structure:	$\ddot{\mathsf{E}}(\mathsf{Y})$	$=$	$\{1,2,3,4,5,6\}$	$\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}})$	$=$	0	
outcome subspace stats. on extended structure:	$\ddot{\mathsf{E}}(\mathsf{Z})$	$=$	$\{0\}$	$\ddot{\mathsf{Var}}(\mathsf{Z};\ddot{\mathsf{E}})$	$=$	$\frac{9}{4}$	$= 2.25$

✎PROOF:

$\dot{\mathsf{C}}(\boldsymbol{G}) = \{①, ②, ③, ④, ⑤, ⑥\}$ — by Example 2.4 (page 37)

$\mathsf{E}(\mathsf{W}) \triangleq \sum_{x\in\mathbb{R}} x\mathsf{P}(x)$ — by definition of E (Definition 1.40 page 24)

$= \frac{7}{2} = 3.5$ — by *fair die* example Example 2.10 (page 52)

$\ddot{\mathsf{Var}}(\mathsf{W};\mathsf{E}) = \mathsf{Var}(\mathsf{X})$ — by Theorem 2.1 page 49

$\triangleq \sum_{x\in\mathbb{R}} [x - \mathsf{E}(\mathsf{X})]^2\mathsf{P}(x)\,\mathsf{d}x$ — by definition of Var (Definition 1.40 page 24)

$= \frac{35}{12} \approx 2.917$ — by *fair die* example Example 2.10 (page 52)

$\ddot{\mathsf{E}}(\mathsf{W}) = \mathsf{E}(\mathsf{W})$ — because on *real line*, P is *symmetric*, and by Theorem 2.2 page 49

$= \{3.5\}$ — by $\mathsf{E}(\mathsf{W})$ result

$\ddot{\mathsf{Var}}(\mathsf{W};\ddot{\mathsf{E}}) = \ddot{\mathsf{Var}}(\mathsf{W};\mathsf{E})$ — because $\ddot{\mathsf{E}}(\mathsf{W}) = \mathsf{E}(\mathsf{W})$

$= \frac{35}{12} \approx 2.917$ — by $\ddot{\mathsf{Var}}(\mathsf{W};\mathsf{E})$ result

$\ddot{\mathsf{E}}(\mathsf{X}) \triangleq \arg\min_{x\in\mathbb{Z}} \max_{y\in\mathbb{Z}} \mathsf{d}(x,y)\,\mathsf{P}(y)$ — by definition of $\ddot{\mathsf{E}}$ (Definition 2.14 page 49)

$\triangleq \arg\min_{x\in\mathbb{Z}} \max_{y\in\mathbb{Z}} |x-y|\frac{1}{6}$ — by definition of *integer line* (Definition 1.36 page 23) and $\boldsymbol{G}$

$= \{3,\ 4\}$ — by *fair die* example Example 2.10 (page 52)

$\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) = \frac{5}{3} \approx 1.667$ — by *fair die* example Example 2.10 (page 52)

$\ddot{\mathsf{E}}(\mathsf{Y}) = \mathsf{Y}\left[\dot{\mathsf{C}}(\boldsymbol{G})\right]$ — because $\boldsymbol{G}$ and H are *isomorphic* under mapping Y

$= \mathsf{Y}\left[\{①, ②, ③, ④, ⑤, ⑥\}\right]$ — by $\dot{\mathsf{C}}(\boldsymbol{G})$ result

$= \{1,2,3,4,5,6\}$ — by definition of Y

$\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) = \ddot{\mathsf{Var}}(\boldsymbol{G}) = 0$ — by *spinner outcome subspace* example (Example 2.4 page 37)

$\ddot{\mathsf{E}}(\mathsf{Z}) \triangleq \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}} \mathsf{d}(x,y)\,\mathsf{P}(y)$ — by definition of $\ddot{\mathsf{E}}$ (Definition 2.14 page 49)

$= \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}\setminus\{0\}} \mathsf{d}(x,y)\,\mathsf{P}(y)$ — because $\mathsf{P}(0) = 0$

$= \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}\setminus\{0\}} \mathsf{d}(x,y)\,\frac{1}{6}$ — by definition of $\boldsymbol{G}$

$= \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}\setminus\{0\}} \mathsf{d}(x,y)$ — because $\mathsf{f}(x) = \frac{1}{6}x$ is *strictly isotone* and by Lemma 1.7 (page 19)

$$= \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}} \left\{\begin{matrix} \mathsf{d}(1,1)\cdots\mathsf{d}(1,6)\\ \mathsf{d}(2,1)\cdots\mathsf{d}(2,6)\\ \mathsf{d}(3,1)\cdots\mathsf{d}(3,6)\\ \mathsf{d}(4,1)\cdots\mathsf{d}(4,6)\\ \mathsf{d}(5,1)\cdots\mathsf{d}(5,6)\\ \mathsf{d}(6,1)\cdots\mathsf{d}(6,6)\\ \mathsf{d}(0,1)\cdots\mathsf{d}(0,6)\end{matrix}\right\} = \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}\setminus\{0\}} \left\{\begin{matrix} 0&1&2&3&2&1\\ 1&0&1&2&3&2\\ 2&1&0&1&2&3\\ 3&2&1&0&1&2\\ 2&3&2&1&0&1\\ 1&2&3&2&1&0\\ \frac{3}{2}&\frac{3}{2}&\frac{3}{2}&\frac{3}{2}&\frac{3}{2}&\frac{3}{2}\end{matrix}\right\} = \arg\min_{x\in\mathsf{H}} \left\{\begin{matrix}3\\3\\3\\3\\3\\3\\\frac{3}{2}\end{matrix}\right\} = \left\{\begin{matrix}\\ \\ \\ \\ \\ \\ 0\end{matrix}\right\}$$

$\ddot{\mathsf{Var}}(\mathsf{Z}) \triangleq \sum_{x\in\mathsf{H}} \mathsf{d}^2\left(\dot{\mathsf{C}}(\boldsymbol{G}), x\right)\,\mathsf{P}(x)$ — by definition of $\ddot{\mathsf{Var}}$ (Definition 2.14 page 49)

$$\begin{aligned}
&= \sum_{x\in \mathsf{H}} \mathsf{d}^2(\{0\},x)\,\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{X}) \text{ result}\\
&= \sum_{x\in \mathsf{H}\setminus\{0\}} \left(\frac{3}{2}\right)^2\frac{1}{6} = |\,\mathsf{H}\setminus\{0\}\,|\left(\frac{3}{2}\right)^2\frac{1}{6} = 6\left(\frac{3}{2}\right)^2\frac{1}{6} = \frac{9}{4}
\end{aligned}$$

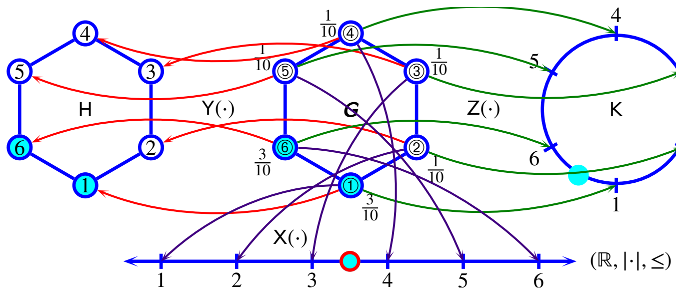


Figure 2.7: weighted spinner mappings (Example 2.16 page 60)

Example 2.16 (weighted spinner mappings). Let $\boldsymbol{G}$ be *weighted spinner outcome subspace* (Example 2.5 page 39) with random variable mappings as illustrated in Figure 2.7 (page 60). This yields the following statistics:

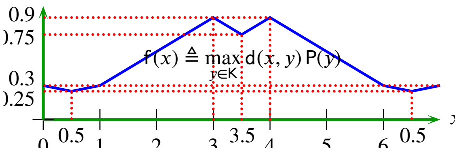
geometry of $\boldsymbol{G}$:	$\dot{\mathsf{C}}(\boldsymbol{G})$	$=$	$\{①, ⑥\}$					
traditional statistics on real line $(\mathbb{R}, \lvert\cdot\rvert, \le)$:	$\mathsf{E}(\mathsf{X})$	$=$	3.5	$\ddot{\mathsf{Var}}(\mathsf{W};\mathsf{E})$	$=$	$\frac{17}{4}$	$\approx$	4.25
outcome subspace statistics on real line $(\mathbb{R}, \lvert\cdot\rvert, \le)$:	$\ddot{\mathsf{E}}(\mathsf{X})$	$=$	$\{3.5\}$	$\ddot{\mathsf{Var}}(\mathsf{W};\ddot{\mathsf{E}})$	$=$	$\frac{17}{4}$	$\approx$	4.25
outcome subspace statistics on isomorphic structure $\boldsymbol{H}$:	$\ddot{\mathsf{E}}(\mathsf{Y})$	$=$	$\{1, 6\}$	$\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}})$	$=$	$\frac{5}{3}$	$\approx$	1.67
outcome subspace statistics on continuous structure $\boldsymbol{K}$:	$\ddot{\mathsf{E}}(\mathsf{Z})$	$=$	$\{0.5\}$	$\ddot{\mathsf{Var}}(\mathsf{Z};\ddot{\mathsf{E}})$	$=$	$\frac{37}{20}$	$=$	1.85

Note that based on the variance values, the statistic $\ddot{\mathsf{E}}(\mathsf{Z})$ on the continuous ring $\boldsymbol{K}$ is arguably a much better statistic than $\ddot{\mathsf{E}}(\mathsf{X})$ on the (continuous) real line $(\mathbb{R}, |\cdot|, \le)$.

✎PROOF:

$$\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &= \{1, 6\} && \text{by \textit{weighted spinner outcome subspace} example (Example 2.5 page 39)}\\
\mathsf{E}(\mathsf{X}) &\triangleq \sum_{x\in\mathbb{Z}} x\,\mathsf{P}(x)\\
&= 1\times\frac{3}{10} + 2\times\frac{1}{10} + 3\times\frac{1}{10} + 4\times\frac{1}{10} + 5\times\frac{1}{10} + 6\times\frac{1}{10}\\
&= \frac{1}{10}(1\times 3 + 2 + 3 + 4 + 5 + 6\times 3) = \frac{35}{10} = \frac{7}{2} = 3.5\\
\ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) &= \mathsf{Var}(\mathsf{X}) && \text{by Theorem 2.1 page 49}\\
&\triangleq \sum_{x\in\mathbb{Z}} [x - \mathsf{E}(\mathsf{X})]^2\mathsf{P}(x) && \text{by definition of } \mathsf{Var} \text{ (Definition 1.40 page 24)}\\
&= \left(1-\frac{7}{2}\right)^2\frac{3}{10} + \left(2-\frac{7}{2}\right)^2\frac{1}{10} + \left(3-\frac{7}{2}\right)^2\frac{1}{10} + \left(4-\frac{7}{2}\right)^2\frac{1}{10} + \left(5-\frac{7}{2}\right)^2\frac{1}{10} + \left(6-\frac{7}{2}\right)^2\frac{3}{10}\\
&= \frac{1}{10}\left[\left(-\frac{5}{2}\right)^2\times 3 + \left(-\frac{3}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{3}{2}\right)^2 + \left(\frac{5}{2}\right)^2\times 3\right]\\
&= \frac{1}{40}[75 + 9 + 1 + 1 + 9 + 75] = \frac{170}{40} = \frac{17}{4} = 4.25\\
\ddot{\mathsf{E}}(\mathsf{X}) &= \mathsf{E}(\mathsf{X}) && \text{because on \textit{real line}, } \mathsf{P} \text{ is \textit{symmetric}, and by Theorem 2.2 page 49}\\
&= \{3.5\} && \text{by } \mathsf{E}(\mathsf{X}) \text{ result}\\
\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) &= \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) && \text{by } \mathsf{E}(\mathsf{X}) \text{ and } \ddot{\mathsf{E}}(\mathsf{X}) \text{ results}\\
&= \frac{17}{4} = 4.25 && \text{by } \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) \text{ result}
\end{aligned}$$

$$\begin{aligned}
\ddot{\mathsf{E}}(\mathsf{Y}) &= \mathsf{Y}\left[\dot{\complement}(\boldsymbol{G})\right] && \text{because } \boldsymbol{G} \text{ and } \mathsf{H} \text{ are } \textit{isomorphic} \text{ under } \mathsf{Y}\\
&= \mathsf{Y}\left[\{①, ⑥\}\right] && \text{by } \dot{\complement}(\boldsymbol{G}) \text{ result}\\
&= \{1, 6\} && \text{by definition of } \mathsf{Y}\\
\ddot{\mathsf{Var}}(\mathsf{Y}; \ddot{\mathsf{E}}) &= \dot{\mathsf{Var}}(\boldsymbol{G}) && \text{because } \boldsymbol{G} \text{ and } \mathsf{H} \text{ are } \textit{isomorphic} \text{ under } \mathsf{Y}\\
&= \frac{5}{3} \approx 1.667 && \text{by } \textit{weighted spinner outcome subspace} \text{ example (Example 2.5 page 39)}\\
\ddot{\mathsf{E}}(\mathsf{Z}) &\triangleq \arg\min_{x\in\mathsf{K}} \max_{y\in\mathsf{K}} \mathsf{d}(x, y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}
\end{aligned}$$



$$\begin{aligned}
&= 0.5\\
\ddot{\mathsf{Var}}(\mathsf{Z}; \ddot{\mathsf{E}}) &\triangleq \sum_{x\in\boldsymbol{K}} \mathsf{d}^2\left(\ddot{\mathsf{E}}(\mathsf{Z}), x\right)\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\boldsymbol{K}} \mathsf{d}^2\left(\frac{1}{2}, x\right)\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{Z}) \text{ result}\\
&= \left(\frac{1}{2}\right)^2\frac{3}{10} + \left(\frac{3}{2}\right)^2\frac{1}{10} + \left(\frac{5}{2}\right)^2\frac{1}{10} + \left(\frac{5}{2}\right)^2\frac{1}{10} + \left(\frac{3}{2}\right)^2\frac{1}{10} + \left(\frac{1}{2}\right)^2\frac{3}{10}\\
&= \frac{1}{40}(3 + 9 + 25 + 25 + 9 + 3) = \frac{74}{40} = \frac{37}{20} = 1.85
\end{aligned}$$

✐

Pseudo-random number generator (PRNG) examples

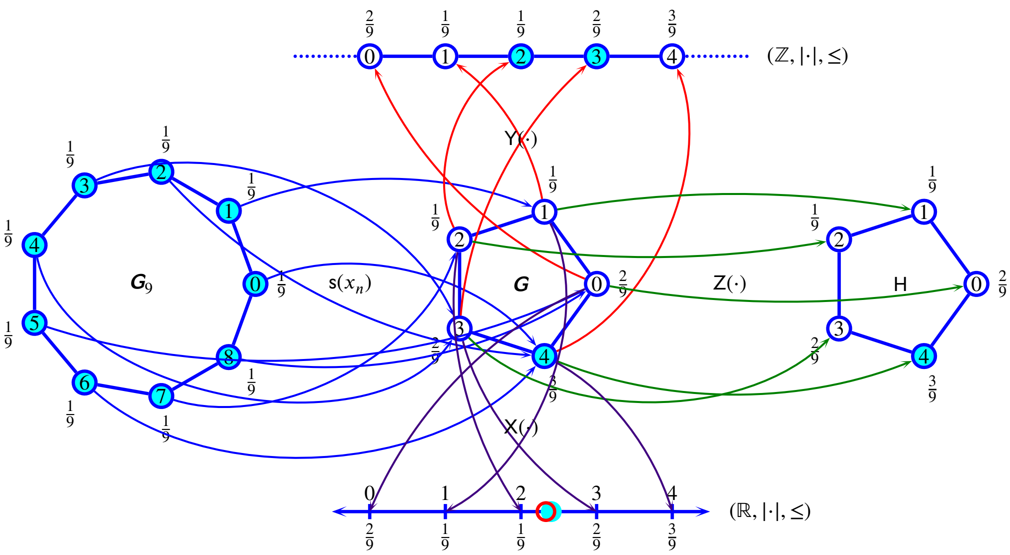


Figure 2.8: LCG mappings to *linear* (X), non-linear discrete (Y) and non-linear continuous (Z) ordered metric spaces (Example 2.17 page 61)

Example 2.17 (LCG mappings, standard ordering).
The equation $x_{n+1} = (7x_n + 5) \mod 9$ with $x_0 = 1$ is a *linear congruential* (LCG) *pseudo-random number generator* (PRNG) that has *full period*[5] of 9 values. These 9 values can be mapped, using a

[5] Hull and Dobell (1962), Jr. and Gentle (1980) page 137 ⟨Theorem 6.1⟩, Severance (2001) page 86 ⟨Hull-Dobell Theorem⟩

surjective (Definition 1.14 page 7) function $\mathsf{s} \in \boldsymbol{G}^{\boldsymbol{G}_9}$ to the 5 element set $\{0, 1, 2, 3, 4\}$ to "shape" the distribution from a *uniform* distribution to *non-uniform*:[6].

n	0	1	2	3	4	5	6	7	8	9	10	11	$\cdots$
x_n	1	3	8	7	0	5	4	6	2	1	3	8	$\cdots$
$y_n \triangleq \mathsf{s}(x_n)$	1	3	0	2	4	1	3	4	4	1	3	0	$\cdots$

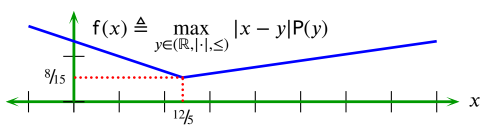
Let $\boldsymbol{G}$ be the *outcome subspace* and X, Y, and Z be the *outcome random variables* illustrated in Figure 2.8 (page 61). This yields the following statistics:

$$\begin{array}{lrclrcll}
\text{geometry of } \boldsymbol{G}_9\text{:} & \dot{\mathsf{C}}(\boldsymbol{G}) & = & \{0,1,2,3,4,5,6,7,8\} & & & & \\
\text{geometry of } \boldsymbol{G}\text{:} & \dot{\mathsf{C}}(\boldsymbol{G}) & = & \{4\} & & & & \\
\text{traditional statistics on real line:} & \mathsf{E}(\mathsf{X}) & = & \frac{7}{3} \approx 2.333 & \ddot{\mathsf{Var}}(\mathsf{X}; \mathsf{E}) & = & \frac{22}{9} & \approx 2.444 \\
\text{outcome subspace statistics on real line:} & \ddot{\mathsf{E}}(\mathsf{X}) & = & \left\{\frac{12}{5} = 2.4\right\} & \ddot{\mathsf{Var}}(\mathsf{X}; \ddot{\mathsf{E}}) & = & \frac{551}{225} & \approx 2.449 \\
\text{outcome subspace statistics on integer line:} & \ddot{\mathsf{E}}(\mathsf{Y}) & = & \{2,\ 3\} & \ddot{\mathsf{Var}}(\mathsf{Y}; \ddot{\mathsf{E}}) & = & \frac{16}{9} & \approx 1.778 \\
\text{outcome subspace statistics on isomorphic structure:} & \ddot{\mathsf{E}}(\mathsf{Z}) & = & \{4\} & \ddot{\mathsf{Var}}(\mathsf{Z}; \ddot{\mathsf{E}}) & = & \frac{4}{3} & \approx 1.333
\end{array}$$

Note that unlike the statistics $\mathsf{E}(\mathsf{X})$ and $\ddot{\mathsf{E}}(\mathsf{X})$ on the *real line*, the statistic $\ddot{\mathsf{E}}(\mathsf{Z})$ on the *isomorphic* structure $\boldsymbol{K}$ yields the *maximally likely* result, and a much smaller variance as well.

✎PROOF:

$$\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}_9) &\triangleq \arg\min_{x \in \boldsymbol{G}_9} \max_{y \in \boldsymbol{G}_9} \mathsf{d}(x, y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)} \\
&= \arg\min_{x \in \boldsymbol{G}_9} \max_{y \in \boldsymbol{G}_9} \mathsf{d}(x, y)\,\frac{1}{9} && \text{by definition of } \boldsymbol{G}_9 \\
&= \arg\min_{x \in \boldsymbol{G}_9} \max_{y \in \boldsymbol{G}_9} \mathsf{d}(x, y) && \text{because } \phi(x) = \frac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.7 page 19} \\
&= \arg\min_{x \in \boldsymbol{G}_9} \{4,\ 4,\ 4,\ 4,\ 4,\ 4,\ 4,\ 4,\ 4\} && \text{because the maximum distance in } \boldsymbol{G}_9 \text{ from any } x \text{ is } 4 \\
&= \{0, 1, 2, \cdots, 8\} && \text{because the distances for values of } x \text{ in } \boldsymbol{G}_9 \text{ are the same} \\
\dot{\mathsf{C}}(\boldsymbol{G}) &= \{4\} && \text{by \textit{weighted ring outcome subspace} example (Example 2.6 page 41)} \\
\mathsf{E}(\mathsf{X}) &\triangleq \sum_{x \in \mathbb{R}} x\mathsf{P}(x) && \text{by definition of } \mathsf{E} \text{ (Definition 1.40 page 24)} \\
&= 0 \times \frac{2}{9} + 1 \times \frac{1}{9} + 2 \times \frac{1}{9} + 3 \times \frac{2}{9} + 4 \times \frac{3}{9} = \frac{21}{9} = \frac{7}{3} \approx 2.333 \\
\ddot{\mathsf{Var}}(\mathsf{X}; \mathsf{E}) &= \mathsf{Var}(\mathsf{X}) && \text{by Theorem 2.1 page 49} \\
&\triangleq \sum_{x \in \mathbb{R}} [x - \mathsf{E}(\mathsf{X})]^2 \mathsf{P}(x) && \text{by definition of } \mathsf{Var} \text{ (Definition 1.40 page 24)} \\
&= \left(0 - \frac{7}{3}\right)^2 \frac{2}{9} + \left(1 - \frac{7}{3}\right)^2 \frac{1}{9} + \left(2 - \frac{7}{3}\right)^2 \frac{1}{9} + \left(3 - \frac{7}{3}\right)^2 \frac{2}{9} + \left(4 - \frac{7}{3}\right)^2 \frac{3}{9} = \frac{198}{81} = \frac{22}{9} \approx 2.457 \\
\ddot{\mathsf{E}}(\mathsf{X}) &\triangleq \arg\min_{x \in \mathbb{R}} \max_{y \in \mathbb{R}} \mathsf{d}(x, y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)} \\
&\triangleq \arg\min_{x \in \mathbb{R}} \max_{y \in \mathbb{R}} |x - y|\mathsf{P}(y) && \text{by definition usual metric on real line (Definition 1.35 page 22)}
\end{aligned}$$



$$\begin{aligned}
&= \arg\min_{x \in \mathbb{R}} \begin{cases} |x - 4|\mathsf{P}(4) & \text{for } x \leq \frac{12}{5} \\ |x - 0|\mathsf{P}(0) & \text{otherwise} \end{cases} \\
&= \left\{\frac{12}{5}\right\} = 2.4 && \text{because expression is minimized at argument } x = \frac{12}{5} \\
\ddot{\mathsf{Var}}(\mathsf{X}; \ddot{\mathsf{E}}) &\triangleq \sum_{x \in \mathbb{R}} \mathsf{d}^2\left(\ddot{\mathsf{E}}(\mathsf{X}), x\right) \mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}
\end{aligned}$$

[6]The sequence $⟅1, 3, 0, 2, 4,\ 1, 3, 0, 2, 4,\ 1, 3, \cdots⟆$ is generated by the equation $y_{n+1} = (y_n + 2) \mod 5$ with $y_0 = 1$

$$
\begin{aligned}
&= \sum_{x\in\mathbb{R}} \mathsf{d}^2\left(\frac{12}{5}, x\right)\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{X}) \text{ result}\\
&= \left(\frac{12}{5}-0\right)^2\frac{2}{9} + \left(\frac{12}{5}-1\right)^2\frac{1}{9} + \left(\frac{12}{5}-2\right)^2\frac{1}{9} + \left(\frac{12}{5}-3\right)^2\frac{2}{9} + \left(\frac{12}{5}-4\right)^2\frac{3}{9}\\
&= \frac{1}{25\times 9}2(12-0)^2 + 1(12-5)^2 + 1(12-10)^2 + 2(12-15)^2 + 3(12-20)^2 = \frac{551}{225} \approx 2.449
\end{aligned}
$$

$$
\begin{aligned}
\ddot{\mathsf{E}}(\mathsf{Y})&\\
&\triangleq \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}} |x-y|\mathsf{P}(y) && \text{by definition of } \textit{integer line} \text{ (Definition 1.36 page 23)}\\
&= \arg\min_{x\in\mathsf{H}}\max_{y\in\mathsf{H}} \frac{1}{9}\begin{Bmatrix} 0\times 2 & 1\times 1 & 2\times 1 & 3\times 2 & 4\times 3\\ 1\times 2 & 0\times 1 & 1\times 1 & 2\times 2 & 3\times 3\\ 2\times 2 & 1\times 1 & 0\times 1 & 1\times 2 & 2\times 3\\ 3\times 2 & 2\times 1 & 1\times 1 & 0\times 2 & 1\times 3\\ 4\times 2 & 3\times 1 & 2\times 1 & 1\times 2 & 0\times 3 \end{Bmatrix} = \arg\min_{x\in\mathsf{H}} \frac{1}{9}\begin{Bmatrix}12\\9\\6\\6\\8\end{Bmatrix} = \arg\min_{x\in\mathsf{H}} \frac{1}{9}\begin{Bmatrix}2\\3\end{Bmatrix}\\
\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) &\triangleq \sum_{x\in\mathbb{Z}} \mathsf{d}^2\big(\ddot{\mathsf{E}}(\mathsf{Y}), x\big)\,\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\mathbb{Z}} \mathsf{d}^2(\{2,3\}, x)\,\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{Y}) \text{ result}\\
&= 2^2\times\frac{2}{9} + 1^2\times\frac{1}{9} + 0^2\times\frac{1}{9} + 0^2\times\frac{2}{9} + 1^2\times\frac{3}{9} = \frac{16}{9} \approx 1.778\\
\ddot{\mathsf{E}}(\mathsf{Z}) &= \mathsf{Z}\big[\dot{\mathsf{C}}(\boldsymbol{G})\big] && \text{because } \boldsymbol{G} \text{ and } \mathsf{H} \text{ are } \textit{isomorphic} \text{ under } \mathsf{Z}\\
&= \{4\} && \text{by } \dot{\mathsf{C}}(\boldsymbol{G}) \text{ result}\\
\ddot{\mathsf{Var}}(\mathsf{Z};\ddot{\mathsf{E}}) &\triangleq \sum_{x\in\mathsf{H}} \mathsf{d}^2\big(\ddot{\mathsf{E}}(\mathsf{Z}), x\big)\,\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\mathsf{H}} \mathsf{d}^2(\{4\}, x)\,\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{Z}) \text{ result}\\
&= 1^2\times\frac{2}{9} + 2^2\times\frac{1}{9} + 2^2\times\frac{1}{9} + 1^2\times\frac{2}{9} + 0^2\times\frac{3}{9} = \frac{12}{9} = \frac{4}{3} \approx 1.333
\end{aligned}
$$

✏

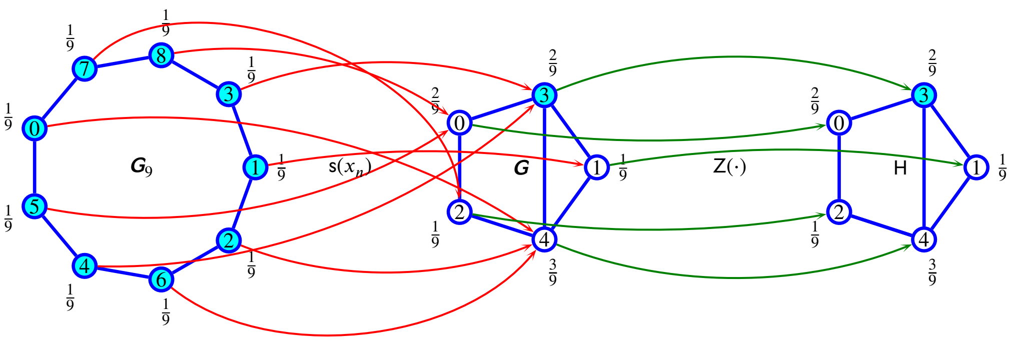


Figure 2.9: sequentially ordered LCG mappings (Example 2.18 page 63)

Example 2.18 (LCG mappings, sequential ordering).
In Example 2.17 (page 61), the structures $\boldsymbol{G}_9$, $\boldsymbol{G}$, and H were ordered as a standard ring of integers ($0 < 1 < 2 < \cdots < 7 < 8 < 0$ for $\boldsymbol{G}_9$). In this current example, as illustrated in Figure 2.9 (page 63), these structures are ordered as they appear in the sequences generated by $x_{n+1} = (7x_n + 5) \mod 9$ and s (see Example 2.17 for sequence description). This yields the following statistics:

geometry of $\boldsymbol{G}_9$:	$\dot{\mathsf{C}}(\boldsymbol{G}) = \{0,1,2,3,4,5,6,7,8\}$			
geometry of $\boldsymbol{G}$:	$\dot{\mathsf{C}}(\boldsymbol{G}) = \{3\}$			
outcome subspace statistics on isomorphic structure:	$\ddot{\mathsf{E}}(\mathsf{Z}) = \{3\}$	$\ddot{\mathsf{Var}}(\mathsf{Z};\ddot{\mathsf{E}}) = \frac{10}{9}$	$\approx$	1.111

Note that a change in ordering structure (from standard ring ordering to sequential ordering) yields a change in statistics ($\ddot{\mathsf{E}}(\mathsf{Z}) = \{3\}$ as opposed to $\ddot{\mathsf{E}}(\mathsf{Z}) = \{4\}$). Intuitively, the sequential ordering of Example 2.18 should yield a better estimate than that of Example 2.17, because it more closely matches the way the PRNG produces a sequence. This intuition is also supported by the variance values ($\ddot{\mathsf{Var}}(\mathsf{Z}) = {}^{12}/_{9}$ for standard ring ordering, $\ddot{\mathsf{Var}}(\mathsf{Z}) = \frac{10}{9}$ for sequential ordering). However, counterintuitively, the sequential ordering no longer yields the maximally likely result of $\{4\}$.

✎PROOF:

$$\begin{aligned}
&\dot{\mathsf{C}}(\boldsymbol{G}_9) \\
&\quad = \{0, 1, 2, \cdots, 8\} && \text{by } \textit{LCG mappings standard ordering} \text{ example (Example 2.17 page 61)} \\
\dot{\mathsf{C}}(\boldsymbol{G}) &= \{3\} && \text{by Example 2.7 (page 43)} \\
\ddot{\mathsf{E}}(\mathsf{Z}) &= \mathsf{Z}\left[\dot{\mathsf{C}}(\boldsymbol{G})\right] && \text{because } \boldsymbol{G} \text{ and } \mathsf{H} \text{ are } \textit{isomorphic} \text{ under } \mathsf{Z} \\
&= \mathsf{Z}[\{3\}] && \text{by } \dot{\mathsf{C}}(\boldsymbol{G}) \text{ result} \\
&= \{3\} && \text{by definition of } \mathsf{Z} \\
\ddot{\mathsf{Var}}(\mathsf{Z}; \ddot{\mathsf{E}}) &= \ddot{\mathsf{Var}}(\boldsymbol{G}) && \text{because } \boldsymbol{G} \text{ and } \mathsf{H} \text{ are } \textit{isomorphic} \text{ under } \mathsf{Z} \\
&= \frac{10}{9} \approx 1.111 && \text{by Example 2.7 (page 43)}
\end{aligned}$$

✏

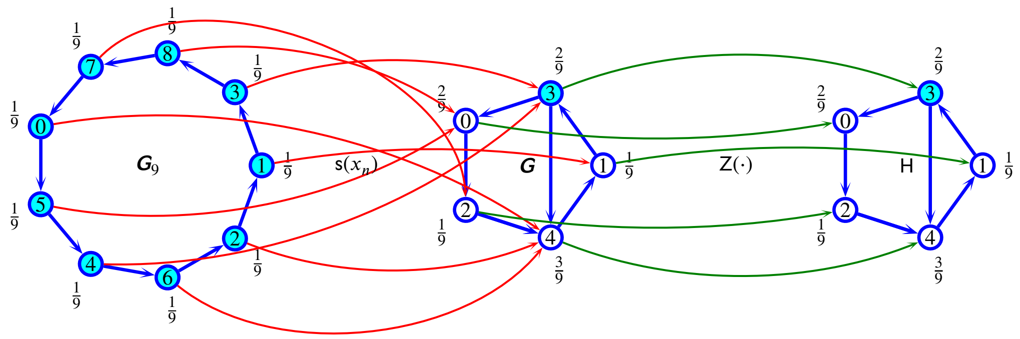


Figure 2.10: LCG mappings to *linear* (X), non-linear discrete (Y) and non-linear continuous (Z) ordered metric spaces (Example 2.19 page 64)

Example 2.19 (LCG mappings, sequential directed graph).
Let $\boldsymbol{G}$, H and Z be a illustrated in Figure 2.10 (page 64). In Example 2.18 (page 63), the outcome values were ordered sequentially *like* a PRNG, but the metrics were *commutative*, which is *unlike* a PRNG. In this example, the outcomes are assigned *quasi-metrics* (Definition D.6 page 163, Remark 1.5 page 22) that are *non-commutative*. For example in the shaped sequence $\mathsf{s}(x_n) = (\!(\ldots, 3, 4, 4, 1, 3, \ldots)\!)$, the "distance" from 3 to 4 is $\mathsf{d}(3,4) = 1$, but from 4 to 3 is $\mathsf{d}(4,3) = 2$. This yields the following statistics:

geometry of $\boldsymbol{G}_9$:	$\dot{\mathsf{C}}(\boldsymbol{G})$	$=$	$\{0, 1, 2, 3, 4, 5, 6, 7, 8\}$				
geometry of $\boldsymbol{G}$:	$\dot{\mathsf{C}}(\boldsymbol{G})$	$=$	$\{3\}$				
outcome subspace statistics on isomorphic structure:	$\ddot{\mathsf{E}}(\mathsf{Z})$	$=$	$\{3\}$	$\ddot{\mathsf{Var}}(\mathsf{Z}; \ddot{\mathsf{E}})$	$=$	$\frac{4}{3}$	≈ 1.333

Note that this technique yields the same estimate $\ddot{\mathsf{E}}(\mathsf{Z}) = \{3\}$ as Example 2.18, but with a larger variance.

✎PROOF:

$$\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}_9) &\triangleq \arg\min_{x \in \boldsymbol{G}_9} \max_{y \in \boldsymbol{G}_9} \mathsf{d}(x, y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)} \\
&= \arg\min_{x \in \boldsymbol{G}_9} \max_{y \in \boldsymbol{G}_9} \mathsf{d}(x, y)\,\frac{1}{9} && \text{by definition of } \boldsymbol{G}_9
\end{aligned}$$

$$
\begin{aligned}
&= \arg\min_{x\in G_9}\max_{y\in G_9} \mathsf{d}(x,y) && \text{because } \phi(x)=\tfrac{1}{9}x \text{ is \textit{strictly isotone} and by Lemma 1.7 page 19}\\
&= \arg\min_{x\in G_9}\{8,8,8,8,8,8,8,8,8\} && \text{because the maximum distance in } G_9 \text{ from any } x \text{ is } 8\\
&= \{0,1,2,\cdots,8\} && \text{because the distances for values of } x \text{ in } G_9 \text{ are the same}\\
\dot{\mathsf{C}}(G) &= \{3\} && \text{by Example 2.8 (page 45)}\\
\ddot{\mathsf{E}}(\mathsf{Z}) &= \mathsf{Z}\left[\dot{\mathsf{C}}(G)\right] && \text{because } G \text{ and H are \textit{isomorphic} under Z}\\
&= \mathsf{Z}[\{3\}] && \text{by } \dot{\mathsf{C}}(G) \text{ result}\\
&= \{3\} && \text{by definition of Z}\\
\ddot{\mathsf{Var}}(\mathsf{Z};\ddot{\mathsf{E}}) &= \dot{\mathsf{Var}}(\dot{\mathsf{C}}(G)) && \text{because } G \text{ and H are \textit{isomorphic} under Z}\\
&= \frac{4}{3} \approx 1.333 && \text{by Example 2.8 (page 45)}
\end{aligned}
$$

✎

Genomic signal processing (GSP) examples

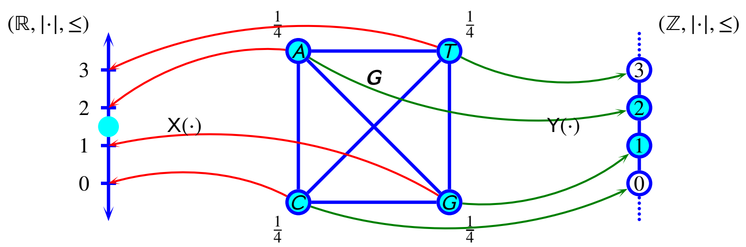


Figure 2.11: DNA random variable mappings to *real line* and *integer line* (Example 2.20 page 65)

Example 2.20 (DNA to linear structures). *Genomic Signal Processing* (*GSP*) analyzes biological sequences called *genome*s. These sequences are constructed over a set of 4 symbols that are commonly referred to as A, T, C, and G, each of which corresponds to a nucleobase (adenine, thymine, cytosine, and guanine, respectively).[7] A typical genome sequence contains a large number of symbols (about 3 billion for humans, 29751 for the SARS virus).[8] Let $G \triangleq (\{A, T, C, G\}, \mathsf{d}, \emptyset, \mathsf{P})$ be the *outcome subspace* (Definition 2.1 page 29) generated by a *genome* where d is the *discrete metric* (Definition D.8 page 166), $\leq = \emptyset$ indicates a completely unordered set (Definition 1.20 page 9), and $\mathsf{P}(A) = \mathsf{P}(T) = \mathsf{P}(C) = \mathsf{P}(G) = \frac{1}{4}$ (uniformly distributed). Let $\mathsf{H} \triangleq (\mathbb{R}, |\cdot|, \leq)$ be the *real line* (Definition 1.35 page 22). This yields the following statistics:

geometry of G:	$\dot{\mathsf{C}}(G)$	=	$\{A, T, C, G\}$					
traditional statistics on real line:	$\mathsf{E}(\mathsf{X})$	=	1.5	$\dot{\mathsf{Var}}(\mathsf{X};\mathsf{E})$	=	$\frac{5}{4}$	=	1.25
outcome subspace statistics on real line:	$\ddot{\mathsf{E}}(\mathsf{X})$	=	1.5	$\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}})$	=	$\frac{5}{4}$	=	1.25
outcome subspace statistics on integer line:	$\ddot{\mathsf{E}}(\mathsf{Y})$	=	$\{1,2\}$	$\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}})$	=	$\frac{1}{2}$	=	0.5

The symbols A, T, C and G in general again have an order structure and a *metric geometry* (Remark 2.2 page 34) that is fundamentally dissimilar from that mapped to by the random variables X and Y.

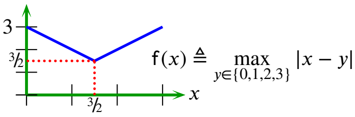
[7] Mendel (1853) ⟨Mendel (1853): gene coding uses discrete symbols⟩, Watson and Crick (1953a) page 737 ⟨Watson and Crick (1953): gene coding symbols are adenine, thymine, cytosine, and guanine⟩, Watson and Crick (1953b) page 965, Pommerville (2013) page 52

[8] GenBank (2014) ⟨http://www.ncbi.nlm.nih.gov/genome/guide/human/⟩⟨Homo sapiens, NC_000001–NC_000022 (22 chromosome pairs), NC_000023 (X chromosome), NC_000024 (Y chromosome), NC_012920 (mitochondria)⟩, GenBank (2014) ⟨http://www.ncbi.nlm.nih.gov/nuccore/30271926⟩⟨SARS coronavirus, NC_004718.3⟩ S. G. Gregory (2006) ⟨homo sapien chromosome 1⟩, Runtao He (2004) ⟨SARS coronavirus⟩

Therefore, statistical inferences made using these random variables will likely lead to results that arguably have little relationship with intuition or reality.

✎PROOF:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &= \{\boxed{A}, \boxed{T}, \boxed{C}, \boxed{G}\} && \text{by Example 2.9 page 47}\\
\mathsf{E}(\mathsf{X}) &\triangleq \int_{\mathbb{R}} x\mathsf{P}(x)\,\mathrm{d}x && \text{by definition of } \mathsf{E} \text{ (Definition 1.40 page 24)}\\
&= \sum_{x\in\mathbb{Z}} x\mathsf{P}(x) && \text{by definition of } \mathsf{P} \text{ and Proposition 1.2 page 24}\\
&= \frac{1}{4}\sum_{x\in\{0,1,2,3\}} x && \text{by definitions of } \boldsymbol{G}, \mathsf{H} \text{ and } \mathsf{X}\\
&= \frac{1}{4}(0+1+2+3) = \frac{6}{4} = \frac{3}{2} = 1.5\\
\ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) &= \mathsf{Var}(\mathsf{X}) && \text{by Theorem 2.1 page 49}\\
&= \int_{\mathbb{R}} [x-\mathsf{E}(\mathsf{X})]^2\mathsf{P}(x) && \text{by definition of } \mathsf{Var} \text{ (Definition 1.40 page 24)}\\
&= \sum_{x\in\mathbb{Z}} [x-\mathsf{E}(\mathsf{X})]^2\mathsf{P}(x) && \text{by definition of } \mathsf{P} \text{ and Proposition 1.2 page 24}\\
&= \frac{1}{4}\sum_{x\in\mathsf{H}}\left[x-\frac{3}{2}\right]^2 && \text{by } \mathsf{E}(\mathsf{X}) \text{ result}\\
&= \frac{1}{4}\left[\left(0-\frac{3}{2}\right)^2+\left(1-\frac{3}{2}\right)^2+\left(2-\frac{3}{2}\right)^2+\left(3-\frac{3}{2}\right)^2\right]\\
&= \frac{1}{4\cdot 2^2}\left[(0-3)^2+(2-3)^2+(4-3)^2+(6-3)^2\right] = \frac{20}{16} = \frac{5}{4} = 1.25\\
\ddot{\mathsf{E}}(\mathsf{X}) &= \mathsf{E}(\mathsf{X}) && \text{because on } \textit{real line}, \mathsf{P} \text{ is } \textit{symmetric}, \text{ and by Theorem 2.2 page 49}\\
&= \frac{3}{2} = 1.5 && \text{by } \mathsf{E}(\mathsf{X}) \text{ result}\\
\ddot{\mathsf{E}}(\mathsf{X}) &\triangleq \arg\min_{x\in\mathbb{R}}\max_{y\in\mathbb{R}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49).} \quad \textbf{(alternate proof)}\\
&= \arg\min_{x\in\mathbb{R}}\max_{y\in\mathbb{R}} |x-y|\mathsf{P}(y) && \text{by definition usual metric on real line}\\
&= \arg\min_{x\in\mathbb{R}}\max_{y\in\{0,1,2,3\}} |x-y|\frac{1}{4} && \text{by definition of } \boldsymbol{G}\\
&= \arg\min_{x\in\mathbb{R}}\max_{y\in\{0,1,2,3\}} |x-y| && \text{because } \mathsf{f}(x) = \frac{1}{4}x \text{ is } \textit{strictly isotone} \text{ and by Lemma 1.7 (page 19)}\\
&= \arg\min_{x\in\mathbb{R}}\left\{\begin{array}{ll} |x-3| & \text{for } x\le\frac{3}{2}\\ |x-0| & \text{otherwise}\end{array}\right\}
\end{aligned}
$$



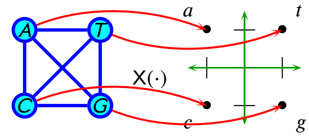
$$
\begin{aligned}
&= \left\{\frac{3}{2}\right\} = 1\frac{1}{2} = 1.5 && \text{because expression is minimized at argument } x = \frac{3}{2}\\
\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) &= \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) && \text{because } \ddot{\mathsf{E}}(\mathsf{X}) = \mathsf{E}(\mathsf{X})\\
&= \frac{5}{4} && \text{by } \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) \text{ result}\\
\ddot{\mathsf{E}}(\mathsf{Y}) &\triangleq \arg\min_{x\in\mathbb{Z}}\max_{y\in\mathbb{Z}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\mathbb{Z}}\max_{y\in\mathbb{Z}} |x-y|\mathsf{P}(y) && \text{by definition of } \textit{integer line} \text{ (Definition 1.36 page 23)}\\
&= \arg\min_{x\in\mathbb{Z}}\max_{y\in\{0,1,2,3\}} |x-y|\frac{1}{4}\\
&= \arg\min_{x\in\mathbb{Z}}\max_{y\in\{0,1,2,3\}} |x-y| && \text{because } \phi(x) = \frac{1}{4}x \text{ is } \textit{strictly isotone} \text{ and by Lemma 1.7 page 19}
\end{aligned}
$$

$$= \arg\min_{x\in\{0,1,2,3\}} \max_{y\in\{0,1,2,3\}} \left\{\begin{matrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{matrix}\right\} = \arg\min_{x\in\{0,1,2,3\}} \left\{\begin{matrix} 3 \\ 2 \\ 2 \\ 3 \end{matrix}\right\} = \arg\min_{x\in\{0,1,2,3\}} \left\{\begin{matrix} 1 \\ 2 \end{matrix}\right\}$$

$$\begin{aligned}
\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) &\triangleq \sum_{x\in\mathbb{Z}} \mathsf{d}^2\big(\ddot{\mathsf{E}}(\mathsf{Y}), x\big)\,\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)} \\
&= \sum_{x\in\mathbb{Z}} \mathsf{d}^2(\{1,2\}, x)\,\mathsf{P}(x) && \text{by } \ddot{\mathsf{E}}(\mathsf{Y}) \text{ result} \\
&= |0-1|^2 \times \frac{1}{4} + |1-1|^2 \times \frac{1}{4} + |2-2|^2 \times \frac{1}{4} + |3-2|^2 \times \frac{1}{4} = \frac{1}{2} = 0.5
\end{aligned}$$

✏

Example 2.21 (GSP to complex plane).



A possible solution for the GSP problem (Example 2.20 page 65) is to map $\{\boxed{\text{A}}, \boxed{\text{T}}, \boxed{\text{C}}, \boxed{\text{G}}\}$ to the *complex plane* (Example 1.10 page 23) rather than the *real line* (Definition 1.35 page 22) such that (see also illustration to the right)

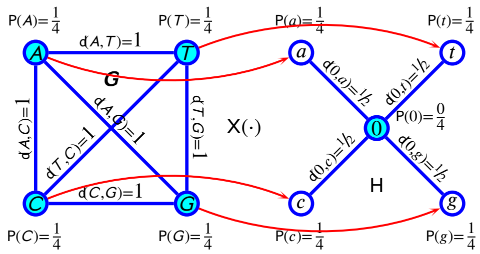
$$\begin{array}{rclcl@{\qquad}rclcl}
a & \triangleq & \mathsf{X}(\boxed{\text{A}}) & = & -1+i & t & \triangleq & \mathsf{X}(\boxed{\text{T}}) & = & 1+i \\
c & \triangleq & \mathsf{X}(\boxed{\text{C}}) & = & -1-i & g & \triangleq & \mathsf{X}(\boxed{\text{G}}) & = & 1-i.
\end{array}$$

However, this solution also is arguably unsatisfactory for two reasons:

1. The order structures are dissimilar. Note that
   $c < a$, but $\boxed{\text{C}}$ and $\boxed{\text{A}}$ are *incomparable* (Definition 1.20 page 9).

2. The metric geometries are dissimilar. Let d be the *discrete metric* and p the *usual metric* in $\mathbb{C}$. Note that
   $\mathsf{d}(\boxed{\text{A}}, \boxed{\text{T}}) = \mathsf{d}(\boxed{\text{A}}, \boxed{\text{C}}) = \mathsf{d}(\boxed{\text{A}}, \boxed{\text{G}}) = 1$, but
   $\mathsf{p}(a,t) = |a-t| = 2 \neq 2\sqrt{2} = |a-g| = \mathsf{p}(a,g)$.

Example 2.22 (DNA mapping with extended range).



Example 2.20 (page 65) presented a mapping from a DNA structure to a linearly ordered lattices, but the order and metric geometry was not preserved. In this example, a different structure is used that does preserve both order and metric geometry (see illustration to the right). This yields the following statistics:

$$\ddot{\mathsf{E}}(\mathsf{X}) = \{0\} \qquad \ddot{\mathsf{Var}}(\mathsf{X}) = \frac{1}{4}$$

✎PROOF:

$$\begin{aligned}
\ddot{\mathsf{E}}(\mathsf{H}) &\triangleq \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)} \\
&= \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}\setminus\{0\}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{because } \mathsf{P}(0) = 0 \\
&= \arg\min_{x\in\mathsf{H}\setminus\{0\}} \max_{y\in\mathsf{H}\setminus\{0\}} \mathsf{d}(x,y)\,\frac{1}{4} && \text{by definition of } \boldsymbol{G} \\
&= \arg\min_{x\in\mathsf{H}\setminus\{0\}} \max_{y\in\mathsf{H}\setminus\{0\}} \mathsf{d}(x,y) && \text{because } \phi(x) = \frac{1}{4}x \text{ is \textit{strictly isotone} and by Lemma 1.7 page 19} \\
&= \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}} \left\{\begin{matrix} \mathsf{d}(1,1) & \mathsf{d}(1,2) & \mathsf{d}(1,3) & \mathsf{d}(1,4) \\ \mathsf{d}(2,1) & \mathsf{d}(2,2) & \mathsf{d}(2,3) & \mathsf{d}(2,4) \\ \mathsf{d}(3,1) & \mathsf{d}(3,2) & \mathsf{d}(3,3) & \mathsf{d}(3,4) \\ \mathsf{d}(4,1) & \mathsf{d}(4,2) & \mathsf{d}(4,3) & \mathsf{d}(4,4) \\ \mathsf{d}(0,1) & \mathsf{d}(0,2) & \mathsf{d}(0,3) & \mathsf{d}(0,4) \end{matrix}\right\} = \arg\min_{x\in\mathsf{H}} \max_{y\in\mathsf{H}} \left\{\begin{matrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{matrix}\right\} = \arg\min_{x\in\mathsf{H}} \left\{\begin{matrix} 1 \\ 1 \\ 1 \\ 1 \\ \frac{1}{2} \end{matrix}\right\} \\
&= \{0\} && \text{because expression is minimized at } x = \{0\} \\
\ddot{\mathsf{Var}}(\mathsf{X}) &\triangleq \sum_{x\in\mathsf{H}} \mathsf{d}^2\big(\ddot{\mathsf{E}}(\mathsf{X}), x\big)\,\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}
\end{aligned}$$

$$= \sum_{x\in\mathsf{H}} \mathsf{d}^2(\{0\},x)\,\mathsf{P}(x) \qquad \text{by } \ddot{\mathsf{E}}(\mathsf{X}) \text{ result}$$

$$= \sum_{x\in\mathsf{H}\setminus\{0\}} \left(\frac{1}{2}\right)^2 \frac{1}{4} = |\,\mathsf{H}\setminus\{0\}\,| \left(\frac{1}{2}\right)^2 \frac{1}{4} = 4\left(\frac{1}{2}\right)^2 \frac{1}{4} = \frac{1}{4}$$

✐

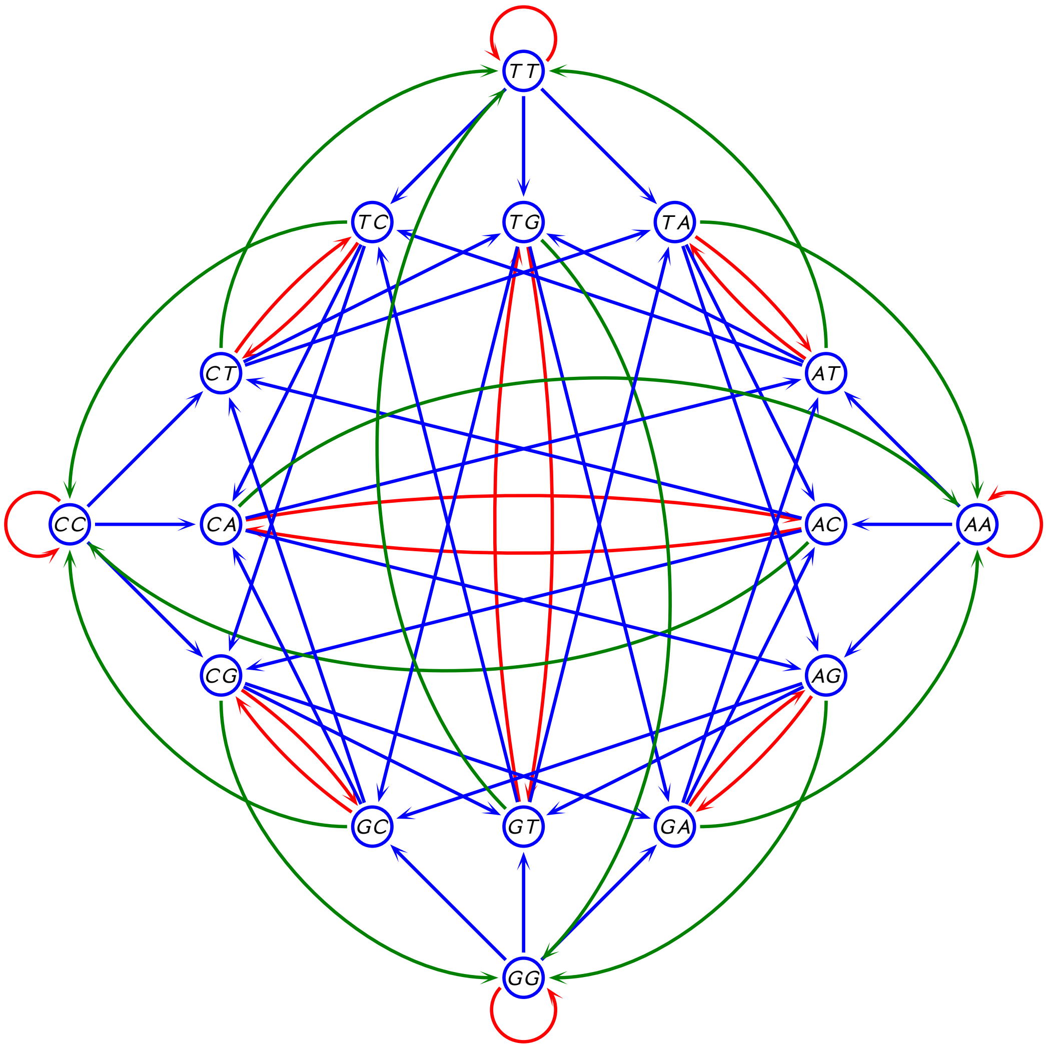
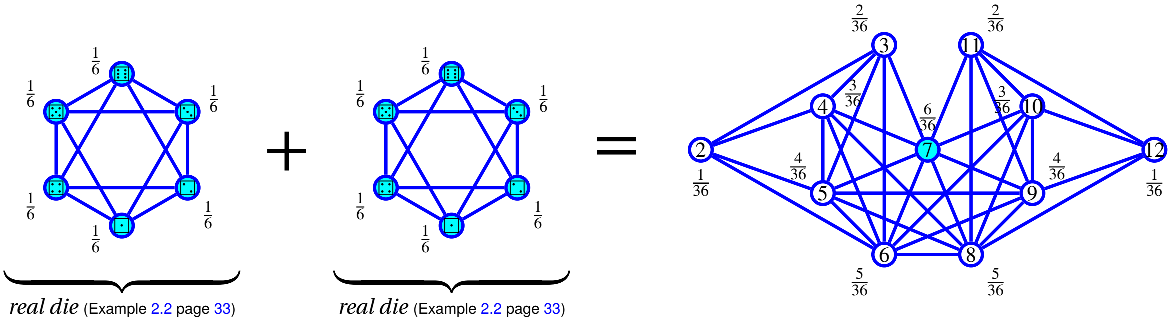


Figure 2.12: DNA with depth-2 Markov modelling (Example 2.23 page 68)

Example 2.23 (GSP with Markov model). Markov probability models have often been used in genomic signal processing (GSP). A change in the statistics in the sequence may in some cases mean a change in function of the genomic sequence (DNA code). Finding such a change in statistics then is very useful in identifying functions of segments of genomic sequences. Let $\boldsymbol{G}$ be an *outcome subspace* (Definition 2.1 page 29) representing a Markov model of depth 2 for a genomic sequence as illustrated in Figure 2.12 (page 68), with joint and conditional probabilities computed over a finite window. Let $\boldsymbol{H}$ be an outcome subspace isomorphic to $\boldsymbol{G}$, and X be a random variable mapping $\boldsymbol{G}$ to H. A change in the value of the statistic $\ddot{\mathsf{E}}(\mathsf{X})$ over the window then may indicate a change in function within the genomic sequence.

2.3 Operations on outcome subspaces

2.3.1 Summation



	2	3	4	5	6	7	8	9	10	11	12
2	0	1	1	1	1	2	2	2	2	3	4
3	1	0	1	1	1	1	2	2	2	2	3
4	1	1	0	1	1	1	1	2	2	2	2
5	1	1	1	0	1	1	1	1	2	2	2
6	1	1	1	1	0	1	1	1	1	2	2
7	2	1	1	1	1	0	1	1	1	1	2
8	2	2	1	1	1	1	0	1	1	1	1
9	2	2	2	1	1	1	1	0	1	1	1
10	2	2	2	2	1	1	1	1	0	1	1
11	3	2	2	2	2	1	1	1	1	0	1
12	4	3	2	2	2	2	1	1	1	1	0

Figure 2.13: metrics based on number of die edges for a *pair of real dice* (Example 2.24 page 69)

Example 2.24 (pair of dice outcome subspace). A *pair of real dice* has a structure as illustrated in Figure 2.13 (page 69). The values represent the standard sum of die faces and thus range from 2 to 12. The table in the figure provides the metric distances between summed values based on the number of edges that must be transversed to move from the first value to the second value. Alternatively, the distance is the number of times the dice must be rotated 90 degrees to move from the first value being face up to the second value being face up. This structure is also illustrated in the undirected graph on the right in Figure 2.13, along with each value's standard probability.

✎PROOF:

$$\dot{\mathsf{C}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}$$

$$= \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \left\{\begin{array}{ccccc} \mathsf{d}(2,2)\mathsf{P}(2) & \mathsf{d}(2,3)\mathsf{P}(3) & \mathsf{d}(2,4)\mathsf{P}(4) & \cdots & \mathsf{d}(2,12)\mathsf{P}(12) \\ \mathsf{d}(3,2)\mathsf{P}(2) & \mathsf{d}(3,3)\mathsf{P}(3) & \mathsf{d}(3,4)\mathsf{P}(4) & \cdots & \mathsf{d}(3,12)\mathsf{P}(12) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathsf{d}(12,2)\mathsf{P}(2) & \mathsf{d}(12,3)\mathsf{P}(3) & \mathsf{d}(12,4)\mathsf{P}(4) & \cdots & \mathsf{d}(12,12)\mathsf{P}(12) \end{array}\right\}$$

$$= \arg\min_{x\in\boldsymbol{G}}\max_{y\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{array}{ccccccccccc}
0\times1 & 1\times2 & 1\times3 & 1\times4 & 1\times5 & 2\times6 & 2\times5 & 2\times4 & 2\times3 & 3\times2 & 4\times1 \\
1\times1 & 0\times2 & 1\times3 & 1\times4 & 1\times5 & 1\times6 & 2\times5 & 2\times4 & 2\times3 & 2\times2 & 3\times1 \\
1\times1 & 1\times2 & 0\times3 & 1\times4 & 1\times5 & 1\times6 & 1\times5 & 2\times4 & 2\times3 & 2\times2 & 2\times1 \\
1\times1 & 1\times2 & 1\times3 & 0\times4 & 1\times5 & 1\times6 & 1\times5 & 1\times4 & 2\times3 & 2\times2 & 2\times1 \\
1\times1 & 1\times2 & 1\times3 & 1\times4 & 0\times5 & 1\times6 & 1\times5 & 1\times4 & 1\times3 & 2\times2 & 2\times1 \\
2\times1 & 1\times2 & 1\times3 & 1\times4 & 1\times5 & 0\times6 & 1\times5 & 1\times4 & 1\times3 & 1\times2 & 2\times1 \\
2\times1 & 2\times2 & 1\times3 & 1\times4 & 1\times5 & 1\times6 & 0\times5 & 1\times4 & 1\times3 & 1\times2 & 1\times1 \\
2\times1 & 2\times2 & 2\times3 & 1\times4 & 1\times5 & 1\times6 & 1\times5 & 0\times4 & 1\times3 & 1\times2 & 1\times1 \\
2\times1 & 2\times2 & 2\times3 & 2\times4 & 1\times5 & 1\times6 & 1\times5 & 1\times4 & 0\times3 & 1\times2 & 1\times1 \\
3\times1 & 2\times2 & 2\times3 & 2\times4 & 2\times5 & 1\times6 & 1\times5 & 1\times4 & 1\times3 & 0\times2 & 1\times1 \\
4\times1 & 3\times2 & 2\times3 & 2\times4 & 2\times5 & 2\times6 & 1\times5 & 1\times4 & 1\times3 & 1\times2 & 0\times1
\end{array}\right\}$$

[8] Many many thanks to Katie L. Greenhoe and Jonathan J. Greenhoe for help computing these values.

$$= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{array}{ccccccccccc} 0 & 2 & 3 & 4 & 5 & 12 & 10 & 8 & 6 & 6 & 4\\ 1 & 0 & 3 & 4 & 5 & 6 & 10 & 8 & 6 & 4 & 3\\ 1 & 2 & 0 & 4 & 5 & 6 & 5 & 8 & 6 & 4 & 2\\ 1 & 2 & 3 & 0 & 5 & 6 & 5 & 4 & 6 & 4 & 2\\ 1 & 2 & 3 & 4 & 0 & 6 & 5 & 4 & 3 & 4 & 2\\ 2 & 2 & 3 & 4 & 5 & 0 & 5 & 4 & 3 & 2 & 2\\ 2 & 4 & 3 & 4 & 5 & 6 & 0 & 4 & 3 & 2 & 1\\ 2 & 4 & 6 & 4 & 5 & 6 & 5 & 0 & 3 & 2 & 1\\ 2 & 4 & 6 & 8 & 5 & 6 & 5 & 4 & 0 & 2 & 1\\ 3 & 4 & 6 & 8 & 10 & 6 & 5 & 4 & 3 & 0 & 1\\ 4 & 6 & 6 & 8 & 10 & 12 & 5 & 4 & 3 & 2 & 0 \end{array}\right\} = \arg\min_{x\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{array}{c} 12\\ 10\\ 8\\ 6\\ 6\\ 5\\ 6\\ 6\\ 8\\ 10\\ 12 \end{array}\right\} = \{7\}$$

$$\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \dot{\mathsf{C}}_{\mathsf{a}} \text{ (Definition 2.4 page 30)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{array}{c} 0+2+3+4+5+12+10+8+6+6+4\\ 1+0+3+4+5+6+10+8+6+4+3\\ 1+2+0+4+5+6+5+8+6+4+2\\ 1+2+3+0+5+6+5+4+6+4+2\\ 1+2+3+4+0+6+5+4+3+4+2\\ 2+2+3+4+5+0+5+4+3+2+2\\ 2+4+3+4+5+6+0+4+3+2+1\\ 2+4+6+4+5+6+5+0+3+2+1\\ 2+4+6+8+5+6+5+4+0+2+1\\ 3+4+6+8+10+6+5+4+3+0+1\\ 4+6+6+8+10+12+5+4+3+2+0 \end{array}\right\} = \arg\min_{x\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{array}{c} 60\\ 50\\ 43\\ 38\\ 34\\ 32\\ 34\\ 38\\ 43\\ 50\\ 60 \end{array}\right\} = \{7\}$$

$$\ddot{\mathsf{Var}}(\boldsymbol{G}) \triangleq \sum_{x\in\boldsymbol{G}} \left[\mathsf{d}\left(\ddot{\mathsf{E}}(\boldsymbol{G}),x\right)\right]^2 \mathsf{P}(x) \qquad \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}$$

$$= \sum_{x\in\boldsymbol{G}} [\mathsf{d}(7,x)]^2 \mathsf{P}(x) \qquad \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}$$

$$= \frac{1}{36}\left(2^2\times1 + 1^2\times2 + 1^2\times3 + 1^2\times4 + 1^2\times5 + 0^2\times6 + 1^2\times5 + 1^2\times4 + 1^2\times3 + 1^2\times2 + 2^2\times1\right)$$

$$= \frac{1}{36}(4+2+3+4+5+0+5+4+3+2+4)$$

$$= \frac{36}{36}$$

$$= 1$$

✏

The next two examples are examples of sums of *outcome subspaces* (Definition 2.1 page 29): Example 2.25 page 70 (sum of dice pair) and Example 2.27 page 73 (sum of spinner pair).

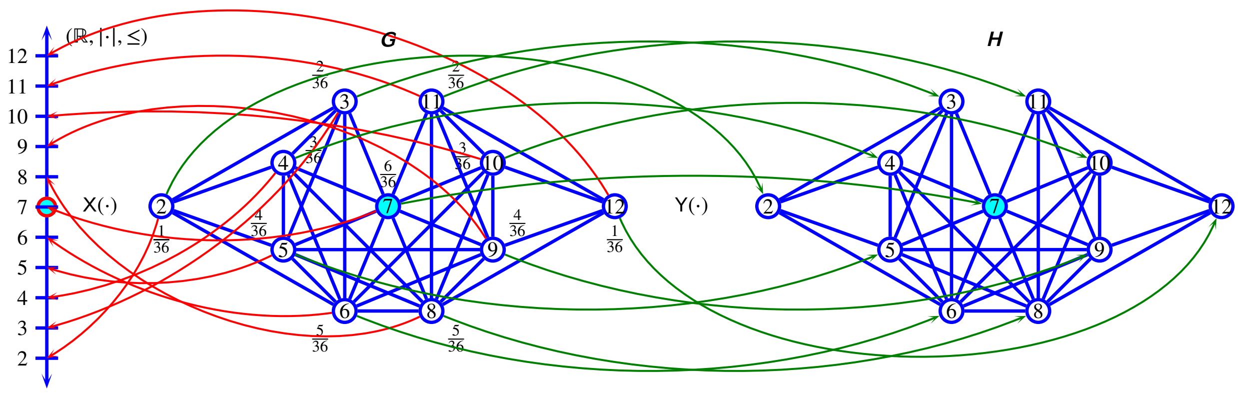


Figure 2.14: *pair of dice* mappings (Example 2.14 page 57)

Example 2.25 (pair of dice and hypothesis testing). Let $\boldsymbol{G}$ be the *pair of dice outcome subspace* (Example 2.24 page 69), $\mathsf{X} \in (\mathbb{R}, |\cdot|, \leq)^{\boldsymbol{G}}$ an *outcome random variable* mapping from $\boldsymbol{G}$ to the *real line* (Definition 1.35 page 22), and $\mathsf{X} \in \mathsf{H}^{\boldsymbol{G}}$ a mapping to a structure H that is *isomorphic* to $\boldsymbol{G}$, as illustrated in Figure 2.14 (page 70). This yields the following statistics:

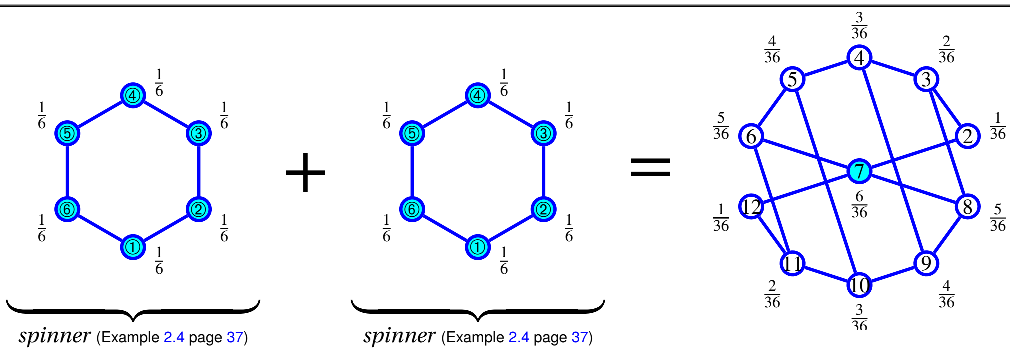
geometry of $\boldsymbol{G}$:	$\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}_a(\boldsymbol{G})$	$=$	$\{7\}$	$\dot{\mathsf{Var}}(\boldsymbol{G})$	$=$	1	
traditional statistics on real line $(\mathbb{R}, \lvert\cdot\rvert, \le)$:	$\mathsf{E}(\mathsf{X})$	$=$	7	$\ddot{\mathsf{Var}}(\mathsf{W}; \mathsf{E})$	$=$	$\frac{35}{6}$	≈ 5.833
outcome subspace statistics on real line $(\mathbb{R}, \lvert\cdot\rvert, \le)$:	$\ddot{\mathsf{E}}(\mathsf{X})$	$=$	$\{7\}$	$\ddot{\mathsf{Var}}(\mathsf{W}; \ddot{\mathsf{E}})$	$=$	$\frac{35}{6}$	≈ 5.833
outcome subspace statistics on isomorphic structure $\boldsymbol{H}$:	$\ddot{\mathsf{E}}(\mathsf{Y})$	$=$	$\{7\}$	$\ddot{\mathsf{Var}}(\mathsf{Y}; \ddot{\mathsf{E}})$	$=$	1	

Although the expected values of both *outcome subspace*s are the essentially the same (7 and $\{7\}$), the isomorphic structure H yields a much smaller variance (a much smaller expected error). This is significant in statistical applications such as hypothesis testing. Suppose for example we have two pair of *real dice* (Example 2.2 page 33), one pair being made of two uniformly distributed die and one pair of weighted die. We want to know which pair is the uniform die. So we roll each pair one time. Suppose the outcome of the first pair is 11 and the the outcome of the second pair is 6. Which pair is more likely to be the uniform pair? Using traditional statistical analysis, the answer is the second pair, because it is closer to the expected value (0.414 standard deviations as opposed to 1.656 standard deviations). However, this result is deceptive, because as can be seen in Figure 2.13 (page 69), the distance from the expected value to the values 11 and 6 are the same ($\mathsf{d}(7, 11) = \mathsf{d}(7, 6) = 1 = 1$ standard deviation). So arguably the outcome of the single roll test would contribute nothing to a good decision algorithm.

✎PROOF:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &= \{7\} && \text{by Example 2.24 (page 69)}\\
\dot{\mathsf{Var}}(\boldsymbol{G}) &= 1 && \text{by Example 2.24 (page 69)}\\
\mathsf{E}(\mathsf{X}) &\triangleq \int_{x\in\mathbb{R}} x\mathsf{P}(x)\,\mathsf{d}x && \text{by definition of } \mathsf{E} \text{ (Definition 1.40 page 24)}\\
&\triangleq \frac{1}{36}\sum_{x\in\mathbb{Z}} x36\mathsf{P}(x) && \text{by definition of } \mathsf{P}\\
&= \frac{1}{36}(2\times1+3\times2+4\times3+5\times4+6\times5+7\times6+8\times5+9\times4+10\times3+11\times2+12\times1)\\
&= \frac{252}{36} = 7\\
\ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) &= \mathsf{Var}(\mathsf{X}) && \text{by Theorem 2.1 page 49}\\
&= \int_{x\in\mathbb{R}} [x-\mathsf{E}(\mathsf{X})]^2\mathsf{P}(x)\,\mathsf{d}x && \text{by definition of } \mathsf{Var} \text{ (Definition 1.40 page 24)}\\
&= \sum_{x\in\mathbb{Z}} [x-\mathsf{E}(\mathsf{X})]^2\mathsf{P}(x) && \text{by definition of } \mathsf{P}\\
&= \sum_{x\in\mathbb{Z}} (x-7)^2\frac{1}{36} && \text{by } \mathsf{E}(\mathsf{X}) \text{ result}\\
&= 2\frac{25\times1+16\times2+9\times3+4\times4+1\times5}{36} = \frac{35}{6} \approx 5.833\\
\ddot{\mathsf{E}}(\mathsf{X}) &= \mathsf{E}(\mathsf{X}) && \text{by Theorem 2.2 (page 49)}\\
&= \{7\} && \text{by } \mathsf{E}(\mathsf{X}) \text{ result}\\
\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) &= \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) && \text{because } \ddot{\mathsf{E}}(\mathsf{X}) \equiv \mathsf{E}(\mathsf{X})\\
&= \frac{35}{6} \approx 5.833 && \text{by } \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) \text{ result}\\
\ddot{\mathsf{E}}(\mathsf{Y}) &= \mathsf{Y}\left[\dot{\mathsf{C}}(\boldsymbol{G})\right] && \text{because } \boldsymbol{G} \text{ and } \mathsf{H} \text{ are \textit{isometric} under } \mathsf{Y}\\
&= \mathsf{Y}[\{7\}] && \text{by } \dot{\mathsf{C}}(\boldsymbol{G}) \text{ result}\\
&= \{7\} && \text{by definition of } \mathsf{Y}\\
\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) &= \dot{\mathsf{Var}}(\boldsymbol{G}) && \text{because } \boldsymbol{G} \text{ and } \mathsf{H} \text{ are \textit{isometric} under } \mathsf{Y}\\
&= 1 && \text{by } \dot{\mathsf{Var}}(\boldsymbol{G}) \text{ result}
\end{aligned}
$$

✏



	2	3	4	5	6	7	8	9	10	11	12
2	0	1	2	3	2	1	2	3	4	3	2
3	1	0	1	2	3	2	1	2	3	4	3
4	2	1	0	1	2	3	2	1	2	3	4
5	3	2	1	0	1	2	3	2	1	2	3
6	2	3	2	1	0	1	2	3	2	1	2
7	1	2	3	2	1	0	1	2	3	2	1
8	2	1	2	3	2	1	0	1	2	3	2
9	3	2	1	2	3	2	1	0	1	2	3
10	4	3	2	1	2	3	2	1	0	1	2
11	3	4	3	2	1	2	3	2	1	0	1
12	2	3	4	3	2	1	2	3	2	1	0

Figure 2.15: metrics based on a pair of spinners (Example 2.26 page 72)

Example 2.26 (pair of spinners). A pair of *spinner*s (Example 2.4 page 37) has a structure as illustrated in Figure 2.15 (page 72). The values represent the standard sum of spinner positions $(1, 2, \ldots, 6)$ and thus range from 2 to 12. The table in the figure provides the metric distances between summed values based on how many positions one must traverse to get from one value to the next (in which ever direction is shortest). This structure is also illustrated in the undirected graph in the upper right of Figure 2.15, along with each value's standard probability.

✎Proof:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \left\{\begin{array}{ccccc}
\mathsf{d}(2,2)\mathsf{P}(2) & \mathsf{d}(2,3)\mathsf{P}(3) & \mathsf{d}(2,4)\mathsf{P}(4) & \cdots & \mathsf{d}(2,12)\mathsf{P}(12)\\
\mathsf{d}(3,2)\mathsf{P}(2) & \mathsf{d}(3,3)\mathsf{P}(3) & \mathsf{d}(3,4)\mathsf{P}(4) & \cdots & \mathsf{d}(3,12)\mathsf{P}(12)\\
\vdots & \vdots & \vdots & \ddots & \vdots\\
\mathsf{d}(12,2)\mathsf{P}(2) & \mathsf{d}(12,3)\mathsf{P}(3) & \mathsf{d}(12,4)\mathsf{P}(4) & \cdots & \mathsf{d}(12,12)\mathsf{P}(12)
\end{array}\right\}\\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{array}{ccccccccccc}
0\times1 & 1\times2 & 2\times3 & 3\times4 & 2\times5 & 1\times6 & 2\times5 & 3\times4 & 4\times3 & 3\times2 & 2\times1\\
1\times1 & 0\times2 & 1\times3 & 2\times4 & 3\times5 & 2\times6 & 1\times5 & 2\times4 & 3\times3 & 4\times2 & 3\times1\\
2\times1 & 1\times2 & 0\times3 & 1\times4 & 2\times5 & 3\times6 & 2\times5 & 1\times4 & 2\times3 & 3\times2 & 4\times1\\
3\times1 & 2\times2 & 1\times3 & 0\times4 & 1\times5 & 2\times6 & 3\times5 & 2\times4 & 1\times3 & 2\times2 & 3\times1\\
2\times1 & 3\times2 & 2\times3 & 1\times4 & 0\times5 & 1\times6 & 2\times5 & 3\times4 & 2\times3 & 1\times2 & 2\times1\\
1\times1 & 2\times2 & 3\times3 & 2\times4 & 1\times5 & 0\times6 & 1\times5 & 2\times4 & 3\times3 & 2\times2 & 1\times1\\
2\times1 & 1\times2 & 2\times3 & 3\times4 & 2\times5 & 1\times6 & 0\times5 & 1\times4 & 2\times3 & 3\times2 & 2\times1\\
3\times1 & 2\times2 & 1\times3 & 2\times4 & 3\times5 & 2\times6 & 1\times5 & 0\times4 & 1\times3 & 2\times2 & 3\times1\\
4\times1 & 3\times2 & 2\times3 & 1\times4 & 2\times5 & 3\times6 & 2\times5 & 1\times4 & 0\times3 & 1\times2 & 2\times1\\
3\times1 & 4\times2 & 3\times3 & 2\times4 & 1\times5 & 2\times6 & 3\times5 & 2\times4 & 1\times3 & 0\times2 & 1\times1\\
2\times1 & 3\times2 & 4\times3 & 3\times4 & 2\times5 & 1\times6 & 2\times5 & 3\times4 & 2\times3 & 1\times2 & 0\times1
\end{array}\right\}\\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{array}{ccccccccccc}
0 & 2 & 6 & 12 & 10 & 6 & 10 & 12 & 12 & 6 & 2\\
1 & 0 & 3 & 8 & 15 & 12 & 5 & 8 & 9 & 8 & 3\\
2 & 2 & 0 & 4 & 10 & 18 & 10 & 4 & 6 & 6 & 4\\
3 & 4 & 3 & 0 & 5 & 12 & 15 & 8 & 3 & 4 & 3\\
2 & 6 & 6 & 4 & 0 & 6 & 10 & 12 & 6 & 2 & 2\\
1 & 4 & 9 & 8 & 5 & 0 & 5 & 8 & 9 & 4 & 1\\
2 & 2 & 6 & 12 & 10 & 6 & 0 & 4 & 6 & 6 & 2\\
3 & 4 & 3 & 8 & 15 & 12 & 5 & 0 & 3 & 4 & 3\\
4 & 6 & 6 & 4 & 10 & 18 & 10 & 4 & 0 & 2 & 2\\
3 & 8 & 9 & 8 & 5 & 12 & 15 & 8 & 3 & 0 & 1\\
2 & 6 & 12 & 12 & 10 & 6 & 10 & 12 & 6 & 2 & 0
\end{array}\right\} = \arg\min_{x\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{array}{c}
12\\ 15\\ 18\\ 15\\ 12\\ 9\\ 12\\ 15\\ 18\\ 15\\ 12
\end{array}\right\} = \{7\}
\end{aligned}
$$

$$\begin{aligned}
\dot{\mathsf{C}}_a(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{G}} \sum_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}}_a \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{matrix}
0+2+6+12+10+6+10+12+12+6+2\\
1+0+3+8+15+12+5+8+9+8+3\\
2+2+0+4+10+18+10+4+6+6+4\\
3+4+3+0+5+12+15+8+3+4+3\\
2+6+6+4+0+6+10+12+6+2+2\\
1+4+9+8+5+0+5+8+9+4+1\\
2+2+6+12+10+6+0+4+6+6+2\\
3+4+3+8+15+12+5+0+3+4+3\\
4+6+6+4+10+18+10+4+0+2+2\\
3+8+9+8+5+12+15+8+3+0+1\\
2+6+12+12+10+6+10+12+6+2+0
\end{matrix}\right\} = \arg\min_{x\in\boldsymbol{G}} \frac{1}{36}\left\{\begin{matrix}78\\72\\66\\60\\56\\54\\56\\60\\66\\72\\78\end{matrix}\right\} = \{7\}\\
\ddot{\mathsf{Var}}(\boldsymbol{G}) &\triangleq \sum_{x\in\boldsymbol{G}} \left[\mathsf{d}\left(\ddot{\mathsf{E}}(\boldsymbol{G}),x\right)\right]^2\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \sum_{x\in\boldsymbol{G}} [\mathsf{d}(7,x)]^2\mathsf{P}(x) && \text{by definition of } \ddot{\mathsf{Var}} \text{ (Definition 2.14 page 49)}\\
&= \frac{1}{36}\left(1^2\times1+2^2\times2+3^2\times3+2^2\times4+1^2\times5+0^2\times6+1^2\times5+2^2\times4+2^2\times3+2^2\times2+1^2\times1\right)\\
&= \frac{1}{36}(2+8+27+16+5+0+5+16+12+8+1)\\
&= \frac{100}{36} = \frac{25}{9} = 2\frac{7}{9} \approx 2.778
\end{aligned}$$

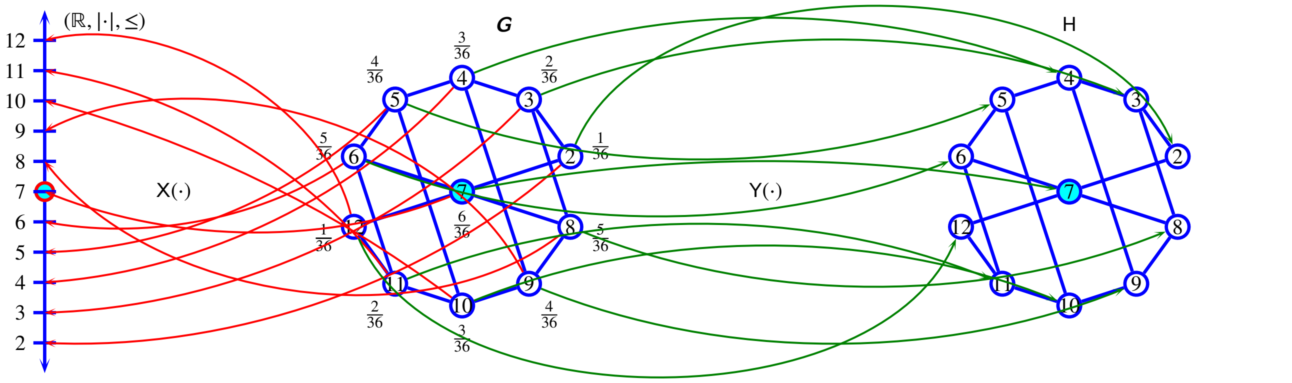


Figure 2.16: *pair of spinner* mappings (Example 2.27 page 73)

Example 2.27 (pair of spinner and hypothesis testing). Let $\boldsymbol{G}$ be a *pair of spinners* (Example 2.26 page 72), X a *random variable* mapping to the *real line* (Definition 1.35 page 22), and Y a *random variable* mapping to an *ordered metric space* (Definition 1.34 page 22) that is *isomorphic* to $\boldsymbol{G}$ under Y, as illustrated in Figure 2.16 (page 73). This yields the following statistics:

geometry of $\boldsymbol{G}$:	$\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}_a(\boldsymbol{G})$	$=$	$\{7\}$	$\dot{\mathsf{Var}}(\boldsymbol{G})$	$=$	$\frac{25}{9}$	$\approx$	2.778
traditional statistics on real line $(\mathbb{R}, \lvert\cdot\rvert, \le)$:	$\mathsf{E}(\mathsf{X})$	$=$	7	$\ddot{\mathsf{Var}}(\mathsf{W};\mathsf{E})$	$=$	$\frac{35}{6}$	$\approx$	5.833
outcome subspace statistics on real line $(\mathbb{R}, \lvert\cdot\rvert, \le)$:	$\ddot{\mathsf{E}}(\mathsf{X})$	$=$	$\{7\}$	$\ddot{\mathsf{Var}}(\mathsf{W};\ddot{\mathsf{E}})$	$=$	$\frac{35}{6}$	$\approx$	5.833
outcome subspace statistics on isomorphic structure $\boldsymbol{H}$:	$\ddot{\mathsf{E}}(\mathsf{Y})$	$=$	$\{7\}$	$\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}})$	$=$	1		

Although the expected value of both *outcome subspace*s are the same ($\mathsf{E}(\mathsf{X}) = \ddot{\mathsf{E}}(\mathsf{Y}) = 7$), the isomorphic outcome subspace $\boldsymbol{H}$ yields a much smaller variance (a much smaller expected error). This is significant in statistical applications such as hypothesis testing. Suppose for example we have two pair of *spinners* (Example 2.4 page 37), one pair being made of two uniformly distributed spinners, and one pair of weighted spinners. We want to estimate which is which. So we spin each pair one time. Suppose the outcome of the first pair is 12 and the the outcome of the second pair is 10. Which pair is more likely to be the uniform pair? Using traditional statistical analysis, the answer is the second pair, because it is closer to the expected value ($\mathsf{d}(7,10) = |7-10| = 3 = 1.8$ standard deviations as opposed to $\mathsf{d}(7,12) = |7-12| = 5 = 3$ standard deviations). However, this result is deceptive, because as can be seen in the table in Figure 2.15 (page 72), 12 is acually closer to the expected value in $\boldsymbol{G}$ than is 10 ($\mathsf{d}(7,12) = 1 < 3 = \mathsf{d}(7,10)$). So arguably the better choice, based on this one trial, is the first pair.

✎ PROOF:

$$
\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{G}) &= \{7\} && \text{by Example 2.26 (page 72)}\\
\dot{\mathsf{C}}_{\mathsf{a}}(\boldsymbol{G}) &= \{7\} && \text{by Example 2.26 (page 72)}\\
\dot{\mathsf{Var}}(\boldsymbol{G}) &= \frac{25}{9} \approx 2.778 && \text{by Example 2.26 (page 72)}\\
\mathsf{E}(\mathsf{X}) &= \frac{252}{36} = 7 && \text{by Example 2.25 (page 70)}\\
\mathsf{Var}(\mathsf{X}) &= \frac{35}{6} \approx 5.833 && \text{by Example 2.25 (page 70)}\\
\ddot{\mathsf{E}}(\mathsf{X}) &= \mathsf{E}(\mathsf{X}) && \text{because on \textit{real line}, P is \textit{symmetric}, and by Theorem 2.2 page 49}\\
&= \{7\} && \text{by E(X) result}\\
\ddot{\mathsf{Var}}(\mathsf{X};\ddot{\mathsf{E}}) &= \ddot{\mathsf{Var}}(\mathsf{X};\mathsf{E}) && \text{because } \ddot{\mathsf{E}}(\mathsf{X}) = \mathsf{E}(\mathsf{X})\\
&= \mathsf{Var}(\mathsf{X}) && \text{by Theorem 2.1 page 49}\\
&= \frac{35}{6} \approx 5.833 &&\\
\ddot{\mathsf{E}}(\mathsf{Y}) &= \mathsf{Y}\big[\dot{\mathsf{C}}(\boldsymbol{G})\big] && \text{because } \boldsymbol{G} \text{ and H are \textit{isomorphic} under Y}\\
&= \mathsf{Y}[\{7\}] && \text{by Example 2.26 (page 72)}\\
&= \{7\} && \text{by definition of X}\\
\ddot{\mathsf{Var}}(\mathsf{Y};\ddot{\mathsf{E}}) &= \ddot{\mathsf{Var}}(\boldsymbol{G}) && \text{because } \boldsymbol{G} \text{ and H are \textit{isomorphic} under Y}\\
&= 1 && \text{by Example 2.26 (page 72)}
\end{aligned}
$$

✏

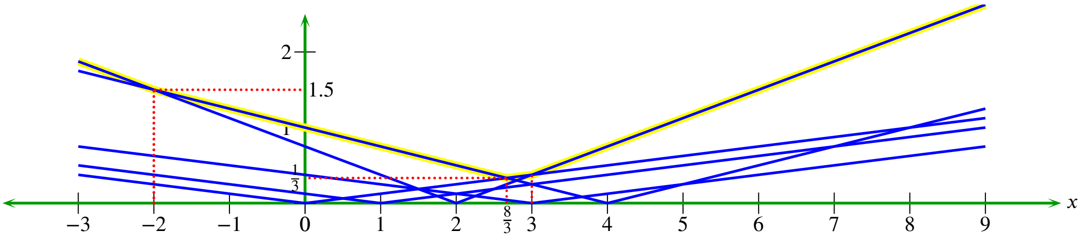


Figure 2.17: real line addition $\arg\min_x \max_y$ calculation graph (Example 2.28 page 74)

Example 2.28 (linear addition). Let X be a *random variable* (Definition 2.13 page 49) mapping to a *real line ordered metric space* (Definition 1.35 page 22) resulting in probability values of

$\mathsf{P}(0) = \mathsf{P}(2) = \frac{1}{2}$, and $\mathsf{P}(x) = 0$ otherwise.

Let Y be a *random variable* mapping to a *real line ordered metric space* resulting in probability values of

$\mathsf{P}(0) = \mathsf{P}(1) = \frac{1}{4}$, $\mathsf{P}(2) = \frac{1}{2}$, and $\mathsf{P}(x) = 0$ otherwise.

Let $\mathsf{Z} \triangleq \mathsf{X} + \mathsf{Y}$ be the random variable mapping to the *outcome subspace* (Definition 2.1 page 29) induced by adding X and Y resulting in probabilities

$\begin{array}{c|ccccc} z & 0 & 1 & 2 & 3 & 4 \\ \hline \mathsf{P}(z) & \frac{1}{8} & \frac{1}{8} & \frac{3}{8} & \frac{1}{8} & \frac{2}{8} \end{array}$, and $\mathsf{P}(z) = 0$ otherwise.

Note that although the traditional expectation E (Definition 1.40 page 24) *distributes* over addition such that

$\mathsf{E}(\mathsf{X}+\mathsf{Y}) = \frac{18}{8} = 1 + \frac{5}{4} = \mathsf{E}(\mathsf{X}) + \mathsf{E}(\mathsf{Y})$,

the alternative expecation $\ddot{\mathsf{E}}$ (Definition 2.14 page 49) does *not*:
$$\ddot{\mathsf{E}}(\mathsf{X}+\mathsf{Y}) = \tfrac{8}{3} \neq \tfrac{7}{3} = 1 + \tfrac{4}{3} = \ddot{\mathsf{E}}(\mathsf{X}) + \ddot{\mathsf{E}}(\mathsf{Y}).$$

✎PROOF:

$$\begin{aligned}
\mathsf{E}(\mathsf{X}) &= \int_{x\in\mathbb{R}} x\mathsf{P}(x)\,\mathsf{d}x && \text{by definition of } \mathsf{E} \text{ (Definition 1.40 page 24)}\\
&= \sum_{x\in\mathbb{Z}} x\mathsf{P}(x)\,\mathsf{d}x && \text{by definition of } \mathsf{P} \text{ and Proposition 1.2 page 24}\\
&= 0\times\frac{1}{2} + 2\times\frac{1}{2} && \text{by definition of } \mathsf{P}\\
&= 1\\
\mathsf{E}(\mathsf{Y}) &= \int_{y\in\mathbb{R}} y\mathsf{P}(y)\,\mathsf{d}y && \text{by definition of } \mathsf{E} \text{ (Definition 1.40 page 24)}\\
&= \sum_{y\in\mathbb{Z}} y\mathsf{P}(y)\,\mathsf{d}y && \text{by definition of } \mathsf{P} \text{ and Proposition 1.2 page 24}\\
&= 0\times\frac{1}{4} + 1\times\frac{1}{4} + 2\times\frac{1}{2} && \text{by definition of } \mathsf{P}\\
&= \frac{5}{4}\\
\mathsf{E}(\mathsf{Z}) &= \int_{z\in\mathbb{R}} z\mathsf{P}(z) && \text{by definition of } \mathsf{E} \text{ (Definition 1.40 page 24)}\\
&= \sum_{z\in\mathbb{Z}} z\mathsf{P}(z) && \text{by definition of } \mathsf{P} \text{ and Proposition 1.2 page 24}\\
&= 0\times\frac{1}{8} + 1\times\frac{1}{8} + 2\times\frac{3}{8} + 3\times\frac{1}{8} + 4\times\frac{2}{8} && \text{by definition of } \mathsf{P}\\
&= \frac{9}{4}\\
\ddot{\mathsf{E}}(\mathsf{X}) &= \arg\min_{x\in\mathbb{R}} \max_{y\in\mathbb{R}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\mathbb{Z}} \max_{y\in\mathbb{Z}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \mathsf{P}(x)\\
&= \arg\min_{x\in\mathbb{Z}} \left\{\begin{array}{ll} |2-x|\frac{1}{2} & \text{for } x\le 1\\ |x|\frac{1}{2} & \text{otherwise}\end{array}\right\}\\
&= \{1\} && \text{because } \max(x) \text{ is minimized at } x=1\\
\ddot{\mathsf{E}}(\mathsf{Y}) &= \arg\min_{x\in\mathbb{R}} \max_{y\in\mathbb{R}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\mathbb{Z}} \max_{y\in\mathbb{Z}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \mathsf{P}(x)\\
&= \arg\min_{x\in\mathbb{Z}} \left\{\begin{array}{ll} |2-x|\frac{1}{2} & \text{for } \frac{4}{3}\ge x\ge 4\\ |x|\frac{1}{4} & \text{otherwise}\end{array}\right\}\\
&= \left\{\frac{4}{3}\right\} && \text{because } \max(y) \text{ is minimized at } y=\frac{4}{3}\\
\ddot{\mathsf{E}}(\mathsf{Z}) &= \arg\min_{x\in\mathbb{R}} \max_{y\in\mathbb{R}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\mathbb{Z}} \max_{y\in\mathbb{Z}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \mathsf{P}(x)\\
&= \arg\min_{x\in\mathbb{Z}} \left\{\begin{array}{ll} |x-2|\frac{3}{8} & \text{for } -2\ge x\ge 3\\ |x-4|\frac{2}{8} & \text{for } -2\le x\le \frac{8}{3}\\ |x|\frac{1}{8} & \text{for } \frac{8}{3}\le x\le 3\end{array}\right\}\\
&= \left\{\frac{8}{3}\right\} && \text{because } \max(z) \text{ is minimized at } z=\frac{8}{3} \text{ (see Figure 2.17 page 74)}
\end{aligned}$$

✏

2.3.2 Multiplication

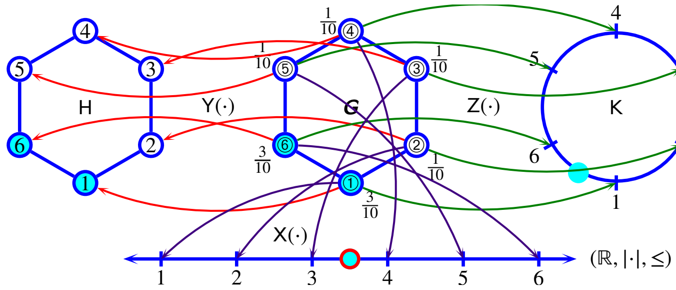


Figure 2.18: *pair of spinner* mappings (Example 2.29 page 76)

Example 2.29 (ring multiplication). Let $\mathsf{X} \in \mathsf{H}^{\boldsymbol{G}}$ be a random variable where $\boldsymbol{G}$ is the *weighted spinners* illustrated in Figure 2.18 (page 76). Note that, in agreement with Corollary 2.1 (page 50),

$$\ddot{\mathsf{E}}(2\mathsf{X}) = \{4\} = \left\{2 \times 5 \mod 6\right\} = 2\{5\} \mod 6 = 2\ddot{\mathsf{E}}(\mathsf{X}) \mod 6\,.$$

✎PROOF:

$$\begin{aligned}
\dot{\complement}(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{H}}\max_{y\in\boldsymbol{H}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\complement} \text{ (Definition 2.3 page 30)}\\
&= \arg\min_{x\in\boldsymbol{H}}\max_{y\in\boldsymbol{H}} \frac{1}{10}\left\{\begin{array}{cccccc} 0\times1 & 1\times1 & 2\times1 & 3\times1 & 2\times1 & 1\times5\\ 1\times1 & 0\times1 & 1\times1 & 2\times1 & 3\times1 & 2\times5\\ 2\times1 & 1\times1 & 0\times1 & 1\times1 & 2\times1 & 3\times5\\ 3\times1 & 2\times1 & 1\times1 & 0\times1 & 1\times1 & 2\times5\\ 2\times1 & 3\times1 & 2\times1 & 1\times1 & 0\times1 & 1\times5\\ 1\times1 & 2\times1 & 3\times1 & 2\times1 & 1\times1 & 0\times5 \end{array}\right\} = \arg\min_{x\in\boldsymbol{H}} \frac{1}{10}\left\{\begin{array}{c}5\\10\\15\\10\\5\\3\end{array}\right\}\\
&= \{5\}\\
\dot{\complement}_a(\boldsymbol{G}) &\triangleq \arg\min_{x\in\boldsymbol{H}}\sum_{y\in\boldsymbol{H}} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \dot{\complement}_a \text{ (Definition 2.4 page 30)}\\
&= \arg\min_{x\in\boldsymbol{H}} \frac{1}{10}\left\{\begin{array}{c} 0\times1+1\times1+2\times1+3\times1+2\times1+1\times5\\ 1\times1+0\times1+1\times1+2\times1+3\times1+2\times5\\ 2\times1+1\times1+0\times1+1\times1+2\times1+3\times5\\ 3\times1+2\times1+1\times1+0\times1+1\times1+2\times5\\ 2\times1+3\times1+2\times1+1\times1+0\times1+1\times5\\ 1\times1+2\times1+3\times1+2\times1+1\times1+0\times5 \end{array}\right\} = \arg\min_{x\in\boldsymbol{H}} \frac{1}{10}\left\{\begin{array}{c}13\\17\\21\\17\\13\\9\end{array}\right\}\\
&= \{5\}\\
\ddot{\mathsf{E}}(\mathsf{X}) &= \mathsf{X}\left[\dot{\complement}(\boldsymbol{G})\right] && \text{because } \boldsymbol{G} \text{ and } \boldsymbol{H}_1 \text{ are } \textit{isomorphic}\\
&= \mathsf{X}[\{5\}] && \text{by } \dot{\complement}(\boldsymbol{G}) \text{ result}\\
&= \{5\} && \text{by definition of } \mathsf{X}\\
\ddot{\mathsf{E}}(2\mathsf{X}) &\triangleq \arg\min_{x\in\boldsymbol{H}_2}\max_{y\in\boldsymbol{H}_2} \mathsf{d}(x,y)\,\mathsf{P}(y) && \text{by definition of } \ddot{\mathsf{E}} \text{ (Definition 2.14 page 49)}\\
&= \arg\min_{x\in\boldsymbol{H}_2}\max_{y\in\boldsymbol{H}_2} \frac{1}{10}\left\{\begin{array}{ccc} 0\times2 & 1\times2 & 1\times6\\ 1\times2 & 0\times2 & 1\times6\\ 1\times2 & 1\times2 & 0\times6 \end{array}\right\} = \arg\min_{x\in\boldsymbol{H}_2} \frac{1}{10}\left\{\begin{array}{c}6\\6\\2\end{array}\right\}\\
&= \{4\}
\end{aligned}$$

✏

2.3.3 Metric transformation

It is possible to use a *metric transform* (Definition D.11 page 166) to transform the structure of an *outcome subspace* (Definition 2.1 page 29) into a completely different *outcome subspace*. This is demonstrated in Example 2.30 (page 77)–Example 2.32 (page 78). Naturally, by doing so one can sometimes even change the geometric *centers* (Definition 2.3 page 30, Definition 2.4 page 30) of the outcome subspaces, and hence also the statistics of random variables that map to/from them. This is demonstrated in Example 2.32 (page 78)–Example 2.33 (page 78).

Theorem 2.4. *Let ϕ be a* METRIC PRESERVING FUNCTION *(Definition D.11 page 166). Let $\boldsymbol{G} \triangleq (\Omega, \dot{\lesssim}, \dot{\mathsf{d}}, \dot{\mathsf{P}})$ and $\boldsymbol{H}$ be* OUTCOME SUBSPACES *(Definition 2.1 page 29).*

THM
$$\left\{\begin{array}{lll} (1). & \phi(\boldsymbol{H}) = \boldsymbol{G} & \text{and} \\ (2). & \phi \text{ is STRICTLY ISOTONE} & \text{and} \\ (3). & \mathsf{P} \text{ is uniform} & \end{array}\right\} \implies \dot{\mathsf{C}}(\boldsymbol{H}) = \dot{\mathsf{C}}(\boldsymbol{G})$$

✎PROOF:

$$\begin{aligned}
\dot{\mathsf{C}}(\boldsymbol{H}) &= \phi\left[\dot{\mathsf{C}}(\boldsymbol{H})\right] && \text{by hypothesis (1)} \\
&\triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \phi[\mathsf{d}(x,y)]\mathsf{P}(y) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)} \\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \phi[\mathsf{d}(x,y)] && \text{by hypothesis (3) and Lemma 1.7 page 19} \\
&= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y) && \text{by hypothesis (2) and Lemma 1.7 page 19} \\
&= \dot{\mathsf{C}}(\boldsymbol{G}) && \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}
\end{aligned}$$

✏

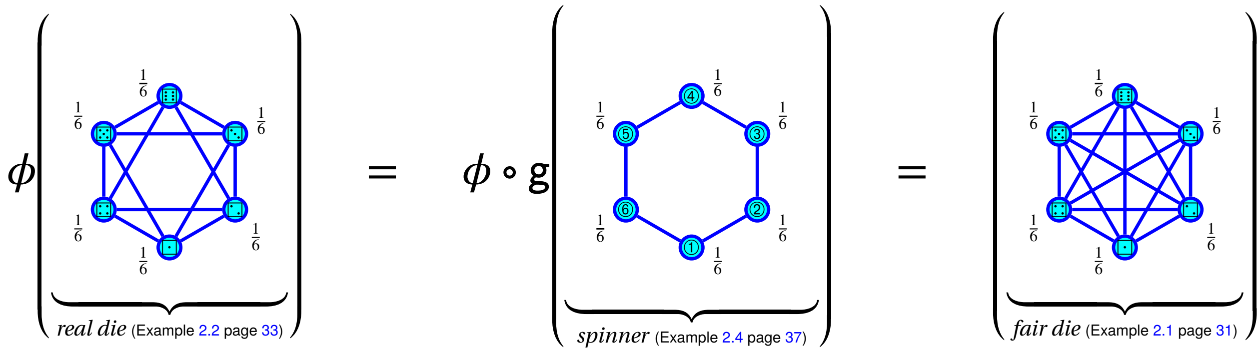


Figure 2.19: *discrete metric preserving function ϕ* on outcome subspaces (Example 2.30 page 77)

Example 2.30 (discrete metric transform on outcome subspaces). Let g be a function (a *pullback function* Theorem 1.7 page 21) such that g(①) = ⚀, g(②) = ⚁, g(③) = ⚂, g(④) = ⚅, g(⑤) = ⚄, and g(⑥) = ⚃. Then under the *dicrete metric preserving function ϕ* (Example D.7 page 168) the *real die outcome subspace* (Example 2.2 page 33) becomes the *fair die outcome subspace* (Example 2.1 page 31), and under $\phi \circ$ g the *spinner outcome subspace* (Example 2.4 page 37) also becomes the *fair die outcome subspace*, as illustrated in Figure 2.19 (page 77). This yields the following geometric statistics:

$\dot{\mathsf{C}}(\boldsymbol{G}) = \dot{\mathsf{C}}(\boldsymbol{H}) = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$.

Example 2.31. Let ϕ_1 be the *metric preserving function* defined in Example D.11 (page 169). Then under ϕ_1, the *spinner outcome subspace* (Example 2.4 page 37) becomes what is here called the *wagon wheel output subspace*, as illustrated on the left in Figure 2.20 (page 78). Let $\boldsymbol{G}$ be the *spinner outcome subspace* and $\boldsymbol{H}$ the *wagon wheel outcome subspace*. This yields the following geometric statistics:

$\dot{\mathsf{C}}(\boldsymbol{G}) = \{④, ⑥\}$ $\qquad$ $\dot{\mathsf{C}}(\boldsymbol{H}) = \{⑤\}$.

Note that the metric transform ϕ_1 also moves the *outcome center* from one that is *not* maximally likely, to one that *is*.

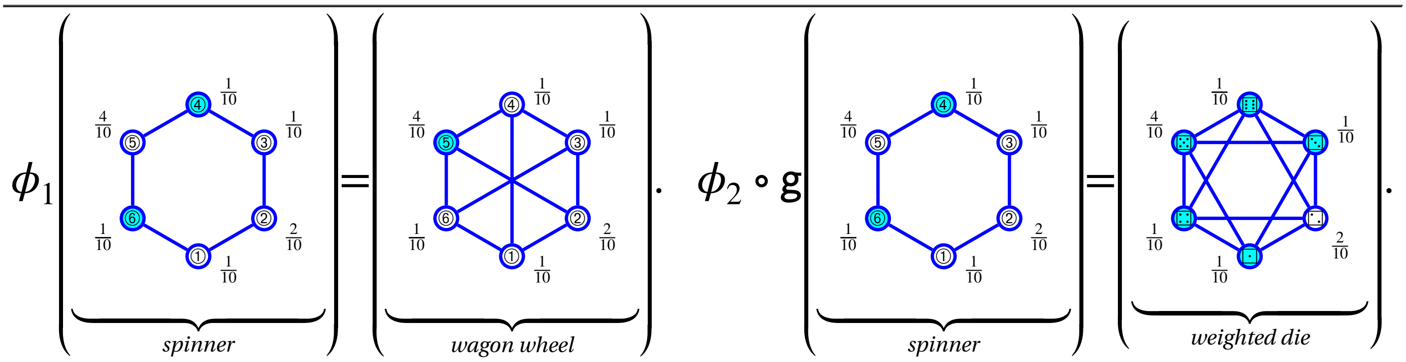


Figure 2.20: Example D.11 (page 169) *metric preserving function* ϕ_1 and Example D.8 (page 168) *metric preserving function* ϕ_2 on *spinner outcome subspace* (Example 2.31 page 77, Example 2.32 page 78)

✎PROOF:

$$\dot{\mathsf{C}}(\boldsymbol{G}) \triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}$$

$$= \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \frac{1}{10}\left\{\begin{array}{cccccc} 0\times1 & 1\times2 & 2\times1 & 3\times1 & 2\times4 & 1\times1\\ 1\times1 & 0\times2 & 1\times1 & 2\times1 & 3\times4 & 2\times1\\ 2\times1 & 1\times2 & 0\times1 & 1\times1 & 2\times4 & 3\times1\\ 3\times1 & 2\times2 & 1\times1 & 0\times1 & 1\times4 & 2\times1\\ 2\times1 & 3\times2 & 2\times1 & 1\times1 & 0\times4 & 1\times1\\ 1\times1 & 2\times2 & 3\times1 & 2\times1 & 1\times4 & 0\times1 \end{array}\right\} = \arg\min_{x\in\boldsymbol{G}} \frac{1}{10}\left\{\begin{array}{c} 8\\ 12\\ 8\\ 4\\ 6\\ 4 \end{array}\right\} = \left\{\begin{array}{c} \\ \\ \\ \textcircled{4}\\ \\ \textcircled{6} \end{array}\right\}$$

$$\dot{\mathsf{C}}(\boldsymbol{H}) \triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}$$

$$= \arg\min_{x\in\boldsymbol{H}} \max_{y\in\boldsymbol{H}} \frac{1}{10}\left\{\begin{array}{cccccc} 0\times1 & 1\times2 & 2\times1 & 1\times1 & 2\times4 & 1\times1\\ 1\times1 & 0\times2 & 1\times1 & 2\times1 & 1\times4 & 2\times1\\ 2\times1 & 1\times2 & 0\times1 & 1\times1 & 2\times4 & 1\times1\\ 1\times1 & 2\times2 & 1\times1 & 0\times1 & 1\times4 & 2\times1\\ 2\times1 & 1\times2 & 2\times1 & 1\times1 & 0\times4 & 1\times1\\ 1\times1 & 2\times2 & 1\times1 & 2\times1 & 1\times4 & 0\times1 \end{array}\right\} = \arg\min_{x\in\boldsymbol{H}} \frac{1}{10}\left\{\begin{array}{c} 8\\ 4\\ 8\\ 4\\ 2\\ 4 \end{array}\right\} = \left\{\begin{array}{c} \\ \\ \\ \\ \textcircled{5}\\ \\ \end{array}\right\}$$

✏

Example 2.32. Let ϕ_2 be the *metric preserving function* defined in Example D.8 (page 168). Let g be the function defined in Example 2.30 (page 77). Then under under $\phi_2 \circ \mathsf{g}$, the *spinner outcome subspace* (Example 2.4 page 37) becomes the *weighted die outcome subspace* (Example 2.3 page 35), as illustrated on the right in Figure 2.20 (page 78). Let $\boldsymbol{G}$ be the *spinner outcome subspace* and $\boldsymbol{H}$ the *weighted die outcome subspace*. These structures have the following geometric statistics:

$$\dot{\mathsf{C}}(\boldsymbol{G}) = \{\textcircled{4}, \textcircled{6}\} \qquad \dot{\mathsf{C}}(\boldsymbol{H}) = \{\textcircled{1}, \textcircled{3}, \textcircled{4}, \textcircled{5}, \textcircled{6}\}.$$

Note that in Example 2.31 (page 77), the metric transform ϕ_1 results in a smaller (smaller *cardinality* Definition 1.13 page 7) *center* ($|\dot{\mathsf{C}}(\boldsymbol{G})| = 2 > 1 = |\dot{\mathsf{C}}(\boldsymbol{H})|$). But here, the metric transform $\phi_2 \circ \mathsf{g}$ results in a larger *center*. ($|\dot{\mathsf{C}}(\boldsymbol{G})| = 2 < 5 = |\dot{\mathsf{C}}(\boldsymbol{H})|$).

✎PROOF:

$$\dot{\mathsf{C}}(\boldsymbol{G}) = \{\textcircled{4}, \textcircled{6}\} \qquad \text{by Example 2.31 (page 77)}$$

$$\dot{\mathsf{C}}(\boldsymbol{H}) \triangleq \arg\min_{x\in\boldsymbol{G}} \max_{y\in\boldsymbol{G}} \mathsf{d}(x,y)\,\mathsf{P}(y) \qquad \text{by definition of } \dot{\mathsf{C}} \text{ (Definition 2.3 page 30)}$$

$$= \arg\min_{x\in\boldsymbol{H}} \max_{y\in\boldsymbol{H}} \frac{1}{10}\left\{\begin{array}{cccccc} 0\times1 & 1\times2 & 1\times1 & 1\times1 & 1\times4 & 2\times1\\ 1\times1 & 0\times2 & 1\times1 & 1\times1 & 2\times4 & 1\times1\\ 1\times1 & 1\times2 & 0\times1 & 2\times1 & 1\times4 & 1\times1\\ 1\times1 & 1\times2 & 2\times1 & 0\times1 & 1\times4 & 1\times1\\ 1\times1 & 2\times2 & 1\times1 & 1\times1 & 0\times4 & 1\times1\\ 2\times1 & 1\times2 & 1\times1 & 1\times1 & 1\times4 & 0\times1 \end{array}\right\} = \arg\min_{x\in\boldsymbol{H}} \frac{1}{10}\left\{\begin{array}{c} 4\\ 8\\ 4\\ 4\\ 4\\ 4 \end{array}\right\} = \left\{\begin{array}{c} \textcircled{1}\\ \\ \textcircled{3}\\ \textcircled{4}\\ \textcircled{5}\\ \textcircled{6} \end{array}\right\}$$

✏

Example 2.33 (linear addition with metric transform). Example 2.28 (page 74) gave the result $\ddot{\mathsf{E}}(\mathsf{X}+\mathsf{Y}) = \frac{8}{3}$ rather than the perhaps more desirable result of $\frac{7}{3}$ (which equals $1 + \frac{4}{3} = \ddot{\mathsf{E}}(\mathsf{X}) + \ddot{\mathsf{E}}(\mathsf{Y})$). Again,

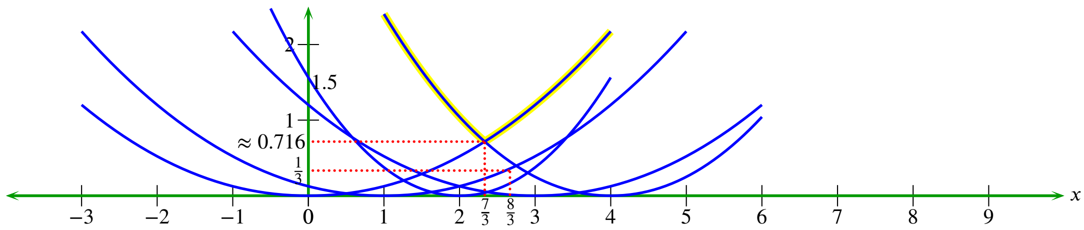


Figure 2.21: real line addition $\arg\min_x \max_y$ calculation graph (Example 2.33 page 78)

we adjust the geometric statistics of an outcome subspace by use of a *metric preserving function* (Definition D.11 page 166). In particular, we use the *power transform/snowflake transform* (Example D.4 page 167) $\mathsf{f}(x) = x^a$. If we let $a = \frac{\ln 2}{\ln 7 - \ln 5} \approx 2.0600427$, then $\ddot{\mathsf{E}}(\mathsf{X}+\mathsf{Y}) = \frac{7}{3}$, as illustrated in Figure 2.21 (page 79).

CHAPTER 3

SYMBOLIC SEQUENCE PROCESSING ON R^{N}

❝. . . *those who assert that the mathematical sciences say nothing of the beautiful or the good are in error. For these sciences say and prove a great deal about them; if they do not expressly mention them, but prove attributes which are their results or definitions, it is not true that they tell us nothing about them. The chief forms of beauty are order and symmetry and definiteness, which the mathematical sciences demonstrate in a special degree.*❞

Aristotle 384 BC – 322 BC, Greek philosopher [1]

3.1 Outcome subspace sequences

3.1.1 Definitions

Definition 3.1. *Let* $\mathbb{D}_1$ *and* $\mathbb{D}_2$ *be* CONVEX SUBSETS (Definition 1.25 page 10) *of* $\mathbb{Z}$.
Let $\mathbb{D} \triangleq \left(\bigwedge \mathbb{D}_1 - \bigvee \mathbb{D}_2 - 1 : \bigvee \mathbb{D}_1 - \bigwedge \mathbb{D}_2 + 1\right)$. *Let* $(\!(x_n)\!)_{\mathbb{D}_1}$ *and* $(\!(y_n)\!)_{\mathbb{D}_2}$ *be* SEQUENCE*s over an* OUTCOME SUBSPACE $\left(\Omega, \dot{\leq}, \dot{\mathsf{d}}, \dot{\mathsf{P}}\right)$.

DEF *The* ***outcome subspace sequence metric*** $\mathsf{p}\left((\!(x_n)\!), (\!(y_n)\!)\right)$ *is defined as*

$$\mathsf{p}\left((\!(x_n)\!), (\!(y_n)\!)\right) \triangleq \sum_{n\in\mathbb{D}} \mathsf{f}(n) \quad \textit{where} \quad \mathsf{f}(n) \triangleq \left\{\begin{array}{ll} \dot{\mathsf{d}}(x_n, y_n) & \textit{if } n \in \mathbb{D}_1 \textit{ and } n \in \mathbb{D}_2 \\ 1 & \textit{if } n \in \mathbb{D}_1 \textit{ but } n \notin \mathbb{D}_2 \\ 1 & \textit{if } n \notin \mathbb{D}_1 \textit{ but } n \in \mathbb{D}_2 \\ 0 & \textit{otherwise} \end{array}\right\} \qquad \forall n\in\mathbb{D}$$

Proposition 3.1. *Let* $(\!(x_n)\!)_{\mathbb{D}_1}$ *and* $(\!(y_n)\!)_{\mathbb{D}_2}$ *be* SEQUENCE*s over an* OUTCOME SUBSPACE $\left(\Omega, \dot{\leq}, \dot{\mathsf{d}}, \dot{\mathsf{P}}\right)$.
Let p *be the* OUTCOME SUBSPACE SEQUENCE METRIC.

PRP $\mathbb{D}_1 \cap \mathbb{D}_2 \neq \varnothing \quad \Longrightarrow \quad \mathsf{p}$ *is a* METRIC

✎PROOF: This follows from the *Fréchet product metric* (Proposition D.9 page 169). In particular, p is a sum of metrics that include the metrics $\dot{\mathsf{d}}(x_n, y_n)$ and the *discrete metric* (Definition D.8 page 166). ✏

[1] quote: Aristotle (330BC?) ⟨Book XIII Part 3⟩
image: http://upload.wikimedia.org/wikipedia/commons/9/98/Sanzio_01_Plato_Aristotle.jpg

In standard signal processing, the *autocorrelation* of a *sequence* $(\!(x_n)\!)_{n\in\mathbb{D}}$ is another *sequence* $(\!(y_n)\!)_{n\in\mathbb{D}}$ defined as $y_n \triangleq \sum_{m\in\mathbb{Z}} x_m x_{m-n}$. However, this definition requires that the sequence $(\!(x_n)\!)_{n\in\mathbb{D}}$ be constructed over a *field*. In an *outcome subspace sequence*, we in general do not have a *field*; for example, in a *die outcome subspace*, the expressions ⚀ + ⚁ and ⚀ × ⚁ are undefined. This paper offers an alternative definition (next) for *autocorrelation* that uses the *distance* $\dot{\mathsf{d}}$ and that does not require a *field*.

Definition 3.2. *Let $(\!(x_n)\!)_{n\in\mathbb{D}}$ and $(\!(y_n)\!)_{n\in\mathbb{D}}$ be* SEQUENCES *over the* OUTCOME SUBSPACE $(\Omega, \dot{\lesssim}, \dot{\mathsf{d}}, \dot{\mathsf{P}})$.
Let $\mathsf{p}\left((\!(x_n)\!), (\!(y_n)\!)\right)$ be the ***outcome subspace sequence metric*** (Definition 3.1 page 81).
The ***cross-correlation*** $\mathsf{R}_{xy}(n)$ *of $(\!(x_n)\!)$ and $(\!(y_n)\!)$ and the* ***autocorrelation*** $\mathsf{R}_{xx}(n)$ *of $(\!(x_n)\!)$ are defined as*

DEF

$$\mathsf{R}_{xy}(n) \triangleq -\sum_{m\in\mathbb{Z}} \mathsf{p}\left((\!(x_{m-n})\!), (\!(y_m)\!)\right) \qquad \text{(\textbf{cross-correlation } } \mathsf{R}_{xy}(n) \text{ of } (\!(x_n)\!) \text{ and } (\!(y_n)\!))$$

$$\mathsf{R}_{xx}(n) \triangleq -\sum_{m\in\mathbb{Z}} \mathsf{p}\left((\!(x_{m-n})\!), (\!(x_m)\!)\right) \qquad \text{(\textbf{autocorrelation } } \mathsf{R}_{xx}(n) \text{ of } (\!(x_n)\!))$$

*Moreover, the M-****offset autocorrelation*** *of $(\!(x_n)\!)$ and $(\!(y_n)\!)$ is here defined as* $\mathsf{R}_{xx}(n) + M$ (Definition 1.44 page 26).

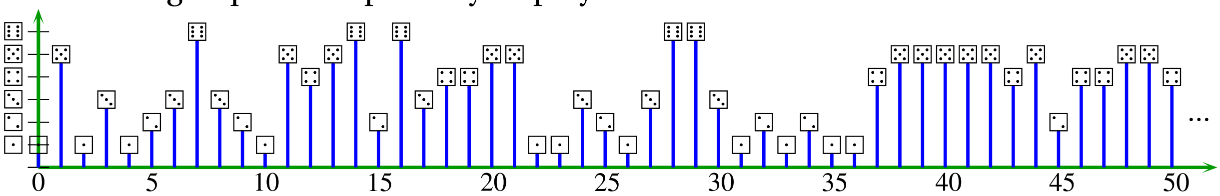
3.1.2 Examples of symbolic sequence statistics

Example 3.1 (fair die sequence). Consider the pseudo-uniformly distributed *fair die* (Definition 2.7 page 30) sequence generated by the C code[2]

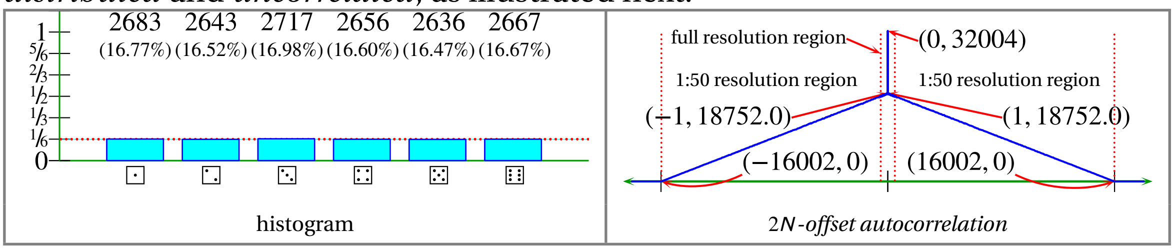
```
#include<stdlib.h>
...
srand(0x5EED);
for(n=0; n<N; n++){x[n] = 'A' + rand()%6;}
```

where 'A' represents ⚀ ,
'B' represents ⚁ ,
⋮
'F' represents ⚅ .

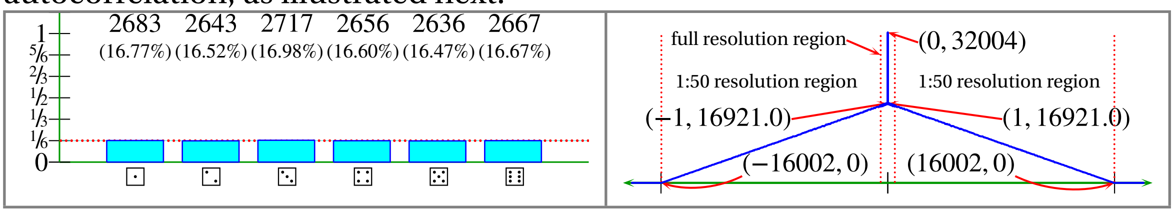
The resulting sequence is partially displayed here:



This sequence constrained to a length of $N = 2667\times 6 = 16002$ elements is approximately *uniformly distributed* and *uncorrelated*, as illustrated next:



Example 3.2 (real die sequence). Consider the pseudo-uniformly distributed *real die* (Definition 2.9 page 31) sequence generated as in Example 3.1 (page 82), but with the real die metric rather than the fair die metric. This change will not affect the distribution of the sequence, but it does affect the autocorrelation, as illustrated next:



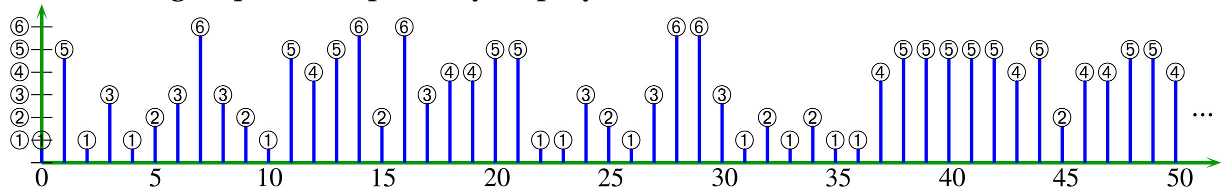
[2] For a more complete source code listing, see Section E.2 (page 192)

Example 3.3 (spinner sequence). Consider the pseudo-uniformly distributed *spinner* (Definition 2.10 page 31) sequence generated by the C code[3]

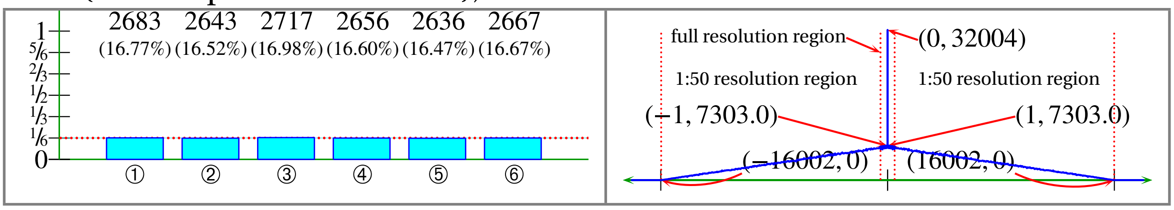
```
#include<stdlib.h>
...
srand(0x5EED);
for(n=0; n<N; n++){x[n] = 'A' + rand()%6;}
```

where 'A' represents ① ,
'B' represents ② ,
⋮
'F' represents ⑥ .

The resulting sequence is partially displayed here:



This sequence is in essence identical to the fair die sequence (Example 3.1 page 82) and real die sequence (Example 3.2 page 82) and thus yields what is essentially an identical histogram. But because the metric is different, the autocorrelation is also different. In particular, because the nodes of the spinner metric are on average farther apart with respect to the spinner metric, the sequence is less correlated (with respect to the metric), as illustrated next:



Example 3.4 (weighted real die sequence). Consider the non-uniformly distributed *weighted real die* (Definition 2.8 page 31) sequence with

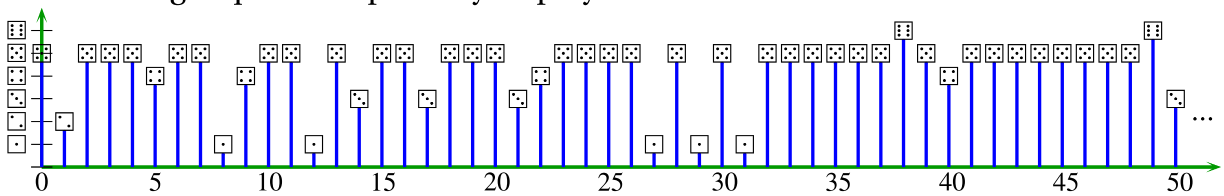
$$\mathsf{P}(⚄) = 0.75 \text{ and } \mathsf{P}(⚀) = \mathsf{P}(⚁) = \mathsf{P}(⚂) = \mathsf{P}(⚃) = \mathsf{P}(⚅) = 0.05,$$

generated by the C code[4]

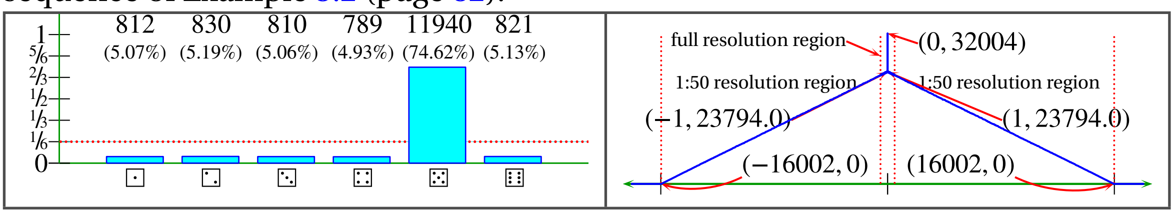
```
srand(0x5EED);
for(n=0; n<N; n++){ u=rand()%100;
  if      (u< 5) x[n]='A'; /* 00-04 */
  else if(u<10) x[n]='B'; /* 05-09 */
  else if(u<15) x[n]='C'; /* 10-14 */
  else if(u<20) x[n]='D'; /* 15-19 */
  else if(u<95) x[n]='E'; /* 20-94 */
  else          x[n]='F'; /* 95-99 */ }
```

where 'A' represents ⚀ ,
'B' represents ⚁ ,
'C' represents ⚂ ,
'D' represents ⚃ ,
'E' represents ⚄ , and
'F' represents ⚅ .

The resulting sequence is partially displayed here:

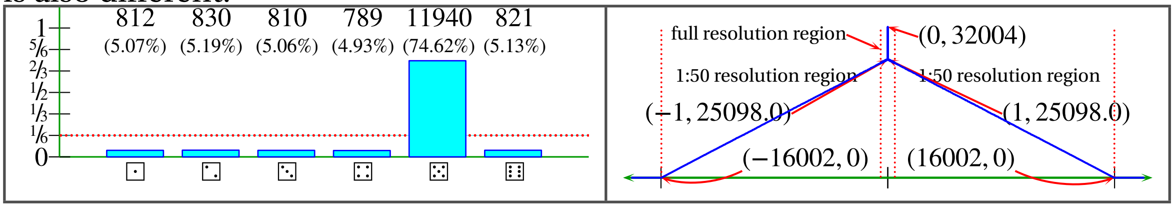


Of course the resulting histogram, as illustrated below on the left, reflects the non-uniform distribution. Also note, as illustrated below on the right, that the weighted sequence is much more correlated (as defined by Definition 3.2 page 82) as compared to the uniformly distributed *real die* sequence of Example 3.2 (page 82).



[3] For a more complete source code listing, see Section E.4 (page 203)
[4] For a more complete source code listing, see Section E.3 (page 197)

Example 3.5 (weighted die sequence). Consider the non-uniformly distributed *weighted die* (Definition 2.6 page 30) sequence generated as in Example 3.4. Of course the resulting histogram is identical to that of Example 3.4, but because the distance function is different, the autocorrelation sequence is also different.



Example 3.6 (weighted spinner sequence). Consider the non-uniformly distributed die sequence with

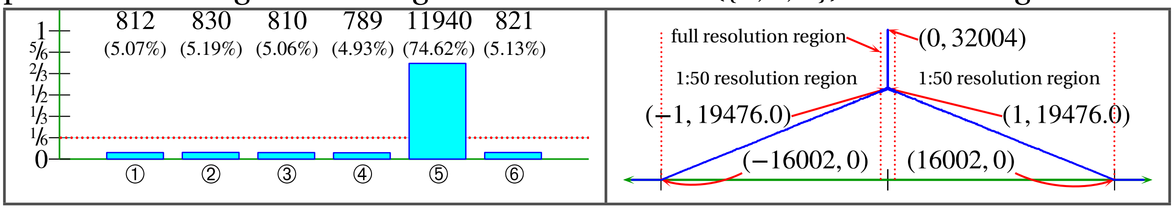
$$P(⑤) = 0.75 \text{ and } P(①) = P(②) = P(③) = P(④) = P(⑥) = 0.05,$$

generated by the C code[5]

```
srand(0x5EED);
for(n=0; n<N; n++){ u=rand()%100;
  if       (u< 5)  x[n]='A';  /* 00-04 */
  else if(u<10) x[n]='B';  /* 05-09 */
  else if(u<15) x[n]='C';  /* 10-14 */
  else if(u<20) x[n]='D';  /* 15-19 */
  else if(u<95) x[n]='E';  /* 20-94 */
  else             x[n]='F';  /* 95-99 */ }
```

where 'A' represents ① ,
'B' represents ② ,
'C' represents ③ ,
'D' represents ④ ,
'E' represents ⑤ , and
'F' represents ⑥ .

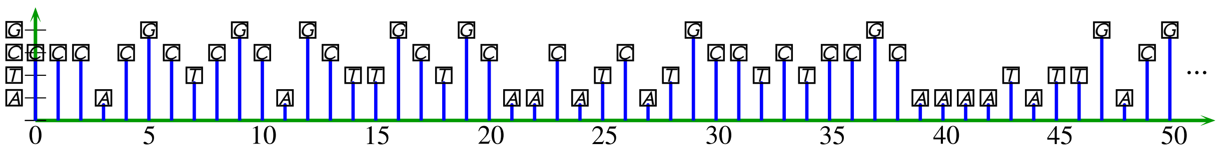
The resulting sequence and histogram is in essence the same as in the *weighted die sequence* example (Example 3.4 page 83). But note, as illustrated below on the right, that the *weighted spinner sequence* of this example is significantly less correlated than the *weighted die sequence* of Example 3.4 (page 83), presumably due to the larger *range* (Definition 1.12 page 7) of the spinner metric ($\{0, 1, 2, 3\}$) as compared to the *range* of the *weighted real die* metric ($\{0, 1, 2\}$) and the *weighted die* metric ($\{0, 1\}$).



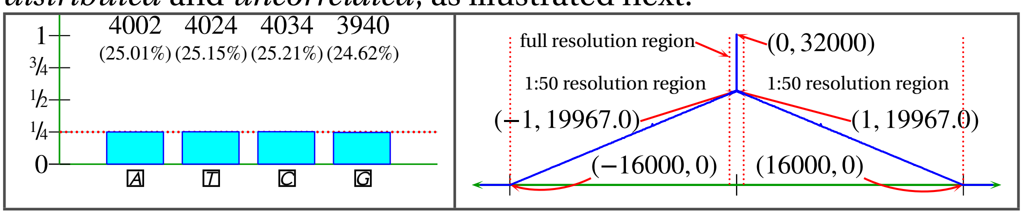
Example 3.7 (Random DNA sequence). Consider the pseudo-uniformly distributed DNA sequence generated by the C code[6]

```
srand(0x5EED);
for(n=0; n<N; n++){   r=rand()%4;
  switch(r){ case 0: x[n]='A'; break;
                 case 1: x[n]='T'; break;
                 case 2: x[n]='C'; break;
                 case 3: x[n]='G'; break; }}
```

where 'A' represents $\boxed{A}$,
'T' represents $\boxed{T}$,
'C' represents $\boxed{C}$, and
'G' represents $\boxed{G}$.



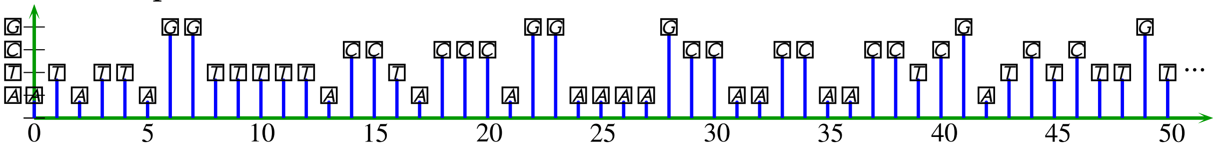
This sequence constrained to a length of $4000 \times 4 = 16000$ elements is approximately *uniformly distributed* and *uncorrelated*, as illustrated next:

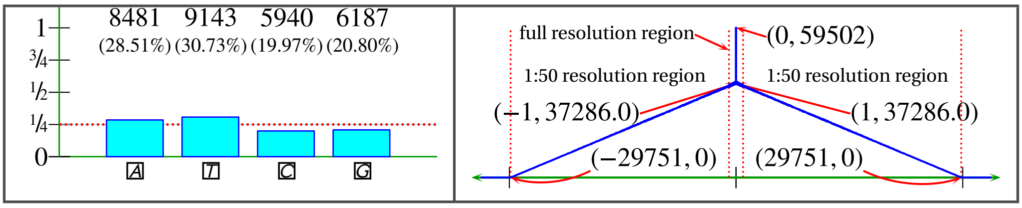


[5] For a more complete source code listing, see Section E.4 (page 203)

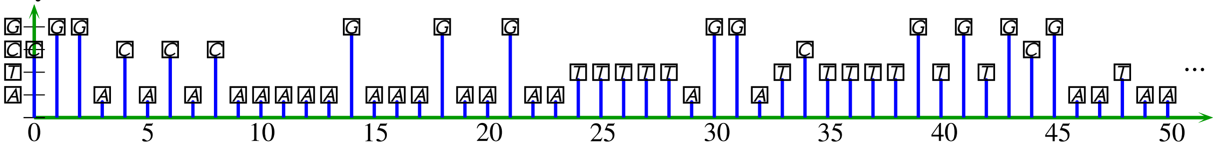
[6] For a more complete source code listing, see Section E.5 (page 209)

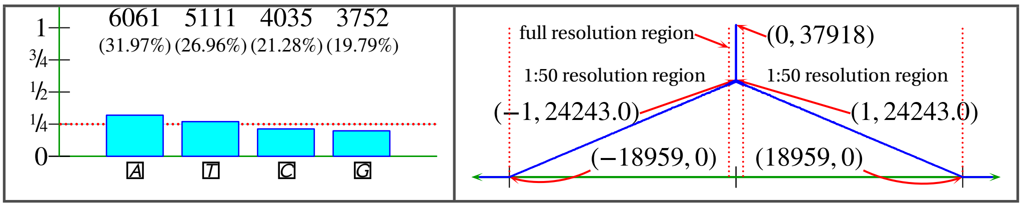
Example 3.8 (SARS coronavirus DNA sequence). Consider the genome sequence (DNA sequence) for the SARS coronavirus with *GenBank* accession number NC_004718.3.[7] This sequence is of length $N = 29751$ and is partially displayed here, followed by its histogram and $2N$*-offset auto-correlation* plots.



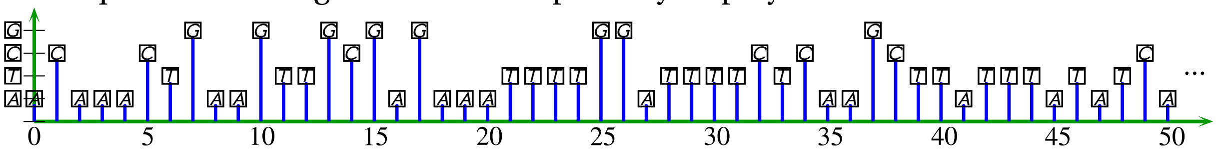


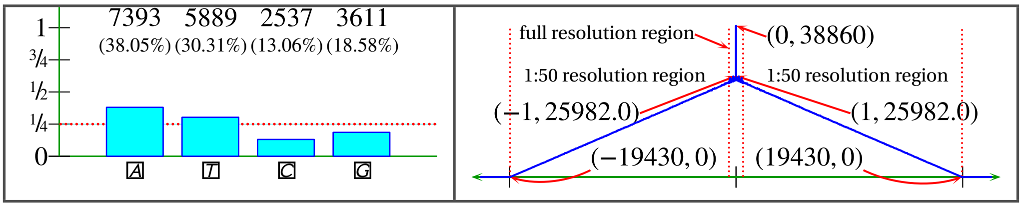
Example 3.9 (Ebola virus DNA sequence). Consider the genome sequence (DNA sequence) for the Ebola virus with GenBank accession AF086833.2.[8] This sequence is of length 18959 and is partially displayed here:





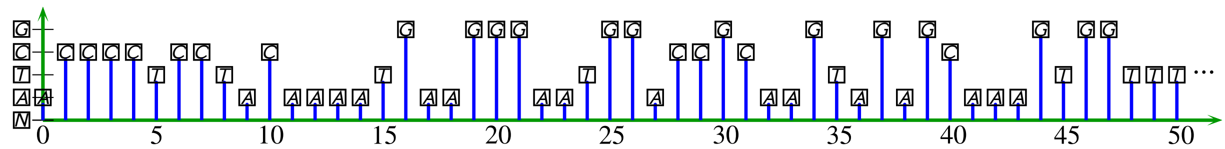
Example 3.10 (Bacterium DNA sequence). Consider the genome sequence (DNA sequence) for the bacterium *Melissococcus plutonius strain 49.3 plasmid pMP19* with GenBank accession NZ_CM003360.1.[9] This sequence is of length 19430 and is partially displayed here:

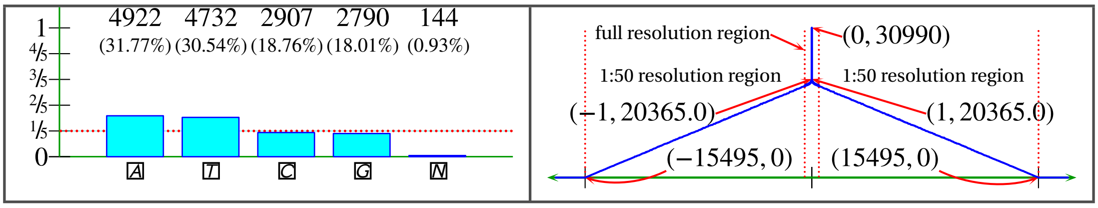




Example 3.11 (Papaya DNA sequence). [10] Consider the genome sequence segment for the fruit *carica papaya* with GenBank accession DS982815.1. This sequence is not a complete genome, rather it is a "genomic scaffold" (Definition 2.12 page 31). As such, there are some elements for which the content is not known. For these locations, the symbol N is used. In this particular sequence, there are 144 N symbols. This sequence is of length 15495 and is partially displayed here:

---
[7] GenBank-NC_004718.3 (2011)
[8] GenBank-AF086833.2 (2013)
[9] GenBank-NZ_CM003360.1 (2015)
[10] GenBank-DS982815.1 (2015)





3.2 Extending to distance linear spaces

3.2.1 Motivation

Section 2.1 (page 29) demonstrated how a stochastic process could be defined as an *outcome subspace* with *order* and *metric* structures. Example 2.13 (page 55) reviewed an example of a *real die outcome subspace* that was mapped through 4 different *random variables* to 4 different *weighted graphs*. Two of these random variables (Y and Z) mapped to structures (*weighted graphs*) that are very similar to the *real die* with respect to order and metric geometry. Two other random variables (W and X) mapped to structures (the *real line* and the *integer line*) that are very dissimilar. The implication of this example is that if we want statistics that closely model the underlying stochastic process, then we should map to a structure that that an order structure and distance geometry similar to that of the underlying stochastic process, and not simply the one that is the most convenient. Ideally, we would like to map to a structure that is *isomorphic* (Definition 1.31 page 14) and *isometric* (Definition 1.33 page 20) to the structure of the stochastic process.

However, for sequence processing using very basic methods such as FIR filtering, Fourier analysis, or wavelet analysis, we would very much like to map into the *real line* $\mathbb{R}^1$ or possibly some higher dimensional space $\mathbb{R}^n$. Because the real line is often very dissimilar to the stochastic process, we are motivated to find structures in $\mathbb{R}^n$ that *are* similar. And that is what this section presents—mapping from a stochastic process $(\Omega, \dot{\le}, \dot{\mathsf{d}}, \dot{\mathsf{P}})$ into an *ordered distance linear space* $(\mathbb{R}^n, \le, \mathsf{d}, +, \cdot, \mathbb{R}, \dot{+}, \dot{\times})$ in which $\mathbb{R}^n$ is an extension of Ω and d is an extension of $\dot{\mathsf{d}}$.

Thus, for sequence processing on an *outcome subspace* $(\Omega, \dot{\le}, \dot{\mathsf{d}}, \dot{\mathsf{P}})$, we would like to define a *random variable* X and an *ordered distance linear space* $(\mathbb{R}^n, \le, \mathsf{d}, +, \cdot, \mathbb{R}, \dot{+}, \dot{\times})$ that satisfy the following constraints:

1. The random variable maps the elements of Ω into $\mathbb{R}^n$ and
2. the order relation $\le$ is an *extension* to $\mathbb{R}^n$ of the order relation $\dot{\le}$ on Ω and
3. the *distance* function d is an *extension* to $\mathbb{R}^n$ of the distance function $\dot{\mathsf{d}}$ on Ω.

3.2.2 Some random variables

In this section, we first define some *random variables* (Definition 2.13 page 49) that are used later in this paper.

Definition 3.3. *The **traditional die random variable*** X *maps from the set* { ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ } *into the set* $\mathbb{R}^1$ *and is defined as*[11]

X(⚀) $\triangleq 1$, X(⚁) $\triangleq 2$, X(⚂) $\triangleq 3$, X(⚃) $\triangleq 4$, X(⚄) $\triangleq 5$, *and* X(⚅) $\triangleq 6$.

Definition 3.4. *The **PAM die random variable*** X *maps from the set* { ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ } *into the set* $\mathbb{R}^1$ *and is defined as*[12]

X(⚀) $\triangleq -2.5$, X(⚁) $\triangleq -1.5$, X(⚂) $\triangleq -0.5$, X(⚃) $\triangleq +0.5$, X(⚄) $\triangleq +1.5$, *and* X(⚅) $\triangleq +2.5$.

Definition 3.5. *The **QPSK die random variable*** X *maps from the set* { ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ } *into the set* $\mathbb{C}^1$ *and is defined as*[13]

$$\begin{array}{lll}
\mathsf{X}(⚀) \triangleq \exp\left(30\times\frac{\pi}{180}i\right), & \mathsf{X}(⚁) \triangleq \exp\left(90\times\frac{\pi}{180}i\right), & \mathsf{X}(⚂) \triangleq \exp\left(150\times\frac{\pi}{180}i\right),\\
\mathsf{X}(⚃) \triangleq \exp\left(210\times\frac{\pi}{180}i\right), & \mathsf{X}(⚄) \triangleq \exp\left(270\times\frac{\pi}{180}i\right), & \mathsf{X}(⚅) \triangleq \exp\left(330\times\frac{\pi}{180}i\right).
\end{array}$$

Definition 3.6. *The* $\mathbb{R}^3$ ***die random variable*** X *maps from the set* { ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ } *into the set* $\mathbb{R}^3$ *and is defined as*

$$\begin{array}{lll}
\mathsf{X}(⚀) \triangleq (+1,0,0), & \mathsf{X}(⚁) \triangleq (0,+1,0), & \mathsf{X}(⚂) \triangleq (0,0,+1),\\
\mathsf{X}(⚃) \triangleq (0,0,-1), & \mathsf{X}(⚄) \triangleq (0,-1,0), & \mathsf{X}(⚅) \triangleq (-1,0,0).
\end{array}$$

Definition 3.7. *The* $\mathbb{R}^6$ ***die random variable*** X *maps from the set* { ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ } *into the set* $\mathbb{R}^6$ *and is defined as*

$$\begin{array}{lll}
\mathsf{X}(⚀) \triangleq (1,0,0,0,0,0), & \mathsf{X}(⚁) \triangleq (0,1,0,0,0,0), & \mathsf{X}(⚂) \triangleq (0,0,1,0,0,0),\\
\mathsf{X}(⚃) \triangleq (0,0,0,1,0,0), & \mathsf{X}(⚄) \triangleq (0,0,0,0,1,0), & \mathsf{X}(⚅) \triangleq (0,0,0,0,0,1).
\end{array}$$

Definition 3.8. *The* $\mathbb{R}^1$ ***spinner random variable*** X *maps from the set* {①, ②, ③, ④, ⑤, ⑥} *into the set* $\mathbb{R}^1$ *and is defined as*[14]

X(①) $\triangleq 1$, X(②) $\triangleq 2$, X(③) $\triangleq 3$, X(④) $\triangleq 4$, X(⑤) $\triangleq 5$, *and* X(⑥) $\triangleq 6$.

Definition 3.9. *The **QPSK spinner random variable*** X *maps from the set* {①, ②, ③, ④, ⑤, ⑥} *into the set* $\mathbb{C}^1$ *and is defined as*

$$\begin{array}{lll}
\mathsf{X}(①) \triangleq \exp\left(-90\times\frac{\pi}{180}i\right), & \mathsf{X}(②) \triangleq \exp\left(-30\times\frac{\pi}{180}i\right), & \mathsf{X}(③) \triangleq \exp\left(30\times\frac{\pi}{180}i\right),\\
\mathsf{X}(④) \triangleq \exp\left(90\times\frac{\pi}{180}i\right), & \mathsf{X}(⑤) \triangleq \exp\left(150\times\frac{\pi}{180}i\right), & \mathsf{X}(⑥) \triangleq \exp\left(210\times\frac{\pi}{180}i\right).
\end{array}$$

Definition 3.10. [15] *The **PAM DNA random variable*** X *maps from the set* $\{\boxed{A}, \boxed{C}, \boxed{G}, \boxed{T}\}$ *into the set* $\mathbb{R}^1$ *and is defined as*[16]

$\mathsf{X}(\boxed{A}) \triangleq -1.5$, $\mathsf{X}(\boxed{C}) \triangleq -0.5$, $\mathsf{X}(\boxed{T}) \triangleq +0.5$, $\mathsf{X}(\boxed{G}) \triangleq +1.5$.

Definition 3.11. [17] *The **QPSK DNA random variable*** X *maps from the set* $\{\boxed{A}, \boxed{C}, \boxed{G}, \boxed{T}\}$ *into the set* $\mathbb{C}^1$ *and is defined as*

$$\begin{array}{ll}
\mathsf{X}(\boxed{A}) \triangleq \exp\left(45\times\frac{\pi}{180}i\right), & \mathsf{X}(\boxed{C}) \triangleq \exp\left(135\times\frac{\pi}{180}i\right),\\
\mathsf{X}(\boxed{G}) \triangleq \exp\left(225\times\frac{\pi}{180}i\right), & \mathsf{X}(\boxed{T}) \triangleq \exp\left(315\times\frac{\pi}{180}i\right).
\end{array}$$

[12] PAM is an acronym for *pulse amplitude modulation* and is a standard technique in the field of digital communications.

[13] QPSK is an acronym for *quadrature phase shift keying* and is a standard technique in the field of digital communications.

[15] Galleani and Garello (2010) page 772

[17] Galleani and Garello (2010) page 772

[17] QPSK is an acronym for *quadrature phase shift keying* and is a standard technique in the field of digital communications.

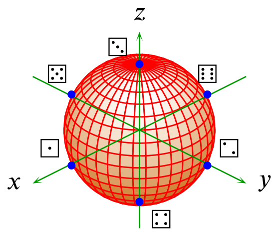
Definition 3.12. *The* $\mathbb{R}^4$ ***DNA random variable*** X *maps from the set* $\{\boxed{A}, \boxed{C}, \boxed{G}, \boxed{T}\}$ *into the set* $\mathbb{R}^4$ *and is defined as*[18]

$$\mathsf{X}(\boxed{A}) \triangleq (1,0,0,0), \quad \mathsf{X}(\boxed{C}) \triangleq (0,1,0,0), \quad \mathsf{X}(\boxed{T}) \triangleq (0,0,1,0), \quad \mathsf{X}(\boxed{G}) \triangleq (0,0,0,1)$$

3.2.3 Some ordered distance linear spaces

Definition 3.13. *The structure* $(\mathbb{R}^1, \le, \mathsf{d})$ *is the* $\mathbb{R}^1$ ***die distance linear space*** *if* $\le$ *is the* STANDARD ORDERING RELATION *on* $\mathbb{R}$, *and* $\mathsf{d}(x,y) \triangleq |x-y|$ *(the* EUCLIDEAN METRIC *on* $\mathbb{R}$, *Definition D.10 page 166).*

Definition 3.14. *The structure* $(\mathbb{R}^3, \le, \mathsf{d})$ *is the* $\mathbb{R}^3$ ***die distance linear space*** *if* $\le = \varnothing$, *and* d *is the* 2*-scaled* LAGRANGE ARC DISTANCE d *defined as follows:* $\mathsf{d}(p,q) \triangleq 2\mathsf{p}(p,q)$
where p *is the* LAGRANGE ARC DISTANCE *(Definition A.1 page 114).*



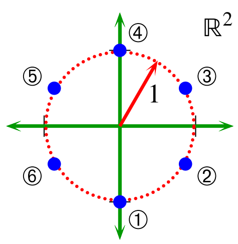
$\dot{\mathsf{d}}(x,y)$	⚀	⚁	⚂	⚃	⚄	⚅
⚀	0	1	1	1	1	2
⚁	1	0	1	1	2	1
⚂	1	1	0	2	1	1
⚃	1	1	2	0	1	1
⚄	1	2	1	1	0	1
⚅	2	1	1	1	1	0

Used together with the $\mathbb{R}^3$ *die random variable* X (Definition 3.6 page 87), the distance d in the $\mathbb{R}^3$ *die distance linear space* (Definition 3.14) is an extension of $\dot{\mathsf{d}}$ in the *real die outcome subspace* $\boldsymbol{G} \triangleq (\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}, \dot{\le}, \dot{\mathsf{d}}, \dot{\mathsf{P}})$ (Definition 2.9 page 31). We can also say that X is an *isometry* (Definition 1.33 page 20) and that the two structures are *isometric*. For example,

$$\begin{aligned} \mathsf{d}[\mathsf{X}(⚀), \mathsf{X}(⚁)] &= \mathsf{d}[(1,0,0),(0,1,0)] &= 1 &= \dot{\mathsf{d}}(⚀, ⚁) \quad \text{and} \\ \mathsf{d}[\mathsf{X}(⚀), \mathsf{X}(⚅)] &= \mathsf{d}[(1,0,0),(-1,0,0)] &= 2 &= \dot{\mathsf{d}}(⚀, ⚅) \quad . \end{aligned}$$

As for order, the mapping X is also *order preserving* (Definition 1.30 page 13), but trivially, because the *real die outcome subspace* is *unordered* (Definition 1.20 page 9). But if we still honor the standard ordering on each dimension $\mathbb{R}$ in $\mathbb{R}^3$, then the two structures are *not isomorphic* (Definition 1.31 page 14) because[19] the inverse X^{-1} is *not order preserving* (Theorem 1.5 page 14)—for example, $\mathsf{X}(⚃) = (0,0,-1) \le (0,0,1) = \mathsf{X}(⚂)$, but ⚃ and ⚂ are *incomparable* (Definition 1.20 page 9) in $\boldsymbol{G}$.

Definition 3.15. *The structure* $(\mathbb{R}^2, \le, \mathsf{d})$ *is the* $\mathbb{R}^2$ ***spinner distance linear space*** *if* $\le = \varnothing$, *and* d *is the* 3*-scaled* LAGRANGE ARC DISTANCE d *defined as follows:* $\mathsf{d}(p,q) \triangleq 3\mathsf{p}(p,q)$
where p *is the* LAGRANGE ARC DISTANCE *(Definition A.1 page 114).*



$\dot{\mathsf{d}}(x,y)$	①	②	③	④	⑤	⑥
①	0	1	2	3	2	1
②	1	0	1	2	3	2
③	2	1	0	1	2	3
④	3	2	1	0	1	2
⑤	2	3	2	1	0	1
⑥	1	2	3	2	1	0

Used together with the *QPSK spinner random variable* X (Definition 3.9 page 87), the distance d in the $\mathbb{R}^2$ *spinner distance linear space* (Definition 3.15 page 88) is an extension of $\dot{\mathsf{d}}$ in the *spinner outcome subspace* $\boldsymbol{G} \triangleq (\{①, ②, ③, ④, ⑤, ⑥\}, \dot{\mathsf{d}}, \dot{\le}, \dot{\mathsf{P}})$ (Definition 2.10 page 31). We can again say that X is an *isometry* and that the two structures are *isometric*. For example,

[18] This type of mapping has previously been used by Voss (1992) in calculating the *Voss Spectrum*, (a kind of Fourier analysis) of DNA sequences. See also Galleani and Garello (2010) page 772.
[19] Note that while X^{-1} (Definition 1.7 page 6) does not exist as a *function*, it does exist as a *relation*.

$$\begin{array}{lclclcll}
\mathsf{d}\big[\mathsf{X}(①),\mathsf{X}(②)\big] &=& \mathsf{d}[(0,-1),(\sqrt{3}/2,-1/2)] &=& 1 &=& \dot{\mathsf{d}}(①,②) & \text{and}\\
\mathsf{d}\big[\mathsf{X}(①),\mathsf{X}(③)\big] &=& \mathsf{d}[(0,-1),(\sqrt{3}/2,+1/2)] &=& 2 &=& \dot{\mathsf{d}}(①,③) & \text{and}\\
\mathsf{d}\big[\mathsf{X}(①),\mathsf{X}(④)\big] &=& \mathsf{d}[(0,-1),(0,1)] &=& 3 &=& \dot{\mathsf{d}}(①,④) & .
\end{array}$$

The mapping X is again trivally *order preserving*. And if we again honor the standard ordering on each dimension $\mathbb{R}$ in $\mathbb{R}^2$, then the two structures are *not isomorphic* (Definition 1.31 page 14) because the inverse X^{-1} is *not order preserving*—for example, $\mathsf{X}(①) = (0,-1) \le (0,1) = \mathsf{X}(④)$, but ① and ④ are *incomparable* in $\boldsymbol{G}$.

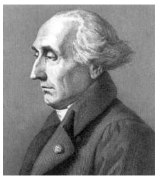
Definition 3.16. *The structure* $(\mathbb{R}^6, \le, \mathsf{d})$ *is the* $\mathbb{R}^6$ ***die distance linear space*** *if* $\le = \emptyset$, *and* d *is defined as* $\mathsf{d}(p,q) \triangleq \sqrt{2}/2\, \mathsf{p}(p,q)$, *where* p *is the* EUCLIDEAN METRIC *on* $\mathbb{R}^6$ *(Definition D.10 page 166).*

Used together with the $\mathbb{R}^6$ *die random variable* X (Definition 3.7 page 87), the distance d in the $\mathbb{R}^6$ *fair die distance linear space* (Definition 3.16 page 89) is an extension of $\dot{\mathsf{d}}$ in the *fair die outcome subspace* $\boldsymbol{G} \triangleq (\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}, \dot{\le}, \dot{\mathsf{d}}, \dot{\mathsf{P}})$ (Definition 2.9 page 31). We can again say that X is an *isometry* and that the two structures are *isometric*. For example,

$$\begin{array}{lclclcll}
\mathsf{d}[\mathsf{X}(⚀),\mathsf{X}(⚁)] &=& \mathsf{d}[(1,0,0,0,0,0),(0,1,0,0,0,0)] &=& 1 &=& \dot{\mathsf{d}}(⚀,⚁) & \text{and}\\
\mathsf{d}[\mathsf{X}(⚀),\mathsf{X}(⚂)] &=& \mathsf{d}[(1,0,0,0,0,0),(0,0,1,0,0,0)] &=& 1 &=& \dot{\mathsf{d}}(⚀,⚂) & \text{and}\\
\mathsf{d}[\mathsf{X}(⚀),\mathsf{X}(⚅)] &=& \mathsf{d}[(1,0,0,0,0,0),(0,0,0,0,0,1)] &=& 1 &=& \dot{\mathsf{d}}(⚀,⚅) & .
\end{array}$$

The mapping X is again trivally *order preserving*, and the inverse X^{-1} is trivally *order preserving* as well. And so unlike the $\mathbb{R}^3$ *die distance linear space* (Definition 3.14) and the $\mathbb{R}^2$ *spinner distance linear space* (Definition 3.15), this pair of structures is *isomorphic*.

3.3 Symbolic sequence processing applications



❝*I regard as quite useless the reading of large treatises of pure analysis: too large a number of methods pass at once before the eyes. It is in the works of applications that one must study them; one judges their ability there and one apprises the manner of making use of them.*❞

Joseph Louis Lagrange (1736-1813), mathematician [20]

3.3.1 Low pass filtering/Smoothing

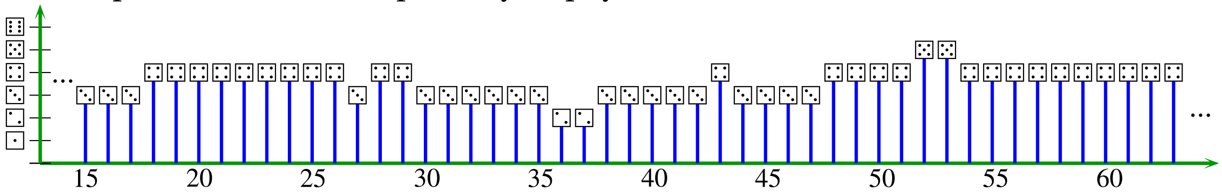
Example 3.12 (low pass filtering of real die sequence).

1. Consider the pseudo-uniformly distributed die sequence presented in Example 3.2 (page 82). Suppose we want to *filter* this sequence with a *low pass sequence* in order to "smooth out" the sequence. But to perform the actual filtering, note that the die sequence must first be mapped into a *linear space* $\mathbb{R}^N$.

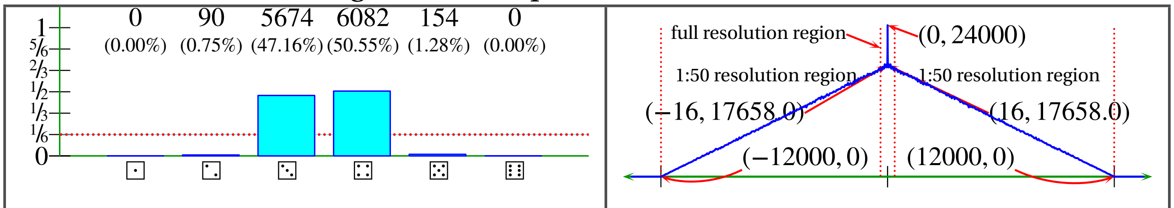
2. Suppose we first use the *traditional die random variable* (Definition 3.3 page 86) to map the die sequence into $\mathbb{R}^1$. *Filtering* (Definition 1.45 page 26) this $\mathbb{R}$-valued sequence using the *length 16 rectangular low pass sequence* (Example 1.18 page 27) in the $\mathbb{R}^1$ *die distance linear space* (Definition 3.13 page 88) and then mapping the result back to a *die sequence* using the *Euclidean metric* (Definition D.10

[20] quote: Stopple (2003), page xi
image: `http://en.wikipedia.org/wiki/Image:Langrange_portrait.jpg`

page 166), produces the result partially dispayed here:



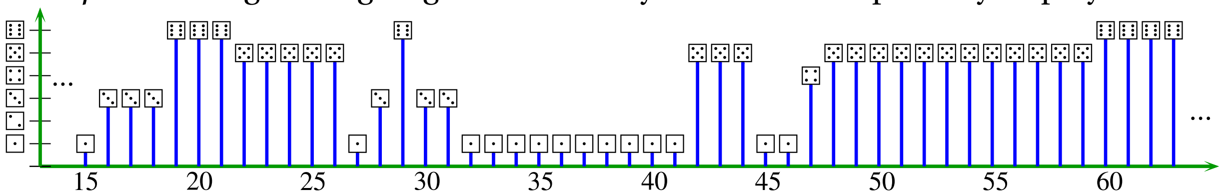
Note that the die sequence has indeed been smoothed out, but it's uniform distribution has been destroyed—almost all of its values are around the "expected value" 3.5, as illustrated below on the left. Of course such filtering also introduces correlation, giving the *autocorrelation* sequence a slightly wider center lobe as illustrated below on the right. Both diagrams are calculated over a length 12000 sequence.



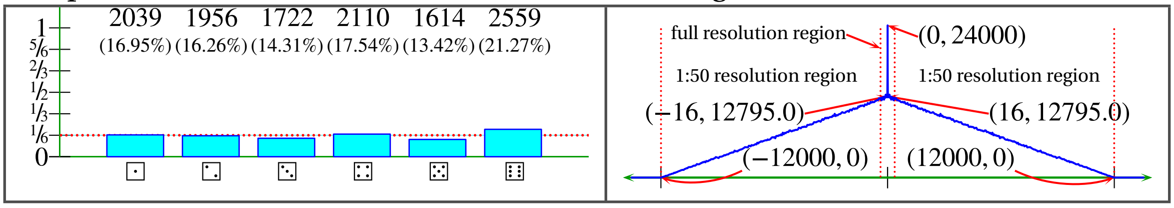
histogram

$2N$*-offset autocorrelation* (Definition 3.2 page 82)

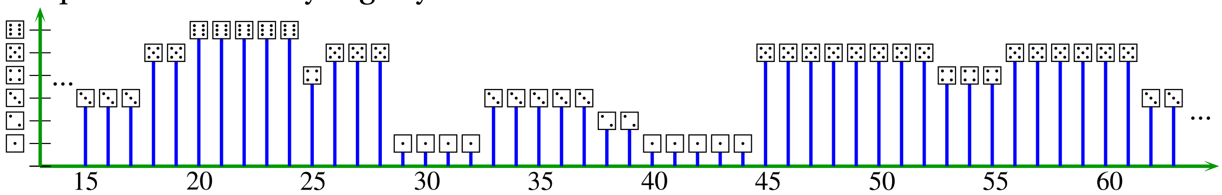
3. Alternatively, suppose we next try using the $\mathbb{R}^3$ *die random variable* (Definition 3.6 page 87) to map the die sequence into $\mathbb{R}^3$. *Filter*ing this new sequence using the *length 16 rectangular low pass sequence* in the $\mathbb{R}^3$ *distance linear space* (Definition 3.14 page 88) and then mapping back to a *die sequence* using the *Lagrange arc distance* yields the result partially displayed here:

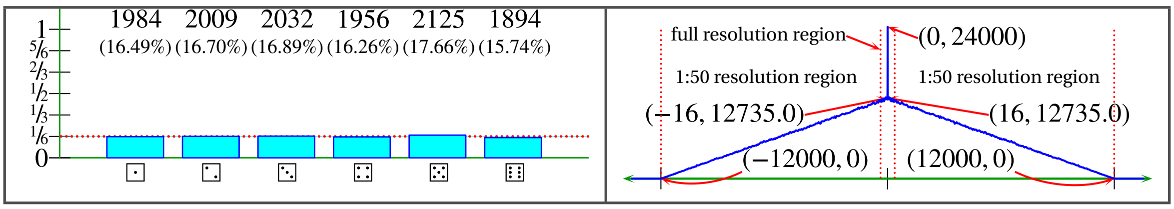


Note that the *die sequence* does appear to be "smoothed out", but this time the distribution is much more uniform, as illustrated below on the left; and is slightly less correlated (12795 compared to 17658), as illustrated below on the right.



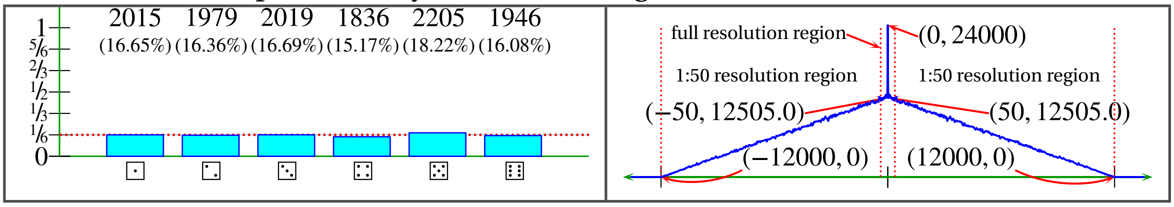
4. Using a *length 16 Hanning low pass sequence* (Definition 1.48 page 27) rather than the *length 16 rectangular low pass sequence* as in item (3) results in a distribution that is more uniform and in a sequence that is very slightly less correlated:



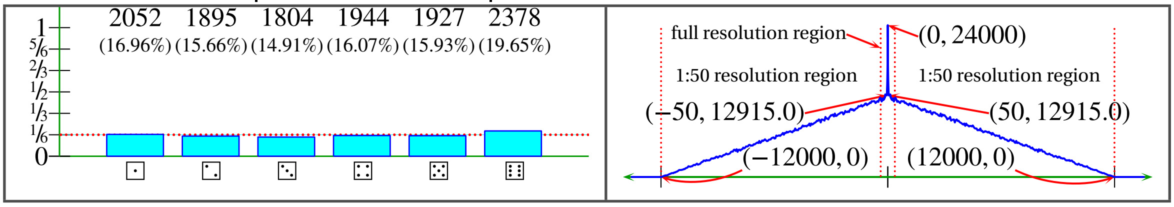


5. Using a *length 50 Hanning low pass sequence* (Example 1.19 page 27) rather than the *length 16 Han-*

ning low pass sequence as in item (4) results in about the same uniformity of distribution, about 1.8% lower side lobes in the autocorrelation sequence ($\frac{12733-12505}{12733} \times 100 \approx 1.8$), but a wider main lobe (presumably due to the longer filter width):



6. Using a *length 50 rectangular low pass sequence* rather than the *length 50 Hanning low pass sequence* as in item (5) results in a distribution that is a little less uniform and about 3.3% more correlated ($\left|\frac{12505-12916}{12505} \times 100\right| \approx 3.3$):



7. Replacing the *Lagrange arc distance* by the *Euclidean metric* in this example has very little effect. More details follow:

   (a) Using the *Euclidean metric* in $\mathbb{R}^3$ rather than the *Lagrange arc distance* in item (3) yields sequences that are **identical**.[21]

   (b) Using the *Euclidean metric* in item (4) rather than the *Lagrange arc distance* yields sequences that **differ** at 6 locations out of $N + M + M - 1 = 12000 + 16 + 16 - 1 = 12031$ locations (differ at approximately 0.05% of the locations):[22]

n	Euclidean	Lagrange
281	⚅	⚃
1630	⚀	⚂
11888	⚅	⚁

   (c) Using the *Euclidean metric* in $\mathbb{R}^3$ rather than the *Lagrange arc distance* as in item (5) (length 50 Hanning filter) yields sequences that are *identical*.[23]

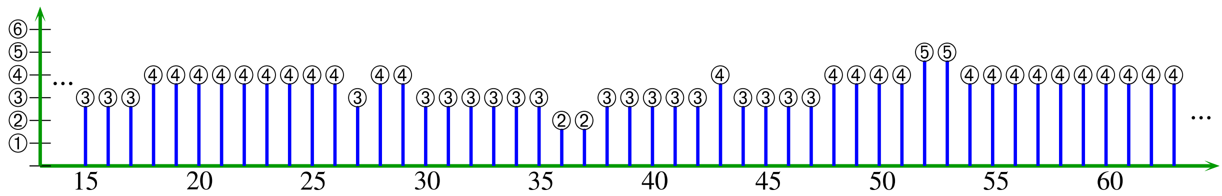
   (d) Using the *Euclidean metric* in $\mathbb{R}^3$ rather than the *Lagrange arc distance* as in item (6) (length 50 rectangular filter) **differ** at 85 locations out of $N + M + M - 1 = 12000 + 50 + 50 - 1 = 12099$ locations (differ at approximately 0.7% of the locations).[24]

Example 3.13 (low pass filtering of spinner sequence).

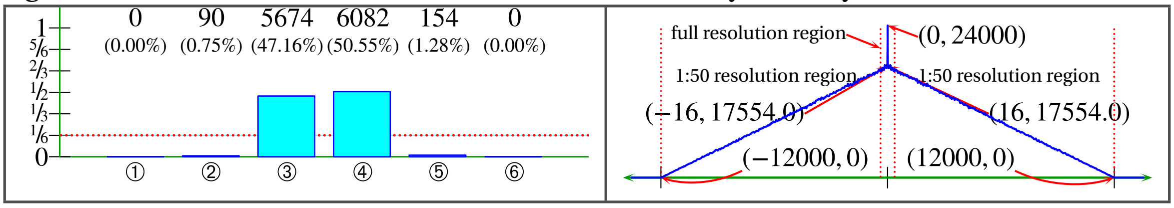
1. Consider the pseudo-uniformly distributed spinner sequence presented in Example 3.3 (page 83). As in Example 3.12 (page 89), suppose we want to *filter* this sequence with a *low pass rectangular sequence* in order to "smooth out" the sequence.

2. Suppose we first use the $\mathbb{R}^1$ *spinner random variable* (Definition 3.8 page 87) to map the spinner sequence into $\mathbb{R}^1$. *Filter*ing this mapped sequence using the *length 16 rectangular low pass sequence* and then mapping the result back to a *spinner sequence* using the *Euclidean metric*, produces the result partially dispayed here (in essense the same as in Example 3.12 page 89):

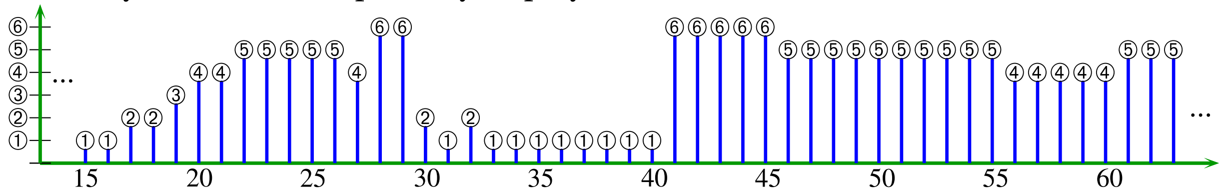
[21] 💾 See experiment log file "rdie_lp_12000m16.xlg" generated by the program "ssp.exe".
[22] 💾 See experiment log file "rdie_lp_12000m16.xlg" generated by the program "ssp.exe".
[23] 💾 See experiment log file "rdie_lp_12000m50.xlg" generated by the program "ssp.exe".
[24] 💾 See experiment log file "rdie_lp_12000m50.xlg" generated by the program "ssp.exe".



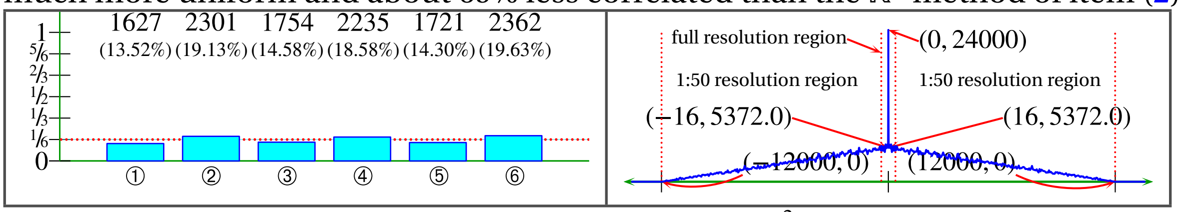
Again, it's uniform distribution has been essentially destroyed.



3. Alternatively, suppose we next try using the *QPSK spinner random variable* (Definition 3.9 page 87) to map the spinner sequence into $\mathbb{C} \triangleq \mathbb{R}^2$. *Filter*ing this new sequence using the *length 16 rectangular low pass sequence* in the $\mathbb{R}^2$ *spinner distance linear space* (Definition 3.15 page 88) and then mapping back to a *sequence* over the *spinner outcome subspace* using the *Lagrange arc distance* yields the result partially displayed here:

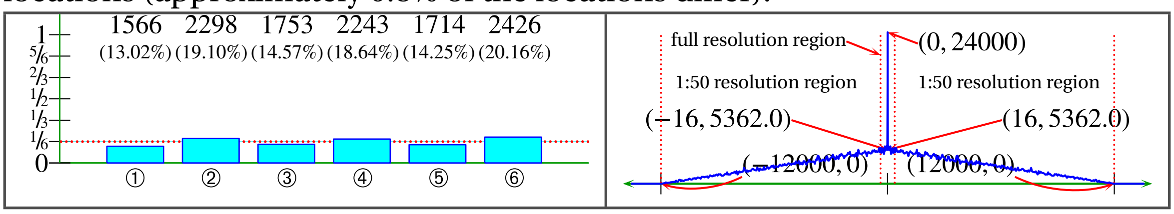


Note that the sequence does appear to be "smoothed out", but this time the distribution is much more uniform and about 69% less correlated than the $\mathbb{R}^1$ method of item (2):

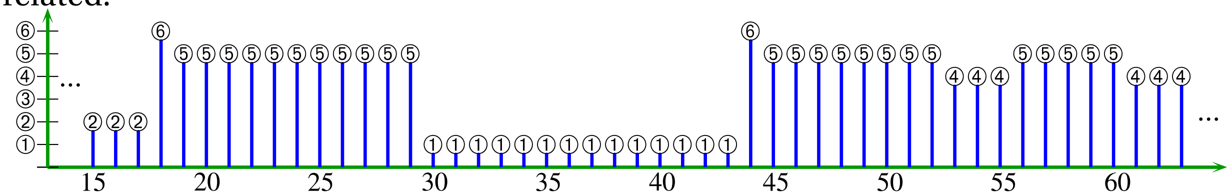


Furthermore, it is about 58% less correlated than the $\mathbb{R}^3$ filtering for the die sequence used in item (3) of Example 3.12 (page 89).

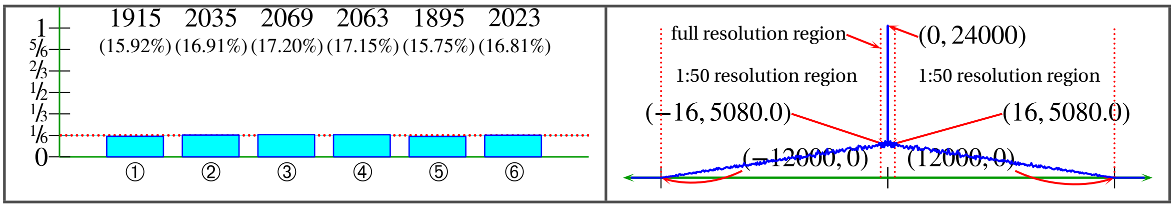
4. Using the *Euclidean metric* rather than the *Lagrange arc distance* as in item (3) results in a sequence that differs at 99 different locations out of $N + M + M - 1 = 12000 + 16 + 15 = 12031$ locations (approximately 0.8% of the locations differ).[25]



5. Using a *length 16 Hanning low pass sequence* rather than the *length 16 Rectangular low pass sequence* as in item (3) results in a distribution that is more uniform and about 5.3% less correlated:



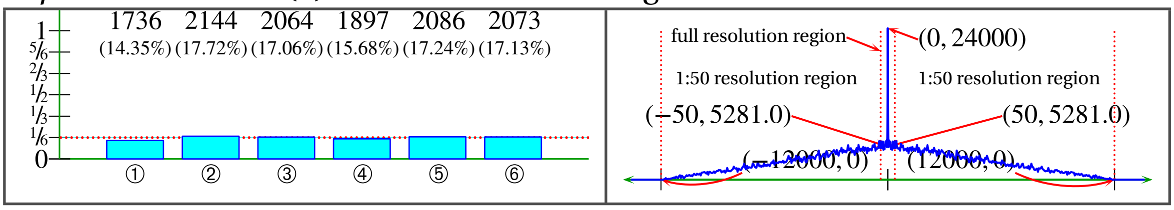
[25] See experiment log file "spin_lp_12000m16.xlg" generated by the program "ssp.exe".



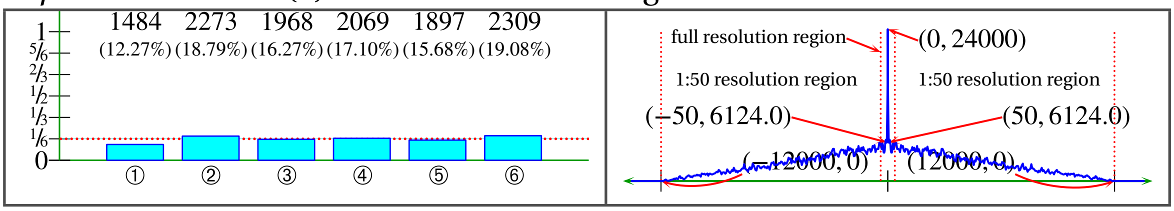
6. Using the *Euclidean metric* rather than the *Lagrange arc distance* as in item (5) results in a sequence that differs at exactly 2 locations (approximately 0.017%) out of 12031 locations:[26]

n	Euclidean	Lagrange
4149	⚃	⚄
5594	⚁	⚀

7. Using a *length 50 Hanning low pass sequence* rather than the *length 16 Hanning low pass sequence* as in item (5) results in the following:



8. Using a *length 50 Rectangular low pass sequence* rather than the *length 50 Hanning low pass sequence* as in item (7) results in the following:

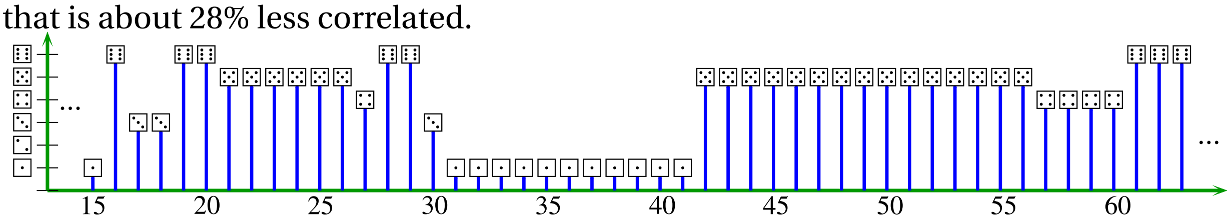


Example 3.14 (low pass filtering of fair die sequence).

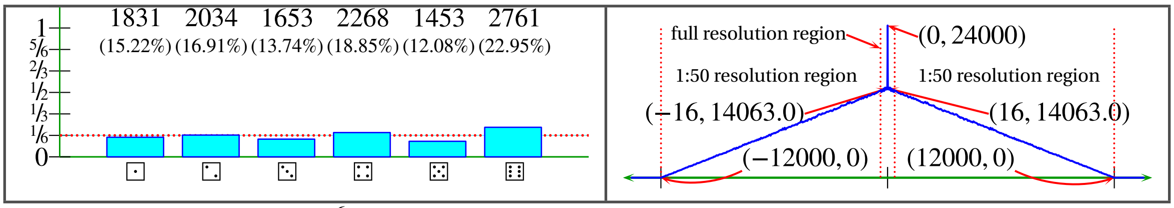
1. Consider the pseudo-uniformly distributed die sequence presented in Example 3.1 (page 82). Suppose we want to *filter* this sequence with a *low pass sequence* in order to "smooth out" the sequence, just as in Example 3.12 (page 89).

2. Suppose we first use the *traditional die random variable* (Definition 3.3 page 86) to map the die sequence into $\mathbb{R}^1$. *Filter*ing this mapped sequence using the *length 16 rectangular low pass sequence* and then mapping the result back to a *die sequence* using the *Euclidean metric*, produces a result identical to that of item (2) (page 89) of Example 3.12.

3. Alternatively, suppose we next use the $\mathbb{R}^6$ *die random variable* (Definition 3.7 page 87) to map the die sequence into $\mathbb{R}^6$. *Filter*ing this new sequence using the *length 16 rectangular low pass sequence* in the $\mathbb{R}^6$ *die distance linear space* (Definition 3.16 page 89) and then mapping back to a *die sequence* using the *Euclidean metric* yields a much more uniform distribution and a sequence that is about 28% less correlated.

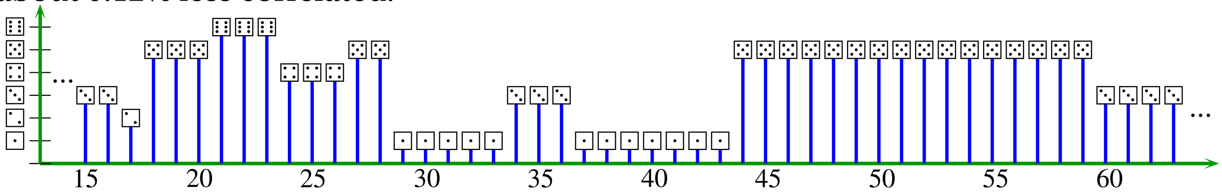


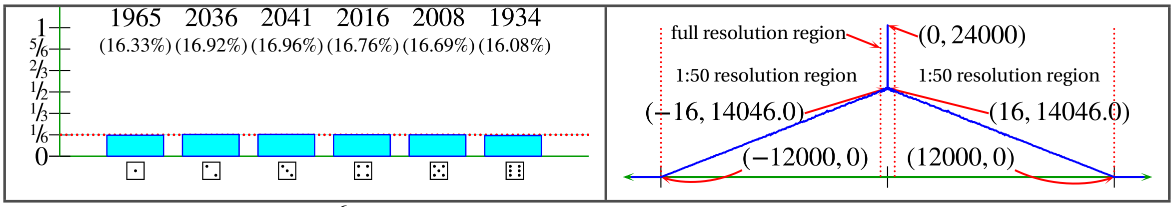
[26] See experiment log file "spin_lp_12000m16.xlg" generated by the program "ssp.exe".



Note further that this $\mathbb{R}^6$ technique yeilds a sequence that is about 9.9% more correlated than yielded by the $\mathbb{R}^3$ technique used in item (3) of Example 3.12 (page 89).

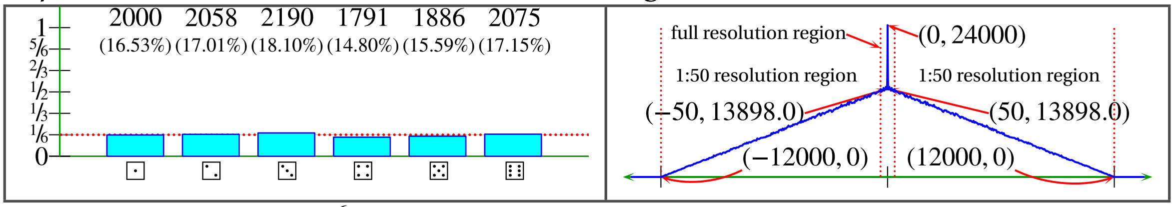
4. Using a *length 16 Hanning low pass sequence* rather than the *length 16 Rectangular low pass sequence* as in item (3) results in a distribution that is more uniform and a sequence that is about 0.12% less correlated:





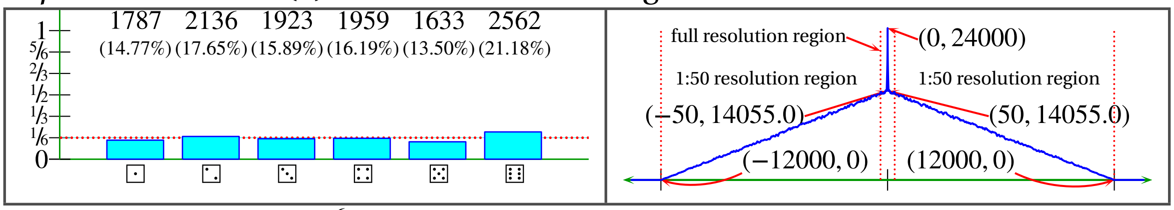
Note further that this $\mathbb{R}^6$ technique yields a sequence that is about 10% more correlated than yielded by the $\mathbb{R}^3$ technique used in item (4) of Example 3.12 (page 89).

5. Using a *length 50 Hanning low pass sequence* rather than the *length 16 Hanning low pass sequence* as in item (4) results in the following:



Note further that this $\mathbb{R}^6$ technique yields a sequence that is about 11% more correlated than yielded by the $\mathbb{R}^3$ technique used in item (5) of Example 3.12 (page 89).

6. Using a *length 50 Rectangular low pass sequence* rather than the *length 50 Hanning low pass sequence* as in item (5) results in the following:



Note further that this $\mathbb{R}^6$ technique yeilds a sequence that is about 8.8% more correlated than yeilded by the $\mathbb{R}^3$ technique used in item (6) of Example 3.12 (page 89).

7. In the *fair die outcome space*, the *Lagrange arc distance* does not seem so appropriate. That being said however, …

   (a) using the *Lagrange arc distance* rather than the *Euclidean metric* in item (3) page 93 yields results that are **identical**[27]

[27] See experiment log file "fdie_lp_12000m16.xlg" generated by the program "ssp.exe".

(b) using the *Lagrange arc distance* rather than the *Euclidean metric* in item (4) page 94 yields results that **differ** at 5 locations (differ at approximately 0.04% of the total possible $N + M + M - 1 = 12000 + 16 + 16 - 1 = 12031$ locations):[28]

n	Euclidean	Lagrange
430	⚀	⚄
2181	⚄	⚂
5055	⚁	⚀

n	Euclidean	Lagrange
8688	⚃	⚂
10866	⚃	⚄

(c) using the *Lagrange arc distance* rather than the *Euclidean metric* in item (5) page 94 yields results that are **identical**[29]

(d) using the *Lagrange arc distance* rather than the *Euclidean metric* in item (6) page 94 yields results that **differ** at 289 locations (differ at approximately 2.4% of the total possible 12031 locations):[30]

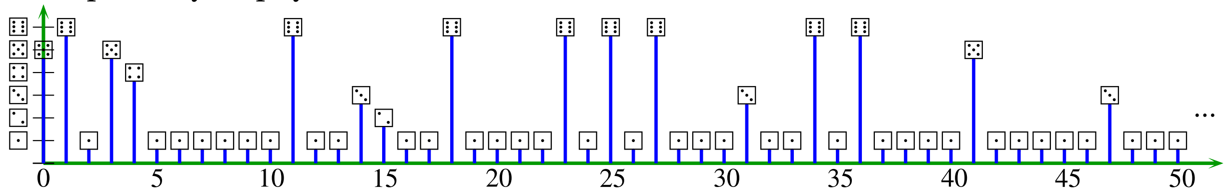
8. Empirical evidence observed in items 3, 4, 5, and 6, suggests that the $\mathbb{R}^6$ technique of this example leads to about 10% more correlation than the $\mathbb{R}^3$ technique of Example 3.12 (page 89).

3.3.2 High pass filtering

Example 3.15 (high pass filtering of weighted real die sequence).

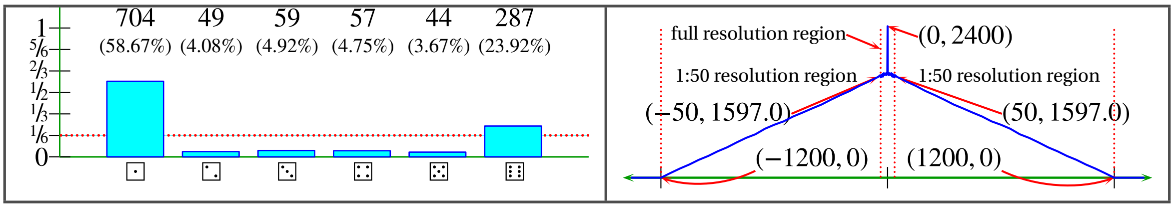
1. Consider a length $50(1200 + 2) - (50 - 1) = 60051$ non-uniformly distributed die sequence generated as described in Example 3.4 (page 83). To remove the strong ⚄ bais, we could map and *filter* (Definition 1.45 page 26) the sequence with the *length 50 high pass rectangular sequence* (Definition 1.47 page 26). Such filtering will obviously introduce correlation into the die sequence. The low pass filtering of Example 3.12 page 89 ("smoothing") also introduced correlation, but wanting a "smooth" sequence informally implies a willingness to accept a highly correlated sequence. However in this current example, we would prefer to have an *uncorrelated* sequence. To negate the correlation introduced by filtering, we *down sample* (Definition 1.42 page 25) the filtered sequence by a factor of 50 and remove the first and last element, leaving a sequence of length 1200.

2. If the filtering and downsampling described in item (1) is performed in the traditional $\mathbb{R}^1$ space, then after mapping back to a *die sequence* using the *Euclidean metric*, we obtain the result partially dispayed here…

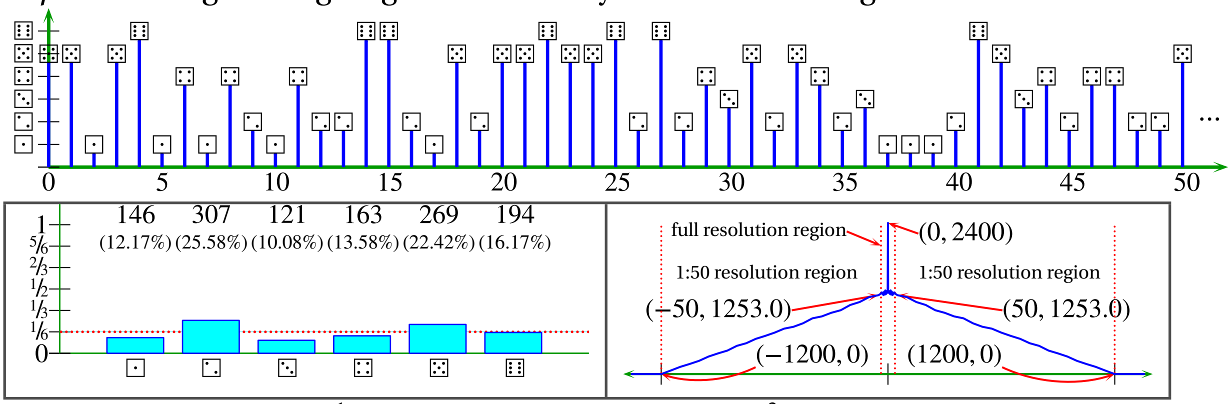


where the bias at ⚄ has been replaced by a new bias at ⚀, as illustrated quantitatively below on the left, calculated over $N = 1200$ elements.

[28] See experiment log file "fdie_lp_12000m16.xlg" generated by the program "ssp.exe".
[29] See experiment log file "fdie_lp_12000m50.xlg" generated by the program "ssp.exe".
[30] See experiment log file "fdie_lp_12000m50.xlg" generated by the program "ssp.exe".

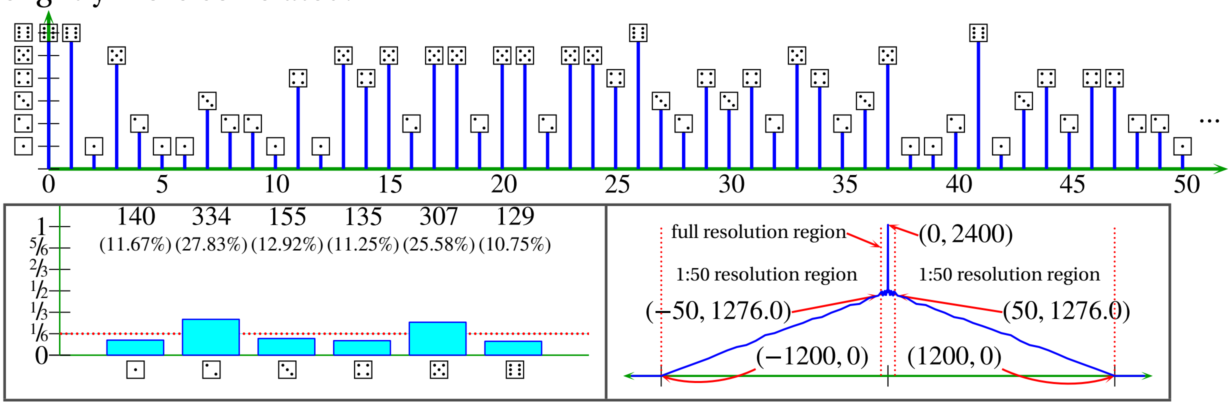


3. Alternatively, suppose we next use the $\mathbb{R}^3$ *die random variable* (Definition 3.6 page 87) to map the die sequence into $\mathbb{R}^3$. *Filter*ing this new sequence using the *length 50 rectangular high pass sequence* in the $\mathbb{R}^3$ *distance linear space* (Definition 3.14 page 88) and then mapping back to a *die sequence* using the *Lagrange arc distance* yields the following results:



Note that neither the $\mathbb{R}^1$ method of item (2) nor the $\mathbb{R}^3$ method of item (3) yields a uniformly distributed sequence; but the $\mathbb{R}^3$ method at least comes significantly closer to this end. Moreover, the $\mathbb{R}^3$ method also yields a sequence that is less correlated.

4. Replacing the *length 50 rectangular high pass filter* in item (3) with the *length 50 Hanning high pass filter* (Definition 1.49 page 27) yields a different sequence with similar distribution but is slightly more correlated:



5. Replacing the *Lagrange arc distance* by the *Euclidean metric* in this example has very little effect, even before downsampling. Before downsampling, the length of each sequence is $M(N+2) = 50(1202) = 60100$ elements. More details follow:

   (a) Using the *Euclidean metric* rather than the *Lagrange arc distance* in item (3) yields results that are **identical**.[31]

   (b) Using the *Euclidean metric* rather than the *Lagrange arc distance* in item (4) yields results that **differ** at 4 locations (approximately 0.007% of all the locations).[32]

6. For the type of sequence processing described in this example, item (5) very informally *suggests* the following:

[31] See experiment log file "wrdie_hp_1200m50.xlg" generated by the program "ssp.exe".
[32] See experiment log file "wrdie_hp_1200m50.xlg" generated by the program "ssp.exe".

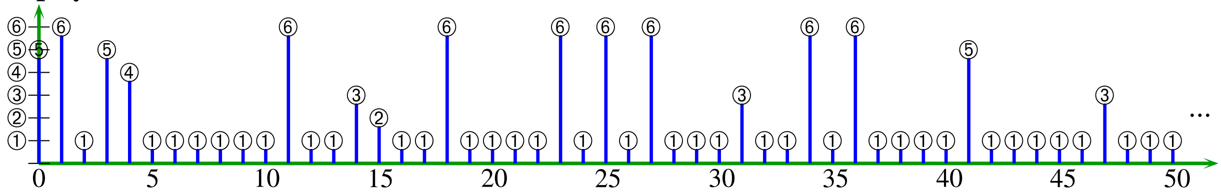
(a) The processing is not highly sensitive to the choice of distance function.

(b) The processing is not heavily dependent on the *triangle inequality*.

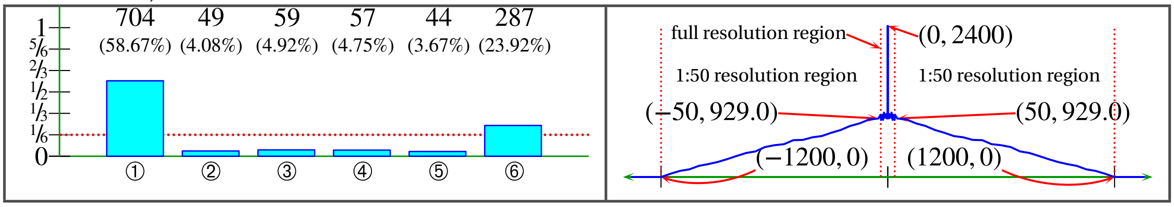
Example 3.16 (high pass filtering of weighted spinner sequence).

1. Consider a length $50(1200+2)-(50-1)=60051$ non-uniformly distributed *spinner sequence* generated as described in Example 3.6 (page 84). To remove the strong ⑤ bais, we could *filter* the *sequence* with the *length 50 high pass rectangular sequence* and down sample the filtered sequence by a factor of 50, as described in Example 3.15 (page 95).

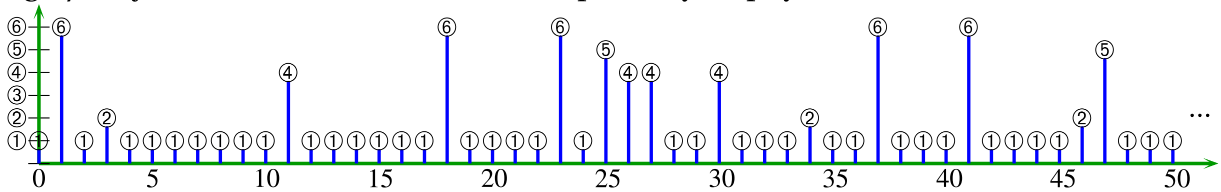
2. If the filtering described in item (1) is performed in the traditional $\mathbb{R}^1$ space, then after mapping back to a *spinner sequence* using the *Euclidean metric*, we obtain the result partially dispayed here…



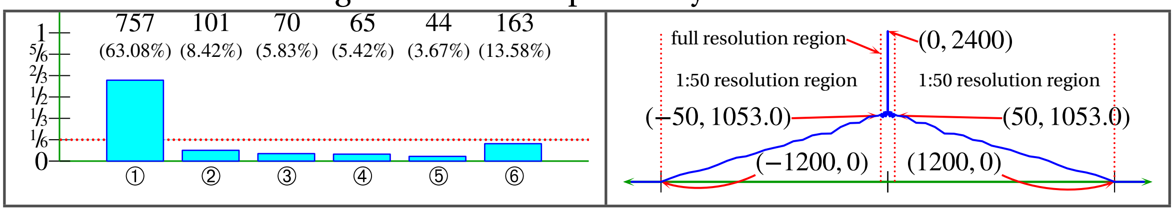
where the bias at ⑤ has been replaced by a new bias at ①, as illustrated quantitatively below on the left, calculated over 1200 elements.



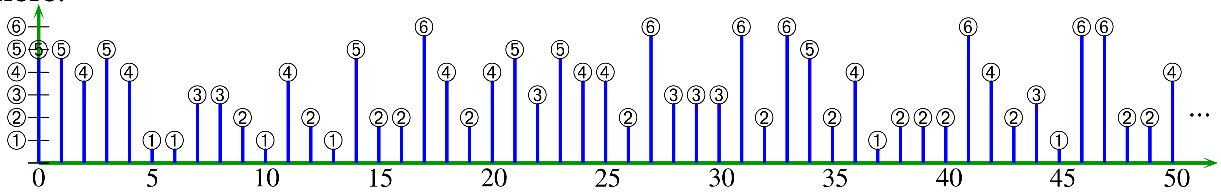
3. If we replace the *length 50 rectangular high pass filter* of item (2) with a *length 50 Hanning high pass filter* then we obtain the result partially dispayed here…

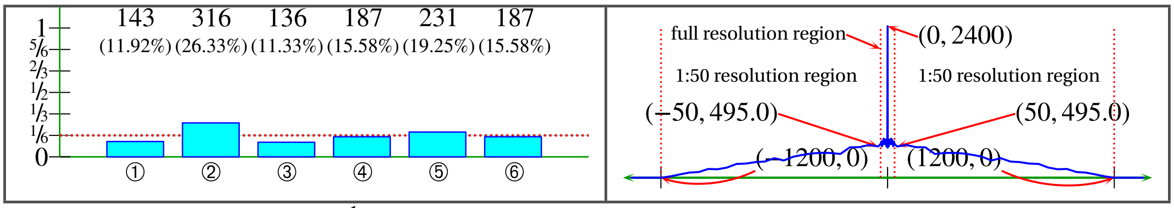


…where the bias at ⑤ again has been replaced by a new bias at ①:



4. If the rectangular filtering in $\mathbb{R}^1$ of item (2) is instead performed in $\mathbb{R}^2$ and mapped back to a *spinner sequence* using the *Lagrange arc distance*, then we obtain the result partially dispayed here:

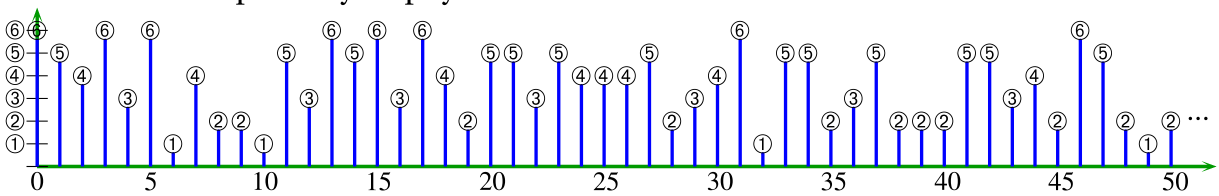


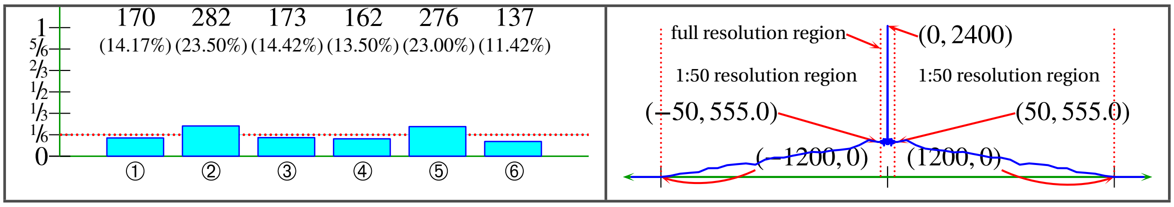


Note that neither the $\mathbb{R}^1$ methods (described in item (2) and item (3)) nor the $\mathbb{R}^2$ method (described in item (4)) yields a uniformly distributed sequence; but the $\mathbb{R}^2$ method at least comes significantly closer to this end.

5. Replacing the *Lagrange arc distance* by the *Euclidean metric* as in item (4) yields a sequence that differs at a total of 272 locations (approximately 0.5% of the locations).[33]

6. If instead of using the rectangular filtering (as in item (4)), we use the Hanning filtering of item (3) in $\mathbb{R}^2$ and map back to a *spinner sequence* using the *Lagrange arc distance*, then we obtain the result partially dispayed here:



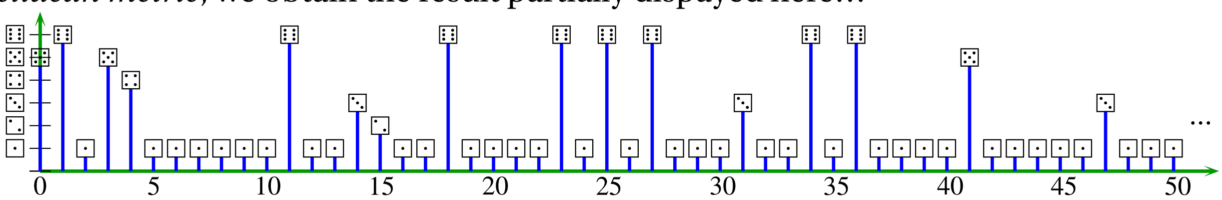


7. Replacing the *Lagrange arc distance* by the *Euclidean metric* in item (6) yields a sequence that differs at a total of 3 locations (approximately 0.005% of the locations).[34]

Example 3.17 (high pass filtering of weighted die sequence).

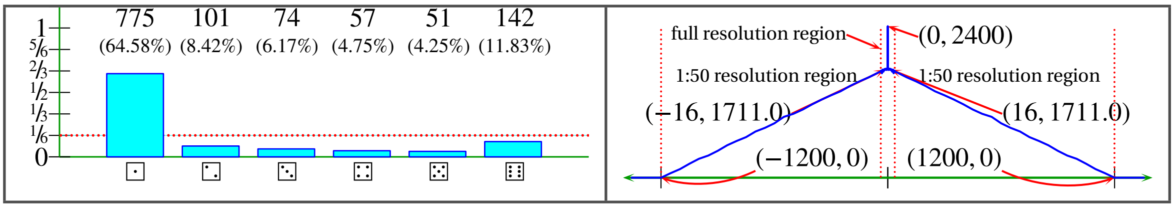
1. Consider a length $50(1200+2)-(50-1) = 60051$ *weighted die sequence* generated as described in Example 3.5 (page 83). To remove the strong ⚄ bais, we could map and *filter* the sequence with the *length 16 high pass rectangular sequence* (Example 1.18 page 27). To negate the correlation introduced by filtering, we *down sample* the filtered sequence by a factor of 16.

2. If the die sequence of item (1) is mapped into $\mathbb{R}^1$ using the *traditional die random variable* (Definition 3.3 page 86), *filter*ed, *down sample*d, and mapped back to a die sequence using the *Euclidean metric*, we obtain the result partially dispayed here…

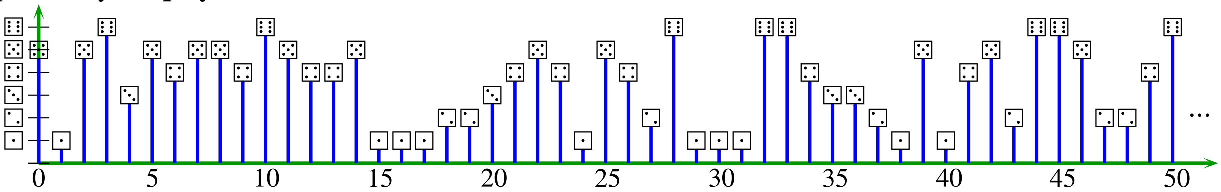


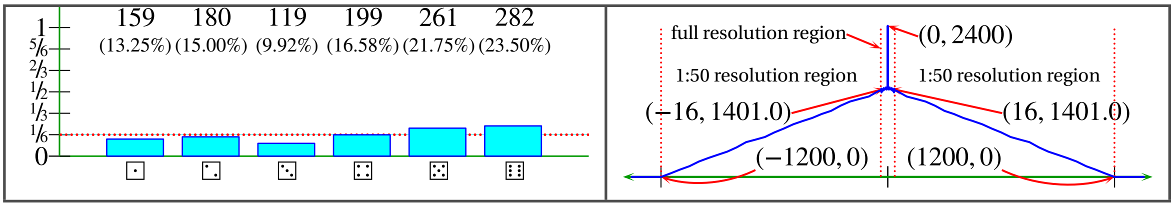
where the bias at ⚄ has been replaced by a new bias at ⚀, as illustrated quantitatively below on the left, calculated over 1200 elements.

[33] See experiment log file "wspin_hp_1200m50.xlg" generated by the program "ssp.exe".
[34] See experiment log file "wspin_hp_1200m50.xlg" generated by the program "ssp.exe".

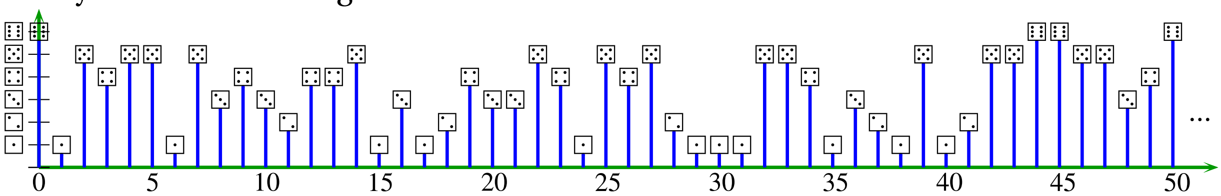


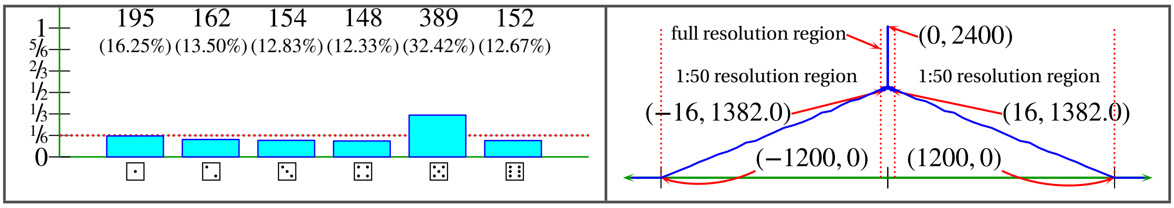
3. But if instead of processing the die sequence in $\mathbb{R}^1$ as in item (2), processing is performed in $\mathbb{R}^6$ and mapped back to a die sequence using the *Euclidean metric*, then we obtain the result partially dispayed here:



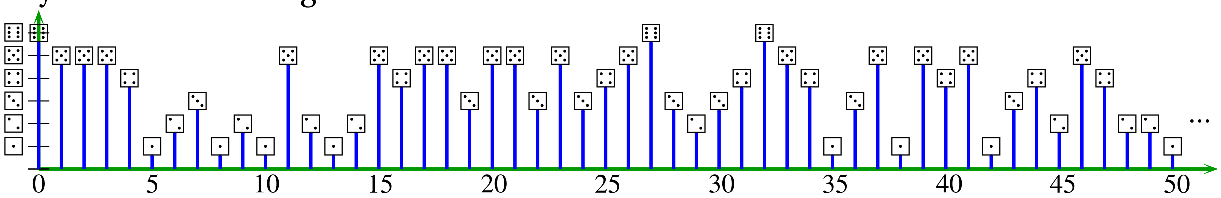


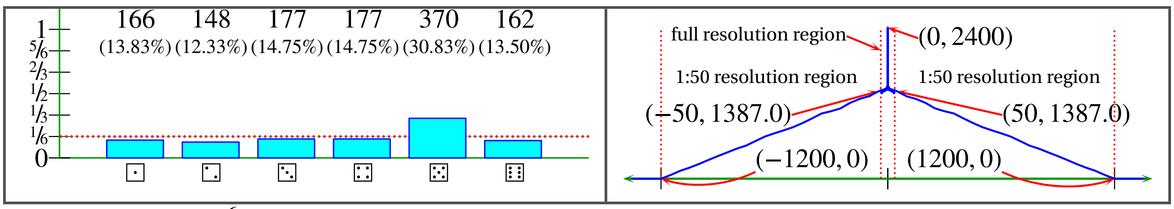
4. Replacing the *length 16 rectangular sequence* in item (3) with a *length 16 Hanning sequence* in $\mathbb{R}^6$ yields the following results:





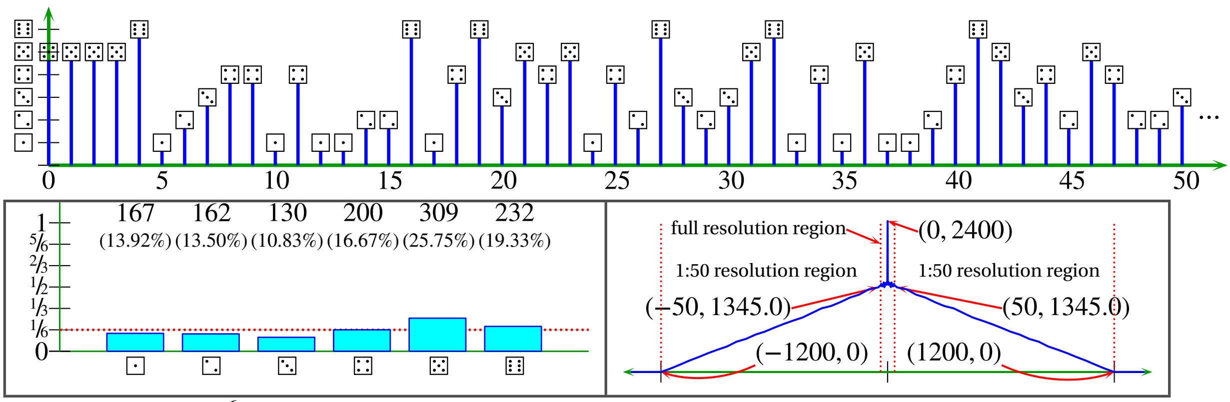
5. Replacing the *length 16 Hanning sequence* in item (4) with a *length 50 Hanning sequence* in $\mathbb{R}^6$ yields the following results:





Note that this $\mathbb{R}^6$ technique yields a sequence that is about 8.7% more correlated than yielded by the $\mathbb{R}^3$ technique used in item (4) of Example 3.15 (page 95).

6. Replacing the *length 50 Hanning sequence* in item (5) with a *length 50 rectangular sequence* in $\mathbb{R}^6$ yields the following results:



Note that this $\mathbb{R}^6$ technique yields a sequence that is about 7.3% more correlated than yielded by the $\mathbb{R}^3$ technique used in item (3) of Example 3.15 (page 95).

7. As in Example 3.14 (page 93), here again the *Lagrange arc distance* does not seem so appropriate. That again being said however, …

   (a) using the *Lagrange arc distance* rather than the *Euclidean metric* in item (3) page 99 yields results that are **identical.**[35]

   (b) using the *Lagrange arc distance* rather than the *Euclidean metric* in item (4) page 99 yields results that **differ** at 17 locations (differ at approximately 0.09% of the total possible $M(N+2) = 16(1200+2) = 19232$ locations).[36]

8. Empirical evidence observed in items item (5) and item (6) suggests that the $\mathbb{R}^6$ technique of this example leads to about 8% more correlation than the $\mathbb{R}^3$ technique of Example 3.15 (page 95).

3.3.3 Fourier Analysis

Example 3.18 (length 1200 non-stationary die sequence with 10Hz oscillating mean).

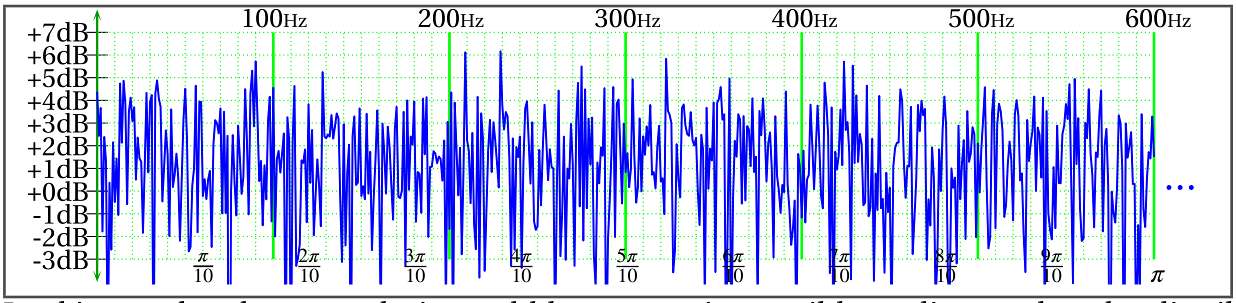
1. Suppose we have a length $N \triangleq 1200$ die sequence $(\!(x_n)\!)$ with the following distribution:

$$\mathsf{P}(⚀) = \mathsf{P}(⚁) = \mathsf{P}(⚃) = \mathsf{P}(⚄) = \mathsf{P}(⚅) = 0.15 \quad \text{and} \quad \mathsf{P}(⚂) = 0.25$$
$$\text{for } n \in \left\{ p + (2m)\frac{M}{2} \middle| p = 0, 1, \ldots, \frac{M}{2} - 1,\ m = 0, 1, 2, \ldots, 9 \right\} \qquad \text{and}$$
$$\mathsf{P}(⚀) = \mathsf{P}(⚁) = \mathsf{P}(⚂) = \mathsf{P}(⚄) = \mathsf{P}(⚅) = 0.15 \quad \text{and} \quad \mathsf{P}(⚃) = 0.25$$
$$\text{for } n \in \left\{ p + (2m+1)\frac{M}{2} \middle| p = 0, 1, \ldots, \frac{M}{2} - 1,\ m = 0, 1, 2, \ldots, 9 \right\}$$

   where $M \triangleq 120$. That is, the distribution of the sequence oscillates every $M/2 = 60$ samples between one that favors ⚂ and one that favors ⚃. Moreover, if we were to evaluate the sequence using a *Discrete Fourier Transform* operator DFT (Definition 1.52 page 27), we might expect to see a strong component at $\frac{N}{M} = 10$ (or 10 Hz—the distribution goes through 10 cycles during the course of the sequence).

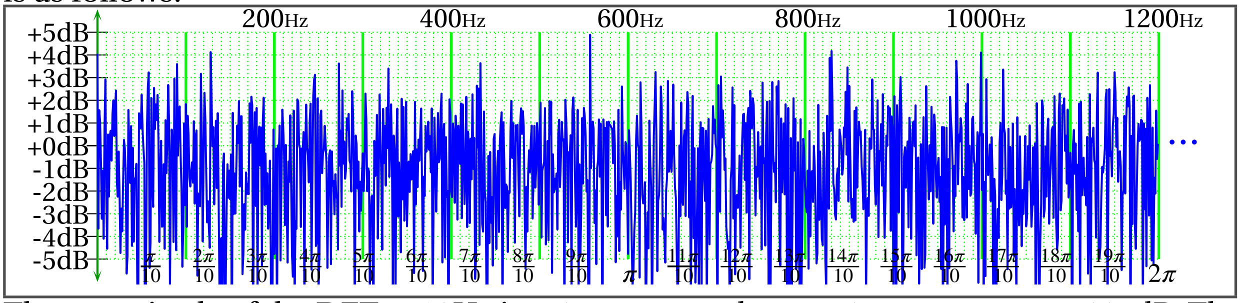
2. Suppose we first use the *PAM die random variable* (Definition 3.4 page 87) to map the sequence of item (1) into $\mathbb{R}^1$. The magnitude of the $\mathsf{DFT} : \mathbb{R}^1 \to \mathbb{C}^1$ of the mapped sequence is as follows:

[35] See experiment log file "wdie_hp_1200m16.xlg" generated by the program "ssp.exe".
[36] See experiment log file "wdie_hp_1200m16.xlg" generated by the program "ssp.exe".



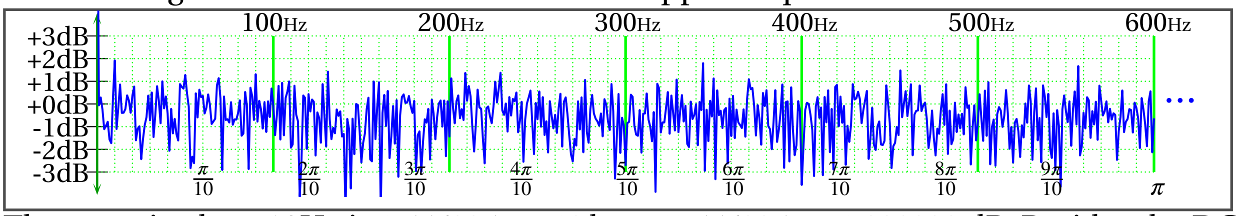
Looking at the above result, it would be next to impossible to discern that the distribution had a significantly strong oscillation of 10 cycles. In fact, the magnitude of the DFT at 10Hz is only 0.895699, or $10\log_{10}(0.895699) = -0.478377$ dB. There are exactly 456 out of a total $N/2 = 600$ values that are greater than the DFT magnitude at 10Hz.[37] That is, to either a human observer or a machine algorithm, the 10Hz component is effectively lost in the noise.

3. Suppose we next use the *QPSK die random variable* (Definition 3.5 page 87) to map the sequence into the complex plane. The magnitude of the DFT : $\mathbb{C}^1 \to \mathbb{C}^1$ operation on the mapped sequence is as follows:



The magnitude of the DFT at 10Hz is 0.589990, or $10\log_{10}(0.589990) = -2.291552$ dB. There are exactly 831 out of a total $N = 1200$ values that are greater than the DFT magnitude at 10Hz.[38] Again, the 10Hz component is effectively lost in the noise.

4. Suppose we next use the $\mathbb{R}^6$ *die random variable* (Definition 3.7 page 87) to map the sequence into $\mathbb{R}^6$. The magnitude of DFT : $\mathbb{R}^6 \to \mathbb{C}^6$ of the mapped sequence is as follows:



The magnitude at 10Hz is 1.556295, or $10\log_{10}(1.556295) = 1.920920$ dB. Besides the DC component (0Hz component), this is the uniquely greatest value of the 600 samples. And in fact, there are only 5 out of a total $N/2 = 600$ samples that are 0.90×1.556295 or greater.[39] Thus, using the $\mathbb{R}^6$ mapping technique of this example, it is much simpler to detect the 10Hz oscillating distribution.

Example 3.19 (length 12000 non-stationary die sequence with 10Hz oscillating mean).

1. Suppose we have a length $N \triangleq 12000$ die sequence $(\!(x_n)\!)$ with the following distribution:

[37] See experiment log file "diedft_1525_1200m120.xlg" generated by the program "ssp.exe".
[38] See experiment log file "diedft_1525_1200m120.xlg" generated by the program "ssp.exe".
[39] See experiment log file "diedft_1525_1200m120.xlg" generated by the program "ssp.exe". The 5 largest values are the points (0, 14.251433), (10, 1.556295), (344, 1.513501), (456, 1.405843) and (557, 1.468970).

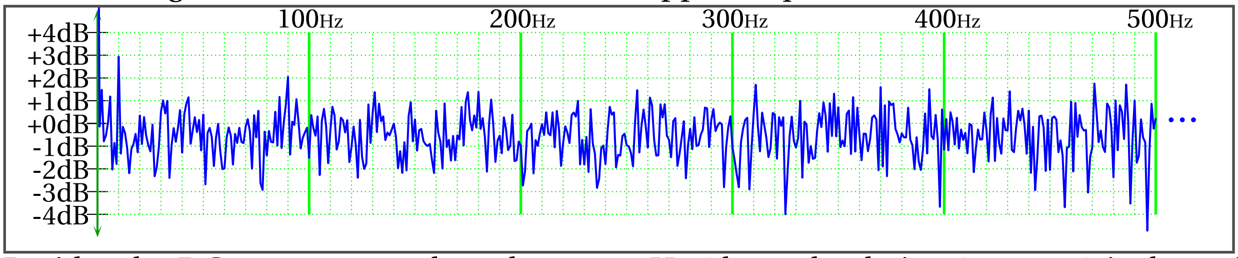
$$\begin{aligned}&\mathsf{P}(⚀) = \mathsf{P}(⚁) = \mathsf{P}(⚃) = \mathsf{P}(⚄) = \mathsf{P}(⚅) = 0.16 \quad\text{and}\quad \mathsf{P}(⚂) = 0.20\\ &\qquad \text{for } n \in \left\{p + (2m)\tfrac{M}{2} \,|\, p = 0, 1, \ldots, \tfrac{M}{2} - 1,\ m = 0, 1, 2, \ldots, 9\right\} \qquad \text{and}\\ &\mathsf{P}(⚀) = \mathsf{P}(⚁) = \mathsf{P}(⚂) = \mathsf{P}(⚄) = \mathsf{P}(⚅) = 0.16 \quad\text{and}\quad \mathsf{P}(⚃) = 0.20\\ &\qquad \text{for } n \in \left\{p + (2m+1)\tfrac{M}{2} \,|\, p = 0, 1, \ldots, \tfrac{M}{2} - 1,\ m = 0, 1, 2, \ldots, 9\right\}\end{aligned}$$

where $M \triangleq 1200$. That is, the distribution of the sequence oscillates every $M/2 = 600$ samples between one that favors ⚂ and one that favors ⚃. If we were to evaluate the sequence using the *Discrete Fourier Transform* operator, we again might expect to see a strong component at $\frac{N}{M} = 10$ (or 10 Hz—the distribution goes through 10 cycles during the course of the sequence).

2. Suppose we first use the *PAM die random variable* (Definition 3.4 page 87) to map the sequence of item (1) into $\mathbb{R}^1$. In the magnitude of $\mathbf{DFT} : \mathbb{R}^1 \to \mathbb{C}^1$ there are 1130 values out of a possible $N/2 = 6000$ values greater than the value at 10Hz (that value being 2.174512).[40] As in Example 3.18 (page 100), the subtle 10Hz component is effectively lost in the noise.

3. Suppose we next use the *QPSK die random variable* (Definition 3.5 page 87) to map the sequence into the complex plane. There are exactly 1932 out of a total $N = 12000$ values that are greater than the DFT value at 10Hz (that value being 1.348693).[41] As in Example 3.18 (page 100), the subtle 10Hz component is effectively lost in the noise.

4. Suppose we next use the $\mathbb{R}^6$ *die random variable* (Definition 3.7 page 87) to map the sequence into $\mathbb{R}^6$. The magnitude of $\mathbf{DFT} : \mathbb{R}^6 \to \mathbb{C}^6$ of the mapped sequence is as follows:



Besides the DC component, the value at 100Hz (that value being (1.965018) is the uniquely greatest value of the $N/2 = 6000$ samples; and it is $10\log_{10}(1.965018/1.660189) = 0.699\cdots$dB larger than the next largest value.[42] Thus, even though the oscillating distribution is very subtle (even more subtle than that of Example 3.18 (page 100)), the $\mathbb{R}^6$ mapping technique and subsequent analysis are able to detect it.

Example 3.20 (length 12000 non-stationary die sequence with 100Hz oscillating mean).

1. Suppose we have a length $N \triangleq 12000$ die sequence $(\!(x_n)\!)$ with the following distribution:

$$\begin{aligned}&\mathsf{P}(⚀) = \mathsf{P}(⚁) = \mathsf{P}(⚃) = \mathsf{P}(⚄) = \mathsf{P}(⚅) = 0.16 \quad\text{and}\quad \mathsf{P}(⚂) = 0.20\\ &\qquad \text{for } n \in \left\{p + (2m)\tfrac{M}{2} \,|\, p = 0, 1, \ldots, \tfrac{M}{2} - 1,\ m = 0, 1, 2, \ldots, 9\right\} \qquad \text{and}\\ &\mathsf{P}(⚀) = \mathsf{P}(⚁) = \mathsf{P}(⚂) = \mathsf{P}(⚄) = \mathsf{P}(⚅) = 0.16 \quad\text{and}\quad \mathsf{P}(⚃) = 0.20\\ &\qquad \text{for } n \in \left\{p + (2m+1)\tfrac{M}{2} \,|\, p = 0, 1, \ldots, \tfrac{M}{2} - 1,\ m = 0, 1, 2, \ldots, 9\right\}\end{aligned}$$

where $M \triangleq 120$. That is, the distribution of the sequence oscillates every $M/2 = 60$ samples between one that favors ⚂ and one that favors ⚃. If we were to evaluate the sequence using the *Discrete Fourier Transform* operator, we might expect to see a strong component at $\frac{N}{M} = 100$ (or 100 Hz—the distribution goes through 100 cycles during the course of the sequence).

[40] 💾 See experiment log file "diedft_1620_12000m1200.xlg" generated by the program "ssp.exe".

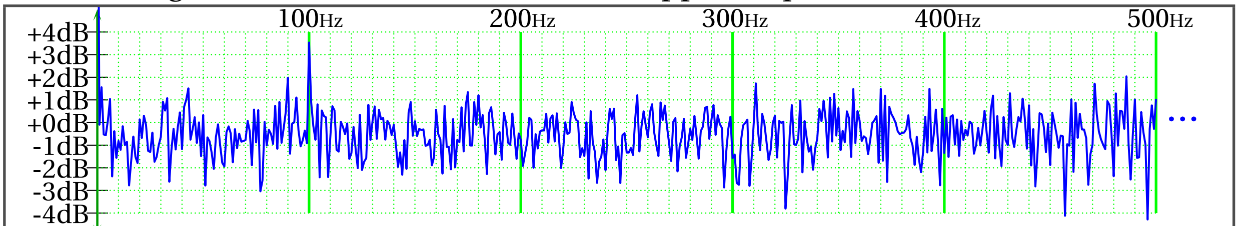
[41] 💾 See experiment log file "diedft_1620_12000m1200.xlg" generated by the program "ssp.exe".

[42] 💾 See experiment log file "diedft_1620_12000m1200.xlg" generated by the program "ssp.exe". The 10 largest values are (0, 44.763194), (10, 1.965018), (90, 1.602474), (1223, 1.660189), (1313, 1.555349), (2385, 1.551028), (3039, 1.550918), (4154, 1.563756), (4187, 1.586362), and (5147, 1.623052).

2. Suppose we first use the *PAM die random variable* (Definition 3.4 page 87) to map the sequence of item (1) into $\mathbb{R}^1$. In the magnitude DFT : $\mathbb{R}^1 \to \mathbb{C}^1$ there are 1320 values out of a possible $N/2 = 6000$ values greater than the value at 100Hz that value being 2.081469).[43] The subtle 100Hz component is effectively lost in the noise.

3. Suppose we next use the *QPSK die random variable* (Definition 3.5 page 87) to map the sequence into the complex plane. There are exactly 1555 out of a total N12000 values that are greater than the DFT value at 100Hz that value being 1.425427).[44] The subtle 100Hz component is effectively lost in the noise.

4. Suppose we next use the $\mathbb{R}^6$ *die random variable* (Definition 3.7 page 87) to map the sequence into $\mathbb{R}^6$. The magnitude of DFT : $\mathbb{R}^6 \to \mathbb{C}^6$ of the mapped sequence is as follows:



Besides the DC component, the value at 100Hz (that value being 2.256927) is the uniquely greatest value of the $N/2 = 6000$. and it is $10\log_{10}(2.256927/1.599335) = 1.495\cdots$dB larger than the next largest value.[45] Thus, even though the oscillating distribution is very subtle, the $\mathbb{R}^6$ mapping technique and subsequent analysis are able to detect it.

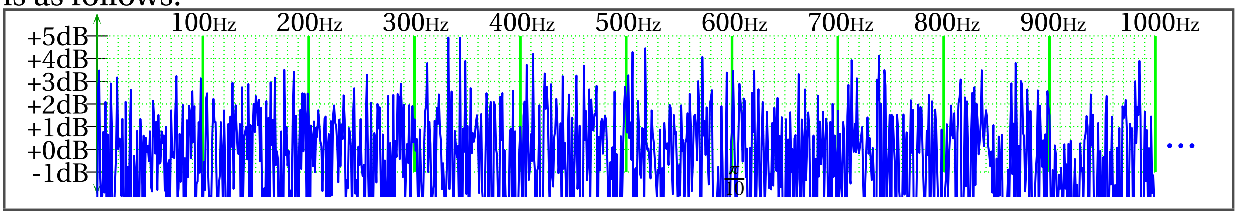
Example 3.21 (length 12000 non-stationary artificial DNA sequence with 10Hz oscillating mean).

1. Suppose we have a length $N \triangleq 12000$ die sequence $(\!(x_n)\!)$ with the following distribution (see also Figure 3.1 page 104):

$$\begin{aligned}
&\mathsf{P}(\boxed{A}) = \mathsf{P}(\boxed{T}) = \mathsf{P}(\boxed{G}) = 0.24 \quad \text{and} \quad \mathsf{P}(\boxed{C}) = 0.28 \\
&\qquad \text{for } n \in \left\{ p + 2m\frac{M}{2} \middle| p = 0,1,\ldots,\frac{M}{2}-1,\ m = 0,1,2,\ldots,9 \right\} \qquad \text{and} \\
&\mathsf{P}(\boxed{A}) = \mathsf{P}(\boxed{C}) = \mathsf{P}(\boxed{G}) = 0.24 \quad \text{and} \quad \mathsf{P}(\boxed{T}) = 0.28 \\
&\qquad \text{for } n \in \left\{ p + (2m+1)\frac{M}{2} \middle| p = 0,1,\ldots,\frac{M}{2}-1,\ m = 0,1,2,\ldots,9 \right\}
\end{aligned}$$

where $M \triangleq 1200$. That is, the distribution of the sequence oscillates every $M/2 = 600$ samples between one that favors $\boxed{C}$ and one that favors $\boxed{T}$. Moreover, if we were to evaluate the sequence using a *Discrete Fourier Transform* (DFT) operator, we might expect to see a strong component at $\frac{N}{M} = 10$ (or 10 Hz—the distribution goes through 10 cycles during the course of the sequence).

2. Suppose we first use the *PAM DNA random variable* (Definition 3.10 page 87) to map the DNA sequence into $\mathbb{R}^1$. The magnitude of DFT : $\mathbb{R}^1 \to \mathbb{C}^1$ of the sequence after applying this mapping is as follows:



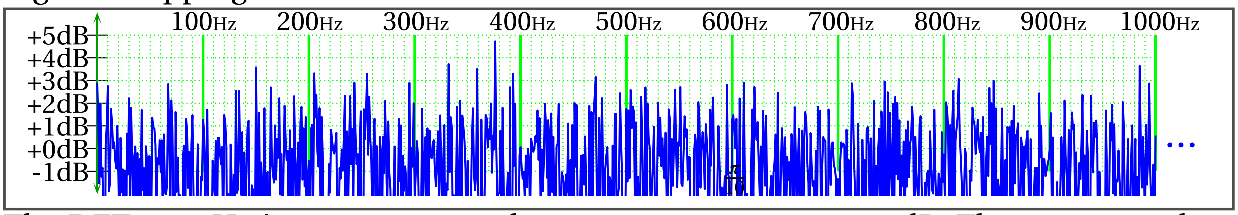
[43] See experiment log file "diedft_1620_12000m120.xlg" generated by the program "ssp.exe".
[44] See experiment log file "diedft_1620_12000m120.xlg" generated by the program "ssp.exe".
[45] See experiment log file "diedft_1620_12000m120.xlg" generated by the program "ssp.exe". The 10 largest values are (0, 44.763060), (90, 1.577597), (100, 2.256927), (486, 1.599335), (1223, 1.585154), (1313, 1.547956), (3039, 1.553522), (3162, 1.561863), (5147, 1.558487), and (5567, 1.533659).

cycle	domain			P(A)	P(C)	P(G)	P(T)
0	0	–	599	0.24	0.28	0.24	0.24
	600	–	1199	0.24	0.24	0.24	0.28
1	1200	–	1799	0.24	0.28	0.24	0.24
	1800	–	2399	0.24	0.24	0.24	0.28
2	2400	–	2999	0.24	0.28	0.24	0.24
	3000	–	3599	0.24	0.24	0.24	0.28
3	3600	–	4199	0.24	0.28	0.24	0.24
	4200	–	4799	0.24	0.24	0.24	0.28
4	4800	–	5399	0.24	0.28	0.24	0.24
	5400	–	5999	0.24	0.24	0.24	0.28
cycle	domain			P(A)	P(C)	P(G)	P(T)
5	6000	–	6599	0.24	0.28	0.24	0.24
	6600	–	7199	0.24	0.24	0.24	0.28
6	7200	–	7799	0.24	0.28	0.24	0.24
	7800	–	8399	0.24	0.24	0.24	0.28
7	8400	–	8999	0.24	0.28	0.24	0.24
	9000	–	9599	0.24	0.24	0.24	0.28
8	9600	–	10199	0.24	0.28	0.24	0.24
	10200	–	10799	0.24	0.24	0.24	0.28
9	10800	–	11399	0.24	0.28	0.24	0.24
	11400	–	11999	0.24	0.24	0.24	0.28

Figure 3.1: Distribution used in Example 3.21 (page 103)

The magnitude of the DFT at 10Hz is only 1.163575 ($10\log_{10}(1.163575) = 0.657944$ dB). There are exactly 2023 out of a total $N/2 = 6000$ values that are greater than the DFT value at 10Hz (that value being 1.163575).[46] Here again, the 10Hz component is effectively lost in the noise.

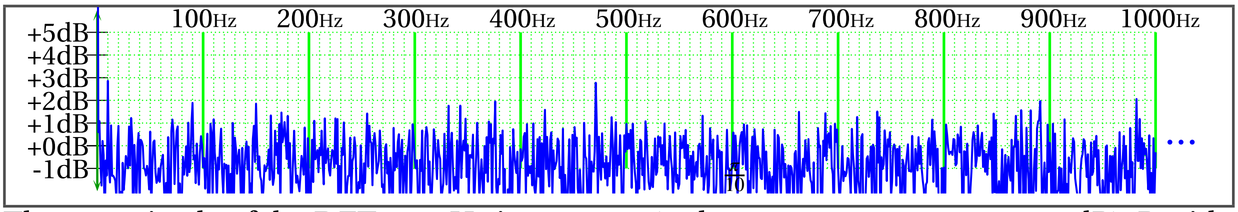
3. Suppose we next use the *QPSK DNA random variable* (Definition 3.5 page 87) to map the DNA sequence into the complex plane. The magnitude of $\mathsf{DFT} : \mathbb{C}^1 \to \mathbb{C}^1$ of the sequence after applying this mapping is as follows:



The DFT at 10Hz is 1.888671, or $10\log_{10}(1.888671) = 2.761563$ dB. There are exactly 343 out of a total $N = 6000$ values that are greater than the DFT value at 10Hz. (that value being 1.888671).[47] Using this mapping it would be difficult to detect the subtle but significant 10Hz component.

4. Suppose we next use the $\mathbb{R}^4$ *DNA random variable* (Definition 3.12 page 88) to map the sequence into $\mathbb{R}^4$. The magnitude of $\mathsf{DFT} : \mathbb{R}^4 \to \mathbb{C}^4$ of the mapped sequence is as follows:

[46] See experiment log file "dnadft_12000m1200.xlg" generated by the program "ssp.exe".
[47] See experiment log file "dnadft_12000m1200.xlg" generated by the program "ssp.exe".

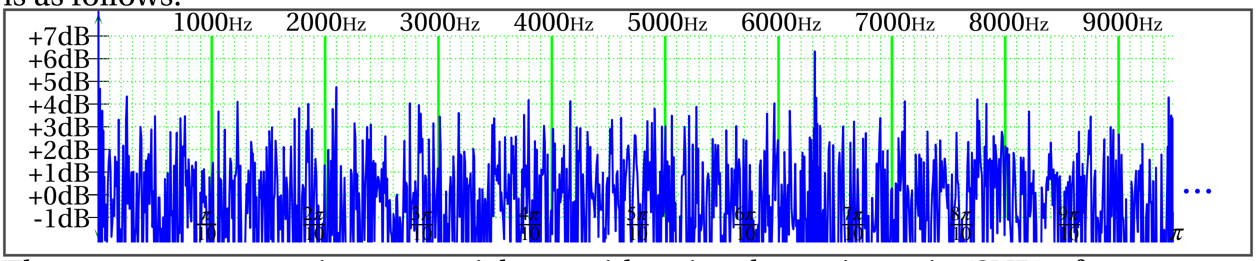


The magnitude of the DFT at 10Hz is 1.932042 ($10\log_{10}(1.932042) = 2.860166$ dB). Besides itself and the DC component, there are only two out of a total $N/2 = 6000$ samples that are greater or equal to this value.[48] Thus, using the $\mathbb{R}^4$ mapping technique and subsequent analysis of this example, it is much simpler to detect the 10Hz oscillation.

Example 3.22 (Fourier analysis of Ebola DNA sequence).

1. Consider the Ebola DNA sequence described in Example 3.9 (page 85). DNA sequences commonly exhibit a strong DFT harmonic component at $2\pi/3$ radians.[49]

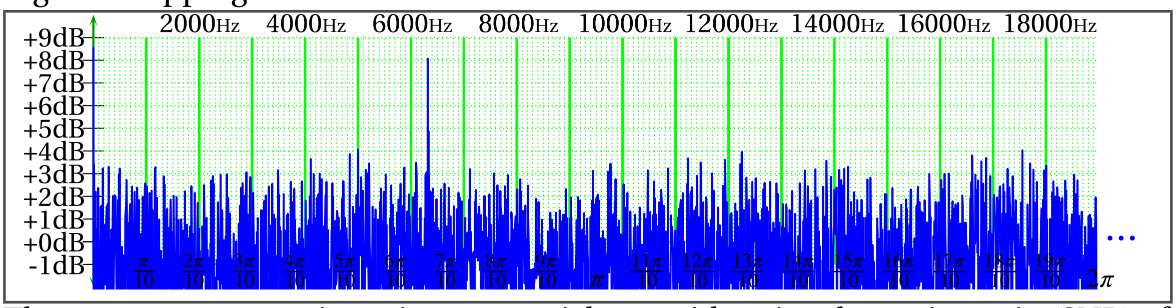
2. Suppose we first use the *PAM DNA random variable* (Definition 3.10 page 87) to map the DNA sequence into $\mathbb{R}^1$. The magnitude of $\mathsf{DFT} : \mathbb{R}^1 \to \mathbb{C}^1$ of the sequence after applying this mapping is as follows:



The component at $2\pi/3$ is easy to pick out with a signal to noise ratio (SNR) of $10\log_{10}(4.290296/1.123163) \approx 5.8$ dB.[50] Here, the noise value 1.123163 is the *RMS* (*root mean square*) of the DFT magnitude sequence from $n = 1$ to $n = N/2 - 1$ computed as follows:

$$\sqrt{\frac{1}{N/2-1}\sum_{n=1}^{n=N/2-1} x_n^2}.$$

3. Suppose we next use the *QPSK DNA random variable* (Definition 3.5 page 87) to map the dna sequence into the complex plane. The magnitude of $\mathsf{DFT} : \mathbb{C}^1 \to \mathbb{C}^1$ of the sequence after applying this mapping is as follows:



The component at $2\pi/3$ is again easy to pick out with a signal to noise ratio (SNR) of $10\log_{10}(6.412578/0.998659) \approx 8.1$ dB.[51] Here, the noise value 0.998659 is the *RMS* of the DFT magnitude sequence from $n = 1$ to $n = N - 1$.

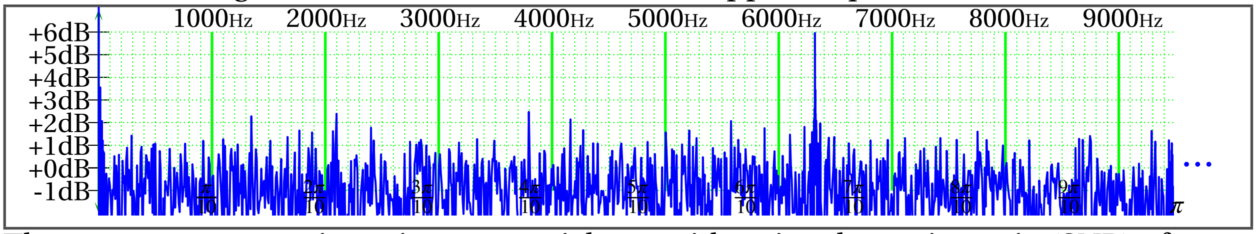
[48] See experiment log file "dnadft_12000m1200.xlg" generated by the program "ssp.exe".
The 4 largest values are at (0, 54.791926), (10, 1.932042), (4187, 1.962836), and (5147, 2.057553).

[49] Galleani and Garello (2010) page 771

[50] See experiment log file "dna_AF086833_ebola_dft.xlg" generated by the program "ssp.exe".

[51] See experiment log file "dna_AF086833_ebola_dft.xlg" generated by the program "ssp.exe".

4. Suppose we next use the $\mathbb{R}^4$ *DNA random variable* (Definition 3.12 page 88) to map the sequence into $\mathbb{R}^4$. The magnitude of DFT : $\mathbb{R}^4 \to \mathbb{C}^4$ of the mapped sequence is as follows:



The component at $^{2\pi}/_3$ is again easy to pick out with a signal to noise ratio (SNR) of $10\log_{10}(3.944811/0.860665) \approx 6.6$ dB.[52] Here, the noise value 0.860665 is the *RMS* of the DFT magnitude sequence from $n = 1$ to $n = \mathsf{N}/2 - 1$.

5. In conclusion, for this application, there is only a small advantage to using the $\mathbb{R}^4$ mapping (item (4)) versus the $\mathbb{R}^1$ mapping (item (2)), and even a demonstrable disadvantage when compared to the $\mathbb{C}^1$ mapping (item (3)).

3.3.4 Wavelet Analysis

In this section, we use what is in *essense* wavelet analysis, but yet is not truly wavelet analysis in the strict sense:

1. For starters, standard *wavelet*s and their associated *scaling functions* are not sequences (Definition 1.41 page 25), but rather are functions with domain $\mathbb{R}$ (not $\mathbb{Z}$ or some convex subset of $\mathbb{Z}$).

2. While it is true that the celebrated *Fast Wavelet Transform* (FWT) does work *internally* with sequences (using *filter banks*),[53] the FWT is actually defined to work on functions with domain $\mathbb{R}$; and so the function to be analyzed by the FWT must first be *sampled* by a *scaling function*, which yields a sequence that can be processed by the *filter bank*s.[54]

3. Wavelet analysis is typically performed by translating the wavelet or scaling function by fixed amounts depending on the "scale" of the given wavelet. For example, a Haar wavelet of length 4000 would typically "jump" in offsets of 4000: 0, 4000, 8000, 12000, ….[55] As one might imagine, this may be reason for concern if you are using this wavelet to perform edge detection (you might jump over and miss detecting the edge). In this section, wavelet sequences are translated by offsets of 1, making an edge harder to miss.

Example 3.23 (statistical edge detection using Haar wavelet on non-stationary die sequence).

1. Suppose we have a length $\mathsf{N} \triangleq 12000$ die sequence $(\!(x_n)\!)$ with the following distribution:

$$\mathsf{P}(⚀) = \mathsf{P}(⚁) = \mathsf{P}(⚂) = \mathsf{P}(⚃) = \mathsf{P}(⚄) = \mathsf{P}(⚅) = \frac{1}{6} \qquad \text{for } n \in [0:3999] \cup [8000:11999]$$

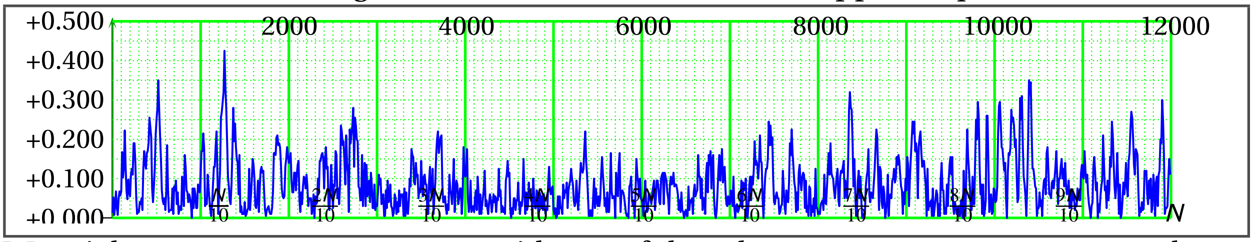
$$\mathsf{P}(⚀) = \mathsf{P}(⚁) = \mathsf{P}(⚂) = \mathsf{P}(⚄) = \mathsf{P}(⚅) = \frac{1}{10} \text{ and } \mathsf{P}(⚃) = \frac{1}{2} \qquad \text{for } n \in [4000:7999]$$

That is, the distribution of the sequence is uniformly distributed in the first and last thirds, but biased towards ⚂ in the middle third. In this example we use a simple statistical edge detector to try to find the statistical "edges" at 4000 and 8000. The edge detector here is a

[52] See experiment log file "dna_AF086833_ebola_dft.xlg" generated by the program "ssp.exe".
[53] Greenhoe (2013) pages 360–364 ⟨J.6 Filter Banks⟩
[54] Greenhoe (2013) pages 369–372 ⟨Appendix L⟩
[55] Greenhoe (2013) pages 27–62 ⟨Chapter 2. The Structure of Wavelets⟩

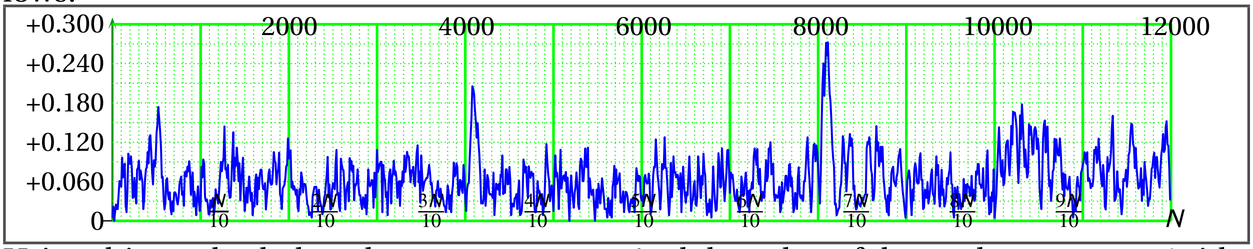
filter operation **W** (Definition 1.45 page 26) using a *length 200 Haar wavelet sequence* (Definition 1.51 page 27).[56]

2. Suppose we first use the *PAM die random variable* (Definition 3.4 page 87) to map the sequence of item (1) into $\mathbb{R}^1$. The magnitude of $\mathbf{W} : \mathbb{R}^1 \to \mathbb{C}^1$ of the mapped sequence is as follows:[57]



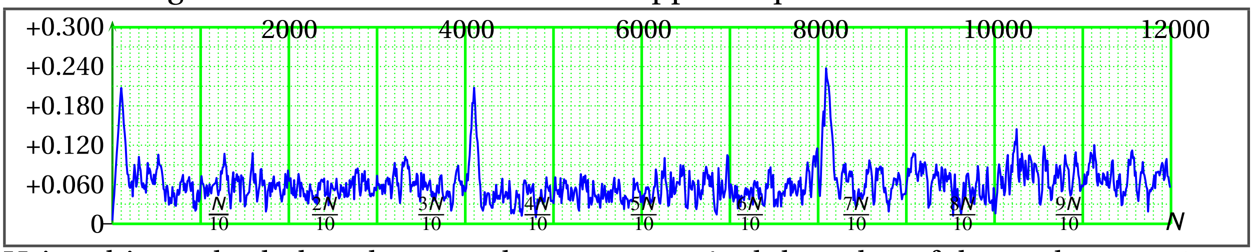
We might expect to see strongest evidence of the edges at $4000 + 200/2 = 4100$ and 8100. But looking at the above result, this is not apparent. In fact, there are a total of 10646 values that are greater than or equal to the value at location 4100 (that value being 0.015).[58]

3. Suppose we next use the *QPSK die random variable* (Definition 3.5 page 87) to map the die sequence into the complex plane. The magnitude of $\mathbf{W} : \mathbb{C}^1 \to \mathbb{C}^1$ of the mapped sequence is as follows:



Using this method, the edges are apparent. And the value of the peak at $n = 4083$ (with value 0.223383) is about $10 \log_{10}(0.223383/0.072741) \approx 4.9$ dB above the noise floor.[59] Here, the noise value 0.072741 is the *RMS* (see item (2) of Example 3.22 page 105) of the DFT magnitude sequence computed over the domain $n = 200 \ldots N - 1$.

4. Suppose we next use the $\mathbb{R}^6$ *die random variable* (Definition 3.7 page 87) to map the sequence into $\mathbb{R}^6$. The magnitude of $\mathbf{W} : \mathbb{R}^6 \to \mathbb{R}^6$ of the mapped sequence is as follows:



Using this method, the edges are also apparent. And the value of the peak at $n = 4102$ (with value 0.209165) is about $10 \log_{10}(0.209165/0.065768) \approx 5.0$ dB above the noise floor.[60] This is only a slight improvement over item (3).

Example 3.24 (statistical edge detection using Haar wavelet on non-stationary artificial DNA sequence).

[56] Empirical evidence due to Singh et al. (1997) suggests that the Haar wavelet performs better than several other common wavelets as an edge detector.

[57] Note that the plot in item (2) has been down sampled by a factor of 10 for practical reasons of displaying the very large data set.

[58] See experiment log file "diehaar_12000m4000_h200_1050.xlg" generated by the program "ssp.exe".

[59] See experiment log file "diehaar_12000m4000_h200_1050.xlg" generated by the program "ssp.exe".

[60] Here the RMS noise value is computed over the domain $n = 200 \ldots N - 1$.
See experiment log file "diehaar_12000m4000_h200_1050.xlg" generated by the program "ssp.exe".

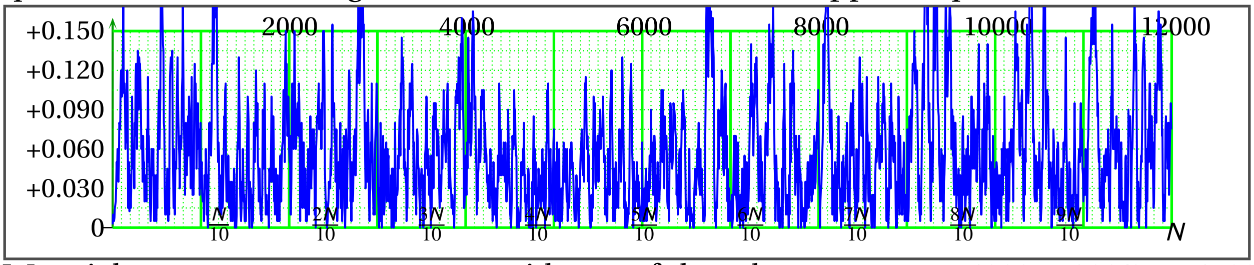
1. Suppose we have a length $N \triangleq 12000$ dna sequence $(\!(x_n)\!)$ with the following distribution:

$$\mathsf{P}(\boxed{A}) = \mathsf{P}(\boxed{C}) = \mathsf{P}(\boxed{G}) = \mathsf{P}(\boxed{T}) = \frac{1}{4} \qquad \text{for } n \in [0 : 3999] \cup [8000 : 11999]$$

$$\mathsf{P}(\boxed{A}) = \mathsf{P}(\boxed{G}) = \mathsf{P}(\boxed{T}) = \frac{17}{100} \text{ and } \mathsf{P}(\boxed{C}) = \frac{49}{100} \qquad \text{for } n \in [4000 : 7999]$$

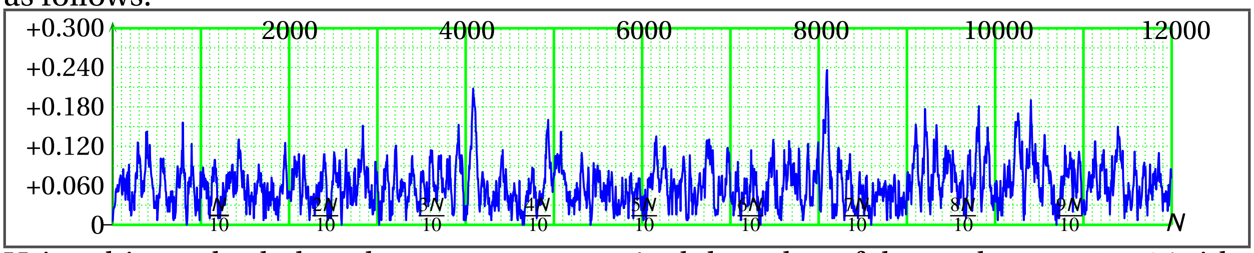
   That is, the distribution of the sequence is uniformly distributed in the first and last thirds, but biased towards $\boxed{C}$ in the middle third. Just as in Example 3.23, we again use a filter operation $\mathbf{W}$ with *length 200 Haar wavelet sequence* as a simple statistical edge detector to try to locate the statistical "edges" at 4000 and 8000.

2. Suppose we first use the *PAM DNA random variable* (Definition 3.10 page 87) to map the DNA sequence into $\mathbb{R}^1$. The magnitude of $\mathbf{W} : \mathbb{R}^1 \to \mathbb{C}^1$ of the mapped sequence is as follows:[61]



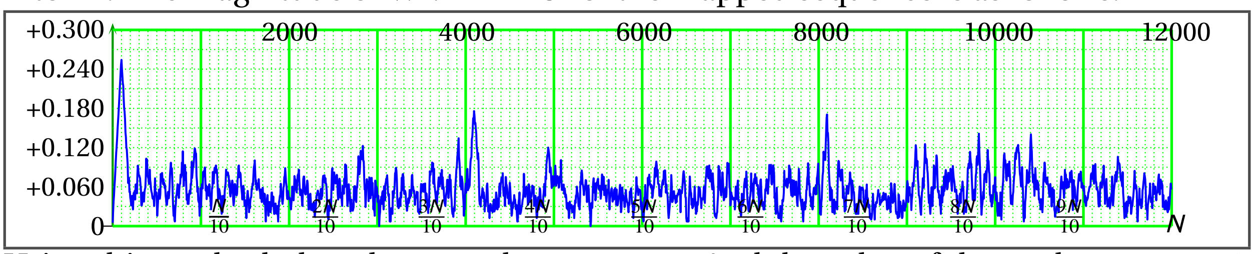
   We might expect to see strongest evidence of the edges at or near $4000 + 200/2 = 4100$ and 8100. In fact, the sequence does have peaks at 4087 (with value 0.185) and at 8087 (with value 0.230). The peak at 4087 is about $10\log_{10}(0.185000/0.071355) \approx 4.1$ dB above the noise floor. However, there are 103 other values not around the $n = 4087$ and $n = 8087$ peaks that are 0.185 or greater. These 102 values represent roughly 11 other peaks, each of which could trigger a "false positive" decision.[62]

3. Suppose we next use the *QPSK DNA random variable* (Definition 3.5 page 87) to map the dna sequence into the complex plane. The magnitude of $\mathbf{W} : \mathbb{C}^1 \to \mathbb{C}^1$ of the mapped sequence is as follows:



   Using this method, the edges are apparent. And the value of the peak at $n = 4086$ (with value 0.215870) is $10\log_{10}(0.215870/0.070068) \approx 4.9$ dB above the noise floor.[63]

4. Suppose we next use the $\mathbb{R}^4$ *DNA random variable* (Definition 3.12 page 88) to map the sequence into $\mathbb{R}^4$. The magnitude of $\mathbf{W} : \mathbb{R}^4 \to \mathbb{C}^4$ of the mapped sequence is as follows:

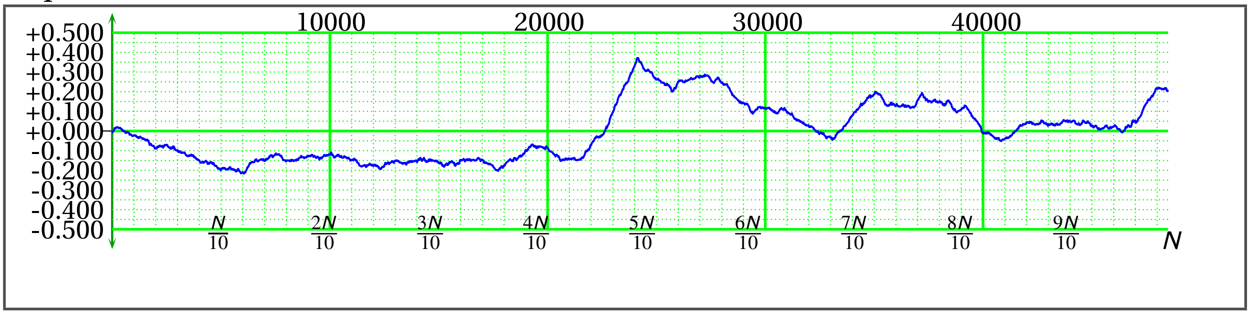


   Using this method, the edges are also apparent. And the value of the peak at $n = 4096$ (with value 0.181246) is $10\log_{10}(0.181246/0.059594) \approx 4.8$ dB above the noise floor.[64] Note that this is a slight decrease in performance as compared to item (3).

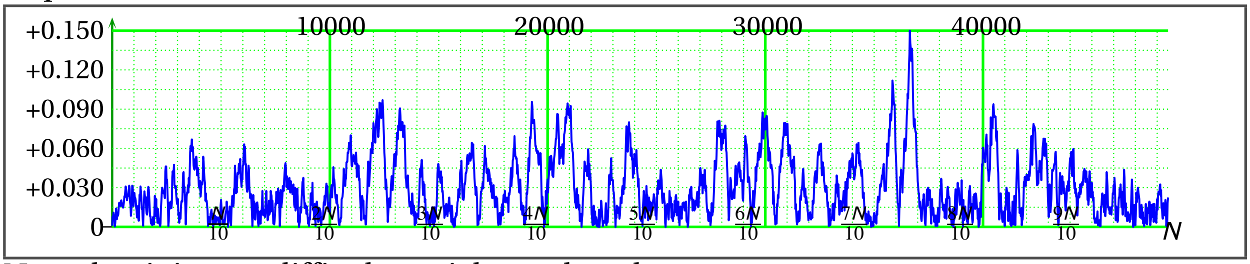
[61] Note that the plot in item (2) has been down sampled by a factor of 10 for practical reasons of displaying the very large data set.
[62] See experiment log file "dnahaar_12000m4000_h200_1749.xlg" generated by the program "ssp.exe".
[63] See experiment log file "dnahaar_12000m4000_h200_1749.xlg" generated by the program "ssp.exe".
[64] See experiment log file "dnahaar_12000m4000_h200_1749.xlg" generated by the program "ssp.exe".

Example 3.25 (Wavelet analysis of Phage Lambda DNA sequence).

1. Consider the Phage Lambda DNA sequence. It has a strong $\boxed{C}\boxed{G}$ bias before $n = 20000$ and a strong $\boxed{A}\boxed{T}$ bias after,[65] as demonstrated next by mapping
   $\boxed{A} \to +1 \qquad \boxed{T} \to +1 \qquad \boxed{C} \to -1 \qquad \boxed{G} \to -1$
   and filtering the resulting sequence in $\mathbb{R}^1$ with a *length 1600 Haar scaling sequence* (Definition 1.50 page 27)—such a filtering operation acts as a kind of "sliding window" histogram of the DNA sequence.

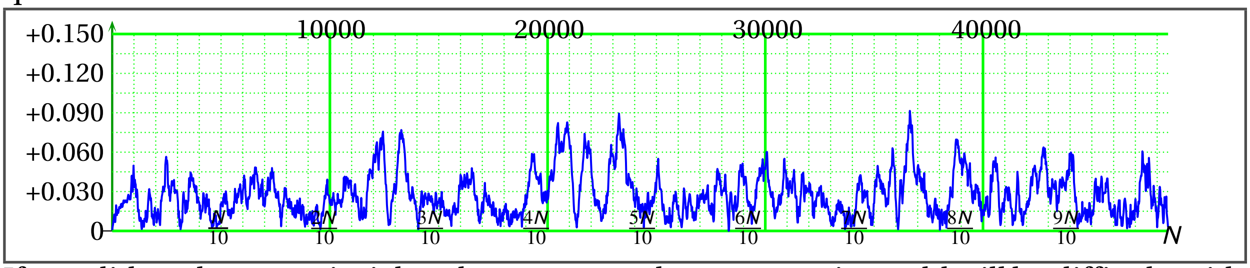


2. Suppose we first use the *PAM DNA random variable* (Definition 3.10 page 87) to map the DNA sequence into $\mathbb{R}^1$. The magnitude of the length 1600 Haar wavelet operation on the mapped sequence is as follows:



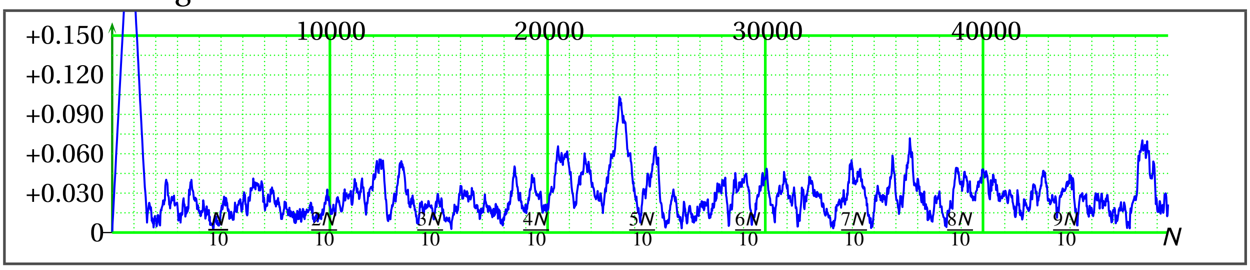
   Note that it is very difficult to pick out the edge at 20000.

3. Suppose we next use the *QPSK DNA random variable* (Definition 3.11 page 87) to map the DNA sequence into the complex plane. The length 1600 Haar wavelet operation on the mapped sequence is as follows:



   If one did not know apriori that there was an edge at 20000, it would still be difficult to identify.

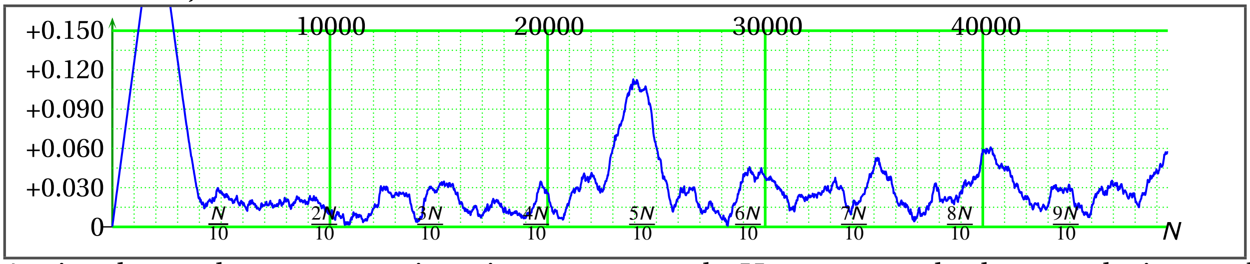
4. Suppose we next use the $\mathbb{R}^4$ *DNA random variable* (Definition 3.12 page 88) to map the DNA sequence into $\mathbb{R}^4$. Filtering the mapped sequence with a length 1600 Haar wavelet sequence results in the following:



[65] Cristianini and Hahn (2007) page 14

Here there is a clear peak near 20000.

5. And here is the same analysis as used in item (4), but at scale 4000 (using a length 4000 Haar wavelet filter):

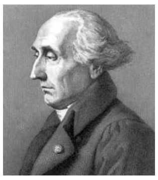


Again, the peak near 20000 is quite pronounced. However, at the low resolution scale (of 4000), it would be difficult to determine precisely where the statistical edge actually was.

APPENDIX A

LAGRANGE ARC DISTANCE

"*Tant que l'Algèbre et la Géométrie ont été séparées, leurs progrès ont été lents et leurs usages bornés; mais lorsque ces deux sciences se sont réunies, elles vers la perfection.*"



"*As long as algebra and geometry have been separated, their progress have been slow and their uses limited; but when these two sciences have been united, they have lent each other mutual forces, and have marched together with a rapid step towards perfection.*"

Joseph-Louis Lagrange (1736–1813, Italian-French mathematician and astronomer [1]

A.1 Introduction

A.1.1 The spherical metric

The *spherical metric*, or *great circle metric*,[2] p_r operates on the *surface of a sphere with radius r* centered at the origin $(0, 0, \cdots, 0)$ in a linear space $\mathbb{R}^N$, where "*surface of a sphere...*" is defined as all the points in $\mathbb{R}^N$ that are a distance r from $(0, 0, \cdots, 0)$ with respect to the *Euclidean metric* (Definition D.10 page 166). Thus, for any pair of points (p, q) *on the surface* of this sphere, (p, q) is in the *domain* of p_r and $\mathsf{p}_r(p, q)$ is the "distance" between those points. However, if p and q *are* both in $\mathbb{R}^N$ but are *not* on the surface of a common sphere centered at the origin, then (p, q) is *not* in the domain of p_r and $\mathsf{p}_r(p, q)$ is simply *undefined*.

In certain applications, however, it would be useful to have an *extension* d of the spherical metric p to the entire space $\mathbb{R}^N$ (rather than just on a surface in $\mathbb{R}^N$). For example, for the points $p \triangleq (0, 1)$ and $q \triangleq (1, 0)$ (which are both on the surface of a common sphere in $\mathbb{R}^2$), we would like d to be compatible with p such that $\mathsf{d}(p, q) = \mathsf{p}(p, q)$. If $r \triangleq (2, 0)$, then the pair (p, r) is *not* in the domain of p, but we still would like it to be in the domain of d such that $\mathsf{d}(p, r)$ is defined—and in this way d would be an *extension* of p.

[1] quote: Lagrange (1795), page 271
translation: Grattan-Guinness (1990) page 254
image: http://en.wikipedia.org/wiki/Joseph_Louis_Lagrange

[2] Ratcliffe (2013) pages 37–38 ⟨The Spherical Metric⟩, Deza and Deza (2014) page 123 ⟨6.4 Non-Euclidean Geometry⟩, Deza and Deza (2006) page 73 ⟨6.4 NON-EUCLIDEAN GEOMETRY⟩, SILVER AND STOKES (2007) PAGE 9

In this text, the *Langrange arc distance* is used in

- the *low pass filtering* of a *real die sequence* (Example 3.12 page 89) and
- the *low pass filtering* of a *spinner sequence* (Example 3.13 page 91) and
- the *high pass filtering* of a *weighted real die sequence* (Example 3.15 page 95) and
- the *high pass filtering* of a *weighted spinner sequence* (Example 3.16 page 97).

A.1.2 Linear interpolation

This paper introduces an extension to the *spherical metric* based on a polar form of *linear interpolation*. *Interpolation* has a very long history with evidence suggesting that it extends possibly all the way back to the Babylonians living around 300BC.[3]

Linear interpolation between two points $p \triangleq (x_1, y_1)$ and $q \triangleq (x_2, y_2)$ in $\mathbb{R}^2$ is conveniently and intuitively expressed in a *Cartesian coordinate system* using what is commonly known as *Lagrange interpolation* (Definition D.21 page 183) in the form

$$y = y_1\left(\frac{x-x_2}{x_1-x_2}\right) + y_2\left(\frac{x-x_1}{x_2-x_1}\right).$$

Newton interpolation (Definition D.22 page 184) yields the same expression, but generally requires more "effort" (back substitution or matrix algebra):

$$y \triangleq \underbrace{\sum_{k=1}^{2}\alpha_k\sum_{m=1}^{k}(x-x_m)}_{\textit{Newton polynomial}\text{ (Definition D.22)}} = \alpha_1\left[x-x_1\right] + \alpha_2\left[(x-x_1)+(x-x_2)\right] = (\alpha_1+\alpha_2)(x-x_1) + \alpha_2(x-x_2)$$

$$y_1 = \alpha_1\left[x_1-x_1\right] + \alpha_2\left[(x_1-x_1)+(x_1-x_2)\right] \implies \alpha_2 = \frac{y_1}{x_1-x_2}$$

$$y_2 = \alpha_1\left[x_2-x_1\right] + \alpha_2\left[(x_2-x_1)+(x_2-x_2)\right] \implies \alpha_1 = \frac{y_1+y_2}{x_2-x_1}$$

$$y = \left[\frac{y_1+y_2}{x_2-x_1} + \frac{y_1}{x_1-x_2}\right](x-x_1) + \left[\frac{y_1}{x_1-x_2}\right](x-x_2) = y_1\left(\frac{x-x_2}{x_1-x_2}\right) + y_2\left(\frac{x-x_1}{x_2-x_1}\right)$$

Of course the 2-point *Lagrange interpolation*/*Newton interpolation* polynomial can also be written in the familiar *slope-intercept* $y = mx + b$ form as

$$y = y_1\left(\frac{x-x_2}{x_1-x_2}\right) + y_2\left(\frac{x-x_1}{x_2-x_1}\right) = y_2\left(\frac{x-x_1}{x_2-x_1}\right) - y_1\left(\frac{x-x_2}{x_2-x_1}\right) = \frac{y_2(x-x_1) - y_1(x-x_2)}{x_2-x_1}$$

$$= \underbrace{\left(\frac{y_2-y_1}{x_2-x_1}\right)}_{\text{slope}} x + \underbrace{\left(\frac{x_2y_1-x_1y_2}{x_2-x_1}\right)}_{y\text{-intercept}}$$

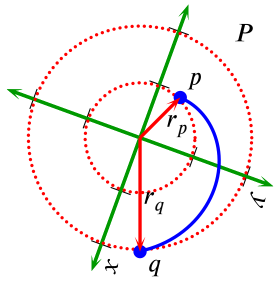
A.1.3 Polar linear interpolation

Linear interpolation is often illustrated in terms of cartesian coordinates (x, y). But there is no reason why the same principles cannot be used in terms of polar coordinates $(\mathsf{r}(\theta), \theta)$. However care does need to be taken where θ may be interpreted to "jump" from 2π to 0 or from $-\pi$ to π.

[3] Meijering (2002) page 320

Here is an expression for 2-point *Lagrange interpolation*/*Newton interpolation* in *polar form*:

$$\mathsf{r}(\theta) \triangleq r_p\left[\frac{\theta-\theta_q}{\theta_p-\theta_q}\right]+r_q\left[\frac{\theta-\theta_p}{\theta_q-\theta_p}\right] \quad \forall\theta\in\left[\theta_p : \theta_q\right]$$



Note the following:

1. The orientation of the axes in plane P is arbitrary, and that without loss of generality we can orient the axes such that p or q is on the positive x-axis and that the other point has a non-negative y value.

2. This means that the length of the arc between p at (r_p, θ_p) and q at (r_q, θ_q) under the original orientation is equal to the length of the arc between the points $(r_p, 0)$ and $(r_q, |\theta_p - \theta_q|)$ in the new orientation.

3. One important reason for the geometrical acrobatics here is that we don't want to have to calculate the values for θ_p and θ_q in a plane P (which we don't even immediately have an algebraic expression for anyways). But calculating the value $\phi \triangleq |\theta_p - \theta_q|$ is quite straightforward because the "*dot product*" $\langle p \mid q\rangle$ of p and q (which is very easy to calculate) in $\mathbb{R}^N$ equals $r_p r_q \cos\phi$ (and so $\phi = \arccos\left(\frac{1}{r_p r_q}\langle p \mid q\rangle\right)$).

4. Actually, $\phi = |\theta_q - \theta_p|$, as demonstrated below:

$$\begin{aligned}
\phi &\triangleq \arccos\left(\frac{1}{r_p r_q}\sum_{n=1}^{N} x_n y_n\right) && \text{by definition of } \phi\\
&\triangleq \arccos\left(\frac{1}{r_p r_q}\langle p \mid q\rangle\right) && \text{a standard definition from the field of "linear algebra"}\\
&= \arccos\left(\frac{1}{r_p r_q}\left[r_p r_q \cos|\theta_q - \theta_p|\right]\right) && \text{a standard result from the field of "linear algebra"}\\
&= \left\{|\theta_q - \theta_p|,\ 2\pi - |\theta_q - \theta_p|\right\} && \text{by definition of } \arccos(x) \text{ and } \cos(x)\\
&= |\theta_q - \theta_p| && \text{by item (1)}
\end{aligned}$$

5. Setting $\theta_p = 0$ and $\theta_q = \phi$ yields the following:

$$\begin{aligned}
\mathsf{r}(\theta) &= r_p\left[\frac{\theta-\theta_q}{\theta_p-\theta_q}\right]+r_q\left[\frac{\theta-\theta_p}{\theta_q-\theta_p}\right] && \text{Langrange form (Definition D.21 page 183)}\\
&= r_p\left[\frac{\theta-\phi}{0-\phi}\right]+r_q\left[\frac{\theta-0}{\phi-0}\right] = \frac{-r_p\theta+r_p\phi+r_q\theta}{\phi}\\
&= \left(\frac{r_q-r_p}{\phi}\right)\theta + r_p && \text{polar slope-intercept form}
\end{aligned}$$

A.1.4 Distance in terms of polar linear interpolation arcs

This paper introduces a new function herein called, for better or for worse,[4] the *Lagrange arc distance* (Definition A.1 page 114) $\mathsf{d}(p,q)$. It's domain is the entire space $\mathbb{R}^N$. It is an extension of the *spherical metric*, which only has as domain the surface of a sphere in $\mathbb{R}^N$.

When p or q is at the origin, or when the polar angle ϕ between p and q is 0, then the *Lagrange arc distance* $\mathsf{d}(p,q)$ is simply a $\frac{1}{\pi}$ scaled *Euclidean metric* (Definition D.10 page 166). In all other cases, $\mathsf{d}(p,q)$ is the $\frac{1}{\pi}$ scaled length of the *Lagrange interpolation* arc extending from p to q.

An equation for the length of an arc in polar coordinates is[5]

$$\mathsf{R}(p,q) = \int_{\theta_p}^{\theta_q} \sqrt{r^2(\theta) + \left(\frac{\mathsf{d}r}{\mathsf{d}\theta}\right)^2}\, \mathsf{d}\theta$$

This integral may look intimidating. Later however, Theorem A.1 (page 117) demonstrates that it has an "easily" computable and straightforward solution only involving *arithmatic operators* $(+, -, \ldots)$, the *absolute value* function $|x|$, the *square root* function $\sqrt{x}$, and the *natural log* function $\ln(x)$.

Finally, note that the extension does come at a cost—the *Lagrange arc distance* is not a *metric* (Definition D.7 page 163), but rather only a *distance* (Definition B.1 page 133, Theorem A.4 page 125). For more details about the impact of this cost, see Theorem B.1 (page 133).

A.2 Definition

Definition A.1. *Let* $p \triangleq (x_1, x_2, \cdots, x_N)$ *and* $q \triangleq (y_1, y_2, \cdots, y_N)$ *be two points in the* SPACE $\mathbb{R}^N$ *with origin* $(0,0,\cdots,0)$. *Let*

$$\underbrace{r_p \triangleq \left(\sum_{n=1}^N x_n^2\right)^{\frac{1}{2}}}_{\text{(magnitude of } p)} \qquad \underbrace{r_q \triangleq \left(\sum_{n=1}^N y_n^2\right)^{\frac{1}{2}}}_{\text{(magnitude of } q)} \qquad \underbrace{\phi \triangleq \arccos\left(\frac{1}{r_p r_q}\sum_{n=1}^N x_n y_n\right)}_{\text{(angle between } p \text{ and } q)}$$

$$\underbrace{\mathsf{r}(\theta) \triangleq \left(\frac{r_q - r_p}{\phi}\right)\theta + r_p}_{\text{(polar interpolation polynomial)}} \qquad \underbrace{\mathsf{R}(p,q) = \int_0^{\phi} \sqrt{r^2(\theta) + \left(\frac{\mathsf{d}r}{\mathsf{d}\theta}\right)^2}\, \mathsf{d}\theta}_{\text{(length of the arc } (\mathsf{r}(\theta),\theta) \text{ between } p \text{ and } q)}$$

DEF: *The* ***Langrange arc distance*** $\mathsf{d}(p,q)$ *is defined as*

$$\mathsf{d}(p,q) = \begin{cases} {}^1\!/\!_\pi \left|r_p - r_q\right| & \text{if } p = (0,0,\cdots,0) \text{ OR } q = (0,0,\cdots,0) \text{ OR } \phi = 0 \\ {}^1\!/\!_\pi \mathsf{R}(p,q) & \text{otherwise} \end{cases} \qquad \forall p,q \in \mathbb{R}^N$$

[4] "for better or for worse": As already pointed out, *Newton interpolation* or simply the slope-intercept form $y = mx + b$ of the line equation can with a little bit of effort give you the same equation as the 2-point *Lagrange interpolation*. So why not name the function $\mathsf{d}(p,q)$ of Definition A.1 "Newton arc distance"? Actually Newton published his interpolation method (for example in his 1711 "Methodus differentialis" Newton (1711)) long before Lagrange (Lagrange (1877)). But besides that, Lagrange was not really the first to discover what is commonly called "Lagrange interpolation". The same result was actually published about 98 years earlier by Edward Waring (Waring (1779)). But in the end, the choice to use the name "Lagrange arc distance" has some justification in that it's form arguably comes more readily using Lagrange interpolation than it does from *Newton interpolation* (which requires back substitution); and even though "Lagrange interpolation" probably should be called "Waring interpolation", the fact is that it's normally called "Lagrange interpolation". So there is some motivation for the choice of the name. And "for better or for worse", the function $\mathsf{d}(p,q)$ is herein called the "Lagrange arc distance". …One last note: for a much fuller historical background of interpolation, see Meijering (2002).

[5] Stewart (2012) page 533 ⟨Section 9.4 Areas and lengths in polar coordinates⟩

A.3 Calculation

The integral in Definition A.1 may look intimidating. However, Theorem A.1 (page 117) demonstrates that it has an "easily" computable and straightforward solution only involving *arithmatic operators* $(+, -, \ldots)$, the *absolute value* function $|x|$, the *square root* function $\sqrt{x}$, and the *natural log* function $\ln(x)$. But first, a lemma (next) to help with the proof of Theorem A.1.

Lemma A.1. [6] *Let* $\sqrt{x} \in \mathbb{R}^{\mathbb{R}}$ *be the* SQUARE ROOT *function, and* $\ln(x) \triangleq \log_e(x) \in \mathbb{R}^{\mathbb{R}}$ *be the* NATURAL LOG *function. Let* ε *be any given value in* $\mathbb{R}$.

LEM

$$\left\{2ax + b + 2\sqrt{a(ax^2+bx+c)} > 0\right\} \implies$$
$$\left\{\int \sqrt{ax^2+bx+c}\,\mathrm{d}x = \frac{2ax+b}{4a}\sqrt{ax^2+bx+c} + \frac{4ac-b^2}{8a^{3/2}}\ln\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right) + \varepsilon\right\}$$

✎PROOF:

1. lemma: (first equality by the *product rule*, and the second equality by the *chain rule*)

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left[\left(\frac{2ax+b}{4a}\right)\sqrt{ax^2+bx+c}\right] &= \left(\frac{2ax+b}{4a}\right)\left(\frac{\mathrm{d}}{\mathrm{d}x}\sqrt{ax^2+bx+c}\right) + \left(\frac{\mathrm{d}}{\mathrm{d}x}\frac{2ax+b}{4a}\right)\left(\sqrt{ax^2+bx+c}\right)\\
&= \left(\frac{2ax+b}{4a}\right)\left(\frac{2ax+b}{2\sqrt{ax^2+bx+c}}\right) + \left(\frac{2a}{4a}\right)\left(\sqrt{ax^2+bx+c}\right)\\
&= \frac{(2ax+b)^2 + 4a(ax^2+bx+c)}{8a\sqrt{ax^2+bx+c}}\\
&= \frac{4a^2x^2 + 4axb + b^2 + 4a^2x^2 + 4abx + 4ac}{8a\sqrt{ax^2+bx+c}}\\
&= \frac{8a^2x^2 + 8abx + b^2 + 4ac}{8a\sqrt{ax^2+bx+c}}
\end{aligned}$$

2. lemma: If $\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right) > 0$ then

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{4ac-b^2}{8a^{3/2}}\ln\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right)\right]\\
&= \left(\frac{4ac-b^2}{8a^{3/2}}\right)\left[\frac{\mathrm{d}}{\mathrm{d}x}\ln\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right)\right] && \text{by } \textit{linearity} \text{ of } \tfrac{\mathrm{d}}{\mathrm{d}x}\\
&= \left(\frac{4ac-b^2}{8a^{3/2}}\right)\left(\frac{1}{2ax+b+2\sqrt{a(ax^2+bx+c)}}\right)\left(2a + \frac{2\sqrt{a}(2ax+b)}{2\sqrt{ax^2+bx+c}}\right) && \text{by } \textit{chain rule}\\
&= \frac{(4ac-b^2)\left[2a\sqrt{ax^2+bx+c} + \sqrt{a}(2ax+b)\right]}{8a^{3/2}\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right)\sqrt{ax^2+bx+c}}\\
&= \frac{(4ac-b^2)\left[\sqrt{a}(2ax+b) + 2a\sqrt{ax^2+bx+c}\right]}{\left(8a\sqrt{ax^2+bx+c}\right)\left[\sqrt{a}(2ax+b) + 2a\sqrt{ax^2+bx+c}\right]}\\
&= \frac{4ac-b^2}{8a\sqrt{ax^2+bx+c}}
\end{aligned}$$

[6] Gradshteyn and Ryzhik (2007) page 94 ⟨2.25 Forms containing $\sqrt{a+bx+cx^2}$, 2.26 Forms containing $\sqrt{a+bx+cx^2}$ and integral powers of x⟩, Jeffrey (1995) page 160 ⟨4.3.4 Integrands containing $(a+bx+cx^2)^{1/2}$⟩, Jeffrey and Dai (2008) pages 172–173 ⟨4.3.4 Integrands containing $(a+bx+cx^2)^{1/2}$⟩

3. lemma: If $\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right)<0$ then

$$\begin{aligned}
&\frac{\mathsf{d}}{\mathsf{d}x}\left[\frac{4ac-b^2}{8a^{3/2}}\ln\left|2ax+b+2\sqrt{a(ax^2+bx+c)}\right|\right]\\
&=\left(\frac{4ac-b^2}{8a^{3/2}}\right)\left[\frac{\mathsf{d}}{\mathsf{d}x}\ln\left(-2ax-b-2\sqrt{a(ax^2+bx+c)}\right)\right] && \text{by } \textit{linearity} \text{ of } \frac{\mathsf{d}}{\mathsf{d}x}\\
&=\left(\frac{4ac-b^2}{8a^{3/2}}\right)\left(\frac{1}{-2ax-b-2\sqrt{a(ax^2+bx+c)}}\right)\left(2a+\frac{2\sqrt{a}(2ax+b)}{2\sqrt{ax^2+bx+c}}\right) && \text{by } \textit{chain rule}\\
&=\frac{(4ac-b^2)\left[2a\sqrt{ax^2+bx+c}+\sqrt{a}(2ax+b)\right]}{8a^{3/2}\left(-2ax-b-2\sqrt{a(ax^2+bx+c)}\right)\sqrt{ax^2+bx+c}}\\
&=\frac{-(4ac-b^2)\left[\sqrt{a}(2ax+b)+2a\sqrt{ax^2+bx+c}\right]}{\left(8a\sqrt{ax^2+bx+c}\right)\left[\sqrt{a}(2ax+b)+2a\sqrt{ax^2+bx+c}\right]}\\
&=\frac{-(4ac-b^2)}{8a\sqrt{ax^2+bx+c}}
\end{aligned}$$

4. Complete the proof: If $\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right)>0$ then

$$\begin{aligned}
&\frac{\mathsf{d}}{\mathsf{d}x}\left[\frac{2ax+b}{4a}\sqrt{ax^2+bx+c}+\frac{4ac-b^2}{8a^{3/2}}\ln\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right)\right]\\
&=\frac{\mathsf{d}}{\mathsf{d}x}\left[\left(\frac{2ax+b}{4a}\right)\sqrt{ax^2+bx+c}\right]+\frac{\mathsf{d}}{\mathsf{d}x}\left[\left(\frac{4ac-b^2}{8a^{3/2}}\right)\ln\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right)\right]\\
&=\frac{8a^2x^2+8abx+b^2+4ac}{8a\sqrt{ax^2+bx+c}}+\frac{\mathsf{d}}{\mathsf{d}x}\left[\left(\frac{4ac-b^2}{8a^{3/2}}\right)\ln\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right)\right] && \text{by item (1)}\\
&=\frac{8a^2x^2+8abx+b^2+4ac}{8a\sqrt{ax^2+bx+c}}+\frac{4ac-b^2}{8a\sqrt{ax^2+bx+c}} && \text{by item (2)}\\
&=\frac{8a(ax^2+bx+c)}{8a\sqrt{ax^2+bx+c}}=\sqrt{ax^2+bx+c}
\end{aligned}$$

5. Note that simply forcing the agruement of ln to be positive as in[7]
$\frac{2ax+b}{4a}\sqrt{ax^2+bx+c}+\frac{4ac-b^2}{8a^{3/2}}\ln\left|2ax+b+2\sqrt{a(ax^2+bx+c)}\right|+\varepsilon$
is *not* a solution to $\int\sqrt{ax^2+bx+c}\,\mathsf{d}x$ when $\left(2ax+b+2\sqrt{a(ax^2+bx+c)}\right)<0$:

$$\begin{aligned}
&\frac{\mathsf{d}}{\mathsf{d}x}\left[\frac{2ax+b}{4a}\sqrt{ax^2+bx+c}+\frac{4ac-b^2}{8a^{3/2}}\ln\left|2ax+b+2\sqrt{a(ax^2+bx+c)}\right|\right]\\
&=\frac{\mathsf{d}}{\mathsf{d}x}\left[\left(\frac{2ax+b}{4a}\right)\sqrt{ax^2+bx+c}\right]+\frac{\mathsf{d}}{\mathsf{d}x}\left[\left(\frac{4ac-b^2}{8a^{3/2}}\right)\ln\left|2ax+b+2\sqrt{a(ax^2+bx+c)}\right|\right]\\
&=\frac{8a^2x^2+8abx+b^2+4ac}{8a\sqrt{ax^2+bx+c}}+\frac{\mathsf{d}}{\mathsf{d}x}\left[\left(\frac{4ac-b^2}{8a^{3/2}}\right)\ln\left|2ax+b+2\sqrt{a(ax^2+bx+c)}\right|\right] && \text{by item (1)}\\
&=\frac{8a^2x^2+8abx+b^2+4ac}{8a\sqrt{ax^2+bx+c}}-\frac{4ac-b^2}{8a\sqrt{ax^2+bx+c}} && \text{by item (3)}\\
&=\frac{8a(ax^2+bx)+2b^2}{8a\sqrt{ax^2+bx+c}}=\frac{8a(ax^2+bx+c)+2b^2-8ac}{8a\sqrt{ax^2+bx+c}}
\end{aligned}$$

[7] The solution $\ln|\cdots|$ is used in Jeffrey (1995) page 160 and Jeffrey and Dai (2008) pages 172–173.

$$= \sqrt{ax^2 + bx + c} + \frac{2b^2 - 8ac}{8a\sqrt{ax^2 + bx + c}} \neq \sqrt{ax^2 + bx + c} \quad \text{for } b^2 \neq 4ac$$

6. Note further that constraining $a > 0$ is also not a solution[8] because it does not guarantee that the argument u of $\ln(u)$ will be positive. Take for example $a = 1$, $b = -3$, $c = 3$ and $x = 1$. Then
$$2ax + b + 2\sqrt{a(ax^2 + bx + c)} = 2 \cdot 1 \cdot 1 - 3 + 2\sqrt{1(1 \cdot 1 - 3 \cdot 1 + 3)} = -1 < 0\,.$$

✏

Theorem A.1. *Let* $\mathsf{R}(p, q)$, r_p, r_q, *and* ϕ *be as defined in Definition A.1 (page 114). Let* $\rho \triangleq r_q - r_p$. *If* $r_p \neq 0$, $r_q \neq 0$ *and* $\phi \neq 0$ *then*

THM
$$\begin{aligned}\mathsf{R}(p, q) = {} & \frac{r_q\sqrt{(r_q\phi)^2 + \rho^2} - r_p\sqrt{(r_p\phi)^2 + \rho^2}}{2\rho} \\ & + \frac{|\rho|}{2\phi}\ln\left(r_q\rho\phi + |\rho|\sqrt{(r_q\phi)^2 + \rho^2}\right) \\ & - \frac{|\rho|}{2\phi}\ln\left(r_p\rho\phi + |\rho|\sqrt{(r_p\phi)^2 + \rho^2}\right)\end{aligned}$$

✎PROOF:

1. Let $\gamma \triangleq r_p\phi$.

2. lemmas:

$$\begin{aligned}\rho\phi + \gamma &= (r_q - r_p)\phi + r_p\phi = r_q\phi \\ \rho^2\phi^2 + 2\rho\gamma\phi + (\gamma^2 + \rho^2) &= (r_q - r_p)^2\phi^2 + 2(r_q - r_p)(r_p\phi)\phi + (r_p\phi)^2 + \rho^2 \\ &= (r_p^2 + r_q^2 - 2r_pr_q)\phi^2 + 2(r_q - r_p)(r_p\phi)\phi + (r_p\phi)^2 + \rho^2 \\ &= (r_p^2 + r_q^2)\phi^2 - 2(r_p\phi)^2 + (r_p\phi)^2 + \rho^2 \\ &= (r_q\phi)^2 + \rho^2\end{aligned}$$

3. lemma: $2\rho^2\theta + 2\rho\gamma + 2\sqrt{\rho^2(\rho^2\theta^2 + 2\rho\gamma\theta + (\gamma^2 + \rho^2))} > 0$. Proof:

$$\begin{aligned}&2\rho^2\theta + 2\rho\gamma + 2\sqrt{\rho^2(\rho^2\theta^2 + 2\rho\gamma\theta + (\gamma^2 + \rho^2))} \\ &\quad > 2\rho^2\theta + 2\rho\gamma + 2\sqrt{\rho^2\gamma^2} && \text{because } \sqrt{x} \text{ is } \textit{strictly monotonically increasing} \\ &\quad = 2\rho^2\theta + 2\rho\gamma + 2|\rho|\gamma && \text{because } \gamma > 0 \\ &\quad = 2\rho^2\theta + 2\gamma(\rho + |\rho|) \\ &\quad \geq 0\end{aligned}$$

[8] The $a > 0$ constraint is used in Gradshteyn and Ryzhik (2007) page 94

4. Completing the proof...

$$
\begin{aligned}
&\mathsf{R}(p,q)\\
&\triangleq \int_{\theta=0}^{\theta=\phi} \sqrt{r^2(\theta) + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta && \text{by def. of } \mathsf{R}(p,q)\\
&= \int_0^\phi \sqrt{\left[\left(\frac{r_q - r_p}{\phi}\right)\theta + r_p\right]^2 + \left[\frac{r_q - r_p}{\phi}\right]^2}\,\mathrm{d}\theta && \text{by def. of } \mathsf{r}(\theta)\\
&= \int_0^\phi \sqrt{\left[\frac{(r_q - r_p)\theta + r_p\phi}{\phi}\right]^2 + \left[\frac{r_q - r_p}{\phi}\right]^2}\,\mathrm{d}\theta\\
&= \frac{1}{\phi}\int_0^\phi \sqrt{(r_q - r_p)^2\theta^2 + 2(r_q - r_p)(r_p\phi)\theta + (r_p\phi)^2 + (r_q - r_p)^2}\,\mathrm{d}\theta\\
&= \frac{1}{\phi}\int_0^\phi \sqrt{\underbrace{\rho^2}_{a}\theta^2 + \underbrace{2\rho\gamma}_{b}\,\theta + \underbrace{(\gamma^2 + \rho^2)}_{c}}\,\mathrm{d}\theta\\
&= \frac{1}{\phi}\left[\frac{2\rho^2\theta + 2\rho\gamma}{4\rho^2}\sqrt{\rho^2\theta^2 + 2\rho\gamma\theta + (\gamma^2 + \rho^2)}\right. && \text{[by item (3) and}\\
&\quad \left. + \frac{4\rho^2(\gamma^2 + \rho^2) - (2\rho\gamma)^2}{8|\rho|^3}\ln\left(2\rho^2\theta + 2\rho\gamma + 2\sqrt{\rho^2(\rho^2\theta^2 + 2\rho\gamma\theta + (\gamma^2 + \rho^2))}\right)\right]_{\theta=0}^{\theta=\phi} && \text{by Lemma A.1]}\\
&= \frac{1}{\phi}\left[\frac{\rho\theta + \gamma}{2\rho}\sqrt{\rho^2\theta^2 + 2\rho\gamma\theta + (\gamma^2 + \rho^2)} + \frac{|\rho|}{2}\ln\left(2\rho^2\theta + 2\rho\gamma + 2|\rho|\sqrt{\rho^2\theta^2 + 2\rho\gamma\theta + (\gamma^2 + \rho^2)}\right)\right]_{\theta=0}^{\theta=\phi}\\
&= \left[\frac{(\rho\phi + \gamma)\sqrt{\rho^2\phi^2 + 2\rho\gamma\phi + (\gamma^2 + \rho^2)}}{2\rho\phi} + \frac{|\rho|}{2\phi}\ln\left(2\rho^2\phi + 2\rho\gamma + 2|\rho|\sqrt{\rho^2\phi^2 + 2\rho\gamma\phi + (\gamma^2 + \rho^2)}\right)\right]\\
&\quad - \left[\frac{\gamma\sqrt{\gamma^2 + \rho^2}}{2\rho\phi} + \frac{|\rho|}{2\phi}\ln\left(2\rho\gamma + 2|\rho|\sqrt{\gamma^2 + \rho^2}\right)\right]\\
&= \frac{r_q\phi\sqrt{(r_q\phi)^2 + \rho^2} - r_p\phi\sqrt{(r_p\phi)^2 + \rho^2}}{2\rho\phi}\\
&\quad + \frac{|\rho|}{2\phi}\left[\ln(2) + \ln\left(\rho^2\phi + \rho\gamma + |\rho|\sqrt{(r_q\phi)^2 + \rho^2}\right) - \ln(2) - \ln\left(\rho\gamma + |\rho|\sqrt{\gamma^2 + \rho^2}\right)\right] && \text{by item (2)}\\
&= \frac{r_q\sqrt{(r_q\phi)^2 + \rho^2} - r_p\sqrt{(r_p\phi)^2 + \rho^2}}{2\rho} + \frac{|\rho|}{2\phi}\left[\ln\left(r_q\rho\phi + |\rho|\sqrt{(r_q\phi)^2 + \rho^2}\right) - \ln\left(r_p\rho\phi + |\rho|\sqrt{(r_p\phi)^2 + \rho^2}\right)\right]
\end{aligned}
$$

A.4 Properties

A.4.1 Arc function R(p,q) properties

If we really want the *Langrange arc distance* $\mathsf{d}(p,q)$ to be an *extension* of the *spherical metric*, then $\mathsf{R}(p,q)$ must equal $r_p\phi$ (the arc length between p and q on a circle centered at the origin) when $r_p = r_q$. This is in fact the case, as demonstrated next.

Proposition A.1 ($\mathsf{R}(p,q)$ on spherical surface). *Let $\mathsf{R}(p,q)$, r_p, r_q, and ϕ be defined as in Definition A.1.*

PRP
$$\left\{\begin{array}{rcl} r_p = r_q & \neq & 0 \quad \textit{and} \\ \phi & \neq & 0 \end{array}\right\} \implies \left\{\mathsf{R}(p,q) = r_p\phi\right\}$$

✎PROOF:

1. lemma:

$$\lim_{\rho\to 0}\frac{r_q\sqrt{(r_q\phi)^2+\rho^2}-r_p\sqrt{(r_p\phi)^2+\rho^2}}{2\rho} = \lim_{\rho\to 0}\frac{r_q\sqrt{(r_q\phi)^2+0}-r_p\sqrt{(r_p\phi)^2+0}}{2\rho} = \lim_{\rho\to 0}\left(\frac{\phi}{2}\right)\frac{r_q^2-r_p^2}{r_q-r_p}$$
$$= \lim_{\rho\to 0}\left(\frac{\phi}{2}\right)\frac{(r_q-r_p)(r_q+r_p)}{r_q-r_p} = \lim_{\rho\to 0}\left(\frac{\phi}{2}\right)(r_q+r_p)$$
$$= r_p\phi$$

2. lemma:

$$\begin{aligned}
&\lim_{\rho\to 0^+}\frac{\rho}{2\phi}\left[\ln\left(r_q\rho\phi+|\rho|\sqrt{(r_q\phi)^2+\rho^2}\right)-\ln\left(r_p\rho\phi+|\rho|\sqrt{(r_p\phi)^2+\rho^2}\right)\right]\\
&=\lim_{\rho\to 0^+}\frac{\rho}{2\phi}\left[\ln\left(r_q\rho\phi+\rho\sqrt{(r_q\phi)^2}\right)-\ln\left(r_p\rho\phi+\rho\sqrt{(r_p\phi)^2}\right)\right]\\
&=\lim_{\rho\to 0^+}\frac{\rho}{2\phi}\left[\ln(\rho)+\ln\left(r_q\phi+\sqrt{(r_q\phi)^2}\right)-\ln(\rho)-\ln\left(r_p\phi+\sqrt{(r_p\phi)^2}\right)\right]\\
&=\lim_{\rho\to 0^+}\frac{\rho}{2\phi}\left[+\ln\left(r_q\phi+\sqrt{(r_q\phi)^2}\right)-\ln\left(r_p\phi+\sqrt{(r_p\phi)^2}\right)\right]\\
&=0
\end{aligned}$$

3. lemma:

$$\begin{aligned}
&\lim_{\rho\to 0^-}\frac{\rho}{2\phi}\left[\ln\left(r_q\rho\phi+|\rho|\sqrt{(r_q\phi)^2+\rho^2}\right)-\ln\left(r_p\rho\phi+|\rho|\sqrt{(r_p\phi)^2+\rho^2}\right)\right]\\
&=\lim_{\rho\to 0^-}\frac{\rho}{2\phi}\left[\ln\left(-r_q|\rho|\phi+|\rho|\sqrt{(r_q\phi)^2}\right)-\ln\left(-r_p|\rho|\phi+|\rho|\sqrt{(r_p\phi)^2}\right)\right]\\
&=\lim_{\rho\to 0^-}\frac{\rho}{2\phi}\left[\ln(\rho)+\ln\left(-r_q\phi+\sqrt{(r_q\phi)^2}\right)-\ln(\rho)-\ln\left(-r_p\phi+\sqrt{(r_p\phi)^2}\right)\right]\\
&=\lim_{\rho\to 0^-}\frac{\rho}{2\phi}\left[\ln\left(-r_q\phi+\sqrt{(r_q\phi)^2}\right)-\ln\left(-r_p\phi+\sqrt{(r_p\phi)^2}\right)\right]\\
&=0
\end{aligned}$$

4. lemma: By item (2), item (3), and by *continuity* …

$$\lim_{\rho\to 0}\frac{\rho}{2\phi}\left[\ln\left(r_q\rho\phi+|\rho|\sqrt{(r_q\phi)^2+\rho^2}\right)-\ln\left(r_p\rho\phi+|\rho|\sqrt{(r_p\phi)^2+\rho^2}\right)\right]=0$$

5. Completing the proof …

$$\begin{aligned}
\lim_{\rho\to 0}\mathsf{R}(p,q) &= \lim_{\rho\to 0}\left[\frac{r_q\sqrt{(r_q\phi)^2+\rho^2}-r_p\sqrt{(r_p\phi)^2+\rho^2}}{2\rho}\right.\\
&\quad\left.+\frac{\rho}{2\phi}\ln\left(r_q\rho\phi+|\rho|\sqrt{(r_q\phi)^2+\rho^2}\right)-\ln\left(r_p\rho\phi+|\rho|\sqrt{(r_p\phi)^2+\rho^2}\right)\right] && \text{by Theorem A.1}\\
&= 0+\lim_{\rho\to 0}\frac{\rho}{2\phi}\left[\ln\left(r_q\rho\phi+|\rho|\sqrt{(r_q\phi)^2+\rho^2}\right)-\ln\left(r_p\rho\phi+|\rho|\sqrt{(r_p\phi)^2+\rho^2}\right)\right] && \text{by item (1)}\\
&= 0+0 && \text{by item (4)}\\
&= 0
\end{aligned}$$

✏

Later in Theorem A.3 (page 124), we want to prove that the *Langrange arc distance* (Definition A.1 page 114) $\mathsf{d}(p,q)$ is indeed, as its name suggests, a *distance function* (Definition B.1 page 133). Proposition A.2 (*symmetry*) and Proposition A.4 (*positivity*) will help. Meanwhile, Proposition A.4 will itself receive help from Proposition A.3 (*monotonicity*).

Proposition A.2 (*symmetry* of R). *Let* $\mathsf{R}(p,q)$ *be defined as in Definition A.1 (page 114).*

PRP

$$\mathsf{R}(p,q) = \mathsf{R}(q,p) \quad \forall p,q \in \mathbb{R}^N \quad \text{(SYMMETRIC)}$$

✎PROOF:

1. dummy variable: Let $\mu \triangleq \phi - \theta$ which implies $\theta = \phi - \mu$ and $\mathsf{d}\theta = -\,\mathsf{d}\mu$.

2. lemma:

$$\begin{aligned}
\mathsf{r}(\mu;q,p) &\triangleq \mathsf{r}(\phi-\theta;q,p) && \text{by item (1)}\\
&= \left(\frac{r_p - r_q}{\phi}\right)(\phi-\theta) + r_q = \left(\frac{r_q - r_p}{\phi}\right)\theta + \left(r_p - r_q\right) + r_q = \left(\frac{r_q - r_p}{\phi}\right)\theta + r_p\\
&= \mathsf{r}(\theta;p,q)
\end{aligned}$$

3. Completing the proof...

$$\begin{aligned}
\mathsf{R}(p,q) &= \int_{\theta=0}^{\theta=\phi} \sqrt{\mathsf{r}^2(\theta;p,q) + \left[\frac{\mathsf{d}r(\theta;p,q)}{\mathsf{d}\theta}\right]^2}\,\mathsf{d}\theta\\
&= \int_{\phi-\mu=0}^{\phi-\mu=\phi} \sqrt{\mathsf{r}^2(\mu;q,p) + \left[\frac{\mathsf{d}r(\mu;q,p)}{-\,\mathsf{d}\mu}\right]^2}\,(-\,\mathsf{d}\mu) && \text{by item (1) and item (2)}\\
&= -\int_{\mu=\phi}^{\mu=0} \sqrt{\mathsf{r}^2(\mu;q,p) + \left[\frac{\mathsf{d}r(\mu;q,p)}{\mathsf{d}\mu}\right]^2}\,\mathsf{d}\mu\\
&= \int_{\mu=0}^{\mu=\phi} \sqrt{\mathsf{r}^2(\mu;q,p) + \left[\frac{\mathsf{d}r(\mu;q,p)}{\mathsf{d}\mu}\right]^2}\,\mathsf{d}\mu && \text{by the \textit{Second Fundamental Theorem of Calculus}[9]}\\
&= \mathsf{R}(q,p)
\end{aligned}$$

✏

Proposition A.3 (*monotonicity* of R). *Let* $\mathsf{R}(p,q)$ *and* ϕ *be defined as in Definition A.1 (page 114). Let* ϕ_1 *be the polar angle between the point pair* (p_1,q_1) *in* $\mathbb{R}^N$ *and Let* ϕ_2 *the polar angle between the point pair* (p_2,q_2) *in* $\mathbb{R}^N$.

PRP

$$\{\phi_1 < \phi_2\} \implies \{\mathsf{R}(p_1,q_1) < \mathsf{R}(p_2,q_2)\} \quad \forall \phi_1,\phi_2 \in (0:\pi] \quad \text{(STRICTLY MONOTONICALLY INCREASING \textit{in} } \phi\text{)}$$

[9] Hijab (2016) page 170 ⟨Theorem 4.4.3 Second Fundamental Theorem of Calculus⟩, Amann and Escher (2008) page 31 ⟨Theorem 4.13 The second fundamental theorem of calculus⟩

✎PROOF:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\phi}\mathsf{R}(p,q) &\triangleq \frac{\mathrm{d}}{\mathrm{d}\phi}\int_0^{\phi}\sqrt{\mathsf{r}^2(\theta)+\left[\frac{\mathrm{d}\mathsf{r}(\theta)}{\mathrm{d}\theta}\right]^2}\,\mathrm{d}\theta && \text{by Definition A.1 (page 114)}\\
&= \frac{\mathrm{d}}{\mathrm{d}\phi}\int_0^{\phi}\sqrt{\left[\left(\frac{r_q-r_p}{\phi}\right)\theta+r_p\right]^2+\left[\frac{r_q-r_p}{\phi}\right]^2}\,\mathrm{d}\theta\\
&= \sqrt{\left[\left(\frac{r_q-r_p}{\phi}\right)\phi+r_p\right]^2+\left[\frac{r_q-r_p}{\phi}\right]^2} && \text{by the \textit{First Fundamental Theorem of Calculus}[10]}\\
&= \sqrt{r_q^2+\left[\frac{r_q-r_p}{\phi}\right]^2}\\
&> 0\\
&\Longrightarrow \mathsf{R}(p,q) \text{ is \textit{strictly monotonically increasing} in } \phi
\end{aligned}$$

✏

Proposition A.4 (*positivity* of R). *Let* R(p,q) *and* ϕ *be defined as in Definition A.1 (page 114).*

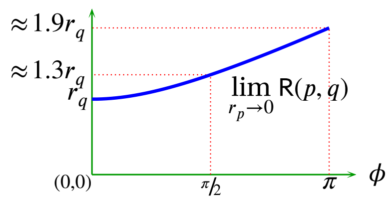
PRP
$$\left\{\begin{array}{ccc}\phi & \in & (0:\pi]\\ (\phi & \neq & 0)\end{array}\right\} \quad\Longrightarrow\quad \left\{\ \mathsf{R}(p,q) \;>\; 0 \quad \forall p,q\in\mathbb{R}^N \quad \text{(POSITIVE)}\ \right\}$$

✎PROOF:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\phi}\mathsf{R}(p,q) &\triangleq \int_0^{\phi}\sqrt{\mathsf{r}^2(\theta)+\left[\frac{\mathrm{d}\mathsf{r}(\theta)}{\mathrm{d}\theta}\right]^2}\,\mathrm{d}\theta && \text{by Definition A.1 (page 114)}\\
&= \left.\int\sqrt{\mathsf{r}^2(\theta)+\left[\frac{\mathrm{d}\mathsf{r}(\theta)}{\mathrm{d}\theta}\right]^2}\,\mathrm{d}\theta\right|_{\phi} - \left.\int\sqrt{\mathsf{r}^2(\theta)+\left[\frac{\mathrm{d}\mathsf{r}(\theta)}{\mathrm{d}\theta}\right]^2}\,\mathrm{d}\theta\right|_{0} && \text{by the Second Fundamental Theorem of Calculus}\\
&> 0 && \text{by Proposition A.3 (page 120)}
\end{aligned}$$

✏

For the sake of *continuity* at the origin of $\mathbb{R}^N$, one might hope that it doesn't matter which "direction" the points p or q approach the origin when computing the limit of R(p,q). This however is *not* the case, as demonstrated next and illustrated to the right and in Example A.1 (page 122). In fact, the limits very much depend on ϕ...resulting in a *discontinuity* at the origin, as demonstrated in Theorem A.2 (page 123).



Proposition A.5 (limit cases of R). *Let* R(p,q), r_p, r_q, *and* ϕ *be defined as in Definition A.1 (page 114).*

THM
$$\begin{aligned}
\lim_{r_p\to 0}\mathsf{R}(p,q) &= \frac{r_q}{2}\left[\sqrt{\phi^2+1}+\frac{\ln\left(\phi+\sqrt{\phi^2+1}\right)}{\phi}\right] && \forall p,q\in\mathbb{R}^N\backslash(0,0,\cdots,0),\ \phi\neq 0 \quad \text{(\textit{p approaching origin})}\\
\lim_{r_q\to 0}\mathsf{R}(p,q) &= \frac{r_p}{2}\left[\sqrt{\phi^2+1}+\frac{\ln\left(\phi+\sqrt{\phi^2+1}\right)}{\phi}\right] && \forall p,q\in\mathbb{R}^N\backslash(0,0,\cdots,0),\ \phi\neq 0 \quad \text{(\textit{q approaching origin})}
\end{aligned}$$

[10] Schechter (1996) page 674 ⟨25.15⟩, Haaser and Sullivan (1991) page 218

✎PROOF:

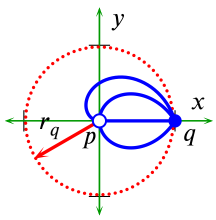
$$\begin{aligned}
&\lim_{r_p\to 0}\mathsf{R}(p,q)\\
&= \lim_{r_p\to 0} \frac{r_q\sqrt{(r_q\phi)^2+\rho^2} - r_p\sqrt{(r_p\phi)^2+\rho^2}}{2\rho} + \frac{|\rho|}{2\phi}\left[\ln\left(r_q\rho\phi + |\rho|\sqrt{(r_q\phi)^2+\rho^2}\right) - \ln\left(r_p\rho\phi+|\rho|\sqrt{(r_p\phi)^2+\rho^2}\right)\right]\\
&\qquad\text{by Theorem A.1 (page 117)}\\
&= \frac{r_q\sqrt{(r_q\phi)^2+r_q^2}-0}{2r_q} + \frac{|r_q|}{2\phi}\left[\ln\left(r_q^2\phi + |r_q|\sqrt{(r_q\phi)^2+r_q^2}\right) - \ln\left(0+|r_q|\sqrt{0+r_q^2}\right)\right] \qquad \text{by } \lim_{r_p\to 0}\text{ operation}\\
&= \frac{r_q\sqrt{\phi^2+1}}{2} + \frac{r_q}{2\phi}\left[\ln\left(r_q^2\phi + r_q^2\sqrt{\phi^2+1}\right) - \ln\left(r_q^2\right)\right]\\
&= \frac{r_q}{2}\left[\sqrt{\phi^2+1} + \frac{\ln(r_q^2) + \ln\left(\phi+\sqrt{\phi^2+1}\right) - \ln\left(r_q^2\right)}{\phi}\right] = \frac{r_q}{2}\left[\sqrt{\phi^2+1} + \frac{\ln\left(\phi+\sqrt{\phi^2+1}\right)}{\phi}\right]
\end{aligned}$$

$$\begin{aligned}
\lim_{r_q\to 0}\mathsf{R}(p,q) &= \lim_{r_q\to 0}\mathsf{R}(q,p) && \text{by Proposition A.2}\\
&= \frac{r_p}{2}\left[\sqrt{\phi^2+1} + \frac{\ln\left(\phi+\sqrt{\phi^2+1}\right)}{\phi}\right] && \text{by previous result}
\end{aligned}$$

✏

Example A.1. Let $\mathsf{R}(p,q)$, ϕ, and r_q be defined as in Definition A.1.

$$\begin{array}{llllll}
\text{If } \phi = 0 & \text{then} & \lim_{p\to 0}\mathsf{R}\left[p,(r_q,0)\right] & = & r_q & \\
\text{If } \phi = \pi/2 & \text{then} & \lim_{p\to 0}\mathsf{R}\left[p,(r_q,0)\right] & = & r_q\times(1.323652\cdots) & \approx 1.3r_q\\
\text{If } \phi = \pi & \text{then} & \lim_{p\to 0}\mathsf{R}\left[p,(r_q,0)\right] & = & r_q\times(1.944847\cdots) & \approx 1.9r_q
\end{array}$$



✎PROOF:

$$\mathsf{R}(p,q)|_{\phi=\pi/2} = \frac{r_q}{2}\left[\sqrt{\left(\frac{\pi}{2}\right)^2+1} + \frac{\ln\left(\frac{\pi}{2}+\sqrt{\left(\frac{\pi}{2}\right)^2+1}\right)}{\frac{\pi}{2}}\right] = r_q\times(1.323652\cdots)$$

$$\mathsf{R}(p,q)|_{\phi=\pi} = \frac{r_q}{2}\left[\sqrt{\pi^2+1} + \frac{\ln\left(\pi+\sqrt{\pi^2+1}\right)}{\pi}\right] = r_q\times(1.944847\cdots)$$

✏

A.4.2 Distance function d(p,q) properties

The *Langrange arc distance* $\mathsf{d}(p,q)$ is defined in two parts: one part being the Euclidean distance $^1\!/\!_\pi\left|r_q - r_p\right|$ and the second part the length of the arc $^1\!/\!_\pi\mathsf{R}(p,q)$. There is risk in creating a multipart definition...with the possible consequences being *discontinuity* at the boundary of the parts. Proposition A.6 (next) demonstrates that when $r_p \neq 0$ and $r_q \neq 0$, there is *continuity* as $\phi \to 0$. However, Theorem A.3 (page 124) demonstrates that in general for values of $\phi > 0$, $\mathsf{d}(p,q)$ is *discontinuous* at the origin.

Proposition A.6. *Let* $\mathsf{R}(p,q)$, r_p, r_q, ϕ, *and* $(0,0,\cdots,0)$ *be defined as in Definition A.1 (page 114).*

PRP

(A). $\displaystyle\lim_{\phi\to 0}\mathsf{R}(p,q) = |r_q - r_p| = \pi\mathsf{d}(p,q) \quad \textit{when } \phi = 0$

(B). $\displaystyle\lim_{\phi\to 0}\lim_{r_p\to 0}\mathsf{R}(p,q) = \lim_{r_p\to 0}\lim_{\phi\to 0}\mathsf{R}(p,q) = r_q = \pi\mathsf{d}((0,0,\cdots,0),q)$

(C). $\displaystyle\lim_{\phi\to 0}\lim_{r_q\to 0}\mathsf{R}(p,q) = \lim_{r_q\to 0}\lim_{\phi\to 0}\mathsf{R}(p,q) = r_p = \pi\mathsf{d}(p,(0,0,\cdots,0))$

✎PROOF:

$$\begin{aligned}
\lim_{\phi\to 0}\mathsf{R}(p,q) &\triangleq \lim_{\phi\to 0}\int_0^{\phi}\sqrt{r^2(\theta)+\left(\frac{\mathsf{d}r}{\mathsf{d}\theta}\right)^2}\,\mathsf{d}\theta && \text{by definition of } \mathsf{R}(p,q) \text{ (Definition A.1 page 114)}\\
&= \lim_{\phi\to 0}\int_0^{\phi}\sqrt{\left[\left(\frac{r_q-r_p}{\phi}\right)\theta + r_p\right]^2 + \left[\frac{r_q-r_p}{\phi}\right]^2}\,\mathsf{d}\theta && \text{by definition of } \mathsf{r}(\theta) \text{ (Definition A.1 page 114)}\\
&= \lim_{\phi\to 0}\left[\frac{1}{\phi}\int_0^{\phi}\sqrt{\left[(r_q-r_p)\theta + r_p\phi\right]^2 + \left[r_q-r_p\right]^2}\,\mathsf{d}\theta\right]\\
&= \frac{\lim_{\phi\to 0}\frac{\mathsf{d}}{\mathsf{d}\phi}\int_0^{\phi}\sqrt{\left[(r_q-r_p)\theta + r_p\phi\right]^2 + \left[r_q-r_p\right]^2}\,\mathsf{d}\theta}{\lim_{\phi\to 0}\frac{\mathsf{d}}{\mathsf{d}\phi}\phi} && \text{by \textit{L'Hôpital's rule}}\\
&= \frac{\lim_{\phi\to 0}\sqrt{\left[(r_q-r_p)\phi + r_p\phi\right]^2 + \left[r_q-r_p\right]^2}}{1} && \text{by \textit{First Fundamental Theorem of Calculus}}\\
&= \sqrt{\left[r_q-r_p\right]^2}\\
&= |r_q-r_p|\\
&= \pi\,\mathsf{d}(p,q)|_{\phi=0} && \text{by Definition A.1 (page 114)}
\end{aligned}$$

$$\begin{aligned}
\lim_{\phi\to 0}\lim_{r_p\to 0}\mathsf{R}(p,q) &= \lim_{\phi\to 0}\frac{r_q}{2}\left[\sqrt{\phi^2+1} + \frac{\ln\left(\phi+\sqrt{\phi^2+1}\right)}{\phi}\right] && \text{by Proposition A.5 (page 121)}\\
&= \frac{r_q}{2}\left[\sqrt{0^2+1} + \frac{\lim_{\phi\to 0}\frac{\mathsf{d}}{\mathsf{d}\phi}\ln\left(\phi+\sqrt{\phi^2+1}\right)}{\lim_{\phi\to 0}\frac{\mathsf{d}}{\mathsf{d}\phi}\phi}\right] && \text{by \textit{L'Hôpital's rule}}\\
&= \frac{r_q}{2}\left[1 + \lim_{\phi\to 0}\frac{1+\frac{2\phi}{2\sqrt{\phi^2+1}}}{\phi+\sqrt{\phi^2+1}}\right] = \frac{r_q}{2}[1+1/1] = r_q = \pi\mathsf{d}((0,0,\cdots,0),q)
\end{aligned}$$

$$\begin{aligned}
\lim_{\phi\to 0}\lim_{r_q\to 0}\mathsf{R}(p,q) &= \lim_{\phi\to 0}\lim_{r_q\to 0}\mathsf{R}(q,p) && \text{by Proposition A.2 (page 120)}\\
&= r_p = \pi\mathsf{d}(p,(0,0,\cdots,0)) && \text{by previous result}\\
\lim_{r_p\to 0}\lim_{\phi\to 0}\mathsf{R}(p,q) &= \lim_{r_p\to 0}|r_q-r_p| && \text{by (A)}\\
&= r_q = \pi\mathsf{d}((0,0,\cdots,0),q)\\
\lim_{r_q\to 0}\lim_{\phi\to 0}\mathsf{R}(p,q) &= \lim_{r_q\to 0}|r_q-r_p| && \text{by (A)}\\
&= r_p = \pi\mathsf{d}(p,(0,0,\cdots,0))
\end{aligned}$$

⇨

Theorem A.2. *Let the* LAGRANGE ARC DISTANCE $\mathsf{d}(p,q)$ *and* ORIGIN *be defined as in Definition A.1.*

THM *The function* $\mathsf{d}(p,q)$ *is* DISCONTINUOUS *at the* ORIGIN *of* $\mathbb{R}^N$, *but is* CONTINUOUS *everywhere else in* $\mathbb{R}^N$.

✎PROOF:

1. Proof for when p and q are *not* at the origin and $\phi \neq 0$:
   (a) In this case, $\mathsf{d}(p,q) = {}^1\!/\!_\pi \mathsf{R}(p,q)$.
   (b) $\mathsf{R}(p,q)$ is continuous everywhere in its domain because its solution, as given by Theorem A.1 (page 117), consists entirely of continuous functions such as $\ln(x)$, $|x|$, etc.
   (c) Therefore, in this case, $\mathsf{d}(p,q)$ is also *continuous*.
2. Proof for when p and q are *not* at the origin and $\phi = 0$:
   This follows from (A) of Proposition A.6 (page 123).
3. Proof for *discontinuity* at origin: This follows from Proposition A.5 (page 121), where it is demonstrated that the limit of $\mathsf{R}(p,q)$ is very much dependent on the "direction" from which p or q approaches the origin. For an illustration of this concept, see Example A.1 (page 122).

✏

Theorem A.3. *Let* $\mathsf{d}(p,q)$ *be* LAGRANGE ARC DISTANCE *(Definition A.1 page 114)*.

THM *The function* $\mathsf{d}(p,q)$ *is a* DISTANCE FUNCTION *(Definition B.1 page 133). In particular,*

(1).	$\mathsf{d}(p,q) \geq 0$	$\forall p,q \in \mathbb{R}^N$	(NON-NEGATIVE)	*and*
(2).	$\mathsf{d}(p,q) = 0 \iff p = q$	$\forall p,q \in \mathbb{R}^N$	(NONDEGENERATE)	*and*
(3).	$\mathsf{d}(p,q) = \mathsf{d}(q,p)$	$\forall p,q \in \mathbb{R}^N$	(SYMMETRIC)	

✎PROOF: The *Langrange arc distance* (Definition A.1 page 114) is simply the *Euclidean metric* (Definition D.10 page 166) if p or q is at the origin, or if $\phi = 0$. In this case, (1)–(3) are satisfied automatically because all *metric*s have these properties (Definition D.7 page 163). What is left to prove is that $\mathsf{R}(p,q)$ has these properties when p and q are not at the origin and $\phi \neq 0$.

1. Proof that $\mathsf{R}(p,q) \geq 0$: If $\phi = 0$, then the Euclidean metric is used.
   For any $\phi > 0$, $\mathsf{R}(p,q) > 0$, as demonstrated by Proposition A.4 (page 121).
2. Proof that $p = q \implies \mathsf{R}(p,q) = 0$: If $p = q$, then $\phi = 0$, and the Euclidean metric is used, not $\mathsf{R}(p,q)$.
3. Proof that $\mathsf{R}(p,q) = 0 \implies p = q$: If $\mathsf{d}(p,q) = 0$ and $\phi = 0$, then the Euclidean metric is used. If $\mathsf{d}(p,q) = 0$ and $\phi > 0$, then $\mathsf{R}(p,q)$ never equals 0 anyways, as demonstrated by Proposition A.4 (page 121).
4. Proof that $\mathsf{R}(p,q) = \mathsf{R}(q,p)$: This is demonstrated by Proposition A.2 (page 120).

✏

The *Lagrange arc distance* is *not* a *metric* because in general the *triangle inequality* property does not hold (next theorem). Furthermore, the *Lagrange arc distance* does not induce a *norm* because it is *not translation invariant* (the *translation invariant* property is a necessary condition for a *metric* to induce a *norm*, Theorem D.5 page 162), and balls in a *Lagrange arc distance space* are in general *not convex* (balls are always *convex* in a *normed linear space*, Theorem D.4 page 160). For more details about *distance space*s, see APPENDIX B (page 131).

Theorem A.4. *In the* LAGRANGE ARC DISTANCE SPACE (X, d) *over a field* $\mathbb{F}$

THM

(1). $\mathsf{d}(p, r) \nleq \mathsf{d}(p, q) + \mathsf{d}(q, r) \quad \forall p,q,r \in X$ (TRIANGLE INEQUALITY FAILS)

(2). $\mathsf{d}(p+r, q+r) \neq \mathsf{d}(p, q) \quad \forall p,q,r \in X$ (NOT TRANSLATION INVARIANT)

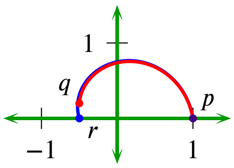
(3). $\mathsf{d}(\alpha p, \alpha q) = |\alpha|\,\mathsf{d}(p, q) \quad \forall p,q,r \in X, \alpha \in \mathbb{R}$ (HOMOGENEOUS)

(4). d *does not induce a norm*

(5). *balls in* (X, d) *are in general* NOT CONVEX

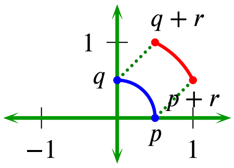
✎PROOF:

1. Proof that the *triangle inequality* property fails to hold in (X, d): Consider the following case[11]...



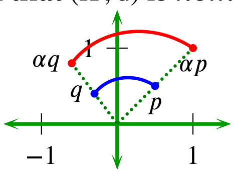
$$\begin{aligned}
\mathsf{d}(p, r) &\triangleq \mathsf{d}((1,0), (-0.5, 0)) &&= 0.767324\cdots \\
&\nleq 0.756406\cdots &&= 0.692330\cdots + 0.064076\cdots \\
&= \mathsf{d}((1,0), (-0.5, 0.2)) &&+ \mathsf{d}((-0.5, 0.2), (-0.5, 0)) \\
&\triangleq \mathsf{d}(p, q) + \mathsf{d}(q, r) \\
\implies & \textit{triangle inequality fails in } (X, \mathsf{d})
\end{aligned}$$

2. Proof that (X, d) is *not translation invariant*:[12] Let $r \triangleq (1/2, 1/2)$. Then...



$$\begin{aligned}
\mathsf{d}(p+r, q+r) &\triangleq \mathsf{d}\left(\left(1, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, 1\right)\right) = 0.229009\cdots \neq \tfrac{1}{2} \\
&= \mathsf{d}\left(\left(\tfrac{1}{2}, 0\right), \left(0, \tfrac{1}{2}\right)\right) \\
&\triangleq \mathsf{d}(p, q) \\
\implies & (X, \mathsf{d}) \textit{ is not translation invariant}
\end{aligned}$$

3. Proof that (X, d) is *homogeneous*:



Let $r_{\alpha p}$ be the magnitude of $\alpha p \triangleq (\alpha x_1, \alpha x_2, \cdots, \alpha x_N)$.
Let $r_{\alpha q}$ be the magnitude of $\alpha q \triangleq (\alpha y_1, \alpha y_2, \cdots, \alpha y_N)$.
Let ϕ_α be the *polar angle* between αp and αq.

(a) If $r_p = 0$ or $r_q = 0$ or $\phi = 0$ then $\mathsf{d}(p, q)$ is the *Euclidean metric*, which is *homogeneous*.

(b) lemmas:

$$\begin{aligned}
r_{\alpha p} &\triangleq \left(\sum_1^N [\alpha x_n]^2\right)^{\frac{1}{2}} &&= |\alpha| \sum_1^N x_n^2 &&\triangleq |\alpha| r_p \\
r_{\alpha q} &\triangleq \left(\sum_1^N [\alpha y_n]^2\right)^{\frac{1}{2}} &&= |\alpha| \sum_1^N y_n^2 &&\triangleq |\alpha| r_q \\
\phi_\alpha &\triangleq \arccos\left(\frac{1}{r_{\alpha p} r_{\alpha q}} \sum_{n=1}^N [\alpha x_n][\alpha y_n]\right) &&= \arccos\left(\frac{1}{r_p r_q} \sum_{n=1}^N x_n y_n\right) &&\triangleq \phi \\
\mathsf{r}(\theta; \alpha p, \alpha q) &\triangleq \left(\frac{r_{\alpha q} - r_{\alpha p}}{\phi_\alpha}\right)\theta + r_{\alpha p} &&= \left(\frac{\alpha r_q - \alpha r_p}{\phi}\right)\theta + \alpha r_p &&= \alpha \mathsf{r}(\theta; p, q)
\end{aligned}$$

(c) If $\mathsf{d}(p, q)$ is *not* the *Euclidean metric* then ...

$$\begin{aligned}
\pi \mathsf{d}(\alpha p, \alpha q) &\triangleq \mathsf{R}(\alpha p, \alpha q) && \text{by definition of } \mathsf{d} \text{ (Definition A.1 page 114)} \\
&\triangleq \int_0^{\phi_\alpha} \sqrt{[\mathsf{r}(\theta; \alpha p, \alpha q)]^2 + \left[\frac{\mathsf{d}\mathsf{r}(\theta; \alpha p, \alpha q)}{\mathsf{d}\theta}\right]^2}\, \mathsf{d}\theta && \text{by definition of } \mathsf{R} \text{ (Definition A.1 page 114)} \\
&= \int_0^{\phi} \sqrt{[\alpha \mathsf{r}(\theta; p, q)]^2 + \left[\tfrac{\mathsf{d}}{\mathsf{d}\theta} \alpha \mathsf{r}(\theta; p, q)\right]^2}\, \mathsf{d}\theta && \text{by item (3b)} \\
&= |\alpha| \int_0^{\phi} \sqrt{[\mathsf{r}(\theta; p, q)]^2 + \left[\tfrac{\mathsf{d}}{\mathsf{d}\theta} \mathsf{r}(\theta; p, q)\right]^2}\, \mathsf{d}\theta && \text{by } \textit{linearity} \text{ of } \int_0^{\phi} \mathsf{d}\theta \text{ operator}
\end{aligned}$$

[11] See experiment log file "lab_larc_distances_R2.xlg" generated by the program "ssp.exe".
[12] See experiment log file "lab_larc_distances_R2.xlg" generated by the program "ssp.exe".

$$\triangleq |\alpha|\mathsf{R}(p, q) \qquad \text{by definition of } \mathsf{R} \text{ (Definition A.1 page 114)}$$

4. Proof that d does *not* induce a norm on X: This follows directly from item (2) and Theorem D.5 (page 162).

5. Proof that *balls* (Definition B.4 page 134) in d are in general *not convex* (Definition 1.25 page 10):
This is demonstrated graphically in Figure A.2 (page 128) and Figure A.3 (page 129).
For an algebraic demonstration, consider the following:[13]

   (a) Let $\mathsf{B}((0,1),1)$ be the *unit ball* in $(\mathbb{R}^2, \mathsf{d})$ centered at $(0,1)$.

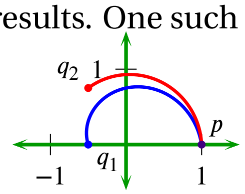
   (b) Let $p \triangleq (-0.70, -1.12)$, $q \triangleq (0.70, -1.12)$, $r \triangleq (0, -1.12)$, and $\lambda = 1/2$.

   (c) Then
$$\begin{aligned}
\mathsf{d}((0,1), p) &= 0.959536\cdots < 1 &&\implies p \in \mathsf{B}((0,1),1) && \text{and}\\
\mathsf{d}((0,1), q) &= 0.959536\cdots < 1 &&\implies q \in \mathsf{B}((0,1),1) && \text{BUT}\\
\mathsf{d}((0,1), r) &= 1.060688\cdots > 1 &&\implies r \notin \mathsf{B}((0,1),1)
\end{aligned}$$

   (d) This implies that the set $\mathsf{B}((0,1),1)$ is *not convex* because
$$\begin{aligned}
\lambda p + (1-\lambda)q &\triangleq \frac{1}{2}(-0.70, -1.12) + \left(1 - \frac{1}{2}\right)(0.70, -1.12) && \text{by item (5b)}\\
&= (0, -1.12)\\
&\triangleq r && \text{by item (5b)}\\
&\notin \mathsf{B}((0,1),1) && \text{by item (5c)}\\
&\implies \text{the set } \mathsf{B}((0,1),1) \text{ is } \textit{not convex} && \text{by Definition 1.25 page 10}
\end{aligned}$$

⇰

Remark A.1 (Lagrange arc distance versus Euclidean metric). As is implied by the metric balls illustrated in Figure A.2 (page 128) and Figure A.3 (page 129), the *Lagrange arc distance* d and *Euclidean metric* p are similar in the sense that they often lead to the same results[14] in determining which of the two points q_1 or q_2 is "closer" to a point p. But in some cases the two metrics lead to two different results. One such case is illustrated as follows:[15]



$$\begin{aligned}
\mathsf{d}(p, q_1) &\triangleq \mathsf{d}((1,0), (-0.5, 0)) &&= 0.767324\cdots\\
\mathsf{d}(p, q_2) &\triangleq \mathsf{d}((1,0), (-0.5, 0.75)) &&= 0.654039\cdots\\
\mathsf{p}(p, q_1) &\triangleq \mathsf{p}((1,0), (-0.5, 0)) &&= 1.5\\
\mathsf{p}(p, q_2) &\triangleq \mathsf{p}((1,0), (-0.5, 0.75)) &&= \sqrt{(1.5)^2 + (0.75)^2} = 1.677050\cdots
\end{aligned}$$

That is, q_2 is closer than q_1 to p with respect to the *Lagrange arc distance*,
but q_1 is closer than q_2 to p with respect to the *Euclidean metric*.

A.5 Examples

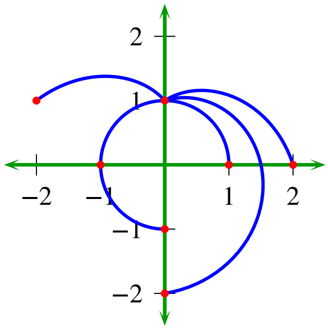
Example A.2 (Lagrange arc distance in $\mathbb{R}^2$). Figure A.1 (page 127) illustrates the *Lagrange arc distance* on some pairs of points in $\mathbb{R}^2$.[16]

Example A.3 (Lagrange arc distance in $\mathbb{R}^3$). Some examples of Lagrange arc distances in $\mathbb{R}^3$ are given in Table A.1 (page 127).[17]

Example A.4 (Lagrange arc distance balls in $\mathbb{R}^2$). Some unit balls in $\mathbb{R}^2$ in with respect to the *Lagrange arc distance* are illustrated in Figure A.2 (page 128).

Example A.5 (Lagrange arc distance balls in $\mathbb{R}^3$). Some unit balls in $\mathbb{R}^3$ with respect to the *Lagrange arc distance* are illustrated in Figure A.3 (page 129).

[13] See experiment log file "lab_larc_distances_R2.xlg" generated by the program "ssp.exe".
[14] For empirical evidence of this, see Greenhoe (2016c).
[15] See experiment log file "lab_larc_distances_R2.xlg" generated by the program "ssp.exe".
[16] See experiment log file "lab_larc_distances_R2.xlg" generated by the program "ssp.exe".
[17] See experiment log file "lab_larc_distances_R3.xlg" generated by the program "ssp.exe".



$$
\begin{aligned}
\mathsf{d}((\;0,1),\;(\;1,\;0)) &= \tfrac{1}{2}\\
\mathsf{d}((\;0,1),\;(-1,\;0)) &= \tfrac{1}{2}\\
\mathsf{d}((\;0,1),\;(\;0,-1)) &= 1\\
\mathsf{d}((\;1,0),\;(\;0,-1)) &= \tfrac{1}{2}\\
\mathsf{d}((\;1,0),\;(-1,\;0)) &= 1\\
\mathsf{d}((-1,0),\;(\;0,-1)) &= \tfrac{1}{2}\\
\mathsf{d}((\;0,1),\;(\;2,\;0)) &= 0.8167968\cdots\\
\mathsf{d}((\;0,1),\;(\;0,-2)) &= 1.5346486\cdots\\
\mathsf{d}((\;0,1),\;(-2,\;1)) &= 0.6966032\cdots
\end{aligned}
$$

Figure A.1: *Lagrange arc distance* examples in $\mathbb{R}^2$

$\mathsf{d}((0,1,0),(1,0,0)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,1,0),(0,0,1)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,1,0),(0,0,-1)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,1,0),(-1,0,0)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,1,0),(0,-1,0)) = 1$	$\phi = \pi$	180°
$\mathsf{d}((1,0,0),(0,0,1)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((1,0,0),(0,0,-1)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((1,0,0),(-1,0,0)) = 1$	$\phi = \pi$	180°
$\mathsf{d}((1,0,0),(0,-1,0)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,0,1),(0,0,-1)) = 1$	$\phi = \pi$	180°
$\mathsf{d}((0,0,1),(-1,0,0)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,0,1),(0,-1,0)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,0,-1),(-1,0,0)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,0,-1),(0,-1,0)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((-1,0,0),(0,-1,0)) = \tfrac{1}{2}$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,1,0),(2,0,0)) = 0.816796\cdots$	$\phi = \tfrac{\pi}{2}$	90°
$\mathsf{d}((0,1,0),(0,-2,0)) = 1.534648\cdots$	$\phi = \pi$	180°
$\mathsf{d}((0,1,0),(-2,1,0)) = 0.696603\cdots$	$\phi \approx 1.107$	63°
$\mathsf{d}((0,1,0),(-1,0,-1)) = 0.617920\cdots$	$\phi = \pi$	90°
$\mathsf{d}((1,1,1),(-\tfrac{1}{2},\tfrac{1}{4},-2)) = 1.366268\cdots$	$\phi \approx 2.2466$	128.72°

Table A.1: Some examples of Lagrange arc distances in $\mathbb{R}^3$ (see Example A.3 page 126)

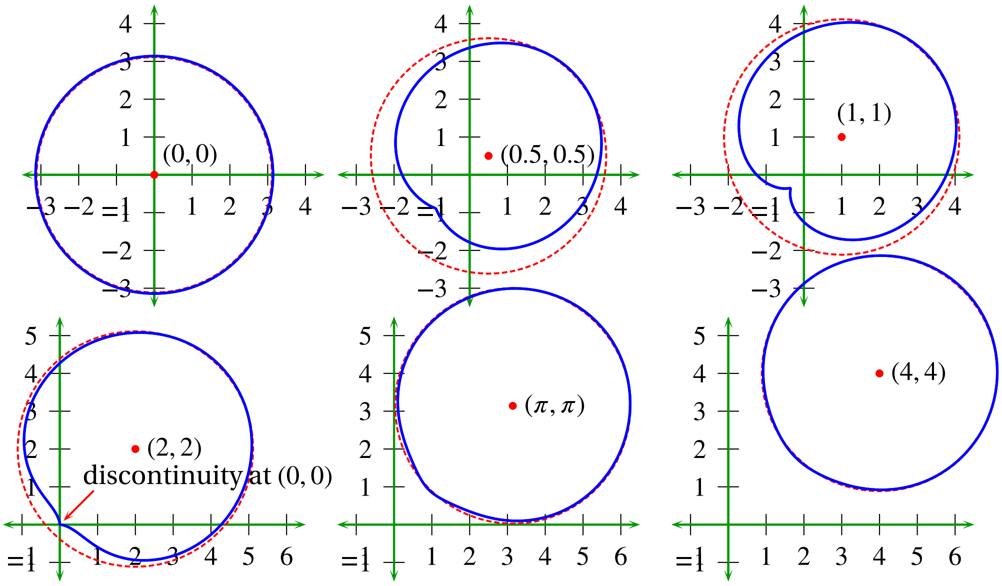


Figure A.2: *Lagrange arc distance* unit balls, and dashed $\frac{1}{\pi}$-*scaled Euclidean metric* unit balls

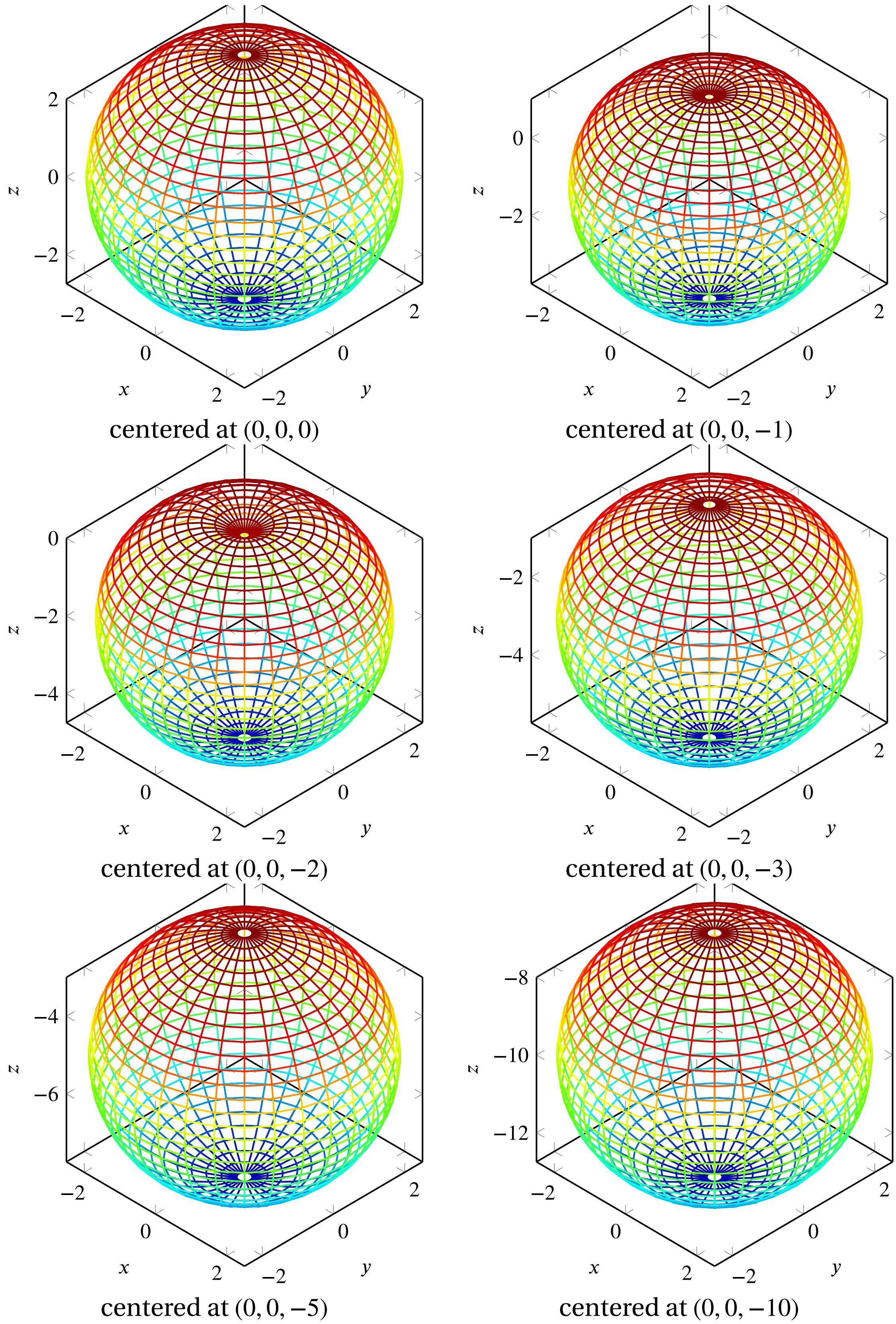


Figure A.3: *unit Lagrange arc distance balls* in $\mathbb{R}^3$

APPENDIX B

DISTANCE SPACES

The Epicureans are wont to ridicule this theorem, saying it is evident even to an ass and needs no proof; it is as much the mark of an ignorant man, they say, to require persuasion of evident truths as to believe what is obscure without question.... That the present theorem is known to an ass they make out from the observation that, if straw is placed at one extremity of the sides, an ass in quest of provender will make his way along the one side and not by way of the two others.

Proclus Lycaeus (412 – 485 AD), Greek philosopher, commenting on the Epicureans view of the *triangle inequality* property. [1]

B.1 Introduction and summary

Metric spaces provide a framework for analysis and have several very useful properties. Many of these properties follow in part from the *triangle inequality*. However, there are several applications[2] in which the triangle inequality does not hold but in which we would still like to perform analysis. So the questions that natually follow are:

Q1. What happens if we remove the *triangle inequality* all together?
Q2. What happens if we replace the *triangle inequality* with a generalized relation?

A *distance space* is a *metric space* without the *triangle inequality* constraint. Section B.2 introduces *distance spaces* and demonstrates that some properties commonly associated with *metric spaces* also hold in any *distance space*:

D1. $\varnothing$ and X are *open* (Theorem B.2 page 134)
D2. the intersection of a finite number of open sets is *open* (Theorem B.2 page 134)
D3. the union of an arbitrary number of open sets is *open* (Theorem B.2 page 134)
D4. every Cauchy sequence is *bounded* (Proposition B.1 page 137)
D5. any subsequence of a *Cauchy* sequence is also *Cauchy* (Proposition B.2 page 137)
D6. the *Cantor Intersection Theorem* holds (Theorem B.5 page 139)

[1] Lycaeus (circa 450), page 251

[2] references for applications in which the *triangle inequality* may not hold: Maligranda and Orlicz (1987) page 54 ⟨"pseudonorm"⟩, Lin (1998) ⟨"similarity measures", Table 6⟩, Veltkamp and Hagedoorn (2000) ⟨"shape similarity measures"⟩, Veltkamp (2001) ⟨"shape matching"⟩, Costa et al. (2004) ⟨"network distance estimation"⟩, Burstein et al. (2005) page 287 ⟨distance matrices for "genome phylogenies"⟩, Jiménez and Yukich (2006) page 224 ⟨"statistical distances"⟩, Szirmai (2007) page 388 ⟨"geodesic ball"⟩, Crammer et al. (2007) page 326 ⟨"decision-theoretic learning"⟩, Crammer et al. (2008) page 1758 ⟨"*approximate* triangle inequality"⟩, Vitányi (2011) page 2455 ⟨"information distance"⟩

The following five properties (M1–M5) *do* hold in any *metric space*. However, the examples from Section B.2 listed below demonstrate that the five properties do *not* hold in all *distance spaces*:

M1.	the *metric function* is *continuous*	fails to hold in	Example B.1–Example B.3
M2.	*open balls* are *open*	fails to hold in	Example B.1 and Example B.2
M3.	the *open balls* form a *base* for a topology	fails to hold in	Example B.1 and Example B.2
M4.	the limits of *convergent sequences* are *unique*	fails to hold in	Example B.1
M5.	*convergent* sequences are *Cauchy*	fails to hold in	Example B.2

Hence, Section B.2 answers question Q1.

APPENDIX C begins to answer question Q2 by first introducing a new function, called the *power triangle function* in a *distance space* (X, d), as $\tau(p,\sigma;x,y,z;\mathsf{d}) \triangleq 2\sigma\left[\frac{1}{2}\mathsf{d}^p(x,z)+\frac{1}{2}\mathsf{d}^p(z,y)\right]^{\frac{1}{p}}$ for some $(p,\sigma)\in\mathbb{R}^*\times\mathbb{R}$. APPENDIX C then goes on to use this function to define a new relation, called the *power triangle inequality* in (X, d), and defined as

$\circledtriangle(p,\sigma;\mathsf{d}) \triangleq \left\{(x,y,z)\in X^3 | \mathsf{d}(x,y) \le \tau(p,\sigma;x,y,z;\mathsf{d})\right\}$.

The *power triangle inequality* is a generalized form of the *triangle inequality* in the sense that the two inequalities coincide at $(p,\sigma)=(1,1)$. Other special values include $(1,\sigma)$ yielding the *relaxed triangle inequality* (and its associated *near metric space*) and (∞,σ) yielding the *σ-inframetric inequality* (and its associated *σ-inframetric space*). Collectively, a distance space with a power triangle inequality is herein called a *power distance space* and denoted (X,d,p,σ).[3]

The *power triangle function*, at $\sigma=\frac{1}{2}$, is a special case of the *power mean* with $N=2$ and $\lambda_1=\lambda_2=\frac{1}{2}$. *Power means* have the elegant properties of being *continuous* and *monontone* with respect to a free parameter p. From this it is easy to show that the *power triangle function* is also *continuous* and *monontone* with respect to both p and σ. Special values of p yield operators coinciding with *maximum*, *minimum*, *mean square*, *arithmetic mean*, *geometric mean*, and *harmonic mean*. *Power means* are briefly described in APPENDIX D.3.3.[4]

Section C.2 investigates the properties of *power distance spaces*. In particular, it shows for what values of (p,σ) the properties M1–M5 hold. Here is a summary of the results in a *power distance space* (X,d,p,σ), for all $x,y,z\in X$:

(M1)	holds for any $(p,\sigma)\in(\mathbb{R}^*\setminus\{0\})\times\mathbb{R}^+$ such that $2\sigma = 2^{\frac{1}{p}}$	(Theorem C.5 page 152)
(M2)	holds for any $(p,\sigma)\in(\mathbb{R}^*\setminus\{0\})\times\mathbb{R}^+$ such that $2\sigma \le 2^{\frac{1}{p}}$	(Corollary C.6 page 150)
(M3)	holds for any $(p,\sigma)\in(\mathbb{R}^*\setminus\{0\})\times\mathbb{R}^+$ such that $2\sigma \le 2^{\frac{1}{p}}$	(Corollary C.5 page 149)
(M4)	holds for any $(p,\sigma)\in\mathbb{R}^*\times\mathbb{R}^+$	(Theorem C.6 page 153)
(M5)	holds for any $(p,\sigma)\in\mathbb{R}^*\times\mathbb{R}^+$	(Theorem C.3 page 151)

APPENDIX D.4 briefly introduces *topological spaces*. The *open balls* of any *metric space* form a *base* for a *topology*. This is largely due to the fact that in a metric space, open balls are *open*. Because of this, in metric spaces it is convenient to use topological structure to define and exploit analytic concepts such as *continuity*, *convergence*, *closed sets*, *closure*, *interior*, and *accumulation point*. For example, in a metric space, the traditional definition of defining continuity using open balls and the topological definition using open sets, coincide with each other. Again, this is largely because the open balls of a metric space are open.[5]

---

[3] *power triangle inequality*: Definition C.3 page 146; *power distance space*: Definition C.2 page 146; examples of *power distance space*: Definition C.4 page 146;

[4] *power triangle function*: Definition C.1 (page 145); *power mean*: Definition D.15 (page 174); power mean is *continuous* and *monontone*: Theorem D.14 (page 175); power triangle function is *continuous* and *monontone*: Corollary C.1 (page 146); Special values of p: Corollary C.2 (page 147), Corollary D.2 (page 177)

[5] *open ball*: Definition B.4 page 134; *metric space*: Definition D.7 page 163; *base*: Definition D.17 page 179; *topology*: Definition D.16 page 179; *open*: Definition B.5 page 134; *continuity* in *topological space*: Definition D.19 page 180; *convergence* in *distance space*: Definition B.7 page 137; *convergence* in *topological space*: Definition D.20 page 181; *closed*

However, this is not the case for all *distance spaces*. In general, the open balls of a distance space are not open, and they are not a base for a topology. In fact, the open balls of a distance space are a base for a topology if and only if the open balls are open. While the open sets in a distance space do induce a topology, it's open balls may not.[6]

A *distance space* (Definition B.1 page 133) can be defined as a *metric space* (Definition D.7 page 163) without the *triangle inequality* constraint. Much of the material in this section about *distance spaces* is standard in *metric spaces*. However, this paper works through this material again to demonstrate "how far we can go", and can't go, without the *triangle inequality*.

B.2 Fundamental structure of distance spaces

B.2.1 Definitions

Definition B.1. [7]

DEF

A function d *in the set* $\mathbb{R}^{X\times X}$ *(Definition 1.6 page 6) is a* ***distance*** *if*

$$\begin{array}{llllll} 1. & \mathsf{d}(x,y) \geq 0 & \forall x,y\in X & (\text{NON-NEGATIVE}) & \text{and} \\ 2. & \mathsf{d}(x,y) = 0 \iff x=y & \forall x,y\in X & (\text{NONDEGENERATE}) & \text{and} \\ 3. & \mathsf{d}(x,y) = \mathsf{d}(y,x) & \forall x,y\in X & (\text{SYMMETRIC}) & \end{array}$$

The pair (X,d) *is a* ***distance space*** *if* d *is a* DISTANCE *on a set* X.

Definition B.2. [8] *Let* (X,d) *be a* DISTANCE SPACE *and* 2^X *be the* POWER SET *of* X *(Definition 1.8 page 6).*

DEF

The ***diameter*** *in* (X,d) *of a set* $A\in 2^X$ *is* $\operatorname{diam} A \triangleq \begin{cases} 0 & \text{for } A=\varnothing \\ \sup\{\mathsf{d}(x,y)\,|\,x,y\in A\} & \text{otherwise} \end{cases}$

Definition B.3. [9] *Let* (X,d) *be a* DISTANCE SPACE. *Let* 2^X *be the* POWER SET *(Definition 1.8 page 6) of* X.

DEF

A set A *is* ***bounded*** *in* (X,d) *if* $A\in 2^X$ *and* $\operatorname{diam} A < \infty$.

B.2.2 Properties

Theorem B.1. [10] *Let* $(\!(x_n)\!)_{n\in\mathbb{Z}}$ *be a* SEQUENCE *in a* DISTANCE SPACE (X,d). *The* DISTANCE SPACE (X,d) *does not necessarily have all the nice properties that a* METRIC SPACE *(Definition D.7 page 163) has. In particular, note the following:*

---

set: Definition D.16 page 179; *closure*, *interior*, *accumulation point*: Definition D.18 page 180; coincidence in all *metric spaces* and some *power distance spaces*: Theorem C.2 page 151;

[6] *if and only if* statement: Theorem B.3 page 136; open sets of a distance space induce a topology: Corollary B.1 page 135;

[7] Menger (1928) page 76 ⟨"Abstand *a b* definiert ist…" (distance from *a* to *b* is defined as…")⟩, Wilson (1931b) page 361 ⟨§1., "distance", "semi-metric space"⟩, Blumenthal (1938) page 38, Blumenthal (1953) page 7 ⟨"DEFINITION 5.1. A distance space is called semimetric provided…"⟩, Galvin and Shore (1984) page 67 ⟨"distance function"⟩, Laos (1998) page 118 ⟨"distance space"⟩, Khamsi and Kirk (2001) page 13 ⟨"semimetric space"⟩, Bessenyei and Pales (2014) page 2 ⟨"semimetric space"⟩, Deza and Deza (2014) page 3 ⟨"**distance** (or **dissimilarity**)"⟩

[8] in *metric space*: Hausdorff (1937), page 166, Copson (1968), page 23, Michel and Herget (1993), page 267, Molchanov (2005) page 389

[9] in *metric space*: Thron (1966), page 154 ⟨definition 19.5⟩, Bruckner et al. (1997) page 356

[10] Greenhoe (2016b)

THM

1.	d *is a* DISTANCE *in* (X, d)	$\not\Longrightarrow$	d *is* CONTINUOUS *in* (X, d)
2.	B *is an* OPEN BALL *in* (X, d)	$\not\Longrightarrow$	B *is* OPEN *in* (X, d)
3.	$\boldsymbol{B}$ *is the set of all* OPEN BALLS *in* (X, d)	$\not\Longrightarrow$	$\boldsymbol{B}$ *is a* BASE *for a topology on* X
4.	$(\!(x_n)\!)$ *is* CONVERGENT *in* (X, d)	$\not\Longrightarrow$	*limit is* UNIQUE
5.	$(\!(x_n)\!)$ *is* CONVERGENT *in* (X, d)	$\not\Longrightarrow$	$(\!(x_n)\!)$ *is* CAUCHY *in* (X, d)

✎PROOF:

1. d is *continuous* in (X, d) — Example B.3 page 142.
2. B is *open* in (X, d) — Example B.2 page 141.
3. $\boldsymbol{B}$ is a *base* for a topology on X — Example B.2 page 141.[11]
4. limit is *unique* — Example B.1 page 140.
5. $(\!(x_n)\!)$ is *Cauchy* in (X, d) — Example B.2 page 141.

✏

B.3 Open sets in distance spaces

B.3.1 Definitions

Definition B.4. [12] *Let* (X, d) *be a* DISTANCE SPACE *(Definition B.1 page 133).*

DEF

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

Definition B.5. *Let* (X, d) *be a* DISTANCE SPACE. *Let* $X \setminus A$ *be the* SET DIFFERENCE *of* X *and a set* A.

DEF

A set U *is **open** in* (X, d) *if* $U \in 2^X$ *and for every* x *in* U *there exists* $r \in \mathbb{R}^+$ *such that* $\mathsf{B}(x, r) \subseteq U$.
A set U *is an **open set** in* (X, d) *if* U *is* OPEN *in* (X, d). *A set* D *is **closed** in* (X, d) *if* $(X \setminus D)$ *is* OPEN.
A set D *is a **closed set** in* (X, d) *if* D *is* CLOSED *in* (X, d).

B.3.2 Properties

Theorem B.2. [13] *Let* (X, d) *be a* DISTANCE SPACE. *Let* N *be any (finite) positive integer. Let* Γ *be a* SET *possibly with an uncountable number of elements.*

THM

1.				X	*is* OPEN.
2.				$\emptyset$	*is* OPEN.
3.	*each element in* $\{U_n \,|\, n=1,2,\dots,N\}$	*is* OPEN	$\Longrightarrow$	$\bigcap_{n=1}^{N} U_n$	*is* OPEN.
4.	*each element in* $\{U_\gamma \in 2^X | \gamma \in \Gamma\}$	*is* OPEN	$\Longrightarrow$	$\bigcup_{\gamma\in\Gamma} U_\gamma$	*is* OPEN.

✎PROOF:

1. Proof that X is *open* in (X, d):
   (a) By definition of *open set* (Definition B.5 page 134), X is *open* $\iff \forall x \in X \quad \exists r$ such that $\mathsf{B}(x, r) \subseteq X$.
   (b) By definition of *open ball* (Definition B.4 page 134), it is always true that $\mathsf{B}(x, r) \subseteq X$ in (X, d).

[11] Heath (1961) page 810 ⟨THEOREM⟩, Galvin and Shore (1984) page 71 ⟨2.3 LEMMA⟩
[12] in *metric space*: Aliprantis and Burkinshaw (1998), page 35
[13] in *metric space*: Dieudonné (1969), pages 33–34, Rosenlicht (1968) page 39

(c) Therefore, X is *open* in (X, d).

2. Proof that $\varnothing$ is *open* in (X, d):

   (a) By definition of *open set* (Definition B.5 page 134), $\varnothing$ is *open* $\iff \forall x \in X \quad \exists r \quad \text{such that} \quad \mathsf{B}(x, r) \subseteq \varnothing$.

   (b) By definition of *empty set* $\varnothing$ (Definition 1.1 page 5), this is always true because no x is in $\varnothing$.

   (c) Therefore, $\varnothing$ is *open* in (X, d).

3. Proof that $\bigcup U_\gamma$ is *open* in (X, d):

   (a) By definition of *open set* (Definition B.5 page 134), $\bigcup U_\gamma$ is *open* $\iff \forall x \in \bigcup U_\gamma \quad \exists r \quad \text{such that} \quad \mathsf{B}(x, r) \subseteq \bigcup U_\gamma$.

   (b) If $x \in \bigcup U_\gamma$, then there is at least one $U \in \bigcup U_\gamma$ that contains x.

   (c) By the left hypothesis in (4), that set U is open and so for that x, $\exists r \quad \text{such that} \quad \mathsf{B}(x, r) \subseteq U \subseteq \bigcup U_\gamma$.

   (d) Therefore, $\bigcup U_\gamma$ is *open* in (X, d).

4. Proof that U_1 and U_2 are *open* $\implies U_1 \cap U_2$ is *open*:

   (a) By definition of *open set* (Definition B.5 page 134), $U_1 \cap U_2$ is *open* $\iff \forall x \in U_1 \cap U_2 \quad \exists r \quad \text{such that} \quad \mathsf{B}(x, r) \subseteq U_1 \cap U_2$.

   (b) By the left hypothesis above, U_1 and U_2 are *open*; and by the definition of *open sets* (Definition B.5 page 134), there exists r_1 and r_2 such that $\mathsf{B}(x, r_1) \subseteq U_1$ and $\mathsf{B}(x, r_2) \subseteq U_2$.

   (c) Let $r \triangleq \min\{r_1, r_2\}$. Then $\mathsf{B}(x, r) \subseteq U_1$ and $\mathsf{B}(x, r) \subseteq U_2$.

   (d) By definition of *set intersection* $\cap$ then, $\mathsf{B}(x, r) \subseteq U_1 \cap U_2$.

   (e) By definition of *open set* (Definition B.5 page 134), $U_1 \cap U_2$ is *open*.

5. Proof that $\bigcap_{n=1}^{N} U_n$ is *open* (by induction):

   (a) Proof for $N = 1$ case: $\bigcap_{n=1}^{N} U_n = \bigcap_{n=1}^{1} U_n = U_1$ is *open* by hypothesis.

   (b) Proof that N case $\implies N+1$ case:

$$\begin{aligned}
\bigcap_{n=1}^{N+1} U_n &= \left(\bigcap_{n=1}^{N} U_n\right) \cap U_{N+1} && \text{by property of } \bigcap \\
&\implies \textit{open} && \text{by "N case" hypothesis and (4) lemma page 135}
\end{aligned}$$

✎

Corollary B.1. *Let* (X, d) *be a* DISTANCE SPACE.

COR: *The set* $\boldsymbol{T} \triangleq \{U \in 2^X | U \text{ is OPEN in } (X, \mathsf{d})\}$ *is a* TOPOLOGY *on* X, *and* $(X, \boldsymbol{T})$ *is a* TOPOLOGICAL SPACE.

✎PROOF: This follows directly from the definition of an *open* set (Definition B.5 page 134), Theorem B.2 (page 134), and the definition of *topology* (Definition D.16 page 179). ✎

Of course it is possible to define a very large number of topologies even on a finite set with just a handful of elements;[14] and it is possible to define an infinite number of topologies even on a *linearly*

[14]For a finite set X with n elements, there are 29 topologies on X if $n = 3$, 6942 topologies on X if $n = 5$, and and 8,977,053,873,043 (almost 9 trillion) topologies on X if $n = 10$. References: Sloane (2014) ⟨http://oeis.org/A000798⟩, Brown and Watson (1996), page 31, Comtet (1974) page 229, Comtet (1966), Chatterji (1967), page 7, Evans et al. (1967), Krishnamurthy (1966), page 157

ordered infinite set like the *real line* $(\mathbb{R}, \leq)$.[15] Be that as it may, Definition B.6 (next definition) defines a single but convenient *topological space* in terms of a *distance space*. Note that every *metric space* conveniently and naturally induces a *topological space* because the *open balls* of the metric space form a *base* for the *topology*. This is not the case for all distance spaces. But if the open balls of a *distance space* are all *open*, then those open balls induce a topology (next theorem).[16]

Definition B.6. *Let* (X, d) *be a* DISTANCE SPACE.

DEF: *The set* $\boldsymbol{T} \triangleq \{U \in 2^X | U \text{ is OPEN in } (X, \mathsf{d})\}$ *is the* ***topology induced by*** (X, d) ***on*** X. *The pair* $(X, \boldsymbol{T})$ *is called the* ***topological space induced by*** (X, d).

For any *distance space* (X, d), no matter how strange, there is guaranteed to be at least one *topological space induced by* (X, d)—and that is the *indiscrete topological space* (Example D.12 page 179) because for any distance space (X, d), $\varnothing$ and X are *open sets* in (X, d) (Theorem B.2 page 134).

Theorem B.3. *Let* $\boldsymbol{B}$ *be the set of all* OPEN BALLS *in a* DISTANCE SPACE (X, d).

THM: $$\{\text{every OPEN BALL in } \boldsymbol{B} \text{ is OPEN}\} \iff \{\boldsymbol{B} \text{ is a BASE for a TOPOLOGY}\}$$

✎PROOF:

$\boxed{\text{every } \textit{open ball} \text{ in } \boldsymbol{B} \text{ is } \textit{open}}$

$\implies$ for every x in $B_y \in \boldsymbol{B}$ there exists $r \in \mathbb{R}^+$ such that $\mathsf{B}(x, r) \subseteq B_y$ — by definition of *open* (Definition B.5 page 134)

$\implies \left\{\begin{array}{ll}\text{for every } x \in X \text{ and for every } B_y \in \boldsymbol{B} \text{ containing } x, & \\ \text{there exists } B_x \in \boldsymbol{B} \text{ such that} & x \in B_x \subseteq B_y.\end{array}\right\}$ — because $\forall (x, r) \in X \times \mathbb{R}^+, \mathsf{B}(x, r) \subseteq X$

$\implies \boxed{\boldsymbol{B} \text{ is a } \textit{base} \text{ for } \boldsymbol{T}}$ — by Theorem D.16 page 179

$\implies \left\{\begin{array}{ll}\text{for every } x \in X \text{ and for every } U \subseteq \boldsymbol{T} \text{ containing } x, & \\ \text{there exists } B_x \in \boldsymbol{B} \text{ such that} & x \in B_x \subseteq U.\end{array}\right\}$ — by Theorem D.16 page 179

$\implies \left\{\begin{array}{ll}\text{for every } x \in X \text{ and for every } B_y \in \boldsymbol{B} \subseteq \boldsymbol{T} \text{ containing } x, & \\ \text{there exists } B_x \in \boldsymbol{B} \text{ such that} & x \in B_x \subseteq B_y.\end{array}\right\}$ — by definition of *base* (Definition D.17 page 179)

$\implies \left\{\begin{array}{ll}\text{for every } x \in B_y \in \boldsymbol{B} \subseteq \boldsymbol{T}, & \\ \text{there exists } B_x \in \boldsymbol{B} \text{ such that} & x \in B_x \subseteq B_y.\end{array}\right\}$

$\implies \boxed{\text{every } \textit{open ball} \text{ in } \boldsymbol{B} \text{ is } \textit{open}}$ — by definition of *open* (Definition B.5 page 134)

⇨

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

[16] *metric space*: Definition D.7 page 163; *open ball*: Definition B.4 page 134; *base*: Definition D.17 page 179; *topology*: Definition D.16 page 179; not all open balls are open in a distance space: Example B.1 (page 140) and Example B.2 (page 141);

B.4 Sequences in distance spaces

B.4.1 Definitions

Definition B.7. [17] *Let* $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ *be a* SEQUENCE *in a* DISTANCE SPACE (X, d).

DEF

The sequence $(\!(x_n)\!)$ ***converges*** *to a* ***limit*** x *if for any* $\varepsilon \in \mathbb{R}^+$, *there exists* $N \in \mathbb{Z}$ *such that for all* $n > N$, $\mathsf{d}(x_n, x) < \varepsilon$.

This condition can be expressed in any of the following forms:

1. *The* ***limit*** *of the sequence* $(\!(x_n)\!)$ *is* x.
2. *The sequence* $(\!(x_n)\!)$ *is* ***convergent*** *with limit* x.
3. $\lim_{n\to\infty} (\!(x_n)\!) = x$.
4. $(\!(x_n)\!) \to x$.

A SEQUENCE *that converges is* ***convergent***.

Definition B.8. [18] *Let* $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ *be a* SEQUENCE *in a* DISTANCE SPACE (X, d).

DEF

The sequence $(\!(x_n)\!)$ *is a* ***Cauchy sequence*** *in* (X, d) *if*

for every $\varepsilon \in \mathbb{R}^+$, *there exists* $N \in \mathbb{Z}$ *such that* $\forall n, m > N,\ \mathsf{d}(x_n, x_m) < \varepsilon$ (CAUCHY CONDITION).

Definition B.9. [19] *Let* $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ *be a* SEQUENCE *in a* DISTANCE SPACE (X, d).

DEF

The sequence $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ *is* ***complete*** *in* (X, d) *if*

$(\!(x_n)\!)$ *is* CAUCHY *in* $(X, \mathsf{d}) \implies (\!(x_n)\!)$ *is* CONVERGENT *in* (X, d).

B.4.2 Properties

Proposition B.1. [20] *Let* $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ *be a* SEQUENCE *in a* DISTANCE SPACE (X, d).

PRP

$\{ (\!(x_n)\!)$ *is* CAUCHY *in* $(X, \mathsf{d}) \} \implies \{ (\!(x_n)\!)$ *is* BOUNDED *in* $(X, \mathsf{d}) \}$

✎PROOF:

$$\begin{aligned}
(\!(x_n)\!) \text{ is } \textit{Cauchy} &\implies \text{for every } \varepsilon \in \mathbb{R}^+,\ \exists N \in \mathbb{Z} \quad \text{such that} \quad \forall n, m > N,\ \mathsf{d}(x_n, x_m) < \varepsilon && \text{(Definition B.8 page 137)}\\
&\implies \exists N \in \mathbb{Z} \quad \text{such that} \quad \forall n, m > N,\ \mathsf{d}(x_n, x_m) < 1 && (\text{arbitrarily choose } \varepsilon \triangleq 1)\\
&\implies \exists N \in \mathbb{Z} \quad \text{such that} \quad \forall n, m \in \mathbb{Z},\ \mathsf{d}(x_n, x_{m+1}) < \max\left\{\{1\} \cup \left\{\mathsf{d}(x_p, x_q)\ |p, q \ngtr N\right\}\right\}\\
&\implies (\!(x_n)\!) \text{ is } \textit{bounded} && \text{(by Definition B.3 page 133)}
\end{aligned}$$

✏

Proposition B.2. [21] *Let* $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ *be a* SEQUENCE *in a* DISTANCE SPACE (X, d). *Let* $\mathsf{f} \in \mathbb{Z}^{\mathbb{Z}}$ *(Definition 1.6 page 6) be a* STRICTLY MONOTONE *function such that* $\mathsf{f}(n) < \mathsf{f}(n+1)$.

PRP

$$\underbrace{(\!(x_n)\!)_{n\in\mathbb{Z}} \text{ is CAUCHY}}_{\textit{sequence is } \text{CAUCHY}} \implies \underbrace{\left(\!\left(x_{\mathsf{f}(n)}\right)\!\right)_{n\in\mathbb{Z}} \text{ is CAUCHY}}_{\textit{subsequence is also } \text{CAUCHY}}$$

[17] in *metric space*: Rosenlicht (1968) page 45, Giles (1987) page 37 ⟨3.2 Definition⟩, Khamsi and Kirk (2001) page 13 ⟨Definition 2.1⟩ "→" symbol: Leathem (1905) page 13 ⟨section III.11⟩

[18] in *metric space*: Apostol (1975) page 73 ⟨4.7⟩, Rosenlicht (1968) page 51

[19] in *metric space*: Rosenlicht (1968) page 52

[20] in *metric space*: Giles (1987) page 49 ⟨Theorem 3.30⟩

[21] in *metric space*: Rosenlicht (1968) page 52

✎PROOF:

$$\begin{aligned}
&(\!(x_n)\!)_{n\in\mathbb{Z}} \text{ is } \textit{Cauchy} \\
\implies &\text{for any given } \varepsilon > 0,\ \exists N \quad \text{such that} \quad \forall n,m > N,\ \mathsf{d}(x_n, x_m) < \varepsilon && \text{by Definition B.8 page 137} \\
\implies &\text{for any given } \varepsilon > 0,\ \exists N' \quad \text{such that} \quad \forall \mathsf{f}(n), \mathsf{f}(m) > N',\ \mathsf{d}(x_{\mathsf{f}(n)}, x_{\mathsf{f}(m)}) < \varepsilon \\
\implies &(\!(x_{\mathsf{f}(n)})\!)_{n\in\mathbb{Z}} \text{ is } \textit{Cauchy} && \text{by Definition B.8 page 137}
\end{aligned}$$

✏

Theorem B.4. [22] *Let (X, d) be a* DISTANCE SPACE. *Let A^- be the* CLOSURE *(Definition D.18 page 180) of a A in a* TOPOLOGICAL SPACE INDUCED BY (X, d).

THM

$$\left\{\begin{array}{ll} 1. & \text{LIMITS } \textit{are} \text{ UNIQUE } \textit{in } (X,\mathsf{d}) \quad \textit{(Definition B.7 page 137)} \quad \textit{and} \\ 2. & (A,\mathsf{d}) \textit{ is} \text{ COMPLETE } \textit{in } (X,\mathsf{d}) \quad \textit{(Definition B.9 page 137)} \end{array}\right\} \implies \underbrace{A \textit{ is} \text{ CLOSED } \textit{in } (X,\mathsf{d})}_{A = A^-}$$

✎PROOF:

1. Proof that $A \subseteq A^-$: by Lemma D.5 page 180
2. Proof that $A^- \subseteq A$ (proof that $x \in A^- \implies x \in A$):
   (a) Let x be a point in A^- ($x \in A^-$).
   (b) Define a *sequence* of open balls $\left(\!\left(\mathsf{B}\left(x, \frac{1}{1}\right), \mathsf{B}\left(x, \frac{1}{2}\right), \mathsf{B}\left(x, \frac{1}{3}\right), \ldots\right)\!\right)$.
   (c) Define a *sequence* of points $(\!(x_1, x_2, x_3, \ldots)\!)$ such that $x_n \in \mathsf{B}\left(x_n, \frac{1}{n}\right) \cap A$.
   (d) Then $(\!(x_n)\!)$ is *convergent* in X with limit x by Definition B.7 page 137
   (e) and $(\!(x_n)\!)$ is *Cauchy* in A by Definition B.8 page 137.
   (f) By the hypothesis 2, $(\!(x_n)\!)$ is therefore also *convergent* in A.
   Let this limit be y. Note that $y \in A$.
   (g) By hypothesis 1, limits are *unique*, so $y = x$.
   (h) Because $y \in A$ (item (2f)) and $y = x$ (item (2g)), so $x \in A$.
   (i) Therefore, $x \in A^- \implies x \in A$ and $A^- \subseteq A$.

✏

Proposition B.3. [23] *Let $(\!(x_n)\!)_{n\in\mathbb{Z}}$ be a sequence in a* DISTANCE SPACE (X, d). *Let $\mathsf{f} : \mathbb{Z} \to \mathbb{Z}$ be a strictly increasing function such that $\mathsf{f}(n) < \mathsf{f}(n+1)$.*

PRP

$$\underbrace{(\!(x_n)\!)_{n\in\mathbb{Z}} \to x}_{\textit{sequence converges to limit } x} \implies \underbrace{(\!(x_{\mathsf{f}(n)})\!)_{n\in\mathbb{Z}} \to x}_{\textit{subsequence converges to the same limit } x}$$

✎PROOF:

$$\begin{aligned}
(\!(x_n)\!)_{n\in\mathbb{Z}} \to x &\implies \forall\varepsilon > 0,\ \exists N \quad \text{such that} \quad \forall n > N,\ \mathsf{d}(x_n, x) < \varepsilon && \text{by Theorem C.2 page 151} \\
&\implies \forall\varepsilon > 0,\ \exists \mathsf{f}(N) \quad \text{such that} \quad \forall \mathsf{f}(n) > \mathsf{f}(N),\ \mathsf{d}(x_{\mathsf{f}(n)}, x) < \varepsilon \\
&\implies (\!(x_{\mathsf{f}(n)})\!)_{n\in\mathbb{Z}} \to x && \text{by Theorem C.2 page 151}
\end{aligned}$$

✏

[22] in *metric space*: Kubrusly (2001) page 128 ⟨Theorem 3.40⟩, Haaser and Sullivan (1991) page 75 ⟨6·10, 6·11 Propositions⟩, Bryant (1985) page 40 ⟨Theorem 3.6, 3.7⟩, Sutherland (1975) pages 123–124
[23] in *metric space*: Rosenlicht (1968) page 46

Theorem B.5 (Cantor intersection theorem). [24] *Let* (X, d) *be a* DISTANCE SPACE *(Definition B.1 page 133),* $(\!(A_n)\!)_{n\in\mathbb{Z}}$ *a* SEQUENCE *with each* $A_n \in 2^X$, *and* $|A|$ *the number of elements in* A.

THM
$$\left\{\begin{array}{llll} 1. & (X,\mathsf{d}) \text{ is COMPLETE} & & \text{(Definition B.9 page 137) and} \\ 2. & A_n \text{ is CLOSED} & \forall n\in\mathbb{N} & \text{(Definition D.16 page 179) and} \\ 3. & \operatorname{diam} A_n \ge \operatorname{diam} A_{n+1} & \forall n\in\mathbb{N} & \text{(Definition B.2 page 133) and} \\ 4. & \operatorname{diam} (\!(A_n)\!)_{n\in\mathbb{Z}} \to 0 & & \text{(Definition B.7 page 137)} \end{array}\right\} \implies \left\{\left|\bigcap_{n\in\mathbb{N}} A_n\right| = 1\right\}$$

✎PROOF:

1. Proof that $\left|\bigcap_{n\in\mathbb{Z}} A_n\right| < 2$:
   (a) Let $A \triangleq \cap A_n$.
   (b) $x \ne y$ and $\{x, y\} \in A \implies \mathsf{d}(x, y) > 0$ and $\{x, y\} \subseteq A_n \forall n$
   (c) $\exists n$ such that $\operatorname{diam} A_n < \mathsf{d}(x, y)$ by left hypothesis 4
   (d) $\implies \exists n$ such that $\sup\{\mathsf{d}(x, y) \,|\, x, y \in A_n\} < \mathsf{d}(x, y)$
   (e) This is a contradiction, so $\{x, y\} \notin A$ and $\left|\bigcap A_n\right| < 2$.
2. Proof that $|\cap A_n| \ge 1$:
   (a) Let $x_n \in A_n$ and $x_m \in A_m$
   (b) $\forall \varepsilon, \exists N \in \mathbb{N}$ such that $A_N < \varepsilon$
   (c) $\forall m, n > N,\ x_n \in A_n \subseteq A_N$ and $x_m \in A_m \subseteq A_N$
   (d) $\mathsf{d}(x_n, x_m) \le \operatorname{diam} A_N < \varepsilon \implies \{x_n\}$ is a Cauchy sequence
   (e) Because $\{x_n\}$ is complete, $x_n \to x$.
   (f) $\implies x \in (A_n)^- = A_n$
   (g) $\implies |A_n| \ge 1$

✏

Definition B.10. [25] *Let* (X, d) *be a* DISTANCE SPACE. *Let* C *be the set of all* CONVERGENT *sequences in* (X, d).

THM
The DISTANCE FUNCTION d *is* ***continuous*** *in* (X, d) *if*
$$(\!(x_n)\!), (\!(y_n)\!) \in C \implies \lim_{n\to\infty} (\!(\mathsf{d}(x_n, y_n))\!) = \mathsf{d}\left(\lim_{n\to\infty} (\!(x_n)\!), \lim_{n\to\infty} (\!(y_n)\!)\right).$$
A DISTANCE FUNCTION *is* ***discontinuous*** *if it is not* CONTINUOUS.

Remark B.1. Rather than defining *continuity* of a *distance function* in terms of the *sequential characterization of continuity* as in Definition B.10 (previous), we could define continuity using an *inverse image characterization of continuity*" (Definition B.6 page 136). Assuming an equivalent *topological space* is used for both characterizations, the two characterizations are equivalent (Theorem D.20 page 182). In fact, one could construct an equivalence such as the following:

REM
$$\left\{\begin{array}{l} \mathsf{d} \text{ is } \textit{continuous} \text{ in } \mathbb{R}^{X^2} \\ \text{(Definition D.19 page 180)} \\ (\textit{inverse image characterization of continuity}) \end{array}\right\} \iff \left\{\begin{array}{l} (\!(x_n)\!), (\!(y_n)\!) \in C \implies \\ \lim\limits_{n\to\infty} (\!(\mathsf{d}(x_n, y_n))\!) = \mathsf{d}\left(\lim\limits_{n\to\infty} (\!(x_n)\!), \lim\limits_{n\to\infty} (\!(y_n)\!)\right) \\ \text{(Definition D.20 page 181)} \\ (\textit{sequential characterization of continuity}) \end{array}\right\}$$

Note that just as $(\!(x_n)\!)$ is a sequence in X, so the ordered pair $\left((\!(x_n)\!), (\!(y_n)\!)\right)$ is a sequence in X^2. The remainder follows from Theorem D.20 (page 182). However, use of the *inverse image characterization* is somewhat troublesome because we would need a topology on X^2, and we don't immediately

[24] in *metric space*: Davis (2005), page 28, Hausdorff (1937), page 150
[25] Blumenthal (1953) page 9 ⟨DEFINITION 6.3⟩

have one defined and ready to use. In fact, we don't even immediately have a distance space on X^2 defined or even open balls in such a distance space. The result is, for the scope of this paper, it is arguably not worthwhile constructing the extra structure, but rather instead this paper uses the *sequential characterization* as a definition (as in Definition B.10).

B.5 Examples

Similar distance functions and several of the observations for the examples in this section can be found in Blumenthal (1953) pages 8–13.

In a *metric space*, all *open balls* are *open*, the *open balls* form a *base* for a *topology*, the limits of *convergent* sequences are *unique*, and the *metric function* is *continuous*. In the *distance space* of the next example, none of these properties hold.

Example B.1. [26] Let (x, y) be an *ordered pair* in $\mathbb{R}^2$. Let $(a : b)$ be an *open interval* and $(a : b]$ a *half-open interval* in $\mathbb{R}$. Let $|x|$ be the *absolute value* of $x \in \mathbb{R}$. The function $\mathsf{d}(x, y) \in \mathbb{R}^{\mathbb{R}\times\mathbb{R}}$ such that

$$\mathsf{d}(x, y) \triangleq \left\{\begin{array}{lll} y & \forall (x, y) \in \{4\} \times (0 : 2] & \textit{(vertical half-open interval)} \\ x & \forall (x, y) \in (0 : 2] \times \{4\} & \textit{(horizontal half-open interval)} \\ |x - y| & \text{otherwise} & \textit{(Euclidean)} \end{array}\right\} \text{ is a } \textit{distance} \text{ on } \mathbb{R}.$$

Note some characteristics of the *distance space* $(\mathbb{R}, \mathsf{d})$:

1. $(\mathbb{R}, \mathsf{d})$ is not a *metric space* because d does not satisfy the *triangle inequality*:
   $$\mathsf{d}(0, 4) \triangleq |0 - 4| = 4 \not\leq 2 = |0 - 1| + 1 \triangleq \mathsf{d}(0, 1) + \mathsf{d}(1, 4)$$

2. Not every *open ball* in $(\mathbb{R}, \mathsf{d})$ is *open*.
   For example, the *open ball* $\mathsf{B}(3, 2)$ is *not open* because $4 \in \mathsf{B}(3, 2)$ *but* for all $0 < \varepsilon < 1$
   $$\mathsf{B}(4, \varepsilon) = (4 - \varepsilon : 4 + \varepsilon) \cup (0 : \varepsilon) \not\subseteq (1 : 5) = \mathsf{B}(3, 2)$$

3. The *open balls* of $(\mathbb{R}, \mathsf{d})$ do not form a *base* for a *topology* on $\mathbb{R}$.
   This follows directly from item (2) and Theorem B.3 (page 136).

4. In the *distance space* $(\mathbb{R}, \mathsf{d})$, limits are *not unique*;
   For example, the sequence $(\!(1/n)\!)_1^\infty$ converges both to the limit 0 and the limit 4 in $(\mathbb{R}, \mathsf{d})$:
   $$\begin{array}{llllll} \lim\limits_{n\to\infty} \mathsf{d}(x_n, 0) & \triangleq \lim\limits_{n\to\infty} \mathsf{d}(1/n, 0) & \triangleq \lim\limits_{n\to\infty} |1/n - 0| & = 0 & \Longrightarrow & (\!(1/n)\!) \to 0 \\ \lim\limits_{n\to\infty} \mathsf{d}(x_n, 4) & \triangleq \lim\limits_{n\to\infty} \mathsf{d}(1/n, 4) & \triangleq \lim\limits_{n\to\infty} (\!(1/n)\!) & = 0 & \Longrightarrow & (\!(1/n)\!) \to 4 \end{array}$$

5. The *topological space* $(X, \boldsymbol{T})$ *induced by* $(\mathbb{R}, \mathsf{d})$ also yields limits of 0 and 4 for the sequence $(\!(1/n)\!)_1^\infty$, just as it does in item (4). This is largely due to the fact that, for small ε, the open balls $\mathsf{B}(0, \varepsilon)$ and $\mathsf{B}(4, \varepsilon)$ are *open*.
   $$\begin{array}{llll} \mathsf{B}(0, \varepsilon) \text{ is } \textit{open} & \Longrightarrow & \text{for each } U \in \boldsymbol{T} \text{ that contains } 0, \exists N \in \mathbb{N} \text{ such that } 1/n \in U \quad \forall n > N \\ & \Longleftrightarrow & (\!(1/n)\!) \to 0 \quad \text{by definition of } \textit{convergence} \text{ (Definition D.20 page 181)} \\ \mathsf{B}(4, \varepsilon) \text{ is } \textit{open} & \Longrightarrow & \text{for each } U \in \boldsymbol{T} \text{ that contains } 4, \exists N \in \mathbb{N} \text{ such that } 1/n \in U \quad \forall n > N \\ & \Longleftrightarrow & (\!(1/n)\!) \to 4 \quad \text{by definition of } \textit{convergence} \text{ (Definition D.20 page 181)} \end{array}$$

[26] A similar distance function d and item (4) page 140 can in essence be found in Blumenthal (1953) page 8. Definitions for Example B.1: (x, y): Definition 1.3 (page 5); $(a : b)$ and $(a : b]$: Definition 1.23 (page 10); $|x|$: Definition 1.24 (page 10); $\mathbb{R}^{\mathbb{R}\times\mathbb{R}}$: Definition 1.6 (page 6); *distance*: Definition B.1 (page 133); *open ball*: Definition B.4 (page 134); *open*: Definition B.5 (page 134); *base*: Definition D.17 (page 179); *topology*: Definition D.16 (page 179); *open set*: Definition B.5 (page 134); *topological space induced by* $(\mathbb{R}, \mathsf{d})$: Definition B.6 (page 136); *discontinuous*: Definition B.10 (page 139);

6. The distance function d is *discontinuous* (Definition B.10 page 139):
$$\lim_{n\to\infty} (\!(\mathsf{d}(1-{}^1\!/\!_n, 4-{}^1\!/\!_n))\!) = \lim_{n\to\infty} (\!(|(1-{}^1\!/\!_n)-(4-{}^1\!/\!_n)|)\!) = |1-4| = 3 \neq 4 = \mathsf{d}(0,4)$$
$$= \mathsf{d}\left(\lim_{n\to\infty}(\!(1-{}^1\!/\!_n)\!), \lim_{n\to\infty}(\!(4-{}^1\!/\!_n)\!)\right)$$

In a *metric space*, all *convergent* sequences are also *Cauchy*. However, this is not the case for all *distance spaces*, as demonstrated next:

Example B.2. [27] The function $\mathsf{d}(x,y) \in \mathbb{R}^{\mathbb{R}\times\mathbb{R}}$ such that
$$\mathsf{d}(x,y) \triangleq \left\{\begin{array}{lll} |x-y| & \text{for } x=0 \text{ or } y=0 \text{ or } x=y & \text{(Euclidean)} \\ 1 & \text{otherwise} & \text{(discrete)} \end{array}\right\} \quad \text{is a } \textit{distance} \text{ on } \mathbb{R}.$$
Note some characteristics of the *distance space* $(\mathbb{R},\mathsf{d})$:

1. (X,d) is not a *metric space* because the *triangle inequality* does not hold:
$$\mathsf{d}\left(\tfrac{1}{4},\tfrac{1}{2}\right) = 1 \not\leq \tfrac{3}{4} = \left|\tfrac{1}{4}-0\right| + \left|0-\tfrac{1}{2}\right| = \mathsf{d}\left(\tfrac{1}{4},0\right) + \mathsf{d}\left(0,\tfrac{1}{2}\right)$$

2. The *open ball* $\mathsf{B}\left(\tfrac{1}{4},\tfrac{1}{2}\right)$ is *not open* because for any $\varepsilon \in \mathbb{R}^+$, no matter how small,
$$\mathsf{B}(0,\varepsilon) = (-\varepsilon : +\varepsilon) \not\subseteq \left\{0, \tfrac{1}{4}\right\} = \left\{x \in X | \mathsf{d}\left(\tfrac{1}{4},x\right) < \tfrac{1}{2}\right\} \triangleq \mathsf{B}\left(\tfrac{1}{4},\tfrac{1}{2}\right)$$

3. Even though not all the *open balls* are *open*, it is still possible to have an *open set* in (X,d). For example, the set $U \triangleq \{1, 2\}$ is *open*:
$$\begin{array}{lllllllll} \mathsf{B}(1,1) & \triangleq & \{x\in X|\mathsf{d}(1,x)<1\} & = & \{1\} & \subseteq & \{1,2\} & \triangleq & U \\ \mathsf{B}(2,1) & \triangleq & \{x\in X|\mathsf{d}(2,x)<1\} & = & \{2\} & \subseteq & \{1,2\} & \triangleq & U \end{array}$$

4. By item (2) and Theorem B.3 (page 136), the *open balls* of $(\mathbb{R},\mathsf{d})$ do not form a *base* for a *topology* on $\mathbb{R}$.

5. Even though the open balls in $(\mathbb{R},\mathsf{d})$ do not induce a topology on X, it is still possible to find a set of *open sets* in (X,d) that *is* a topology. For example, the set $\{\varnothing, \{1,2\}, \mathbb{R}\}$ is a topology on $\mathbb{R}$.

6. In $(\mathbb{R},\mathsf{d})$, limits of *convergent* sequences are *unique*:
$$(\!(x_n)\!) \to x \quad\Longrightarrow\quad \lim_{n\to\infty}\mathsf{d}(x_n,x) = \left\{\begin{array}{rcll} \lim|x_n-0| & = & 0 & \text{for } x=0 & \text{OR} \\ |x-x| & = & 0 & \text{for constant } (\!(x_n)\!) \text{ for } n>N & \text{OR} \\ 1 & \neq & 0 & \text{otherwise} & \end{array}\right\}$$
which says that there are only two ways for a sequence to converge: either $x=0$ or the sequence eventually becomes constant (or both). Any other sequence will *diverge*. Therefore we can say the following:

   (a) If $x=0$ and the sequence is not constant, then the limit is *unique* and 0.

   (b) If $x=0$ and the sequence is constant, then the limit is *unique* and 0.

   (c) If $x\neq 0$ and the sequence is constant, then the limit is *unique* and x.

   (d) If $x\neq 0$ and the sequence is not constant, then the sequence diverges and there is no limit.

7. In $(\mathbb{R},\mathsf{d})$, a *convergent* sequence is not necessarily *Cauchy*. For example,

   (a) the sequence $(\!({}^1\!/\!_n)\!)_{n\in\mathbb{N}}$ is *convergent* with limit 0: $\lim_{n\to\infty}\mathsf{d}({}^1\!/\!_n,0) = \lim_{n\to\infty}{}^1\!/\!_n = 0$

[27] The distance function d and item (7) page 141 can in essence be found in Blumenthal (1953) page 9

(b) However, even though $(\!(1/n)\!)$ is *convergent*, it is *not Cauchy*: $\lim_{n,m\to\infty} \mathsf{d}(1/n, 1/m) = 1 \neq 0$

8. The *distance function* d is *discontinuous* in (X, d):
$$\lim_{n\to\infty} (\!(\mathsf{d}(1/n, 2 - 1/n))\!) = 1 \neq 2 = \mathsf{d}(0,2) = \mathsf{d}\left(\lim_{n\to\infty}(\!(1/n)\!), \lim_{n\to\infty}(\!(2-1/n)\!)\right).$$

Example B.3. [28] The function $\mathsf{d}(x,y) \in \mathbb{R}^{\mathbb{R}\times\mathbb{R}}$ such that
$$\mathsf{d}(x,y) \triangleq \left\{\begin{array}{lll} 2|x-y| & \forall (x,y) \in \{(0,1), (1,0)\} & \textit{(dilated Euclidean)} \\ |x-y| & \text{otherwise} & \textit{(Euclidean)} \end{array}\right\} \text{ is a } \textit{distance} \text{ on } \mathbb{R}.$$
Note some characteristics of the *distance space* $(\mathbb{R}, \mathsf{d})$:

1. $(\mathbb{R}, \mathsf{d})$ is *not* a *metric space* because d does *not* satisfy the *triangle inequality*:
$$\mathsf{d}(0,1) \triangleq 2|0-1| = 2 \nleq 1 = |0 - 1/2| + |1/2 - 1| \triangleq \mathsf{d}(0, 1/2) + \mathsf{d}(1/2, 1)$$

2. The function d is *discontinuous*:
$$\lim_{n\to\infty} (\!(\mathsf{d}(1-1/n, 1/n))\!) \triangleq \lim_{n\to\infty} (\!(|1 - 1/n - 1/n|)\!) = 1 \neq 2 = 2|0-1| \triangleq \mathsf{d}(0,1) = \mathsf{d}\left(\lim_{n\to\infty}(\!(1-1/n)\!), \lim_{n\to\infty}(\!(1/n)\!)\right).$$

3. In (X, d), *open balls* are *open*:

   (a) $\mathsf{p}(x,y) \triangleq |x-y|$ is a *metric* and thus all open balls in that do not contain both 0 and 1 are *open*.

   (b) By Example D.3 (page 167), $\mathsf{q}(x,y) \triangleq 2|x-y|$ is also a *metric* and thus all open balls containing 0 and 1 only are *open*.

   (c) The only question remaining is with regards to open balls that contain 0, 1 and some other element(s) in $\mathbb{R}$. But even in this case, open balls are still open. For example:
   $$\mathsf{B}(-1,2) = (-1 : 2) = (-1 : 1) \cup (1 : 2)$$
   Note that both $(-1 : 1)$ and $(1 : 2)$ are *open*, and thus by Theorem B.2 (page 134), $\mathsf{B}(-1,2)$ is *open* as well.

4. By item (3) and Theorem B.3 (page 136), the *open balls* of $(\mathbb{R}, \mathsf{d})$ *do* form a *base* for a *topology* on $\mathbb{R}$.

5. In (X, d), the limits of *convergent* sequences are *unique*. This is demonstrated in Example C.3 (page 154) using additional structure developed in APPENDIX C.

6. In (X, d), *convergent* sequences are *Cauchy*. This is also demonstrated in Example C.3 (page 154).

The *distance functions* in Example B.1 (page 140)–Example B.3 (page 142) were all *discontinuous*. In the absence of the *triangle inequality* and in light of these examples, one might try replacing the *triangle inequality* with the weaker requirement of *continuity*. However, as demonstrated by the next example, this also leads to an arguably disastrous result.

Example B.4. [29] The function $\mathsf{d} \in \mathbb{R}^{\mathbb{R}\times\mathbb{R}}$ such that $\mathsf{d}(x,y) \triangleq (x-y)^2$ is a *distance* on $\mathbb{R}$.
Note some characteristics of the *distance space* $(\mathbb{R}, \mathsf{d})$:

1. $(\mathbb{R}, \mathsf{d})$ is *not* a *metric space* because the *triangle inequality* does not hold:
$$\mathsf{d}(0,2) \triangleq (0-2)^2 = 4 \nleq 2 = (0-1)^2 + (1-2)^2 \triangleq \mathsf{d}(0,1) + \mathsf{d}(1,2)$$

2. The *distance function* d is *continuous* in (X, d). This is demonstrated in the more general setting of APPENDIX C in Example C.4 (page 155).

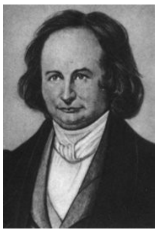
[28] The distance function d and item (2) page 142 can in essence be found in Blumenthal (1953) page 9
[29] Blumenthal (1953) pages 12–13, Laos (1998) pages 118–119

3. Calculating the length of curves in (X, d) leads to a paradox:[30]
   (a) Partition $[0:1]$ into 2^N consecutive line segments connected at the points $\left(\!\left(0, \frac{1}{2^N}, \frac{2}{2^N}, \frac{3}{2^N}, \ldots, \frac{2^{N-1}1}{2^N}, 1\right)\!\right)$
   (b) Then the distance, as measured by d, between any two consecutive points is $\mathsf{d}(p_n, p_{n+1}) \triangleq (p_n - p_{n+1})^2 = \left(\frac{1}{2^N}\right)^2 = \frac{1}{2^{2N}}$
   (c) But this leads to the paradox that the total length of $[0:1]$ is 0:
   $$\lim_{N\to\infty}\sum_{n=0}^{2^N-1}\frac{1}{2^{2N}} = \lim_{N\to\infty}\frac{2^N}{2^{2N}} = \lim_{N\to\infty}\frac{1}{2^N} = 0$$

[30] This is the method of "inscribed polygons" for calculating the length of a curve and goes back to Archimedes: Brunschwig et al. (2003) page 26, Walmsley (1920), page 200 ⟨§158⟩,

APPENDIX C

POWER DISTANCE SPACES



❝*Man muss immer generalisieren.*❞
❝*One should always generalize.*❞
attributed to Karl Gustav Jakob Jacobi (1804–1851), mathematician [1]

C.1 Definitions

This paper introduces a new relation called the *power triangle inequality* (Definition C.2 page 146). It is a generalization of other common relations, including the *triangle inequality* (Definition C.3 page 146). The *power triangle inequality* is defined in terms of a function herein called the *power triangle function* (next definition). This function is a special case of the *power mean* with $N = 2$ and $\lambda_1 = \lambda_2 = \frac{1}{2}$ (Definition D.15 page 174). *Power means* have the attractive properties of being *continuous* and *strictly monontone* with respect to a free parameter $p \in \mathbb{R}^*$ (Theorem D.14 page 175). This fact is inherited and exploited by the *power triangle inequality* (Corollary C.1 page 146).

Definition C.1. *Let* (X, d) *be a* DISTANCE SPACE *(Definition B.1 page 133). Let* $\mathbb{R}^+$ *be the set of all* POSITIVE REAL NUMBERS *and* $\mathbb{R}^*$ *be the set of* EXTENDED REAL NUMBERS *(Definition 1.2 page 5).*

DEF *The* ***power triangle function*** τ *on* (X, d) *is defined as*

$$\tau(p, \sigma; x, y, z; \mathsf{d}) \triangleq 2\sigma\left[\frac{1}{2}\mathsf{d}^p(x, z) + \frac{1}{2}\mathsf{d}^p(z, y)\right]^{\frac{1}{p}} \quad \forall (p,\sigma)\in\mathbb{R}^*\times\mathbb{R}^{\vdash}, \quad x,y,z\in X$$

Remark C.1. [2] In the field of *probabilistic metric spaces*, a function called he *triangle function* was introduced by Sherstnev in 1962. However, the *power triangle function* as defined in this present paper is *not* a special case of (is not compatible with) the *triangle function* of Sherstnev. Another definition of *triangle function* has been offered by Bessenyei in 2014 with special cases of $\Phi(u, v) \triangleq$

[1] quote: Davis and Hersh (1999) page 134
image: `http://en.wikipedia.org/wiki/Carl_Gustav_Jacobi`

[2] Sherstnev (1962), page 4, Schweizer and Sklar (1983) page 9 ⟨(1.6.1)–(1.6.4)⟩, Bessenyei and Pales (2014) page 2

$c(u+v)$ and $\Phi(u,v) \triangleq (u^p + v^p)^{\frac{1}{p}}$, which *are* similar to the definition of *power triangle function* offered in this present paper.

Definition C.2. *Let* (X, d) *be a* DISTANCE SPACE. *Let* 2^{XXX} *be the set of all trinomial* RELATIONS *(Definition 1.5 page 6) on* X.

DEF

A relation $\circledtriangle(p, \sigma; \mathsf{d})$ *in* 2^{XXX} *is a **power triangle inequality** on* (X, d) *if*

$$\circledtriangle(p,\sigma;\mathsf{d}) \triangleq \left\{(x,y,z) \in X^3 | \mathsf{d}(x,y) \le \tau(p,\sigma;x,y,z;\mathsf{d})\right\} \qquad \text{for some } (p,\sigma) \in \mathbb{R}^* \times \mathbb{R}^+.$$

The tupple $(X, \mathsf{d}, p, \sigma)$ *is a **power distance space** and* d *a **power distance** or **power distance function** if* (X, d) *is a* DISTANCE SPACE *in which the* TRIANGLE RELATION $\circledtriangle(p, \sigma; \mathsf{d})$ *holds.*

The *power triangle function* can be used to define some standard inequalities (next definition). See Corollary C.2 (page 147) for some justification of the definitions.

Definition C.3. [3] *Let* $\circledtriangle(p, \sigma; \mathsf{d})$ *be a* POWER TRIANGLE INEQUALITY *on a* DISTANCE SPACE (X, d).

DEF

1. $\circledtriangle(\infty, \sigma/2; \mathsf{d})$ *is the **σ-inframetric inequality***
2. $\circledtriangle(\infty, \frac{1}{2}; \mathsf{d})$ *is the **inframetric inequality***
3. $\circledtriangle(2, \sqrt{2}/2; \mathsf{d})$ *is the **quadratic inequality***
4. $\circledtriangle(1, \sigma; \mathsf{d})$ *is the **relaxed triangle inequality***
5. $\circledtriangle(1, 1; \mathsf{d})$ *is the **triangle inequality***
6. $\circledtriangle(1/2, 2; \mathsf{d})$ *is the **square mean root inequality***
7. $\circledtriangle(0, \frac{1}{2}; \mathsf{d})$ *is the **geometric inequality***
8. $\circledtriangle(-1, \frac{1}{4}; \mathsf{d})$ *is the **harmonic inequality***
9. $\circledtriangle(-\infty, \frac{1}{2}; \mathsf{d})$ *is the **minimal inequality***

Definition C.4. [4] *Let* (X, d) *be a* DISTANCE SPACE *(Definition B.1 page 133).*

DEF

1.	(X, d) *is a*	***metric space***	*if the* TRIANGLE INEQUALITY	*holds in* X.
2.	(X, d) *is a*	***near metric space***	*if the* RELAXED TRIANGLE INEQUALITY	*holds in* X.
3.	(X, d) *is an*	***inframetric space***	*if the* INFRAMETRIC INEQUALITY	*holds in* X.
4.	(X, d) *is a*	***σ-inframetric space***	*if the* σ-INFRAMETRIC INEQUALITY	*holds in* X.

C.2 Properties

C.2.1 Relationships of the power triangle function

Corollary C.1. *Let* $\tau(p, \sigma; x, y, z; \mathsf{d})$ *be the* POWER TRIANGLE FUNCTION *(Definition C.1 page 145) in the* DISTANCE SPACE *(Definition B.1 page 133)* (X, d). *Let* $(\mathbb{R}, |\cdot|, \le)$ *be the* ORDERED METRIC SPACE *with the usual ordering relation* $\le$ *and usual metric* $|\cdot|$ *on* $\mathbb{R}$.

COR

The function $\tau(p, \sigma; x, y, z; \mathsf{d})$ *is* CONTINUOUS *and* STRICTLY MONOTONE *in* $(\mathbb{R}, |\cdot|, \le)$ *with respect to both the variables* p *and* σ.

---

[3] Bessenyei and Pales (2014) page 2, Czerwik (1993) page 5 ⟨*b-metric*; (1),(2),(5)⟩, Fagin et al. (2003b), Fagin et al. (2003a) ⟨Definition 4.2 (Relaxed metrics)⟩, Xia (2009) page 453 ⟨Definition 2.1⟩, Heinonen (2001) page 109 ⟨14.1 Quasimetric spaces.⟩, Kirk and Shahzad (2014) page 113 ⟨Definition 12.1⟩, Deza and Deza (2014) page 7, Hoehn and Niven (1985) page 151, Gibbons et al. (1977) page 51 ⟨*square-mean-root* (*SMR*) (2.4.1)⟩, Euclid (circa 300BC) ⟨triangle inequality—Book I Proposition 20⟩

[4] **metric space**: Dieudonné (1969), page 28, Copson (1968), page 21, Hausdorff (1937) page 109, Fréchet (1928), Fréchet (1906) page 30 **near metric space**: Czerwik (1993) page 5 ⟨*b-metric*; (1),(2),(5)⟩, Fagin et al. (2003b), Fagin et al. (2003a) ⟨Definition 4.2 (Relaxed metrics)⟩, Xia (2009) page 453 ⟨Definition 2.1⟩, Heinonen (2001) page 109 ⟨14.1 Quasimetric spaces.⟩, Kirk and Shahzad (2014) page 113 ⟨Definition 12.1⟩, Deza and Deza (2014) page 7

✎PROOF:

1. Proof that $\tau(p,\sigma;x,y,z;\mathsf{d})$ is *continuous* and *strictly monotone* with respect to p: This follows directly from Theorem D.14 (page 175).
2. Proof that $\tau(p,\sigma;x,y,z;\mathsf{d})$ is *continuous* and *strictly monotone* with respect to σ:

$$\begin{aligned}
\tau(p,\sigma;x,y,z;\mathsf{d}) &\triangleq 2\sigma\underbrace{\left[\frac{1}{2}\mathsf{d}^p(x,z)+\frac{1}{2}\mathsf{d}^p(z,y)\right]^{\frac{1}{p}}}_{\mathsf{f}(p,x,y,z)} && \text{by definition of } \tau \text{ (Definition C.1 page 145)}\\
&= 2\sigma\mathsf{f}(p,x,y,z) && \text{where f is defined as above}\\
&\implies \tau \text{ is \textit{affine} with respect to } \sigma\\
&\implies \tau \text{ is \textit{continuous} and \textit{strictly monotone} with respect to } \sigma:
\end{aligned}$$

✏

Corollary C.2. *Let* $\tau(p,\sigma;x,y,z;\mathsf{d})$ *be the* POWER TRIANGLE FUNCTION *in the* DISTANCE SPACE *(Definition B.1 page 133)* (X,d).

COR

$$\tau(p,\sigma;x,y,z;\mathsf{d}) = \begin{cases}
2\sigma\max\{\mathsf{d}(x,z),\ \mathsf{d}(z,y)\} & \text{for } p=\infty, & \text{(MAXIMUM, \textit{corresponds to} INFRAMETRIC SPACE)}\\
2\sigma\left[\tfrac{1}{2}\mathsf{d}^2(x,z)+\tfrac{1}{2}\mathsf{d}^2(z,y)\right]^{\frac{1}{2}} & \text{for } p=2, & \text{(QUADRATIC MEAN)}\\
\sigma[\mathsf{d}(x,z)+\mathsf{d}(z,y)] & \text{for } p=1, & \text{(ARITHMETIC MEAN, \textit{corresponds to} NEAR METRIC SPACE)}\\
2\sigma\sqrt{\mathsf{d}(x,z)}\sqrt{\mathsf{d}(z,y)} & \text{for } p=0 & \text{(GEOMETRIC MEAN)}\\
4\sigma\left[\frac{1}{\mathsf{d}(x,z)}+\frac{1}{\mathsf{d}(z,y)}\right]^{-1} & \text{for } p=-1 & \text{(HARMONIC MEAN)}\\
2\sigma\min\{\mathsf{d}(x,z),\ \mathsf{d}(z,y)\} & \text{for } p=-\infty, & \text{(MINIMUM)}
\end{cases}$$

✎PROOF: These follow directly from Theorem D.14 (page 175). ✏

Corollary C.3. *Let* (X,d) *be a* DISTANCE SPACE.

COR

$$\begin{aligned}
2\sigma\min\{\mathsf{d}(x,z),\ \mathsf{d}(z,y)\} &\le 4\sigma\left[\frac{1}{\mathsf{d}(x,z)}+\frac{1}{\mathsf{d}(z,y)}\right]^{-1} &&\le 2\sigma\sqrt{\mathsf{d}(x,z)}\sqrt{\mathsf{d}(z,y)}\\
&\le \sigma[\mathsf{d}(x,z)+\mathsf{d}(z,y)] &&\le 2\sigma\max\{\mathsf{d}(x,z),\ \mathsf{d}(z,y)\}
\end{aligned}$$

✎PROOF: These follow directly from Corollary D.2 (page 177). ✏

C.2.2 Properties of power distance spaces

The *power triangle inequality* property of a *power distance space* axiomatically endows a metric with an upper bound. Lemma C.1 (next) demonstrates that there is a complementary lower bound somewhat similar in form to the *power triangle inequality* upper bound. In the special case where $2\sigma = 2^{\frac{1}{p}}$, the lower bound helps provide a simple proof of the *continuity* of a large class of *power distance functions* (Theorem C.5 page 152). The inequality $2\sigma \le 2^{\frac{1}{p}}$ is a special relation in this text and appears repeatedly in this appendix; it appears as an inequality in Lemma C.2 (page 150), Corollary C.5 (page 149) and Corollary C.6 (page 150), and as an equality in Lemma C.1 (next) and Theorem C.5 (page 152). It is plotted in Figure C.1 (page 148).

Lemma C.1. [5] *Let* (X,d,p,σ) *be a* POWER TRIANGLE TRIANGLE SPACE *(Definition C.2 page 146). Let* $|\cdot|$ *be the* ABSOLUTE VALUE *function (Definition 1.24 page 10). Let* $\max\{x,y\}$ *be the maximum and* $\min\{x,y\}$ *the*

[5] in *metric space* ($(p,\sigma)=(1,1)$): Dieudonné (1969) page 28, Michel and Herget (1993) page 266, Berberian (1961) page 37 ⟨Theorem II.4.1⟩

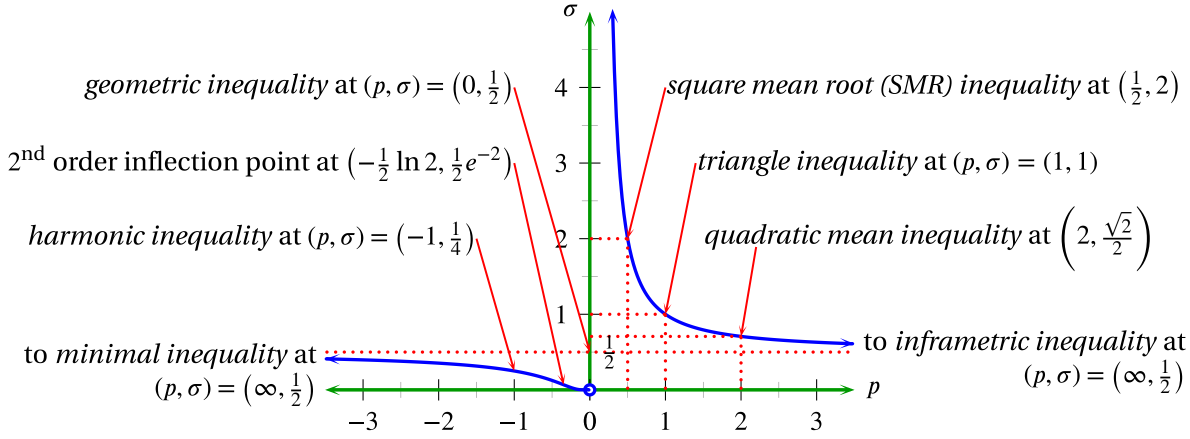


Figure C.1: $\sigma=\frac{1}{2}(2^{\frac{1}{p}})=2^{\frac{1}{p}-1}$ or $p=\frac{\ln 2}{\ln(2\sigma)}$ (see Lemma C.1 page 147, Lemma C.2 page 150, Corollary C.6 page 150, Corollary C.5 page 149, and Theorem C.5 page 152).

minimum of any $x,y\in\mathbb{R}^*$. *Then, for all* $(p,\sigma)\in\mathbb{R}^*\times\mathbb{R}^+$,

LEM

1. $$\mathsf{d}^p(x,y) \geq \max\left\{0,\ \frac{2}{(2\sigma)^p}\mathsf{d}^p(x,z)-\mathsf{d}^p(z,y),\ \frac{2}{(2\sigma)^p}\mathsf{d}^p(y,z)-\mathsf{d}^p(z,x)\right\} \quad \forall x,y,z\in X \quad \text{and}$$
2. $$\mathsf{d}(x,y) \geq |\mathsf{d}(x,z)-\mathsf{d}(z,y)| \quad \textit{if } p\neq 0 \textit{ and } 2\sigma=2^{\frac{1}{p}} \quad \forall x,y,z\in X.$$

✎PROOF:

1. lemma: $\frac{2}{(2\sigma)^p}\mathsf{d}^p(x,z)-\mathsf{d}^p(z,y)\leq \mathsf{d}^p(x,y) \quad \forall (p,\sigma)\in\mathbb{R}^*\times\mathbb{R}^+$: Proof:

$$\begin{aligned}
\frac{2}{(2\sigma)^p}\mathsf{d}^p(x,z)-\mathsf{d}^p(z,y) &\leq \frac{2}{(2\sigma)^p}\left[2\sigma\left[{}^1\!/\!_2\mathsf{d}^p(x,y)+{}^1\!/\!_2\mathsf{d}^p(y,z)\right]^{\frac{1}{p}}\right]^p-\mathsf{d}^p(z,y) && \text{by \textit{power triangle inequality}}\\
&= \frac{2(2\sigma)^p}{(2\sigma)^p}\left[{}^1\!/\!_2\mathsf{d}^p(x,y)+{}^1\!/\!_2\mathsf{d}^p(y,z)\right]-\mathsf{d}^p(z,y)\\
&= \left[\mathsf{d}^p(x,y)+\mathsf{d}^p(y,z)\right]-\mathsf{d}^p(y,z) && \text{by \textit{symmetric} property of } \mathsf{d}\\
&= \mathsf{d}^p(x,y)
\end{aligned}$$

2. Proof for $(p,\sigma)\in\mathbb{R}^*\times\mathbb{R}^+$ case:

$$\begin{aligned}
\mathsf{d}^p(x,y) & & &\geq \frac{2}{(2\sigma)^p}\mathsf{d}^p(x,z)-\mathsf{d}^p(z,y) && \text{by (1) lemma}\\
\mathsf{d}^p(x,y) &= \mathsf{d}^p(y,x) & &\geq \frac{2}{(2\sigma)^p}\mathsf{d}^p(y,z)-\mathsf{d}^p(z,x) && \text{by \textit{commutative} property of } \mathsf{d} \text{ and (1) lemma}\\
\mathsf{d}^p(x,y) & & &\geq 0 && \text{by \textit{non-negative} property of } \mathsf{d} \text{ (Definition B.1 page 133)}
\end{aligned}$$

The rest follows because $\mathsf{g}(x)\triangleq x^{\frac{1}{p}}$ is *strictly monotone* in $\mathbb{R}^{\mathbb{R}}$.

3. Proof for $2\sigma=2^{\frac{1}{p}}$ case:

$$\begin{aligned}
\mathsf{d}(x,y) &\geq \max\left\{0,\ \frac{2}{(2\sigma)^p}\mathsf{d}^p(x,z)-\mathsf{d}^p(z,y),\ \frac{2}{(2\sigma)^p}\mathsf{d}^p(y,z)-\mathsf{d}^p(z,x)\right\}^{\frac{1}{p}} && \text{by item (2) (page 148)}\\
&= \max\{0,\ \mathsf{d}(x,z)-\mathsf{d}(z,y),\ \mathsf{d}(y,z)-\mathsf{d}(z,x)\} && \text{by } 2\sigma=2^{\frac{1}{p}} \text{ hypothesis} \iff \frac{2}{(2\sigma)^p}=1\\
&= \max\{0,\ (\mathsf{d}(x,z)-\mathsf{d}(z,y)),\ -(\mathsf{d}(x,z)-\mathsf{d}(z,y))\} && \text{by \textit{symmetric} property of } \mathsf{d}\\
&= |(\mathsf{d}(x,z)-\mathsf{d}(z,y))|
\end{aligned}$$

✏

Theorem C.1. *Let* (X,d,p,σ) *be a* POWER DISTANCE SPACE *(Definition C.2 page 146). Let* B *be an* OPEN BALL *(Definition B.4 page 134) on* (X,d). *Then for all* $(p,\sigma)\in(\mathbb{R}^*\backslash\{0\})\times\mathbb{R}^+$,

THM

$$\left\{\begin{array}{lll} \text{A.} & 2\sigma\leq 2^{\frac{1}{p}} & \text{and}\\ \text{B.} & q\in\mathsf{B}(\theta,r) & \end{array}\right\} \implies \left\{\begin{array}{ll} \text{1.} & \exists r_q\in\mathbb{R}^+ \text{ such that}\\ & \mathsf{B}(q,r_q)\subseteq\mathsf{B}(\theta,r)\end{array}\right\} \implies \left\{\text{B.}\quad q\in\mathsf{B}(\theta,r)\right\}$$

✎PROOF:

1. lemma:

$$\begin{aligned}
q \in \mathsf{B}(\theta, r) &\iff \mathsf{d}(\theta, q) < r && \text{by definition of } \textit{open ball} \text{ (Definition B.4 page 134)}\\
&\iff 0 < r - \mathsf{d}(\theta, q) && \text{by field property of real numbers}\\
&\iff \exists r_q \in \mathbb{R}^+ \quad \text{such that} \quad 0 < r_q < r - \mathsf{d}(\theta, q) && \text{by } \textit{The Archimedean Property}^{6}
\end{aligned}$$

2. Proof that $(A), (B) \implies (1)$:

$$\begin{aligned}
\mathsf{B}(q, r_q) &\triangleq \{x \in X | \mathsf{d}(q, x) < r_q\} && \text{by definition of } \textit{open ball} \text{ (Definition B.4 page 134)}\\
&= \{x \in X | \mathsf{d}^p(q, x) < r_q^p \in \mathbb{R}^+\} && \text{because } \mathsf{f}(x) \triangleq x^p \text{ is } \textit{monotone}\\
&\subseteq \{x \in X | \mathsf{d}^p(q, x) < r^p - \mathsf{d}^p(\theta, q)\} && \text{by hypothesis B and (1) lemma page 149}\\
&= \{x \in X | \mathsf{d}^p(\theta, q) + \mathsf{d}^p(q, x) < r^p\} && \text{by field property of real numbers}\\
&= \left\{x \in X | \left[\mathsf{d}^p(\theta, q) + \mathsf{d}^p(q, x)\right]^{\frac{1}{p}} < r\right\} && \text{because } \mathsf{f}(x) \triangleq x^{\frac{1}{p}} \text{ is } \textit{monotone}\\
&\subseteq \left\{x \in X | 2^{1-1/p}\sigma\left[\mathsf{d}^p(\theta, q) + \mathsf{d}^p(q, x)\right]^{\frac{1}{p}} < r\right\} && \text{by hypothesis A which implies } 2^{1-1/p}\sigma \le 1\\
&= \left\{x \in X | 2\sigma\left[\tfrac{1}{2}\mathsf{d}(\theta, q)\, x + \tfrac{1}{2}\mathsf{d}^p(q, x)\right]^{\frac{1}{p}} < r\right\} && \text{because } 2^{1-1/p}\sigma = 2\sigma(\tfrac{1}{2})^{\frac{1}{p}}\\
&\triangleq \{x \in X | \tau(p, \sigma, \theta, x, q) < r\} && \text{by definition of } \tau \text{ (Definition C.1 page 145)}\\
&\subseteq \{x \in X | \mathsf{d}(\theta, x) < r\} && \text{by definition of } (X, \mathsf{d}, p, \sigma) \text{ (Definition C.2 page 146)}\\
&\triangleq \mathsf{B}(\theta, r) && \text{by definition of } \textit{open ball} \text{ (Definition B.4 page 134)}
\end{aligned}$$

3. Proof that $(B) \impliedby (1)$:

$$\begin{aligned}
q &\in \{x \in X | \mathsf{d}(q, x) = 0\} && \text{by } \textit{nondegenerate} \text{ property (Definition B.1 page 133)}\\
&\subseteq \{x \in X | \mathsf{d}(q, x) < r_q\} && \text{because } r_q > 0\\
&\triangleq \mathsf{B}(q, r_q) && \text{by definition of } \textit{open ball} \text{ (Definition B.4 page 134)}\\
&\subseteq \mathsf{B}(\theta, r) && \text{by hypothesis 2}
\end{aligned}$$

✏

Corollary C.4. *Let* $(X, \mathsf{d}, p, \sigma)$ *be a* POWER DISTANCE SPACE. *Then for all* $(p, \sigma) \in (\mathbb{R}^* \setminus \{0\}) \times \mathbb{R}^+$,

THM
$$\left\{2\sigma \le 2^{\frac{1}{p}}\right\} \implies \left\{\textit{every OPEN BALL in } (X, \mathsf{d}) \textit{ is OPEN}\right\}$$

✎PROOF: This follows from Theorem C.1 (page 148) and Theorem B.3 (page 136). ✏

Corollary C.5. *Let* $(X, \mathsf{d}, p, \sigma)$ *be a* POWER DISTANCE SPACE. *Let* $\boldsymbol{B}$ *be the set of all* OPEN BALLs *in* (X, d). *Then for all* $(p, \sigma) \in (\mathbb{R}^* \setminus \{0\}) \times \mathbb{R}^+$,

COR
$$\left\{2\sigma \le 2^{\frac{1}{p}}\right\} \implies \left\{\ \boldsymbol{B} \textit{ is a BASE for } (X, \boldsymbol{T})\ \right\}$$

✎PROOF:

[6] Aliprantis and Burkinshaw (1998) page 17 ⟨Theorem 3.3 (“*The Archimedean Property*”) and Theorem 3.4⟩, Zorich (2004) page 53 ⟨6° (“*The principle of Archimedes*”) and 7°⟩

1. The set of all *open balls* in (X, d) is a *base* for $(X, \boldsymbol{T})$ by Corollary C.4 (page 149) and Theorem D.16 (page 179).
2. $\boldsymbol{T}$ is a topology on X by Definition D.17 (page 179).

✏

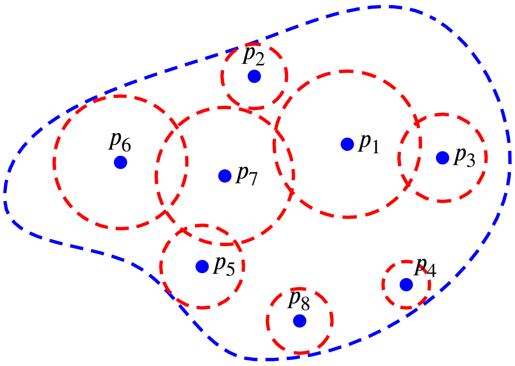


Figure C.2: *open set* (see Lemma C.2 page 150)

Lemma C.2 (next) demonstrates that every point in an open set is contained in an open ball that is contained in the original open set (see also Figure C.2 page 150).

Lemma C.2. *Let* $(X, \mathsf{d}, p, \sigma)$ *be a* POWER DISTANCE SPACE. *Let* B *be an* OPEN BALL *on* (X, d). *Then for all* $(p, \sigma) \in (\mathbb{R}^* \setminus \{0\}) \times \mathbb{R}^+$,

LEM

$$\left\{ \begin{array}{ll} \text{A.} & 2\sigma \le 2^{\frac{1}{p}} \quad \textit{and} \\ \text{B.} & U \textit{ is } \text{OPEN } \textit{in } (X, \mathsf{d}) \end{array} \right\} \implies \left\{ \begin{array}{ll} \text{1.} & \forall x \in U,\ \exists r \in \mathbb{R}^+ \quad \text{such that} \\ & \mathsf{B}(x, r) \subseteq U \end{array} \right\} \implies \left\{ \begin{array}{ll} \text{B.} & U \textit{ is} \\ & \text{OPEN } \textit{in } (X, \mathsf{d}) \end{array} \right\}$$

✎PROOF:

1. Proof that for $((A), (B) \implies (1))$:

$$\begin{aligned} U &= \bigcup \left\{ \mathsf{B}\left(x_\gamma, r_\gamma\right) \mid \mathsf{B}\left(x_\gamma, r_\gamma\right) \subseteq U \right\} && \text{by left hypothesis and Corollary C.5 page 149} \\ &\supseteq \mathsf{B}(x, r) && \text{because } x \text{ must be in one of those balls in } U \end{aligned}$$

2. Proof that $((B) \impliedby (1))$ case:

$$\begin{aligned} U &= \bigcup \{x \in X \mid x \in U\} && \text{by definition of union operation } \bigcup \\ &= \bigcup \{\mathsf{B}(x, r) \mid x \in U \text{ and } \mathsf{B}(x, r) \subseteq U\} && \text{by hypothesis (1)} \\ &\implies U \text{ is } \textit{open} && \text{by Corollary C.5 page 149 and Corollary B.1 page 135} \end{aligned}$$

✏

Corollary C.6. [7] *Let* $(X, \mathsf{d}, p, \sigma)$ *be a* POWER DISTANCE SPACE. *Let* B *be an* OPEN BALL *on* (X, d). *Then for all* $(p, \sigma) \in (\mathbb{R}^* \setminus \{0\}) \times \mathbb{R}^+$,

COR

$$\left\{ 2\sigma \le 2^{\frac{1}{p}} \right\} \implies \left\{ \textit{every } \text{OPEN BALL } \mathsf{B}(x, r) \textit{ in } (X, \mathsf{d}) \textit{ is } \text{OPEN} \right\}$$

[7] in *metric space* $((p, \sigma) = (1, 1))$: Rosenlicht (1968) pages 40–41, Aliprantis and Burkinshaw (1998) page 35

✎PROOF:

The union of any set of open balls is open — by Corollary C.5 page 149
$\implies$ the union of a set of just one open ball is open
$\implies$ every open ball is open.

✏

Theorem C.2. [8] *Let* $(X, \mathsf{d}, p, \sigma)$ *be a* POWER DISTANCE SPACE. *Let* $(X, \boldsymbol{T})$ *be a* TOPOLOGICAL SPACE INDUCED BY (X, d). *Let* $(\!(x_n \in X)\!)_{n\in\mathbb{Z}}$ *be a sequence in* (X, d).

THM
$$\underbrace{(\!(x_n)\!) \text{ converges to a limit } x}_{\text{(Definition D.20 page 181)}} \iff \left\{\begin{array}{l}\text{for any } \varepsilon \in \mathbb{R}^+, \text{ there exists } N \in \mathbb{Z} \\ \text{such that for all } n > N, \quad \mathsf{d}(x_n, x) < \varepsilon\end{array}\right\}$$

✎PROOF:

$$\begin{aligned}
(\!(x_n)\!) \to x &\iff x_n \in U \quad \forall U \in N_x,\ n > N && \text{by Definition D.20 page 181}\\
&\iff \exists \mathsf{B}(x,\varepsilon) \quad \text{such that} \quad x_n \in \mathsf{B}(x,\varepsilon)\,\forall n > N && \text{by Lemma C.2 page 150}\\
&\iff \mathsf{d}(x_n, x) < \varepsilon && \text{by Definition B.4 page 134}
\end{aligned}$$

✏

In *distance spaces* (Definition B.1 page 133), not all *convergent* sequences are *Cauchy* (Example B.2 page 141). However in a distance space with any *power triangle inequality* (Definition C.2 page 146), all *convergent* sequences are *Cauchy* (next theorem).

Theorem C.3. [9] *Let* $(X, \mathsf{d}, p, \sigma)$ *be a* POWER DISTANCE SPACE. *Let* B *be an* OPEN BALL *on* (X, d).

THM
For any $(p, \sigma) \in \mathbb{R}^* \times \mathbb{R}^+$,
$$\left\{\begin{array}{l}(\!(x_n)\!) \text{ is CONVERGENT} \\ \text{in } (X, \mathsf{d})\end{array}\right\} \implies \left\{\begin{array}{l}(\!(x_n)\!) \text{ is CAUCHY} \\ \text{in } (X, \mathsf{d})\end{array}\right\} \implies \left\{\begin{array}{l}(\!(x_n)\!) \text{ is BOUNDED} \\ \text{in } (X, \mathsf{d})\end{array}\right\}$$

✎PROOF:

1. Proof that *convergent* $\implies$ *Cauchy*:

$$\begin{aligned}
\mathsf{d}(x_n, x_m) &\le \tau(p, \sigma; x_n, x_m, x) && \text{by definition of } \textit{power triangle inequality} \text{ (Definition C.2 page 146)}\\
&\triangleq 2\sigma\left[\frac{1}{2}\mathsf{d}^p(x_n, x) + \frac{1}{2}\mathsf{d}^p(x_m, x)\right]^{\frac{1}{p}} && \text{by definition of } \textit{power triangle function} \text{ (Definition C.1 page 145)}\\
&< 2\sigma\left[\frac{1}{2}\varepsilon^p + \frac{1}{2}\varepsilon^p\right]^{\frac{1}{p}} && \text{by } \textit{convergence} \text{ hypothesis (Definition D.20 page 181)}\\
&= 2\sigma\varepsilon && \text{by definition of } \textit{convergence} \text{ (Definition D.20 page 181)}\\
&\implies \textit{Cauchy} && \text{by definition of } \textit{Cauchy} \text{ (Definition B.8 page 137)}\\
\mathsf{d}(x_n, x_m) &\le \tau(\infty, \sigma; x_n, x_m, x) && \text{by definition of } \textit{power triangle inequality} \text{ at } p = \infty\\
&= 2\sigma \max\{\mathsf{d}(x_n, x),\ \mathsf{d}(x_m, x)\} && \text{by Corollary C.2 (page 147)}\\
&= 2\sigma \max\{\varepsilon,\ \varepsilon\} && \text{by } \textit{convergent} \text{ hypothesis (Definition D.20 page 181)}\\
&= 2\sigma\varepsilon && \text{by definition of max}\\
\mathsf{d}(x_n, x_m) &\le \tau(-\infty, \sigma; x_n, x_m, x) && \text{by definition of } \textit{power triangle inequality} \text{ at } p = -\infty\\
&= 2\sigma \min\{\mathsf{d}(x_n, x),\ \mathsf{d}(x_m, x)\} && \text{by Corollary C.2 (page 147)}
\end{aligned}$$

[8] in *metric space*: Rosenlicht (1968) page 45, Giles (1987) page 37 ⟨3.2 Definition⟩
[9] in *metric space*: Giles (1987) page 49 ⟨Theorem 3.30⟩, Rosenlicht (1968) page 51, Apostol (1975) pages 72–73 ⟨Theorem 4.6⟩

$$\begin{aligned}
&= 2\sigma \min\{\varepsilon, \varepsilon\} && \text{by \textit{convergent} hypothesis (Definition D.20 page 181)}\\
&= 2\sigma\varepsilon && \text{by definition of min}\\
\mathsf{d}(x_n, x_m) &\le \tau(0, \sigma; x_n, x_m, x) && \text{by definition of \textit{power triangle inequality} at } p = 0\\
&= 2\sigma\sqrt{\mathsf{d}(x_n, x)}\sqrt{\mathsf{d}(x_m, x)} && \text{by Corollary C.2 (page 147)}\\
&= 2\sigma\sqrt{\varepsilon}\sqrt{\varepsilon} && \text{by \textit{convergent} hypothesis (Definition D.20 page 181)}\\
&= 2\sigma\varepsilon && \text{by property of } \mathbb{R}
\end{aligned}$$

2. Proof that *Cauchy* $\Longrightarrow$ *bounded*: by Proposition B.1 (page 137).

✏

Theorem C.4. [10] *Let* $(X, \mathsf{d}, p, \sigma)$ *be a* POWER DISTANCE SPACE. *Let* $\mathsf{f} \in \mathbb{Z}^{\mathbb{Z}}$ *be a* STRICTLY MONOTONE *function such that* $\mathsf{f}(n) < \mathsf{f}(n+1)$.

THM *For any* $(p, \sigma) \in \mathbb{R}^* \times \mathbb{R}^+$

$$\left\{\begin{array}{lll} 1. & (x_n)_{n\in\mathbb{Z}} \text{ is CAUCHY} & \text{and}\\ 2. & \left(x_{\mathsf{f}(n)}\right)_{n\in\mathbb{Z}} \text{ is CONVERGENT} & \end{array}\right\} \implies \left\{(x_n)_{n\in\mathbb{Z}} \text{ is CONVERGENT.}\right\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{d}(x_n, x) &= \mathsf{d}(x, x_n) && \text{by \textit{symmetric} property of d}\\
&\le \tau(p, \sigma; x, x_n, x_{\mathsf{f}(n)}) && \text{by definition of \textit{power triangle inequality} (Definition C.2 page 146)}\\
&\triangleq 2\sigma\left[\frac{1}{2}\mathsf{d}^p(x, x_{\mathsf{f}(n)}) + \frac{1}{2}\mathsf{d}^p(x_{\mathsf{f}(n)}, x_n)\right]^{\frac{1}{p}} && \text{by definition of \textit{power triangle function} (Definition C.1 page 145)}\\
&= 2\sigma\left[\frac{1}{2}\varepsilon + \frac{1}{2}\mathsf{d}^p(x_{\mathsf{f}(n)}, x_n)\right]^{\frac{1}{p}} && \text{by left hypothesis 2}\\
&= 2\sigma\left[\frac{1}{2}\varepsilon^p + \frac{1}{2}\varepsilon^p\right]^{\frac{1}{p}} && \text{by left hypothesis 1}\\
&= 2\sigma\varepsilon\\
&\implies \textit{convergent} && \text{by definition of \textit{convergent} (Definition D.20 page 181)}
\end{aligned}$$

✏

Theorem C.5. [11] *Let* $(X, \mathsf{d}, p, \sigma)$ *be a* POWER DISTANCE SPACE. *Let* $(\mathbb{R}, \mathsf{q})$ *be a metric space of real numbers with the usual metric* $\mathsf{q}(x, y) \triangleq |x - y|$. *Then*

THM
$$\left\{2\sigma = 2^{\frac{1}{p}}\right\} \implies \left\{\ \mathsf{d} \text{ is CONTINUOUS in } (\mathbb{R}, \mathsf{q})\ \right\}$$

✎PROOF:

$$\begin{aligned}
\left|\mathsf{d}(x, y) - \mathsf{d}(x_n, y_n)\right| &\le \left|\mathsf{d}(x, y) - \mathsf{d}(x_n, y)\right| + \left|\mathsf{d}(x_n, y) - \mathsf{d}(x_n, y_n)\right| && \text{by \textit{triangle inequality} of } (\mathbb{R}, |x-y|)\\
&= \left|\mathsf{d}(x, y) - \mathsf{d}(y, x_n)\right| + \left|\mathsf{d}(y, x_n) - \mathsf{d}(x_n, y_n)\right| && \text{by \textit{commutative} property of d (Definition B.1 page 133)}\\
&\le \mathsf{d}(x, x_n) + \mathsf{d}(y, y_n) && \text{by } 2\sigma = 2^{\frac{1}{p}} \text{ and Lemma C.1 (page 147)}\\
&= 0 \qquad \text{as } n \to \infty
\end{aligned}$$

✏

In *distance spaces* and *topological spaces*, limits of convergent sequences are in general *not unique* (Example B.1 page 140, Example D.16 page 181). However Theorem C.6 (next) demonstrates that, in a *power distance space*, limits *are* unique.

[10] in *metric space*: Rosenlicht (1968) page 52
[11] in *metric space* ($(p, \sigma) = (1, 1)$ case): Berberian (1961) page 37 ⟨Theorem II.4.1⟩

Theorem C.6 (Uniqueness of limit). [12] *Let* $(X, \mathsf{d}, p, \sigma)$ *be a* POWER DISTANCE SPACE. *Let* $x, y, \in X$ *and let* $(\!(x_n \in X)\!)$ *be an* X*-valued sequence.*

$$\textbf{THM}\quad \left\{\begin{array}{ll} 1. & \{((\!(x_n)\!), (\!(y_n)\!)) \to (x, y)\} \quad \text{and} \\ 2. & (p, \sigma) \in \mathbb{R}^* \times \mathbb{R}^+ \end{array}\right\} \implies \{x = y\}$$

✎PROOF:

1. lemma: Proof that for all $(p, \sigma) \in \mathbb{R}^* \times \mathbb{R}^+$ and for any $\varepsilon \in \mathbb{R}^+$, there exists N such that $\mathsf{d}(x, y) < 2\sigma\varepsilon$:

$$\begin{aligned}
\mathsf{d}(x, y) &\le \tau(p, \sigma; x, y, x_n) && \text{by definition of } \textit{power triangle inequality} \text{ (Definition C.2 page 146)}\\
&\triangleq 2\sigma\left[\frac{1}{2}\mathsf{d}^p(x, x_n) + \frac{1}{2}\mathsf{d}^p(x_n, y)\right]^{\frac{1}{p}} && \text{by definition of } \textit{power triangle function} \text{ (Definition C.1 page 145)}\\
&< 2\sigma\left[\frac{1}{2}\varepsilon^p + \frac{1}{2}\varepsilon^p\right]^{\frac{1}{p}} && \text{by left hypothesis and for } p \in \mathbb{R}^* \backslash \{-\infty, 0, \infty\}\\
&= 2\sigma\varepsilon\\
\mathsf{d}(x, y) &\le \tau(\infty, \sigma; x, y, x_n) && \text{by definition of } \textit{power triangle inequality} \text{ at } p = \infty\\
&= 2\sigma \max\{\mathsf{d}(x, x_n), \mathsf{d}(x_n, y)\} && \text{by Corollary C.2 (page 147)}\\
&< 2\sigma\varepsilon && \text{by left hypothesis}\\
\mathsf{d}(x, y) &\le \tau(-\infty, \sigma; x, y, x_n) && \text{by definition of } \textit{power triangle inequality} \text{ at } p = -\infty\\
&= 2\sigma \min\{\mathsf{d}(x, x_n), \mathsf{d}(x_n, y)\} && \text{by Corollary C.2 (page 147)}\\
&< 2\sigma\varepsilon && \text{by left hypothesis}\\
\mathsf{d}(x, y) &\le \tau(0, \sigma; x, y, x_n) && \text{by definition of } \textit{power triangle inequality} \text{ at } p = 0\\
&= 2\sigma\sqrt{\mathsf{d}(x, x_n)}\sqrt{\mathsf{d}(x_n, y)} && \text{by Corollary C.2 (page 147)}\\
&= 2\sigma\sqrt{\varepsilon}\sqrt{\varepsilon} && \text{by left hypothesis}\\
&< 2\sigma\varepsilon && \text{by property of real numbers}
\end{aligned}$$

2. Proof that $x = y$ (proof by contradiction):

$$\begin{aligned}
x \neq y &\implies \mathsf{d}(x, y) \neq 0 && \text{by the } \textit{nondegenerate} \text{ property of } \mathsf{d} \text{ (Definition B.1 page 133)}\\
&\implies \mathsf{d}(x, y) > 0 && \text{by } \textit{non-negative} \text{ property of } \mathsf{d} \text{ (Definition B.1 page 133)}\\
&\implies \exists\varepsilon \quad \text{such that} \quad \mathsf{d}(x, y) > 2\sigma\varepsilon\\
&\implies \textit{contradiction} \text{ to (1) lemma page 153}\\
&\implies \mathsf{d}(x, y) = 0\\
&\implies x = y
\end{aligned}$$

✏

C.3 Examples

It is not always possible to find a *triangle relation* (Definition C.2 page 146) $\triangle(p, \sigma; \mathsf{d})$ that holds in every *distance space* (Definition B.1 page 133), as demonstrated by Example C.1 and Example C.2 (next two examples).

Example C.1. Let $\mathsf{d}(x, y) \in \mathbb{R}^{\mathbb{R}\times\mathbb{R}}$ be defined such that

$$\mathsf{d}(x, y) \triangleq \left\{\begin{array}{lll} y & \forall (x, y) \in \{4\} \times (0 : 2] & \textit{(vertical half-open interval)}\\ x & \forall (x, y) \in (0 : 2] \times \{4\} & \textit{(horizontal half-open interval)}\\ |x - y| & \text{otherwise} & \textit{(Euclidean)} \end{array}\right\}.$$

Note the following about the pair $(\mathbb{R}, \mathsf{d})$:

[12] in *metric space*: Rosenlicht (1968) page 46, Thomson et al. (2008) page 32 ⟨Theorem 2.8⟩

1. By Example B.1 (page 140), $(\mathbb{R}, \mathsf{d})$ is a *distance space*, but not a *metric space*—that is, the *triangle relation* $\triangle(1, 1; \mathsf{d})$ does not hold in $(\mathbb{R}, \mathsf{d})$.

2. Observe further that $(\mathbb{R}, \mathsf{d})$ is *not* a *power distance space*. In particular, the *triangle relation* $\triangle(p, \sigma; \mathsf{d})$ does not hold in $(\mathbb{R}, \mathsf{d})$ for any finite value of σ (does not hold for any $\sigma \in \mathbb{R}^+$):

$$\begin{aligned}\mathsf{d}(0,4) = 4 \nleq 0 = \lim_{\varepsilon\to 0} 2\sigma\varepsilon = \lim_{\varepsilon\to 0} 2\sigma\left[{}^{1}\!/\!{}_{2}|0-\varepsilon|^p + {}^{1}\!/\!{}_{2}\varepsilon^p\right]^{\frac{1}{p}} \\ \triangleq \lim_{\varepsilon\to 0} 2\sigma\left[{}^{1}\!/\!{}_{2}\mathsf{d}^p(0,\varepsilon) + {}^{1}\!/\!{}_{2}\mathsf{d}^p(\varepsilon,4)\right]^{\frac{1}{p}} \triangleq \lim_{\varepsilon\to 0} \triangle(p,\sigma;0,4,\varepsilon;\mathsf{d})\end{aligned}$$

Example C.2. Let $\mathsf{d}(x, y) \in \mathbb{R}^{\mathbb{R}\times\mathbb{R}}$ be defined such that

$$\mathsf{d}(x,y) \triangleq \left\{\begin{array}{lll} |x-y| & \text{for } x=0 \text{ or } y=0 \text{ or } x=y & (\textit{Euclidean}) \\ 1 & \text{otherwise} & (\textit{discrete}) \end{array}\right\}.$$

Note the following about the pair $(\mathbb{R}, \mathsf{d})$:

1. By Example B.2 (page 141), $(\mathbb{R}, \mathsf{d})$ is a *distance space*, but not a *metric space*—that is, the *triangle relation* $\triangle(1, 1; \mathsf{d})$ does not hold in $(\mathbb{R}, \mathsf{d})$.

2. Observe further that $(\mathbb{R}, \mathsf{d})$ is *not* a *power distance space*—that is, the *triangle relation* $\triangle(p, \sigma; \mathsf{d})$ does not hold in $(\mathbb{R}, \mathsf{d})$ for any value of $(p, \sigma) \in \mathbb{R}^* \times \mathbb{R}^+$.

   (a) Proof that $\triangle(p, \sigma; \mathsf{d})$ does not hold for any $(p, \sigma) \in \{\infty\} \times \mathbb{R}^+$:

$$\begin{aligned}
\lim_{n,m\to\infty} \mathsf{d}({}^{1}\!/\!{}_{n}, {}^{1}\!/\!{}_{m}) &\triangleq 1 \nleq 0 = 2\sigma \max\{0, 0\} && \text{by definition of } \mathsf{d} \\
&= 2\sigma \lim_{n,m\to\infty} \max\{\mathsf{d}({}^{1}\!/\!{}_{n}, 0), \mathsf{d}(0, {}^{1}\!/\!{}_{m})\} && \text{by Corollary C.2 (page 147)} \\
&\geq \lim_{n,m\to\infty} 2\sigma\left[{}^{1}\!/\!{}_{2}\mathsf{d}^p({}^{1}\!/\!{}_{n}, 0) + {}^{1}\!/\!{}_{2}\mathsf{d}^p(0, {}^{1}\!/\!{}_{m})\right]^{\frac{1}{p}} && \text{by Corollary C.1 (page 146)} \\
&\triangleq \lim_{n,m\to\infty} \tau(p, \sigma, {}^{1}\!/\!{}_{n}, {}^{1}\!/\!{}_{m}, 0) && \text{by definition of } \tau \text{ (Definition C.1 page 145)}
\end{aligned}$$

   (b) Proof that $\triangle(p, \sigma; \mathsf{d})$ does not hold for any $(p, \sigma) \in \mathbb{R}^* \times \mathbb{R}^+$: By Corollary C.1 (page 146), the *triangle function* (Definition C.1 page 145) $\tau(p, \sigma; x, y, z; \mathsf{d})$ is *continuous* and *strictly monotone* in $(\mathbb{R}, |\cdot|, \leq)$ with respect to the variable p. Item 2a demonstrates that $\triangle(p, \sigma; \mathsf{d})$ fails to hold at the best case of $p = \infty$, and so by Corollary C.1, it doesn't hold for any other value of $p \in \mathbb{R}^*$ either.

Example C.3. Let d be a function in $\mathbb{R}^{\mathbb{R}\times\mathbb{R}}$ such that

$$\mathsf{d}(x,y) \triangleq \left\{\begin{array}{lll} 2|x-y| & \forall (x,y) \in \{(0,1), (1,0)\} & (\textit{dilated Euclidean}) \\ |x-y| & \text{otherwise} & (\textit{Euclidean}) \end{array}\right\}.$$

Note the following about the pair $(\mathbb{R}, \mathsf{d})$:

1. By Example B.3 (page 142), $(\mathbb{R}, \mathsf{d})$ is a *distance space*, but not a *metric space*—that is, the *triangle relation* $\triangle(1, 1; \mathsf{d})$ does not hold in $(\mathbb{R}, \mathsf{d})$.

2. But observe further that $(\mathbb{R}, \mathsf{d}, 1, 2)$ *is* a *power distance space*:

   (a) Proof that $\triangle(1, 2; \mathsf{d})$ (Definition C.2 page 146) holds for all $(x, y) \in \{(0, 1), (1, 0)\}$:

$$\begin{aligned}
\mathsf{d}(1,0) = \mathsf{d}(0,1) &\triangleq 2|0-1| = 2 && \text{by definition of } \mathsf{d} \\
&\leq 2 \leq 2(|0-z| + |z-1|) \quad \forall z\in\mathbb{R} && \text{by definition of } |\cdot| \text{ (Definition 1.24 page 10)} \\
&= 2\sigma\left({}^{1}\!/\!{}_{2}|0-z|^p + {}^{1}\!/\!{}_{2}|z-1|^p\right)^{\frac{1}{p}} \quad \forall z\in\mathbb{R} && \text{for } (p,\sigma) = (1,2) \\
&\triangleq 2\sigma\left({}^{1}\!/\!{}_{2}\mathsf{d}^p(0,z) + \mathsf{d}^p(z,1)\right)^{\frac{1}{p}} \quad \forall z\in\mathbb{R} && \text{for } (p,\sigma) = (1,2) \text{ and by definition of } \mathsf{d} \\
&\triangleq \tau(1, 2; 0, 1, z) && \text{by definition of } \tau \text{ (Definition C.1 page 145)}
\end{aligned}$$

(b) Proof that $\circledtriangle(1,2;\mathsf{d})$ holds for all other $(x,y) \in \mathbb{R}^* \times \mathbb{R}^+$:

$$\begin{aligned}
\mathsf{d}(x,y) &\triangleq 2|x-y| && \text{by definition of } \mathsf{d}\\
&\le (|x-z| + |z-y|) && \text{by property of \textit{Euclidean metric spaces}}\\
&= 2\sigma\left(\tfrac{1}{2}|0-z|^p + \tfrac{1}{2}|z-1|^p\right)^{\frac{1}{p}} && \text{for } (p,\sigma) = (1,1)\\
&\triangleq \tau(1,1;x,y,z) && \text{by definition of } \tau \text{ (Definition C.1 page 145)}\\
&\le \tau(1,2;x,y,z) && \text{by Corollary C.1 (page 146)}
\end{aligned}$$

3. In (X,d), the limits of *convergent* sequences are *unique*. This follows directly from the fact that $(\mathbb{R},\mathsf{d},1,2)$ is a *power distance space* (item (2) page 154) and by Theorem C.6 page 153.

4. In (X,d), *convergent* sequences are *Cauchy*. This follows directly from the fact that $(\mathbb{R},\mathsf{d},1,2)$ is a *power distance space* (item (2) page 154) and by Theorem C.3 page 151.

Example C.4. Let d be a function in $\mathbb{R}^{\mathbb{R}\times\mathbb{R}}$ such that $\mathsf{d}(x,y) \triangleq (x-y)^2$. Note the following about the pair $(\mathbb{R},\mathsf{d})$:

1. It was demonstrated in Example B.4 (page 142) that $(\mathbb{R},\mathsf{d})$ is a *distance space*, but that it is *not* a *metric space* because the *triangle inequality* does not hold.

2. However, the tuple $(\mathbb{R},\mathsf{d},p,\sigma)$ *is* a *power distance space* (Definition C.2 page 146) for any $(p,\sigma) \in \mathbb{R}^* \times [2:\infty)$: In particular, for all $x,y,z \in \mathbb{R}$, the *power triangle inequality* (Definition C.2 page 146) must hold. The "worst case" for this is when a third point z is exactly "halfway between" x and y in $\mathsf{d}(x,y)$; that is, when $z = \frac{x+y}{2}$:

$$\begin{aligned}
(x-y)^2 &\triangleq \mathsf{d}(x,y) && \text{by definition of } \mathsf{d}\\
&\le \tau(p,\sigma;x,y,z;\mathsf{d}) && \text{by definition \textit{power triangle inequality}}\\
&\triangleq 2\sigma\left[\tfrac{1}{2}\mathsf{d}^p(x,z) + \tfrac{1}{2}\mathsf{d}^p(z,y)\right]^{\frac{1}{p}} && \text{by definition } \tau \text{ (Definition C.1 page 145)}\\
&\triangleq 2\sigma\left[\tfrac{1}{2}(x-z)^{2p} + \tfrac{1}{2}(z-y)^{2p}\right]^{\frac{1}{p}} && \text{by definition of } \mathsf{d}\\
&= 2\sigma\left[\tfrac{1}{2}|x-z|^{2p} + \tfrac{1}{2}|z-y|^{2p}\right]^{\frac{1}{p}} && \text{because } (x)^2 = |x|^2 \text{ for all } x \in \mathbb{R}\\
&= 2\sigma\left[\tfrac{1}{2}\left|x - \frac{x+y}{2}\right|^{2p} + \tfrac{1}{2}\left|\frac{x+y}{2} - y\right|^{2p}\right]^{\frac{1}{p}} && \text{because } z = \frac{x+y}{2} \text{ is the ``worst case'' scenario}\\
&= 2\sigma\left[\tfrac{1}{2}\left|\frac{y-x}{2}\right|^{2p} + \tfrac{1}{2}\left|\frac{x-y}{2}\right|^{2p}\right]^{\frac{1}{p}}\\
&= 2\sigma\left[\left|\frac{x-y}{2}\right|^{2p}\right]^{\frac{1}{p}} = \frac{2\sigma}{4}|x-y|^2\\
&\implies (p,\sigma) \in \mathbb{R}^* \times [2:\infty)
\end{aligned}$$

3. The *power distance function* d is *continuous* in $(\mathbb{R},\mathsf{d},p,\sigma)$ for any (p,σ) such that $\sigma \ge 2$ and $2\sigma = p^{\frac{1}{p}}$. This follows directly from Theorem C.5 (page 152).

APPENDIX D

SOME MATHEMATICAL TOOLS

❝*Dirichlet alone, not I, nor Cauchy, nor Gauss knows what a completely rigorous proof is. Rather we learn it first from him. When Gauss says he has proved something it is clear; when Cauchy says it, one can wager as much pro as con; when Dirichlet says it, it is certain.*❞

Carl Gustav Jacob Jacobi (1804–1851), Jewish-German mathematician [1]

D.1 Linear spaces

D.1.1 Structure

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

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*

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

Definition D.2. [3] *Let* $\boldsymbol{\Omega} \triangleq \big(X,\ +,\ \cdot,\ (\mathbb{F},\ \dot{+},\ \dot{\times})\big)$ *be a* LINEAR SPACE (Definition D.1 page 157).

---

[1] quote: http://lagrange.math.trinity.edu/aholder/misc/quotes.shtml
Schubring (2005), page 558
image: http://en.wikipedia.org/wiki/Carl_Gustav_Jakob_Jacobi

[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

[3] Mitrinović et al. (2010) page 1, van de Vel (1993) pages 5–6, Bollobás (1999), page 2

DEF: *A set $D \subseteq X$ is **convex** in $\boldsymbol{\Omega}$ if*

$$\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in D \quad \forall \boldsymbol{x},\boldsymbol{y} \in D \quad \text{and} \quad \forall \lambda \in (0,1)$$

*A set is **concave** in $\boldsymbol{\Omega}$ if it is* NOT CONVEX *in $\boldsymbol{\Omega}$.*

D.1.2 Metric Linear Spaces

Metric space stucture can be added to a linear space resulting in a *metric linear space* (next definition). One key difference between metric linear spaces and normed linear spaces is that the balls in a *normed linear space* (Definition D.4 page 159) are always *convex* (Definition D.2 page 157); this is not true for all metric linear spaces (Theorem D.4 page 160).[4]

Definition D.3. [5] *Let $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d}\right)$.*

DEF: *The tuple $\boldsymbol{\Omega}$ is a **metric linear space** if*

1. *if $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ is a* LINEAR SPACE *and*
2. *d is a* METRIC *in $\mathbb{R}^{\boldsymbol{X}}$ and*
3. *$\mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z}) = \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \quad \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \in X$ (*TRANSLATION INVARIANT*)[6] and*
4. *$\alpha_n \to \alpha$ and $\boldsymbol{x}_n \to \boldsymbol{x} \implies \alpha_n \boldsymbol{x}_n \to \alpha \boldsymbol{x}$*

Theorem D.1. [7] *Let $\boldsymbol{\Omega} \triangleq \left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d}\right)$ be a metric linear space.*

THM:
$$\underbrace{\mathsf{d}(\theta, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \le \lambda \mathsf{d}(\theta, \boldsymbol{x}) + (1-\lambda)\mathsf{d}(\theta, \boldsymbol{y})}_{\mathsf{d} \text{ is a CONVEX function}} \implies \left\{ \begin{array}{l} \mathsf{B}(\theta, r) \in \boldsymbol{\Omega} \text{ is CONVEX} \\ \forall \theta \in X, r \in \mathbb{R}^+ \end{array} \right\}$$

✎PROOF:

$$\begin{aligned}
\mathsf{d}(\theta, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &\le \lambda \mathsf{d}(\theta, \boldsymbol{x}) + (1-\lambda)\mathsf{d}(\theta, \boldsymbol{y}) && \text{by convexity hypothesis} \\
&\le \lambda r + (1-\lambda) r && \forall \boldsymbol{x}, \boldsymbol{y} \in \mathsf{B}(\theta, r) \\
&= r \\
\implies \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} &\in \mathsf{B}(\theta, r) && \forall \boldsymbol{x}, \boldsymbol{y} \in \mathsf{B}(\theta, r) \\
\implies \mathsf{B}(\theta, r) &\in (\boldsymbol{X}, \mathsf{d}) \text{ is convex} && \forall \theta \in \boldsymbol{X}
\end{aligned}$$

✏

Theorem D.2. [8] *Let $\left(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \mathsf{d}\right)$ be a real metric linear space.*

THM:
$$\left\{ \begin{array}{lllll} 1. & \mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z}) = \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) & \forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \in \boldsymbol{X} & \text{(TRANSLATION INVARIANT)} & \text{and} \\ 2. & \mathsf{d}(\lambda\boldsymbol{x}, \lambda\boldsymbol{y}) = \lambda\, \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) & \forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{X}, \lambda \in [0,1] & \text{(HOMOGENEOUS)} & \end{array} \right\}$$
$$\implies \{\mathsf{B}(\theta, r) \in (\boldsymbol{X}, \mathsf{d}) \quad \text{is CONVEX} \quad \forall \theta \in \boldsymbol{X}, r \in \mathbb{R}^+\}$$

---

[4] Bruckner et al. (1997) page 478

[5] Maddox (1989) page 90, Bruckner et al. (1997) page 477 ⟨Definition 12.3⟩, Rolewicz (1985) page 1, Loève (1977) page 79

[6] Some authors do not require the *translation invariant* property for the definition of the *metric linear space*, as indicated by the following references: Maddox (1989) page 90 ⟨"Some authors…do not include translation invariance in the definition of metric linear space, since they use a theorem of Kakutani to show that a non-translation invariant metric may be replaced by a translation invariant metric which yields the same topolgy."⟩, Friedman (1970) page 125 ⟨Definition 4.1.4⟩, Dobrowolski and Mogilski (1995) page 86

[7] Norfolk (1991), page 5

[8] Norfolk (1991) pages 5–6, `http://groups.google.com/group/sci.math/msg/a6f0a7924027957d`

✎PROOF:

$$\begin{aligned}
&\mathsf{d}(\boldsymbol{\theta}, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\
&= \mathsf{d}(\mathbb{0}, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} - \boldsymbol{\theta}) && \text{by translation invariance hypothesis} \\
&= \mathsf{d}(\mathbb{0}, \lambda(\boldsymbol{x}-\boldsymbol{\theta}) + (1-\lambda)(\boldsymbol{y}-\boldsymbol{\theta})) \\
&\le \mathsf{d}(\mathbb{0}, \lambda(\boldsymbol{x}-\boldsymbol{\theta})) + \mathsf{d}(\lambda(\boldsymbol{x}-\boldsymbol{\theta}), \lambda(\boldsymbol{x}-\boldsymbol{\theta}) + (1-\lambda)(\boldsymbol{y}-\boldsymbol{\theta})) && \text{by } \textit{subadditive} \text{ property} \\
&= \mathsf{d}(\mathbb{0}, \lambda(\boldsymbol{x}-\boldsymbol{\theta})) + \mathsf{d}(\mathbb{0}, \mathbb{0} + (1-\lambda)(\boldsymbol{y}-\boldsymbol{\theta})) && \text{by translation invariance hypothesis} \\
&= \lambda\mathsf{d}(\mathbb{0}, \boldsymbol{x}-\boldsymbol{\theta}) + (1-\lambda)\mathsf{d}(\mathbb{0}, \boldsymbol{y}-\boldsymbol{\theta}) && \text{by homogeneous hypothesis} \\
&= \lambda\mathsf{d}(\boldsymbol{\theta}, \boldsymbol{x}) + (1-\lambda)\mathsf{d}(\boldsymbol{\theta}, \boldsymbol{y}) && \text{by translation invariance hypothesis} \\
&\le \lambda r + (1-\lambda) r && \forall \boldsymbol{x}, \boldsymbol{y} \in \mathsf{B}(\theta, r) \\
&= r \\
\Longrightarrow\; & \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in \mathsf{B}(\theta, r) && \forall \boldsymbol{x}, \boldsymbol{y} \in \mathsf{B}(\theta, r) \\
\Longrightarrow\; & \mathsf{B}(\theta, r) \in (\boldsymbol{X}, \mathsf{d}) \text{ is convex} && \forall \theta \in \boldsymbol{X}
\end{aligned}$$

✏

D.1.3 Normed Linear Spaces

Definition D.4. [9] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}))$ *be a* LINEAR SPACE (Definition D.1 page 157) *and* $|\cdot| \in \mathbb{R}^{\mathbb{F}}$ *the* ABSOLUTE VALUE *function.*

DEF

A functional $\|\cdot\|$ *in* $\mathbb{R}^{\boldsymbol{X}}$ *is a* ***norm*** *if*

1. $\|\boldsymbol{x}\| \geq 0 \quad \forall \boldsymbol{x} \in X$ (STRICTLY POSITIVE) *and*
2. $\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = \mathbb{0} \quad \forall \boldsymbol{x} \in X$ (NONDEGENERATE) *and*
3. $\|\alpha\boldsymbol{x}\| = |\alpha|\,\|\boldsymbol{x}\| \quad \forall \boldsymbol{x} \in X, \alpha \in \mathbb{C}$ (HOMOGENEOUS) *and*
4. $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \quad \forall \boldsymbol{x}, \boldsymbol{y} \in X$ (SUBADDITIVE/TRIANGLE INQUALITY).

A ***normed linear space*** *is the tuple* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$.

Example D.1 (The usual norm). [10] 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** $|\cdot| \in \mathbb{R}^{\mathbb{R}}$ is a *norm.*

Example D.2 (l_p norms). Let $(x_n)_{n\in\mathbb{Z}}$ be a *sequence* of real numbers.

EX

$$\|(x_n)\|_p \triangleq \left(\sum_{n\in\mathbb{Z}} |x_n|^p\right)^{\frac{1}{p}} \qquad \text{is a norm for } p \in [1 : \infty]$$

D.1.4 Relationship between metrics and norms

Metrics generated by norms

Theorem D.3. [11] *Let* $\mathsf{d} \in \mathbb{R}^{\boldsymbol{X}\times\boldsymbol{X}}$ *be a function on a* REAL *normed linear space* $(X, +, \cdot, (\mathbb{R}, \dot{+}, \dot{\times}), \|\cdot\|)$. *Let* $\mathsf{B}(\boldsymbol{x}, r) \triangleq \{\boldsymbol{y} \in \boldsymbol{X} | \, \|\boldsymbol{y}-\boldsymbol{x}\| < r\}$ *be the* OPEN BALL (Definition B.4 page 134) *of radius* r *centered at a point*

[9] Aliprantis and Burkinshaw (1998), pages 217–218, Banach (1932a), page 53, Banach (1932b), page 33, Banach (1922) page 135
[10] Giles (1987) page 3
[11] Michel and Herget (1993), page 344, Banach (1932a) page 53

$\boldsymbol{x}$.

THM $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \|\boldsymbol{x} - \boldsymbol{y}\|$ *is a metric on* $\boldsymbol{X}$

The next definition defines this metric formally.

Definition D.5. [12] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a* NORMED LINEAR SPACE *(Definition D.4 page 159).*

DEF *The* ***metric induced by the norm*** $\|\cdot\|$ *is the function* $\mathsf{d} \in \mathbb{R}^{\boldsymbol{X}}$ *such that* $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) \triangleq \|\boldsymbol{x} - \boldsymbol{y}\| \quad \forall x, y \in X$.

Corollary D.1. [13] *Let* $\boldsymbol{\Omega} \triangleq (X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \|\cdot\|)$ *be a* NORMED LINEAR SPACE *(Definition D.4 page 159).*

COR *The norm* $\|\cdot\|$ *is* CONTINUOUS *in* $\boldsymbol{\Omega}$.

Theorem D.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 D.4. [14] *Let* $(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d})$ *be a* METRIC LINEAR SPACE *(Definition D.3 page 158). Let* B *be an* OPEN BALL *(Definition B.4 page 134).*

THM
$$\left.\underbrace{\begin{array}{l}\exists \|\cdot\| \in \mathbb{R}^{\boldsymbol{X}} \quad \text{such that} \\ \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{x}\|\end{array}}_{\mathsf{d} \text{ is generated by a norm}}\right\} \implies \left\{\begin{array}{lll}1. & \mathsf{B}(\boldsymbol{x}, r) = \boldsymbol{x} + B(0, r) & \text{and} \\ 2. & \mathsf{B}(\mathbb{0}, r) = r\,\mathsf{B}(\mathbb{0}, 1) & \text{and} \\ 3. & \mathsf{B}(\boldsymbol{x}, r) \text{ is CONVEX} & \text{and} \\ 4. & \boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) \iff -\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) & (\text{SYMMETRIC})\end{array}\right.$$

✎PROOF:

1. Proof that $\mathsf{d}(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} + v\boldsymbol{z}) = \mathsf{d}(\boldsymbol{x}, \boldsymbol{y})$ (invariant):

$$\begin{aligned}\mathsf{d}(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} + v\boldsymbol{z}) &= \|(\boldsymbol{y} + v\boldsymbol{z}) - (\boldsymbol{x} + \boldsymbol{z})\| && \text{by left hypothesis} \\ &= \|\boldsymbol{y} - \boldsymbol{x}\| \\ &= \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) && \text{by left hypothesis}\end{aligned}$$

2. Proof that $\mathsf{B}(\boldsymbol{x}, r) = \boldsymbol{x} + B(0, r)$:

$$\begin{aligned}\mathsf{B}(\boldsymbol{x}, r) &= \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) < r\} && \text{by definition of open ball } B \\ &= \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\boldsymbol{y} - \boldsymbol{y}, \boldsymbol{y} - \boldsymbol{x}) < r\} && \text{by right result 1.} \\ &= \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{y} - \boldsymbol{x}) < r\} \\ &= \{\boldsymbol{u} + \boldsymbol{x} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\} && \text{let } \boldsymbol{u} \triangleq \boldsymbol{y} - \boldsymbol{x} \\ &= \boldsymbol{x} + \{\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\} \\ &= \boldsymbol{x} + B(0, r) && \text{by definition of open ball } B\end{aligned}$$

[12] Giles (2000) page 1 ⟨1.1 Definition⟩
[13] Giles (2000) page 2
[14] Giles (2000) page 2 ⟨1.2 Remarks⟩, Giles (1987) pages 22–26 ⟨2.4 Theorem, 2.11 Theorem⟩

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

$$\begin{aligned}
\mathsf{B}(\mathbb{0}, r) &= \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{y}) < r\} && \text{by definition of open ball } B\\
&= \left\{\boldsymbol{y} \in \boldsymbol{X} | \frac{1}{r}\mathsf{d}(\mathbb{0}, \boldsymbol{y}) < 1\right\}\\
&= \left\{\boldsymbol{y} \in \boldsymbol{X} | \frac{1}{r}\,\|\boldsymbol{y} - \mathbb{0}\| < 1\right\} && \text{by left hypothesis}\\
&= \left\{\boldsymbol{y} \in \boldsymbol{X} | \left\|\frac{1}{r}\boldsymbol{y} - \frac{1}{r}\mathbb{0}\right\| < 1\right\} && \text{by homogeneous property of } \|\cdot\| \text{ page 159}\\
&= \left\{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}\left(\frac{1}{r}\mathbb{0}, \frac{1}{r}\boldsymbol{y}\right) < 1\right\} && \text{by left hypothesis}\\
&= \{r\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < 1\} && \text{let } \boldsymbol{u} \triangleq \frac{1}{r}\boldsymbol{y}\\
&= r\,\{\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < 1\}\\
&= r\,B(\mathbb{0}, 1) && \text{by definition of open ball } B
\end{aligned}$$

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.

$$\begin{aligned}
\mathsf{d}(\boldsymbol{p}, \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \|\boldsymbol{p} - \lambda\boldsymbol{x} - (1-\lambda)\boldsymbol{y}\| && \text{by left hypothesis}\\
&= \left\|\underbrace{\lambda\boldsymbol{p} + (1-\lambda)\boldsymbol{p}}_{\boldsymbol{p}} - \lambda\boldsymbol{x} - (1-\lambda)\boldsymbol{y}\right\|\\
&= \|\lambda\boldsymbol{p} - \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{p} - (1-\lambda)\boldsymbol{y}\|\\
&\le \|\lambda\boldsymbol{p} - \lambda\boldsymbol{x}\| + \|(1-\lambda)\boldsymbol{p} - (1-\lambda)\boldsymbol{y}\| && \text{by subadditivity property of } \|\cdot\| \text{ page 159}\\
&= |\lambda|\,\|\boldsymbol{p} - \boldsymbol{x}\| + |1-\lambda|\,\|\boldsymbol{p} - \boldsymbol{y}\| && \text{by homogeneous property of } \|\cdot\| \text{ page 159}\\
&= \lambda\,\|\boldsymbol{p} - \boldsymbol{x}\| + (1-\lambda)\,\|\boldsymbol{p} - \boldsymbol{y}\| && \text{because } 0 \le \lambda \le 1\\
&\le \lambda r + (1-\lambda)r && \text{because } \boldsymbol{x}, \boldsymbol{y} \text{ are in the ball } \mathsf{B}(\boldsymbol{p}, r)\\
&= r
\end{aligned}$$

5. Proof that $\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) \iff -\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r)$ (symmetric):

$$\begin{aligned}
\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r) &\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{y}) < r\} && \text{by definition of open ball } B\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} |\, \|\boldsymbol{y} - \mathbb{0}\| < r\} && \text{by left hypothesis}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} |\, \|\boldsymbol{y}\| < r\}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} |\, \|(-1)(-\boldsymbol{y})\| < r\}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} |\, |-1|\,\|-\boldsymbol{y}\| < r\} && \text{by homogeneous property of } \|\cdot\| \text{ page 159}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} |\, \|-\boldsymbol{y} - \mathbb{0}\| < r\}\\
&\iff \boldsymbol{x} \in \{\boldsymbol{y} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, -\boldsymbol{y}) < r\} && \text{by left hypothesis}\\
&\iff \boldsymbol{x} \in \{-\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\} && \text{let } \boldsymbol{u} \triangleq -\boldsymbol{y}\\
&\iff \boldsymbol{x} \in (-\{\boldsymbol{u} \in \boldsymbol{X} | \mathsf{d}(\mathbb{0}, \boldsymbol{u}) < r\})\\
&\iff \boldsymbol{x} \in (-\mathsf{B}(\mathbb{0}, r))\\
&\iff -\boldsymbol{x} \in \mathsf{B}(\mathbb{0}, r)
\end{aligned}$$

✏

Theorem D.4 (page 160) demonstrates that if a metric d in a metric space $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d}\right)$ is generated by a norm, then the ball $\mathsf{B}(x, r)$ 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.

Norms generated by metrics

Every normed linear space is also a metric linear space (Theorem D.3 page 159). 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 provided by Theorem D.5 (page 162).

Lemma D.1. [15] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times}), \mathsf{d}\right)$ *be a* METRIC LINEAR SPACE. *Let* $\|x\| \triangleq \mathsf{d}(x, \mathbb{0})\ \forall x \in X$.

LEM

$$\underbrace{\mathsf{d}(x+z, y+z) = \mathsf{d}(x, y) \quad \forall x,y,z \in X}_{\text{TRANSLATION INVARIANT}} \implies \left\{\begin{array}{llcll} 1. & \|x\| & = & \|-x\| & \forall x \in X \quad \text{and} \\ 2. & \|x\| & = & 0 \iff x = 0 & \forall x \in X \quad \text{and} \\ 3. & \|x+y\| & \le & \|x\| + \|y\| & \forall x,y \in X \end{array}\right.$$

✎PROOF:

1. Proof that $\|x\| = \|-x\|$:

$$\begin{aligned} \|x\| &= \mathsf{d}(x, \mathbb{0}) && \text{by definition of } \|\cdot\| \\ &= \mathsf{d}(x - x, \mathbb{0} - x) && \text{by translation invariance hypothesis} \\ &= \mathsf{d}(\mathbb{0}, -x) \\ &= \|-x\| && \text{by definition of } \|\cdot\| \end{aligned}$$

2a. Proof that $\|x\| = 0 \implies x = 0$:

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

2b. Proof that $\|x\| = 0 \impliedby x = 0$:

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

3. Proof that $\|x + y\| \le \|x\| + \|y\|$:

$$\begin{aligned} \|x+y\| &= \mathsf{d}(x+y, \mathbb{0}) && \text{by definition of } \|\cdot\| \\ &= \mathsf{d}(x+y-y, \mathbb{0}-y) && \text{by translation invariance hypothesis} \\ &= \mathsf{d}(x, -y) \\ &\le \mathsf{d}(x, \mathbb{0}) + \mathsf{d}(\mathbb{0}, y) && \text{by property of metrics} \\ &= \mathsf{d}(x, \mathbb{0}) + \mathsf{d}(y, \mathbb{0}) && \text{by property of metrics} \\ &= \|x\| + \|y\| && \text{by definition of } \|\cdot\| \end{aligned}$$

✏

Theorem D.5. [16] *Let* $\left(X, +, \cdot, (\mathbb{F}, \dot{+}, \dot{\times})\right)$ *be a* LINEAR SPACE. *Let* $\mathsf{d}(x, y) \triangleq \|x - y\|\ \forall x, y \in X$.

THM

$$\left.\begin{array}{llll} 1. & \mathsf{d}(x+z, y+z) = \mathsf{d}(x, y) & \forall x,y,z \in X & \text{(TRANSLATION INVARIANT)} \quad \text{and} \\ 2. & \mathsf{d}(\alpha x, \alpha y) = |\alpha| \mathsf{d}(x, y) & \forall x,y \in X, \alpha \in \mathbb{F} & \text{(HOMOGENEOUS)} \end{array}\right\} \iff \|\cdot\| \text{ is a NORM}$$

✎PROOF:

[15] Oikhberg and Rosenthal (2007) page 599
[16] Bollobás (1999), page 21

1. Proof of $\Longrightarrow$ assertion:
   (a) Proof that $\|\cdot\|$ is *strictly positive*: This follows directly from the definition of d.
   (b) Proof that $\|\cdot\|$ is *nondegenerate*: This follows directly from Lemma D.1 (page 162).
   (c) Proof that $\|\cdot\|$ is *homogeneous*: This follows from the second left hypothesis.
   (d) Proof that $\|\cdot\|$ satisfies the *triangle-inequality*: This follows directly from Lemma D.1 (page 162).
2. Proof of $\Longleftarrow$ assertion:

$$\begin{aligned}
\mathsf{d}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z}) &= \|(\boldsymbol{x}+\boldsymbol{z})-(\boldsymbol{y}+\boldsymbol{z})\| && \text{by definition of } \mathsf{d}\\
&= \|\boldsymbol{x}-\boldsymbol{y}\|\\
&= \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) && \text{by definition of } \mathsf{d}\\
\mathsf{d}(\alpha\boldsymbol{x}, \alpha\boldsymbol{y}) &= \|(\alpha\boldsymbol{x})-(\alpha\boldsymbol{y})\| && \text{by definition of } \mathsf{d}\\
&= \|\alpha(\boldsymbol{x}-\boldsymbol{y})\|\\
&= |\alpha|\,\|\boldsymbol{x}-\boldsymbol{y}\| && \text{by definition of } \|\cdot\| \text{ page 159}\\
&= |\alpha|\mathsf{d}(\boldsymbol{x}, \boldsymbol{y}) && \text{by definition of } \mathsf{d}
\end{aligned}$$

✎

D.2 Metric spaces

D.2.1 Algebraic structure

Definition D.6. [17]

DEF

A function $\mathsf{d} \in \mathbb{R}^{\vdash X \times X}$ (Definition 1.6 page 6) *is a* ***quasi-metric*** *on a* SET X *if*

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

The pair (X,d) *is a* ***quasi-metric space*** *if* d *is a* QUASI-METRIC *on* X. *A* QUASI-METRIC *is also called an* ***asymmetric metric*** *and a* ***directed metric***.

Definition D.7. [18] *Let* X *be a set and* $\mathbb{R}^{\vdash}$ *the set of non-negative real numbers.*

DEF

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

1. $\mathsf{d}(x,y) \geq 0 \quad \forall x,y \in X$ (NON-NEGATIVE) *and*
2. $\mathsf{d}(x,y) = 0 \iff x = y \quad \forall x,y \in X$ (NONDEGENERATE) *and*
3. $\mathsf{d}(x,y) = \mathsf{d}(y,x) \quad \forall x,y \in X$ (SYMMETRIC) *and*
4. $\mathsf{d}(x,y) \leq \mathsf{d}(x,z) + \mathsf{d}(z,y) \quad \forall x,y,z \in X$ (SUBADDITIVE/TRIANGLE INEQUALITY).[19]

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

Actually, it is possible to significantly simplify the definition of a metric to an equivalent statement requiring only half as many conditions. These equivalent conditions (a "*characterization*") are stated in Theorem D.6 (next).

---

[17] Deza and Deza (2014) pages 6–7, Deza and Deza (2006) page 4, Wilson (1931a) page 675 ⟨§1.⟩, Ribeiro (1943), Kelly (1963) page 71 ⟨Introduction⟩, Patty (1967), Stoltenberg (1969) page 65, Grabiec et al. (2006) page 3 ⟨Introduction⟩, Euclid (circa 300BC) ⟨triangle inequality—Book I Proposition 20⟩

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

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

Theorem D.6 (metric characterization). [20] *Let* d *be a function in* $(\mathbb{R}^{\vdash})^{X\times X}$.

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

✎PROOF:

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

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

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

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

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

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

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

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

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

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

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

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

✎

The *triangle inequality* property stated in the definition of metrics (Definition D.7 page 163) axiomatically endows a metric with an upper bound. Lemma D.2 (next) demonstrates that there is a complementary lower bound similar in form to the triangle-inequality upper bound.

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

Lemma D.2. [21] *Let* (X, d) *be a* METRIC SPACE *(Definition D.7 page 163). Let* $|\cdot|$ *be the* ABSOLUTE VALUE *function (Definition 1.24 page 10).*

LEM

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

✎PROOF:

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

$$\begin{aligned}
|\mathsf{d}(x,p) - \mathsf{d}(p,y)| &\le |\mathsf{d}(x,y) + \mathsf{d}(y,p) - \mathsf{d}(p,y)| && \text{by } \textit{subadditive} \text{ property (Definition D.7 page 163)}\\
&= |\mathsf{d}(x,y) + \mathsf{d}(p,y) - \mathsf{d}(p,y)| && \text{by } \textit{symmetry} \text{ property of metrics (Definition D.7 page 163)}\\
&= |\mathsf{d}(x,y) + 0|\\
&= \mathsf{d}(x,y) && \text{by } \textit{non-negative} \text{ property of metrics (Definition D.7 page 163)}
\end{aligned}$$

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

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

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

$$\begin{aligned}
|\mathsf{d}(x,p) - \mathsf{d}(p,y)| &\le 0 && \text{by left hypothesis}\\
&\le \mathsf{d}(x,y) && \text{by } \textit{non-negative} \text{ property of metrics (Definition D.7 page 163)}
\end{aligned}$$

✏

The *triangle inequality* property stated in the definition of metrics (Definition D.7 page 163) can be extended from two to any finite number of metrics (next).

Proposition D.1. [22] *Let* (X, d) *be a* METRIC SPACE *(Definition D.7 page 163) and* $(\!(x_n \in X)\!)_1^N$ *an* N-TUPLE *(Definition 1.11 page 7) on* X.

PRP

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

✎PROOF: Proof by induction:

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

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

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

$$\begin{aligned}
\mathsf{d}(x_1, x_{N+1}) &\le \mathsf{d}(x_1, x_N) + \mathsf{d}(x_N, x_{N+1}) && \text{by } \textit{subadditive} \text{ property (Definition D.7 page 163)}\\
&\le \left(\sum_{n=1}^{N-1} \mathsf{d}(x_N, x_{N+1}) \right) + \mathsf{d}(x_N, x_{N+1}) && \text{by } \{N \text{ case}\} \text{ hypothesis}\\
&= \sum_{n=1}^{N} \mathsf{d}(x_n, x_{n+1})
\end{aligned}$$

✏

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

Definition D.8. [23] *Let X be a set and $\mathsf{d} \in \mathbb{R}^{X\times X}$. The function d is the **discrete metric** on $\mathbb{R}^{X\times X}$ if*

$$\mathsf{d}(x,y) \triangleq \left\{\begin{array}{ll} 0 & \text{if } x=y \\ 1 & \text{otherwise} \end{array}\right\} \qquad \forall x,y\in X$$

Definition D.9 (usual metric). [24] *Let $|\cdot| \in \mathbb{R}^{\vdash R}$ be an* ABSOLUTE VALUE *function on a* RING *R.*

DEF: *The function* $\mathsf{d}(x,y) \triangleq |x-y|$ *is a* METRIC *on $\mathbb{R}$, called the **usual metric**.*

Definition D.10. *Let X be a set and $\mathsf{d} \in \mathbb{R}^{\mathbb{R}^N}$. The function d is the **Euclidean metric** on $\mathbb{R}^N$ if*

$$\mathsf{d}\left((x_1,x_2,\cdots,x_N),(y_1,y_2,\cdots,y_N)\right) \triangleq \sqrt{\sum_{n=1}^{N}(x_n-y_n)^2} \qquad \forall (x_1,x_2,\cdots,x_N),(y_1,y_2,\cdots,y_N)\in\mathbb{R}^N$$

D.2.2 Metric preserving functions

Definition D.11. [25] *Let $\mathbb{M}$ be the set of all* METRIC SPACES *(Definition D.7 page 163) on a set X.*

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

Theorem D.7 (necessary conditions). [26] *Let $\mathcal{R}\phi$ be the* RANGE *of a function ϕ.*

THM:
$$\left\{\begin{array}{l} \phi \text{ is a} \\ \text{METRIC PRESERVING FUNCTION} \\ \text{(Definition D.11 page 166)} \end{array}\right\} \implies \left\{\begin{array}{lll} 1. & \phi^{-1}(0)=\{0\} & \text{and} \\ 2. & \mathcal{R}\phi \subseteq \mathbb{R}^{\vdash} & \text{and} \\ 3. & \phi(x+y)\le\phi(x)+\phi(y) & (\phi \text{ is SUBADDITIVE}) \end{array}\right\}$$

✎PROOF:

1. Proof that ϕ is a *metric preserving function* $\implies \phi^{-1}(0)=\{0\}$:
   (a) Suppose that the statement is not true and $\phi^{-1}(0)=\{0,a\}$.
   (b) Then $\phi(a)=0$ and for some x,y such that $x\neq y$ and $\mathsf{d}(x,y)=a$ we have

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

   (c) But this contradicts the original hypothesis, and so it must be that $\phi^{-1}(0)=\{0\}$.
2. Proof that $\mathcal{R}\phi \subseteq \mathbb{R}^{\vdash}$:

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

3. Proof that ϕ is a metric preserving function $\implies \phi$ is *subadditive*:

[23] Busemann (1955) page 4 ⟨COMMENTS ON THE AXIOMS⟩, Giles (1987), page 13, Copson (1968), page 24, Khamsi and Kirk (2001) page 19 ⟨Example 2.1⟩
[24] Davis (2005) page 16
[25] Vallin (1999), page 849 ⟨Definition 1.1⟩, Corazza (1999), page 309, Deza and Deza (2009) page 80
[26] Corazza (1999), page 310 ⟨Proposition 2.1⟩, Deza and Deza (2009) page 80

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

$$\begin{aligned}
\phi(x) + \phi(y) &= \phi(|(x + y) - x|) + \phi(|x - 0|) && \text{by definition of } |\cdot| \\
&= (\phi \circ \mathsf{d})(x + y, x) + (\phi \circ \mathsf{d})(x, 0) && \text{by definition of } \mathsf{d} \\
&\geq (\phi \circ \mathsf{d})(x + y, 0) && \text{by left hypothesis and Definition D.7 page 163} \\
&= \phi(|(x + y) - 0|) && \text{by definition of } \mathsf{d} \\
&= \phi(x + y) && \text{because } x, y \in \mathbb{R}^{\vdash}
\end{aligned}$$

✑

Theorem D.8 (next theorem) presents some sufficient conditions for a function to be metric preserving.

Theorem D.8 (sufficient conditions).[27] *Let ϕ be a function in $\mathbb{R}^{\mathbb{R}}$.*

THM

$$\left\{\begin{array}{llll}
1. & x \geq y \implies \phi(x) \geq \phi(y) \quad \forall x,y \in \mathbb{R}^{\vdash} & \text{(ISOTONE)} & \textit{and} \\
2. & \phi(0) = 0 & & \textit{and} \\
3. & \phi(x + y) \leq \phi(x) + \phi(y) \quad \forall x,y \in \mathbb{R}^{\vdash} & \text{(SUBADDITIVE)} &
\end{array}\right\} \implies \left\{\begin{array}{l}
\phi \textit{ is a } \text{METRIC} \\
\text{PRESERVING FUNCTION} \\
\text{(Definition D.11 page 166).}
\end{array}\right\}$$

✎PROOF:

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

$$\begin{aligned}
\phi \circ \mathsf{d}(x, y) = 0 &\implies \mathsf{d}(x, y) = 0 && \text{by } \phi \text{ hypothesis 2} \\
&\implies x = y && \text{by } \textit{nondegenerate} \text{ property page 163}
\end{aligned}$$

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

$$\begin{aligned}
\phi \circ \mathsf{d}(x, y) &= \phi \circ \mathsf{d}(x, x) && \text{by } x = y \text{ hypothesis} \\
&= \phi(0) && \text{by } \textit{nondegenerate} \text{ property page 163} \\
&= 0 && \text{by } \phi \text{ hypothesis 2}
\end{aligned}$$

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

$$\begin{aligned}
\phi \circ \mathsf{d}(x, y) &\leq \phi(\mathsf{d}(x, z) + \mathsf{d}(z, y)) && \text{by } \phi \text{ hypothesis 1 and triangle inequality page 163} \\
&\leq \phi(\mathsf{d}(z, x) + \mathsf{d}(z, y)) && \text{by symmetric property of } \mathsf{d} \text{ page 163} \\
&\leq \phi \circ \mathsf{d}(z, x) + \phi \circ \mathsf{d}(z, y) && \text{by } \phi \text{ hypothesis 3}
\end{aligned}$$

✑

Example D.3 (α-scaled metric/dilated metric).[28] Let (X, d) be a *metric space* (Definition D.7 page 163).

EX $\phi(x) \triangleq \alpha x, \ \alpha \in \mathbb{R}^{+}$ is a *metric preserving function* (Figure D.1 page 168 (A))

✎PROOF: The proofs for Example D.3–Example D.8 (page 168) follow from Theorem D.8 (page 167). ✑

Example D.4 (power transform metric/snowflake transform metric).[29] Let (X, d) be a *metric space* (Definition D.7 page 163).

EX $\phi(x) \triangleq x^{\alpha}, \ \alpha \in (0 : 1]$, is a *metric preserving function* (see Figure D.1 page 168 (B))

[27] Corazza (1999) ⟨Proposition 2.3⟩, Deza and Deza (2009) page 80, Kelley (1955) page 131 ⟨Problem C⟩
[28] Deza and Deza (2006) page 44
[29] Deza and Deza (2009) page 81, Deza and Deza (2006) page 45

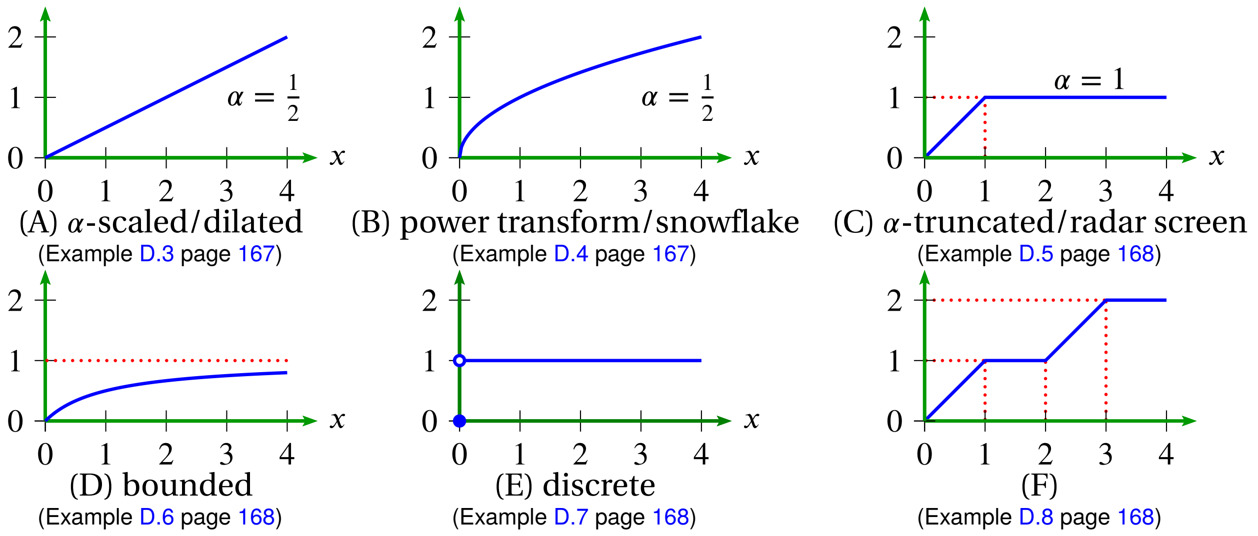


Figure D.1: metric preserving functions

Example D.5 (α-truncated metric/radar screen metric). [30] Let (X, d) be a *metric space* (Definition D.7 page 163).

EX	$\phi(x) \triangleq \min\{\alpha,\ x\},\ \alpha \in \mathbb{R}^+$ is a *metric preserving function* (see Figure D.1 page 168 (C)).

✎PROOF: by Theorem D.8 (page 167). ⇒

Example D.6 (bounded metric). [31] Let (X, d) be a *metric space* (Definition D.7 page 163).

EX	$\phi(x) \triangleq \dfrac{x}{1+x}$ is a *metric preserving function* (see Figure D.1 page 168 (D)).

✎PROOF: by Theorem D.8 (page 167). ⇒

Example D.7 (discrete metric preserving function). [32] Let ϕ be a function in $\mathbb{R}^{\mathbb{R}}$.

EX	$\phi(x) \triangleq \begin{cases} 0 & \text{for } x \le 0 \\ 1 & \text{otherwise} \end{cases}$ is a *metric preserving function* (see Figure D.1 page 168 (E)).

✎PROOF: by Theorem D.8 page 167. ⇒

Example D.8. Let ϕ be a function in $\mathbb{R}^{\mathbb{R}}$.

EX	$\phi(x) \triangleq \begin{cases} x & \text{for } 0 \le x < 1, & 1 & \text{for } 1 \le x \le 2, \\ x-1 & \text{for } 2 < x < 3, & 2 & \text{for } x \ge 3 \end{cases}$ is a *metric preserving function* (see Figure D.1 page 168 (F)).

✎PROOF: by Theorem D.8 page 167. ⇒

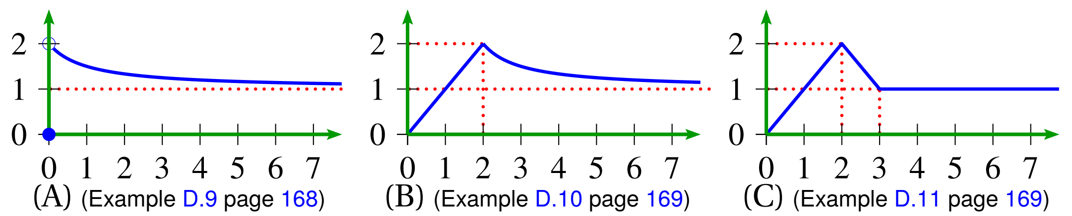


Figure D.2: *non-monotone metric preserving functions*

Example D.9. [33]

[30] Giles (1987), page 33, Deza and Deza (2006) pages 242–243
[31] Vallin (1999), page 849, Aliprantis and Burkinshaw (1998) page 39
[32] Corazza (1999), page 311
[33] Greenhoe (2015), pages 10–11 ⟨Theorem 4.16⟩

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

EX	$\phi(x) \triangleq \left\{ \begin{array}{ll} 0 & \text{for } x = 0 \\ 1 + \frac{1}{x+1} & \text{for } x > 0 \end{array} \right\}$ is a *metric preserving function* (see Figure D.2 page 168 (A)).

Example D.10. [34] Let ϕ be a function in $\mathbb{R}^{\mathbb{R}}$.

EX	$\phi(x) \triangleq \left\{ \begin{array}{ll} x & \text{for } x \leq 2 \\ 1 + \frac{1}{x-1} & \text{for } x > 2 \end{array} \right\}$ is a *metric preserving function* (see Figure D.2 page 168 (B)).

Example D.11. Let ϕ be a function in $\mathbb{R}^{\mathbb{R}}$.

EX	$\phi(x) \triangleq \left\{ \begin{array}{ll} x & \text{for } 0 \leq x \leq 2 \\ -x+4 & \text{for } 2 < x < 3 \\ 1 & \text{for } x \geq 3 \end{array} \right\}$ is a *metric preserving function* (see Figure D.2 page 168 (C)).

D.2.3 Product metrics

Theorem D.9 (Fréchet product metric). [35] *Let X be a set.*

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

✎PROOF:

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

$$\begin{aligned} \mathsf{d}(x, y) &= \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(x, y) && \text{by definition of } \mathsf{d} \\ &= \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(x, x) && \text{by left hypothesis} \\ &= \sum_{n=1}^{N} 0 && \text{by } \textit{nondegenerate} \text{ property of metrics (Definition D.7 page 163)} \\ &= 0 \end{aligned}$$

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

$$\begin{aligned} 0 &= \mathsf{d}(x, y) && \text{by right hypothesis} \\ &= \sum_{n=1}^{N} \alpha_n \mathsf{p}_n(x, y) && \text{by definition of } \mathsf{d} \\ \implies & \mathsf{p}_n(x, y) = 0 \quad \forall x, y \in \boldsymbol{X} && \text{by metric properties page 163} \\ \implies & x = y \quad \forall x, y \in \boldsymbol{X} && \text{by non-degenerate property of metrics page 163} \end{aligned}$$

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

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

$$\begin{aligned}
\mathsf{d}(x,y) &= \sum_{n=1}^{N} \alpha_n\, \mathsf{p}_n(x,y) && \text{by definition of } \mathsf{d}\\
&\le \sum_{n=1}^{N} \alpha_n \Big[\mathsf{p}_n(x,z) + \mathsf{p}_n(z,y)\Big] && \text{by \textit{subadditive} property (Definition D.7 page 163)}\\
&= \sum_{n=1}^{N} \alpha_n \Big[\mathsf{p}_n(z,x) + \mathsf{p}_n(z,y)\Big] && \text{by \textit{symmetry} property (Definition D.7 page 163)}\\
&= \sum_{n=1}^{N} \alpha_n\, \mathsf{p}_n(z,x) + \sum_{n=1}^{N} \alpha_n\, \mathsf{p}_n(z,y)\\
&= \mathsf{d}(z,x) + \mathsf{d}(z,y) && \text{by definition of } \mathsf{d}
\end{aligned}$$

✏

Theorem D.10 (Power mean metrics). *Let X be a set. Let $\langle x_n \in X \rangle_1^N$ and $\langle y_n \in X \rangle_1^N$ be N-tuples on X.*

THM

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

Moreover, if $r = \infty$, then $\mathsf{d}\big(\langle x_n\rangle, \langle y_n\rangle\big) = \max_{n=1,\ldots,N} \mathsf{p}\,(x_n, y_n)$.

✎PROOF:

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

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

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

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

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

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

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

$$\begin{aligned}
&\mathsf{d}\left(\langle\!\langle x_n\rangle\!\rangle, \langle\!\langle y_n\rangle\!\rangle\right)\\
&\triangleq \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(x_n, y_n)\right)^{\frac{1}{r}} && \text{by definition of } \mathsf{d}\\
&\le \left(\sum_{n=1}^{N} \lambda_n \left[\mathsf{p}(z_n, x_n) + \mathsf{p}(z_n, y_n)\right]^r\right)^{\frac{1}{r}} && \text{by \textit{subaddtive} property (Definition D.7 page 163)}\\
&= \left(\sum_{n=1}^{N} \left[\lambda_n^{\frac{1}{r}} \mathsf{p}(z_n, x_n) + \lambda_n^{\frac{1}{r}} \mathsf{p}(z_n, y_n)\right]^r\right)^{\frac{1}{r}} && \text{by \textit{subaddtive} property (Definition D.7 page 163)}\\
&\le \left(\sum_{n=1}^{N} \left[\lambda_n^{\frac{1}{r}} \mathsf{p}(z_n, x_n)\right]^r\right)^{\frac{1}{r}} + \left(\sum_{n=1}^{N} \left[\lambda_n^{\frac{1}{r}} \mathsf{p}(z_n, y_n)\right]^r\right)^{\frac{1}{r}} && \text{by \textit{Minkowski's inequality}}\\
&\le \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(z_n, x_n)\right)^{\frac{1}{r}} + \left(\sum_{n=1}^{N} \lambda_n \mathsf{p}^r(z_n, y_n)\right)^{\frac{1}{r}}\\
&\triangleq \mathsf{d}\left(\langle\!\langle z_n\rangle\!\rangle, \langle\!\langle x_n\rangle\!\rangle\right) + \mathsf{d}\left(\langle\!\langle z_n\rangle\!\rangle, \langle\!\langle y_n\rangle\!\rangle\right) && \text{by definition of } \mathsf{d}
\end{aligned}$$

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

   (a) Proof that $\mathsf{d}\left(\langle\!\langle x_n\rangle\!\rangle, \langle\!\langle y_n\rangle\!\rangle\right) = \max \langle\!\langle x_n\rangle\!\rangle$: by Theorem D.14 page 175

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

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

(c) Proof that $\llparenthesis x_n \rrparenthesis = \llparenthesis y_n \rrparenthesis \iff \mathsf{d}(\llparenthesis x_n \rrparenthesis, \llparenthesis y_n \rrparenthesis) = 0$:

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

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

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

6. And so by Theorem D.6 (page 164), d is a *metric* for $r \in [1 : \infty]$.

⇨

D.3 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 thought to be from Harald Bohr (1887–1951), Danish mathematician and pictured to the left. [36]

D.3.1 Summation

Definition D.12. [37] *Let $+$ be an addition operator on a tuple $\llparenthesis x_n \rrparenthesis_m^N$.*

DEF

*The **summation** of $\llparenthesis x_n \rrparenthesis$ from index m to index N with respect to $+$ is*

$$\sum_{n=m}^{N} x_n \triangleq \begin{cases} 0 & \text{for } N < m\\ \left(\displaystyle\sum_{n=m}^{N-1} x_n\right) + x_N & \text{for } N \ge m \end{cases}$$

Theorem D.11 (Generalized associative property). [38] *Let $+$ be an addition operator on a tuple $\llparenthesis x_n \rrparenthesis_m^N$.*

---

[36] quote: Hardy (1929), page 64
image: http://www-history.mcs.st-andrews.ac.uk/PictDisplay/Bohr_Harald.html

[37] Berberian (1961) page 8 ⟨Definition I.3.1⟩, Fourier (1820) page 280 ⟨"$\sum$" notation⟩

[38] Berberian (1961) pages 9–10 ⟨Theorem I.3.1⟩

THM

$+$ *is* ASSOCIATIVE $\implies$

$$\underbrace{\sum_{n=m}^{L} x_n + \left(\sum_{n=L+1}^{M} x_n + \sum_{n=M+1}^{N} x_n\right) = \left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \sum_{n=M+1}^{N} x_n \qquad \text{for } m < L < M \le N}_{\sum_{n=m}^{N} \text{ is ASSOCIATIVE}}$$

✎PROOF:

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:

$$\begin{aligned}
\sum_{n=m}^{m+2} x_n &= \left(\sum_{n=m}^{m+1} x_n\right) + x_{m+2} && \text{by Definition D.12 page 172}\\
&= \left(x_m + x_{m+1}\right) + x_{m+2} && \text{by item (3)}\\
&= x_m + \left(x_{m+1} + x_{m+2}\right) && \text{by left hypothesis}
\end{aligned}$$

5. Proof that N case $\implies$ $N+1$ case:

$$\begin{aligned}
\sum_{n=m}^{N+1} x_n &= \underbrace{\left(\sum_{n=m}^{N} x_n\right)}_{\text{associative}} + x_{N+1} && \text{by Definition D.12 page 172}\\
&= \left(\sum_{n=m}^{L} x_n + \left(\sum_{n=L+1}^{M} x_n + \sum_{n=M+1}^{N} x_n\right)\right) + x_{N+1}\\
&= \left(\left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \sum_{n=M+1}^{N} x_n\right) + x_{N+1}\\
&= \left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \left(\sum_{n=M+1}^{N} x_n + x_{N+1}\right)\\
&= \left(\sum_{n=m}^{L} x_n + \sum_{n=L+1}^{M} x_n\right) + \left(\sum_{n=M+1}^{N+1} x_n\right)
\end{aligned}$$

✏

D.3.2 Convexity

Definition D.13. [39]

[39] Simon (2011) page 2, Barvinok (2002) page 2, Bollobás (1999), page 3, Jensen (1906), page 176

DEF
A function $\mathsf{f} \in \mathbb{R}^{\mathbb{R}}$ *is **convex** if*

$$\mathsf{f}\left(\lambda x + [1-\lambda]y\right) \le \lambda\mathsf{f}(x) + (1-\lambda)\,\mathsf{f}(y) \qquad \forall x,y \in \mathbb{R} \text{ and } \forall \lambda \in (0:1)$$

A function $\mathsf{g} \in \mathbb{R}^{\mathbb{R}}$ *is **strictly convex** if*

$$\mathsf{g}\left(\lambda x + [1-\lambda]y\right) = \lambda\mathsf{g}(x) + (1-\lambda)\,\mathsf{g}(y) \qquad \forall x,y \in D,\ x \ne y,\ \text{and } \forall \lambda \in (0:1)$$

A function $\mathsf{f} \in \mathbb{R}^{\mathbb{R}}$ *is **concave** if* $-\mathsf{f}$ *is* CONVEX.
A function $\mathsf{f} \in \mathbb{R}^{\mathbb{R}}$ *is **affine** if* f *is* CONVEX *and* CONCAVE.

Theorem D.12 (Jensen's Inequality). [40] *Let* $\mathsf{f} \in \mathbb{R}^{\mathbb{R}}$ *be a function.*

THM
$$\left\{\begin{array}{ll} 1. & \mathsf{f} \text{ is CONVEX} \quad \text{and} \\ 2. & \displaystyle\sum_{n=1}^{N} \lambda_n = 1 \end{array}\right\} \implies \left\{ \mathsf{f}\left(\sum_{n=1}^{N} \lambda_n x_n\right) \le \sum_{n=1}^{N} \lambda_n \mathsf{f}(x_n) \qquad \forall x_n \in D,\ N \in \mathbb{N} \right\}$$

D.3.3 Power means

Definition D.14. [41]

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* $\displaystyle\sum_{n=1}^{N} \lambda_n = 1$.

Lemma D.3. [42] *Let* $\mathsf{M}_\phi\left((\!(x_n)\!)\right)$ *be the* $(\!(\lambda_n)\!)_1^N$ *weighted* ϕ*-mean and* $\mathsf{M}_\psi\left((\!(x_n)\!)\right)$ *the* $(\!(\lambda_n)\!)_1^N$ *weighted* ψ*-mean of a tuple* $(\!(x_n)\!)_1^N$.

LEM
$$\begin{array}{llllll}
\phi\psi^{-1} \text{ is CONVEX} & \text{and} & \phi \text{ is INCREASING} & \implies & \mathsf{M}_\phi\left((\!(x_n)\!)\right) \ge \mathsf{M}_\psi\left((\!(x_n)\!)\right) \\
\phi\psi^{-1} \text{ is CONVEX} & \text{and} & \phi \text{ is DECREASING} & \implies & \mathsf{M}_\phi\left((\!(x_n)\!)\right) \le \mathsf{M}_\psi\left((\!(x_n)\!)\right) \\
\phi\psi^{-1} \text{ is CONCAVE} & \text{and} & \phi \text{ is INCREASING} & \implies & \mathsf{M}_\phi\left((\!(x_n)\!)\right) \le \mathsf{M}_\psi\left((\!(x_n)\!)\right) \\
\phi\psi^{-1} \text{ is CONCAVE} & \text{and} & \phi \text{ is DECREASING} & \implies & \mathsf{M}_\phi\left((\!(x_n)\!)\right) \ge \mathsf{M}_\psi\left((\!(x_n)\!)\right)
\end{array}$$

One of the most well known inequalities in mathematics is *Minkowski's Inequality*. In 1946, H.P. Mulholland submitted a result that generalizes Minkowski's Inequality to an equal weighted ϕ-mean.[43] And Milovanović and Milovanovć (1979) generalized this even further to a *weighted* ϕ-mean (next).

Theorem D.13. [44] *Let* ϕ *be a function in* $\mathbb{R}^{\mathbb{R}}$.

THM
$$\left\{\begin{array}{llll} 1. & \phi \text{ is CONVEX} \quad \text{and} & 2. & \phi \text{ is STRICTLY MONOTONE} \quad \text{and} \\ 3. & \phi(0) = 0 \quad \text{and} & 4. & \log\circ\phi\circ\exp \text{ is CONVEX} \end{array}\right\}$$

$$\implies \left\{ \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(x_n + y_n)\right) \le \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(x_n)\right) + \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(y_n)\right) \right\}$$

Definition D.15. [45] *Let* $\mathsf{M}_{\phi(x;p)}\left((\!(x_n)\!)\right)$ *be the* $(\!(\lambda_n)\!)_1^N$ *weighted* ϕ*-mean of a* NON-NEGATIVE *tuple* $(\!(x_n)\!)_1^N$.

---

[40] Mitrinović et al. (2010) page 6, Bollobás (1999) page 3, Jensen (1906) pages 179–180
[41] Bollobás (1999) page 5
[42] Pečarić et al. (1992) page 107, Bollobás (1999) page 5, Hardy et al. (1952) page 75
[43] Minkowski (1910) page 115, Mulholland (1950), Hardy et al. (1952) ⟨Theorem 24⟩, Tolsted (1964) page 7, Maligranda (1995) page 258, Carothers (2000), page 44, Bullen (2003) page 179
[44] Milovanović and Milovanović (1979), Bullen (2003) page 306 ⟨Theorem 9⟩
[45] Bullen (2003) page 175, Bollobás (1999) page 6

DEF *A mean $\mathsf{M}_{\phi(x;p)}\left(\lparen x_n \rparen\right)$ is a **power mean** with parameter p if $\phi(x) \triangleq x^p$. That is,*

$$\mathsf{M}_{\phi(x;p)}\left(\lparen x_n \rparen\right) = \left(\sum_{n=1}^{N} \lambda_n (x_n)^p\right)^{\frac{1}{p}}$$

Theorem D.14. [46] *Let $\mathsf{M}_{\phi(x;p)}\left(\lparen x_n \rparen\right)$ be the* POWER MEAN *with parameter p of an N-tuple $\lparen x_n \rparen_1^N$ in which the elements are* NOT *all equal.*

THM

$$\mathsf{M}_{\phi(x;p)}\left(\lparen x_n \rparen\right) \triangleq \left(\sum_{n=1}^{N} \lambda_n (x_n)^p\right)^{\frac{1}{p}} \quad \text{is CONTINUOUS and STRICTLY MONOTONE in } \mathbb{R}^*.$$

$$\mathsf{M}_{\phi(x;p)}\left(\lparen x_n \rparen\right) = \begin{cases} \max\limits_{n=1,2,\ldots,N} \lparen x_n \rparen & \text{for } p = +\infty \\ \prod\limits_{n=1}^{N} x_n^{\lambda_n} & \text{for } p = 0 \\ \min\limits_{n=1,2,\ldots,N} \lparen x_n \rparen & \text{for } p = -\infty \end{cases}$$

✎PROOF:

1. Proof that $M_{\phi(x;p)}$ is *strictly monotone* in p:
   (a) Let p and s be such that $-\infty < p < s < \infty$.
   (b) Let $\phi_p \triangleq x^p$ and $\phi_s \triangleq x^s$. Then $\phi_p\phi_s^{-1} = x^{\frac{p}{s}}$.
   (c) The composite function $\phi_p\phi_s^{-1}$ is *convex* or *concave* depending on the values of p and s:

	$p < 0$ (ϕ_p decreasing)	$p > 0$ (ϕ_p increasing)
$s < 0$	*convex*	(not possible)
$s > 0$	*convex*	*concave*

   (d) Therefore by Lemma D.3 (page 174),
$$-\infty < p < s < \infty \implies \mathsf{M}_{\phi(x;p)}\left(\lparen x_n \rparen\right) < \mathsf{M}_{\phi(x;s)}\left(\lparen x_n \rparen\right).$$

2. Proof that $\mathsf{M}_{\phi(x;p)}$ is continuous in p for $p \in \mathbb{R}\backslash 0$: The sum of continuous functions is continuous. For the cases of $p \in \{-\infty, 0, \infty\}$, see the items that follow.

3. Lemma: $\mathsf{M}_{\phi(x;-p)}\left(\lparen x_n \rparen\right) = \left\{\mathsf{M}_{\phi(x;p)}\left(\lparen x_n^{-1} \rparen\right)\right\}^{-1}$. Proof:

$$\begin{aligned} \left\{\mathsf{M}_{\phi(x;p)}\left(\lparen x_n^{-1} \rparen\right)\right\}^{-1} &= \left\{\left(\sum_{n=1}^{N} \lambda_n (x_n^{-1})^p\right)^{\frac{1}{p}}\right\}^{-1} && \text{by definition of } \mathsf{M}_\phi \\ &= \left(\sum_{n=1}^{N} \lambda_n (x_n)^{-p}\right)^{\frac{1}{-p}} \\ &= \mathsf{M}_{\phi(x;-p)}\left(\lparen x_n \rparen\right) && \text{by definition of } \mathsf{M}_\phi \end{aligned}$$

4. Proof that $\lim\limits_{p\to\infty} \mathsf{M}_\phi\left(\lparen x_n \rparen\right) = \max\limits_{n\in\mathbb{Z}} \lparen x_n \rparen$:
   (a) Let $x_m \triangleq \max\limits_{n\in\mathbb{Z}} \lparen x_n \rparen$

[46] Bullen (2003) pages 175–177 ⟨see also page 203⟩, Bollobás (1999) pages 6–8, Bullen (1990) page 250, Besso (1879), Bienaymé (1840) page 68, Brenner (1985) page 160

(b) Note that $\lim_{p\to\infty} \mathsf{M}_\phi \le \max_{n\in\mathbb{Z}} ⦇x_n⦈$ because

$$\begin{aligned}
\lim_{p\to\infty} \mathsf{M}_\phi\left(⦇x_n⦈\right) &= \lim_{p\to\infty}\left(\sum_{n=1}^{N}\lambda_n x_n^p\right)^{\frac{1}{p}} && \text{by definition of } \mathsf{M}_\phi\\
&\le \lim_{p\to\infty}\left(\sum_{n=1}^{N}\lambda_n x_m^p\right)^{\frac{1}{p}} && \text{by definition of } x_m \text{ in item (4a) and because } \phi(x)\triangleq x^p \text{ and } \phi^{-1} \text{ are both increasing or both decreasing}\\
&= \lim_{p\to\infty}\left(x_m^p \underbrace{\sum_{n=1}^{N}\lambda_n}_{1}\right)^{\frac{1}{p}} && \text{because } x_m \text{ is a constant}\\
&= \lim_{p\to\infty}\left(x_m^p\cdot 1\right)^{\frac{1}{p}}\\
&= x_m\\
&= \max_{n\in\mathbb{Z}} ⦇x_n⦈ && \text{by definition of } x_m \text{ in item (4a)}
\end{aligned}$$

(c) But also note that $\lim_{p\to\infty} \mathsf{M}_\phi \ge \max_{n\in\mathbb{Z}} ⦇x_n⦈$ because

$$\begin{aligned}
\lim_{p\to\infty} \mathsf{M}_\phi\left(⦇x_n⦈\right) &= \lim_{p\to\infty}\left(\sum_{n=1}^{N}\lambda_n x_n^p\right)^{\frac{1}{p}} && \text{by definition of } \mathsf{M}_\phi\\
&\ge \lim_{p\to\infty}\left(w_m x_m^p\right)^{\frac{1}{p}} && \text{by definition of } x_m \text{ in item (4a) and because } \phi(x)\triangleq x^p \text{ and } \phi^{-1} \text{ are both increasing or both decreasing}\\
&= \lim_{p\to\infty} w_m^{\frac{1}{p}} x_m^{\frac{p}{p}}\\
&= x_m\\
&= \max_{n\in\mathbb{Z}} ⦇x_n⦈ && \text{by definition of } x_m \text{ in item (4a)}
\end{aligned}$$

(d) Combining items (b) and (c) we have $\lim_{p\to\infty} \mathsf{M}_\phi = \max_{n\in\mathbb{Z}} ⦇x_n⦈$.

5. Proof that $\lim_{p\to-\infty} \mathsf{M}_\phi\left(⦇x_n⦈\right) = \min_{n\in\mathbb{Z}} ⦇x_n⦈$:

$$\begin{aligned}
\lim_{p\to-\infty} \mathsf{M}_{\phi(x;p)}\left(⦇x_n⦈\right) &= \lim_{p\to\infty} \mathsf{M}_{\phi(x;-p)}\left(⦇x_n⦈\right) && \text{by change of variable } p\\
&= \lim_{p\to\infty}\left\{\mathsf{M}_{\phi(x;p)}\left(⦇x_n^{-1}⦈\right)\right\}^{-1} && \text{by Lemma in item (3) page 175}\\
&= \lim_{p\to\infty}\frac{1}{\mathsf{M}_{\phi(x;p)}\left(⦇x_n^{-1}⦈\right)}\\
&= \frac{\lim_{p\to\infty} 1}{\lim_{p\to\infty}\mathsf{M}_{\phi(x;p)}\left(⦇x_n^{-1}⦈\right)} && \text{by property of lim [47]}\\
&= \frac{1}{\max_{n\in\mathbb{Z}} ⦇x_n^{-1}⦈} && \text{by item (4)}\\
&= \frac{1}{\left(\min_{n\in\mathbb{Z}} ⦇x_n⦈\right)^{-1}}\\
&= \min_{n\in\mathbb{Z}} ⦇x_n⦈
\end{aligned}$$

6. Proof that $\lim_{p\to 0} \mathsf{M}_\phi(\lbag x_n \rbag) = \prod_{n=1}^{N} x_n^{\lambda_n}$:

$$\begin{aligned}
\lim_{p\to0}\mathsf{M}_\phi(\lbag x_n\rbag) &= \lim_{p\to0}\exp\left\{\ln\left\{\mathsf{M}_\phi(\lbag x_n\rbag)\right\}\right\} \\
&= \lim_{p\to0}\exp\left\{\ln\left\{\left(\sum_{n=1}^{N}\lambda_n(x_n^p)\right)^{\frac{1}{p}}\right\}\right\} && \text{by definition of } \mathsf{M}_\phi \\
&= \exp\left\{\frac{\frac{\partial}{\partial p}\ln\left(\sum_{n=1}^{N}\lambda_n(x_n^p)\right)}{\frac{\partial}{\partial p}p}\right\}_{p=0} && \text{by l'Hôpital's rule[48]} \\
&= \exp\left\{\frac{\sum_{n=1}^{N}\lambda_n\frac{\partial}{\partial p}(x_n^p)}{\sum_{n=1}^{N}\lambda_n(x_n^p)}\right\}_{p=0} = \exp\left\{\frac{\sum_{n=1}^{N}\lambda_n\frac{\partial}{\partial p}\exp\left(\ln(x_n^p)\right)}{\sum_{n=1}^{N}\lambda_n}\right\}_{p=0} \\
&= \exp\left\{\frac{\sum_{n=1}^{N}\lambda_n\frac{\partial}{\partial p}\exp\left(r\ln(x_n)\right)}{1}\right\}_{p=0} = \exp\left\{\sum_{n=1}^{N}\lambda_n\frac{\partial}{\partial p}\exp\left(p\ln(x_n)\right)\right\}_{p=0} \\
&= \exp\left\{\sum_{n=1}^{N}\lambda_n\exp\{p\ln x_n\}\ln(x_n)\right\}_{p=0} = \exp\left\{\sum_{n=1}^{N}\lambda_n\ln(x_n)\right\} \\
&= \exp\left\{\sum_{n=1}^{N}\ln\left(x_n^{\lambda_n}\right)\right\} = \exp\left\{\ln\prod_{n=1}^{N}x_n^{\lambda_n}\right\} = \prod_{n=1}^{N}x_n^{\lambda_n}
\end{aligned}$$

✏

Corollary D.2. [49] *Let* $\lbag x_n\rbag_1^N$ *be a tuple. Let* $\lbag\lambda_n\rbag_1^N$ *be a tuple of weighting values such that* $\sum_{n=1}^{N}\lambda_n = 1$.

COR
$$\min\lbag x_n\rbag \le \underbrace{\left(\sum_{n=1}^{N}\lambda_n\frac{1}{x_n}\right)^{-1}}_{\text{harmonic mean}} \le \underbrace{\prod_{n=1}^{N}x_n^{\lambda_n}}_{\text{geometric mean}} \le \underbrace{\sum_{n=1}^{N}\lambda_n x_n}_{\text{arithmetic mean}} \le \max\lbag x_n\rbag$$

✎PROOF:

47 Rudin (1976) page 85 ⟨4.4 Theorem⟩
48 Rudin (1976) page 109 ⟨5.13 Theorem⟩
49 Bullen (2003) page 71, Bollobás (1999) page 5, Cauchy (1821) pages 457–459 ⟨Note II, theorem 17⟩, Jensen (1906) page 183, Hoehn and Niven (1985) page 151

1. These five means are all special cases of the *power mean* $\mathsf{M}_{\phi(x\,:\,p)}$ (Definition D.15 page 174):
   $p = \infty$: $\max (\!| x_n |\!)$
   $p = 1$: arithmetic mean
   $p = 0$: geometric mean
   $p = -1$: harmonic mean
   $p = -\infty$: $\min (\!| x_n |\!)$

2. The inequalities follow directly from Theorem D.14 (page 175).

3. Generalized AM-GM inequality: If one is only concerned with the arithmetic mean and geometric mean, their relationship can be established directly using *Jensen's Inequality*:

$$\begin{aligned}
\sum_{n=1}^{N} \lambda_n x_n &= b^{\log_b\left(\sum_{n=1}^{N} \lambda_n x_n\right)} \\
&\ge b^{\left(\sum_{n=1}^{N} \lambda_n \log_b x_n\right)} && \text{by } \textit{Jensen's Inequality} \text{ (Theorem D.12 page 174)} \\
&= \prod_{n=1}^{N} b^{(\lambda_n \log_b x_n)} = \prod_{n=1}^{N} b^{(\log_b x_n)\lambda_n} = \prod_{n=1}^{N} x_n^{\lambda_n}
\end{aligned}$$

✏

D.3.4 Inequalities

Lemma D.4 (Young's Inequality). [50]

LEM

$$\begin{aligned}
xy &< \frac{x^p}{p} + \frac{y^q}{q} \quad \textit{with } \tfrac{1}{p} + \tfrac{1}{q} = 1 \quad \forall 1 < p < \infty,\ x, y \ge 0, \quad \textit{but } y \ne x^{p-1} \\
xy &= \frac{x^p}{p} + \frac{y^q}{q} \quad \textit{with } \tfrac{1}{p} + \tfrac{1}{q} = 1 \quad \forall 1 < p < \infty,\ x, y \ge 0, \quad \textit{and } y = x^{p-1}
\end{aligned}$$

Theorem D.15 (Minkowski's Inequality for sequences). [51] *Let* $(\!| x_n \in \mathbb{C} |\!)_1^N$ *and* $(\!| y_n \in \mathbb{C} |\!)_1^N$ *be complex* N*-tuples.*

THM

$$\left(\sum_{n=1}^{N} |x_n + y_n|^p\right)^{\frac{1}{p}} \le \left(\sum_{n=1}^{N} |x_n|^p\right)^{\frac{1}{p}} + \left(\sum_{n=1}^{N} |y_n|^p\right)^{\frac{1}{p}} \qquad \forall\ 1 < p < \infty$$

[50] Young (1912) page 226, Hardy et al. (1952) ⟨Theorem 24⟩, Tolsted (1964) page 5, Maligranda (1995) page 257, Carothers (2000), page 43

[51] Minkowski (1910), page 115, Hardy et al. (1952) ⟨Theorem 24⟩, Maligranda (1995) page 258, Tolsted (1964), page 7, Carothers (2000), page 44, Bullen (2003) page 179

D.4 Topological Spaces

❝*Nevertheless I should not pass over in silence the fact that today the feeling among mathematicians is beginning to spread that the fertility of these abstracting methods is approaching exhaustion. The case is this: that all these nice general concepts do not fall into our laps by themselves. But definite concrete problems were first conquered in their undivided complexity, singlehanded by brute force, so to speak. Only afterwards the axiomaticians came along and stated: Instead of breaking the door with all your might and bruising your hands, you should have constructed such and such a key of skill, and by it you would have been able to open the door quite smoothly. But they can construct the key only because they are able, after the breaking in was successful, to study the lock from within and without. Before you can generalize, formalize, and axiomatize, there must be a mathematical substance.*❞
Hermann Weyl (1885–1955); mathematician, theoretical physicist, and philosopher [52]

Definition D.16. [53] *Let Γ be a set with an arbitrary (possibly uncountable) number of elements. Let 2^X be the* POWER SET *of a set X* (Definition 1.8 page 6).

DEF: *A family of sets $\boldsymbol{T} \subseteq 2^X$ is a* ***topology*** *on X if*

1. $\varnothing \in \boldsymbol{T}$ *and*
2. $X \in \boldsymbol{T}$ *and*
3. $U, V \in \boldsymbol{T} \implies U \cap V \in \boldsymbol{T}$ *and*
4. $\{U_\gamma | \gamma \in \Gamma\} \subseteq \boldsymbol{T} \implies \bigcup_{\gamma\in\Gamma} U_\gamma \in \boldsymbol{T}$.

The ordered pair $(X, \boldsymbol{T})$ is a TOPOLOGICAL SPACE *if $\boldsymbol{T}$ is a* TOPOLOGY *on X. A set U is* ***open*** *in $(X, \boldsymbol{T})$ if U is any element of $\boldsymbol{T}$. A set D is* ***closed*** *in $(X, \boldsymbol{T})$ if D^c is* OPEN *in $(X, \boldsymbol{T})$.*

Just as the power set 2^X and the set $\{\varnothing, X\}$ are algebras of sets on a set X, so also are these sets topologies on X (next example):

Example D.12. [54] Let $\mathcal{T}(X)$ be the set of topologies on a set X and 2^X the *power set* (Definition 1.8 page 6) on X.

EX:

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

Definition D.17. [55] *Let $(X, \boldsymbol{T})$ be a* TOPOLOGICAL SPACE.

DEF: *A set $\boldsymbol{B} \subseteq 2^X$ is a* ***base*** *for $\boldsymbol{T}$ if*

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

Theorem D.16. [56] *Let $(X, \boldsymbol{T})$ be a* TOPOLOGICAL SPACE. *Let $\boldsymbol{B}$ be a subset of 2^X such that $\boldsymbol{B} \subseteq 2^X$.*

THM:

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

[52] quote: Weyl (1935) page 14 ⟨H. Weyl, quoting himself from "a conference on topology and abstract algebra as two ways of mathematical understanding, in 1931"⟩. image: https://en.wikipedia.org/wiki/File:Hermann_Weyl_ETH-Bib_Portr_00890.jpg: "This work is free and may be used by anyone for any purpose."

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

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

[55] Joshi (1983) page 92 ⟨(3.1) Definition⟩, Davis (2005) page 46 ⟨Definition 4.15⟩

[56] Joshi (1983) pages 92–93 ⟨(3.2) Proposition⟩, Davis (2005) page 46

Theorem D.17. [57] *Let $(X,\boldsymbol{T})$ be a* TOPOLOGICAL SPACE *(Definition D.16 page 179) and $\boldsymbol{B} \subseteq 2^X$.*

THM

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

Example D.13. [58] Let (X,d) be a *metric space*.

EX

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

Example D.14 (the standard topology on the real line). [59]

EX

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

Definition D.18. [60] *Let $(X,\boldsymbol{T})$ be a* TOPOLOGICAL SPACE *(Definition D.16 page 179). Let 2^X be the* POWER SET *of X.*

DEF

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

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

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

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

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

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

Proposition D.2. [61] *Let $(X,\boldsymbol{T})$ be a* TOPOLOGICAL SPACE *(Definition D.16 page 179). Let A^- be the* CLOSURE*, A° the* INTERIOR*, and ∂A the* BOUNDARY *of a set A. Let 2^X be the* POWER SET *of X.*

PRP

1. A^- is CLOSED $\quad \forall A \in 2^X$.
2. A° is OPEN $\quad \forall A \in 2^X$.

Lemma D.5. [62] *Let A^- be the* CLOSURE*, A° the* INTERIOR*, and ∂A the* BOUNDARY *of a set A in a topological space $(X,\boldsymbol{T})$. Let 2^X be the* POWER SET *of X.*

LEM

1. $A^\circ \subseteq A \subseteq A^- \quad \forall A \in 2^X$.
2. $A = A^\circ \iff A$ is OPEN $\quad \forall A \in 2^X$.
3. $A = A^- \iff A$ is CLOSED $\quad \forall A \in 2^X$.

Definition D.19. [63] *Let $(X,\boldsymbol{T}_x)$ and $(Y,\boldsymbol{T}_y)$ be* TOPOLOGICAL SPACES *(Definition D.16 page 179). Let f be a function in Y^X. A function $\mathsf{f} \in Y^X$ is **continuous** if*

$$\underbrace{U \in \boldsymbol{T}_y}_{\text{OPEN in } (Y,\boldsymbol{T}_y)} \implies \underbrace{\mathsf{f}^{-1}(U) \in \boldsymbol{T}_x}_{\text{OPEN in } (X,\boldsymbol{T}_x)}.$$

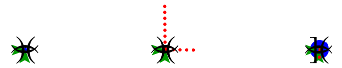
*A function is **discontinuous** in $(X,\boldsymbol{T}_y)^{(X,\boldsymbol{T}_x)}$ if it is not* CONTINUOUS *in $(X,\boldsymbol{T}_y)^{(X,\boldsymbol{T}_x)}$.*

Example D.15. Some *continuous/discontinuous* functions are illustrated in Figure D.3 (page 181).

Definition D.19 (previous definition) defines continuity using open sets. Continuity can alternatively be defined using closed sets or closure (next theorem).

---

[57] Bollobás (1999) page 19
[58] Davis (2005) page 46 ⟨Example 4.16⟩
[59] Munkres (2000) page 81, Davis (2005) page 46 ⟨Example 4.16⟩
[60] Gemignani (1972) pages 55–56 ⟨Definition 3.5.7⟩, McCarty (1967) page 90, Munkres (2000) page 95 ⟨§Closure and Interior of a Set⟩, Thron (1966), pages 21–22 ⟨definition 4.8, defintion 4.9⟩, Kelley (1955) page 42, Kubrusly (2001) pages 115–116
[61] McCarty (1967) page 90 ⟨IV.1 THEOREM⟩
[62] McCarty (1967) pages 90–91 ⟨IV.1 THEOREM⟩, ALIPRANTIS AND BURKINSHAW (1998) PAGE 59
[63] Davis (2005) page 34



continuous *continuous* *discontinuous*

Figure D.3: *continuous*/ *discontinuous* functions (Example D.15 page 180)

Theorem D.18. [64] *Let* (X,T) *and* (Y,S) *be topological spaces. Let* f *be a function in* Y^X.

THM

The following are equivalent:

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

Remark D.1. A word of warning about defining *continuity* in terms of topological spaces—*continuity* is defined in terms of a pair of *topological spaces*, and whether function is *continuous* or *discontinuous* in general depends very heavily on the selection of these spaces. This is illustrated in Proposition D.3 (next). The ramification of this is that when declaring a function to be *continuous* or *discontinuous*, one must make clear the assumed *topological spaces*.

Proposition D.3. [65] *Let* (X,T) *and* (Y,S) *be* TOPOLOGICAL SPACES*. Let* f *be a* FUNCTION *in* $(Y,S)^{(X,T)}$.

PRP

1. T *is the* DISCRETE TOPOLOGY $\implies$ f *is* CONTINUOUS $\forall \mathsf{f}\in(Y,S)^{(X,T)}$
2. S *is the* INDISCRETE TOPOLOGY $\implies$ f *is* CONTINUOUS $\forall \mathsf{f}\in(Y,S)^{(X,T)}$

Definition D.20. [66] *Let* (X,T) *be a* TOPOLOGICAL SPACE *(Definition D.16 page 179).*

DEF

A sequence $(\!(x_n)\!)_{n\in\mathbb{Z}}$ ***converges*** *in* (X,T) *to a point* x *if for each* OPEN SET *(Definition D.16 page 179)* $U\in T$ *that contains* x *there exists* $N\in\mathbb{N}$ *such that*

$$x_n \in U \text{ for all } n > N.$$

This condition can be expressed in any of the following forms:

1. *The **limit** of the sequence* $(\!(x_n)\!)$ *is* x.
2. *The sequence* $(\!(x_n)\!)$ *is **convergent** with limit* x.
3. $\lim_{n\to\infty} (\!(x_n)\!) = x$.
4. $(\!(x_n)\!) \to x$.

*A sequence that converges is **convergent**. A sequence that does not converge is said to **diverge**, or is **divergent**. An element* $x \in A$ *is a **limit point** of* A *if it is the limit of some* A*-valued sequence* $(\!(x_n\in A)\!)$.

Example D.16. [67]

EX

Let (X,T_{31}) be a *topological space* where $X \triangleq \{x,y,z\}$ and

$$T_{31} \triangleq \{\varnothing,\ \{x\},\ \{x,y\},\ \{x,z\},\ \{x,y,z\}\}.$$

In this space, the sequence $(\!(x,x,x,\ldots)\!)$ converges to x. But this sequence also converges to both y and z because x is in every *open set* (Definition D.16 page 179) that contains y and x is in every *open set* that contains z. So, the *limit* (Definition D.20 page 181) of the sequence is *not unique*.

Example D.17. In contrast to the low resolution topological space of Example D.16, the limit of the sequence $(\!(x,x,x,\ldots)\!)$ *is* unique in a *topological space* with sufficiently high resolution with respect to y and z such as the following:

EX

Define a *topological space* (X,T_{56}) where $X \triangleq \{x,y,z\}$ and

$$T_{56} \triangleq \{\varnothing,\ \{y\},\ \{z\},\ \{x,y\},\ \{y,z\},\ \{x,y,z\}\}.$$

In this space, the sequence $(\!(x,x,x,\ldots)\!)$ converges to x only. The sequence does *not* converge to y or z because there are *open sets* (Definition D.16 page 179) containing y or z that do not contain x (the open sets $\{y\}$, $\{z\}$, and $\{y,z\}$).

[64] McCarty (1967) pages 91–92 ⟨IV.2 THEOREM⟩, Searcóid (2006) PAGE 130 ⟨"THEOREM 8.3.1 (CRITERIA FOR CONTINUITY)", SET IN *metric spaces*⟩

[65] Crossley (2006) page 18 ⟨Proposition 3.9⟩, Ponnusamy (2002) page 98 ⟨2.64. Theorem.⟩

[66] Joshi (1983) page 83 ⟨(3.1) Definition⟩, Leathem (1905), page 13 ⟨"→" symbol, section III.11⟩

[67] Munkres (2000) page 98 ⟨Hausdorff Spaces⟩

Theorem D.19 (The Closed Set Theorem). [68] *Let (X, T) be a* TOPOLOGICAL SPACE. *Let A be a subset of X $(A \subseteq X)$. Let A^- be the* CLOSURE (Definition D.18 page 180) *of A in (X, T).*

$$\text{THM} \quad \underbrace{A \text{ is CLOSED in } (X, T)}_{(A = A^-)} \iff \left\{ \begin{array}{l} \text{Every } A\text{-valued sequence } (\!(x_n \in A)\!)_{n \in \mathbb{Z}} \\ \text{that CONVERGES in } (X, T) \text{ has its LIMIT in } A \end{array} \right\}$$

Theorem D.20. [69] *Let (X, T) and (Y, S) be a* TOPOLOGICAL SPACES. *Let f be a function in $(Y, S)^{(X,T)}$.*

$$\text{THM} \quad \underbrace{\left\{ \begin{array}{c} f \text{ is CONTINUOUS in } (Y, S)^{(X,T)} \\ \text{(Definition D.19 page 180)} \end{array} \right\}}_{\text{INVERSE IMAGE CHARACTERIZATION OF CONTINUITY}} \iff \underbrace{\left\{ \begin{array}{c} (\!(x_n)\!) \to x \implies f((\!(x_n)\!)) \to f(x) \\ \text{(Definition D.20 page 181)} \end{array} \right\}}_{\text{SEQUENTIAL CHARACTERIZATION OF CONTINUITY}}$$

✎PROOF:

1. Proof for the $\implies$ case (proof by contradiction):

   (a) Let U be an *open set* in (Y, T) that contains $f(x)$ but for which there exists no N such that $f(x_n) \in U$ for all $n > N$.

   (b) Note that the set $f^{-1}(U)$ is also *open* by the *continuity* hypothesis.

   (c) If $(\!(x_n)\!) \to x$, then

$$\begin{aligned} f((\!(x_n)\!)) \not\to f(x) &\implies \text{there exists no } N \text{ such that } f(x_n) \in U \text{ for all } n > N && \text{by Definition D.20} \\ &\implies \text{there exists no } M \text{ such that } x_n \in f^{-1}(U) \text{ for all } n > M && \text{by definition of } f^{-1} \\ &\implies (\!(x_n)\!) \not\to x && \text{by } \textit{continuity} \text{ hyp. and def. of } \textit{convergence} \text{ (Definition D.20 page 181)} \\ &\implies \text{contradiction of } (\!(x_n)\!) \to x \text{ hypothesis} \\ &\implies f((\!(x_n)\!)) \to f(x) \end{aligned}$$

2. Proof for the $\impliedby$ case (proof by contradiction):

   (a) Let D be a *closed* set in (Y, S).

   (b) Suppose $f^{-1}(D)$ is *not closed*...

   (c) then by the *closed set theorem* (Theorem D.19 page 182), there must exist a *convergent* sequence $(\!(x_n)\!)$ in (X, T), but with limit x *not* in $f^{-1}(D)$.

   (d) Note that $f(x)$ must be in D. Proof:

      i. by definition of D and f, $f((\!(x_n)\!))$ is in D

      ii. by left hypothesis, the sequence $f((\!(x_n)\!))$ is *convergent* with limit $f(x)$

      iii. by *closed set theorem* (Theorem D.19 page 182), $f(x)$ must be in D.

   (e) Because $f(x) \in D$, it must be true that $x \in f^{-1}(D)$.

   (f) But this is a contradiction to item (2c) (page 182), and so item (2b) (page 182) must be wrong, and $f^{-1}(D)$ must be *closed*.

   (g) And so by Theorem D.18 (page 181), f is *continuous*.

✎

[68] Kubrusly (2001) page 118 ⟨Theorem 3.30⟩, Haaser and Sullivan (1991) page 75 ⟨6·9 Proposition⟩, Rosenlicht (1968) pages 47–48

[69] Ponnusamy (2002) pages 94–96 ⟨"2.59. Proposition."; in the context of *metric spaces*; includes the "*inverse image characterization of continuity*" and "*sequential characterization of continuity*" terminology; this terminology does not seem to be widely used in the literature in general, but has been adopted for use in this text⟩

D.5 Polynomial interpolation

D.5.1 Lagrange interpolation

Definition D.21. [70] *The **Lagrange polynomial** $L_{P,n}(x)$ with respect to the $n+1$ points* $P = \{(x_k, y_k) | k = 0, 1, 2, \dots, n\}$ *is defined as*

DEF
$$L_{P,n}(x) \triangleq \sum_{k=0}^{n} y_k \prod_{m \neq n} \frac{x - x_m}{x_k - x_m}$$

Proposition D.4. *Let $L_{P,n}(x)$ be the Lagrange polynomial with respect to the points* $P = \{(x_k, y_k) | k = 0, 1, 2, \dots, n\}$.

PRP
1. *$L_{P,n}(x)$ is an nth order polynomial.*
2. *$L_{P,n}(x)$ intersects all $n+1$ points in P.*

Example D.18 (Lagrange interpolation). The Lagrange polynomial $L_{P,3}(x)$ with respect to the 4 points
$P = \{(-2, 1),\ (-1, 3),\ (3, 2),\ (5, 4)\}$ is

EX
$$L_{P,3}(x) = \frac{79}{840}x^3 + \frac{-378}{840}x^2 + \frac{-7}{840}x + \frac{2970}{840}$$

✎PROOF:

$$\begin{aligned}
L_{P,3}(x) &= \sum_{k=0}^{n} y_k \prod_{m \neq n} \frac{x - x_m}{x_k - x_m} \qquad \text{by Definition D.21} \\
&= y_0 \frac{(x+1)(x-3)(x-5)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)} + y_1 \frac{(x+2)(x-3)(x-5)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)} \\
&\quad + y_2 \frac{(x+2)(x+1)(x-5)}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3)} + y_3 \frac{(x+2)(x+1)(x-3)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)} \\
&= 1 \frac{(x+1)(x-3)(x-5)}{(-2+1)(-2-3)(-2-5)} + 3 \frac{(x+2)(x-3)(x-5)}{(-1+2)(-1-3)(-1-5)} \\
&\quad + 2 \frac{(x+2)(x+1)(x-5)}{(3+2)(3+1)(3-5)} + 4 \frac{(x+2)(x+1)(x-3)}{(5+2)(5+1)(5-3)} \\
&= \underbrace{1 \frac{x^3 - 7x^2 + 7x + 15}{-35}}_{\text{roots}=-1,3,5} + \underbrace{3 \frac{x^3 - 6x^2 - x + 30}{24}}_{\text{roots}=-2,3,5} + \underbrace{2 \frac{x^3 - 2x^2 - 13x - 10}{-40}}_{\text{roots}=-2,-1,5} + \underbrace{4 \frac{x^3 - 7x - 6}{84}}_{\text{roots}=-2,-1,3} \\
&= -\frac{x^3 - 7x^2 + 7x + 15}{35} + \frac{x^3 - 6x^2 - x + 30}{8} - \frac{x^3 - 2x^2 - 13x - 10}{20} + \frac{x^3 - 7x - 6}{21} \\
&= x^3 \left(\frac{-8 \cdot 20 \cdot 21 + 35 \cdot 20 \cdot 21 - 35 \cdot 8 \cdot 21 + 35 \cdot 8 \cdot 20}{35 \cdot 8 \cdot 20 \cdot 21} \right) \\
&\quad + x^2 \left(\frac{7 \cdot 8 \cdot 20 \cdot 21 - 6 \cdot 35 \cdot 20 \cdot 21 + 2 \cdot 35 \cdot 8 \cdot 21 + 0 \cdot 35 \cdot 8 \cdot 20}{35 \cdot 8 \cdot 20 \cdot 21} \right) \\
&\quad + x \left(\frac{-7 \cdot 8 \cdot 20 \cdot 21 - 35 \cdot 20 \cdot 21 + 13 \cdot 35 \cdot 8 \cdot 21 - 7 \cdot 35 \cdot 8 \cdot 20}{35 \cdot 8 \cdot 20 \cdot 21} \right) \\
&\quad + \left(\frac{-15 \cdot 8 \cdot 20 \cdot 21 + 30 \cdot 35 \cdot 20 \cdot 21 + 10 \cdot 35 \cdot 8 \cdot 21 - 6 \cdot 35 \cdot 8 \cdot 20}{35 \cdot 8 \cdot 20 \cdot 21} \right) \\
&= \frac{11060}{117600}x^3 + \frac{-52920}{117600}x^2 + \frac{-980}{117600}x + \frac{415800}{117600} \\
&= \frac{79}{840}x^3 + \frac{-378}{840}x^2 + \frac{-7}{840}x + \frac{2970}{840}
\end{aligned}$$

✏

[70] Waring (1779) page 60, Euler (1783) page 165 ⟨§. 10. Problema 2. Corollarium 3.⟩, Gauss (1866), Lagrange (1877), Matthews and Fink (1992), page 206, Meijering (2002) ⟨historical background⟩

D.5.2 Newton interpolation

Definition D.22. [71] *The **Newton polynomial** $N_{P,n}(x)$ with respect to the $n+1$ points $P = \{(x_k, y_k) | k = 0, 1, 2, \ldots, n\}$ is defined as*

DEF
$$N_{P,n}(x) \triangleq \sum_{k=0}^{n} \alpha_k \prod_{m=0}^{k} (x - x_m)$$

Proposition D.5. *Let $N_{P,n}(x)$ be the Newton polynomial with respect to the points $P = \{(x_k, y_k) | k = 0, 1, 2, \ldots, n\}$.*

PRP
1. *$N_{P,n}(x)$ is an nth order polynomial.*
2. *$N_{P,n}(x)$ intersects all $n+1$ points in P.*

Example D.19 (Newton polynomial interpolation).

EX
The Newton polynomial $N_{P,3}(x)$ with respect to the 4 points $P = \{(-2,1), (-1,3), (3,2), (5,4)\}$ is

$$N_{P,3}(x) = \frac{79}{840}x^3 + \frac{-378}{840}x^2 + \frac{-7}{840}x + \frac{2970}{840}$$

✎PROOF:

$$\begin{aligned}
N_{P,3}(x) &= \sum_{k=0}^{n} \alpha_k \prod_{m=1}^{k} (x - x_m) \\
&= \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)(x - x_1) + \alpha_3(x - x_0)(x - x_1)(x - x_2) \\
&= \alpha_0 + \alpha_1(x + 2) + \alpha_2(x + 2)(x + 1) + \alpha_3(x + 2)(x + 1)(x - 3) \\
&= \alpha_0 + \alpha_1(x + 2) + \alpha_2(x^2 + 3x + 2) + \alpha_3(x^3 - 7x - 6) \\
&= x^3(\alpha_3) + x^2(\alpha_2) + x(-7\alpha_3 + 3\alpha_2 + \alpha_1) + (-6\alpha_3 + 2\alpha_2 + 2\alpha_1 + \alpha_0) \\
&= \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ -6 & -7 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix} &= \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & (x_1 - x_0) & 0 & 0 \\ 1 & (x_2 - x_0) & (x_2 - x_0)(x_2 - x_1) & 0 \\ 1 & (x_3 - x_0) & (x_3 - x_0)(x_3 - x_1) & (x_3 - x_0)(x_3 - x_1)(x_3 - x_2) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & (-1 + 2) & 0 & 0 \\ 1 & (3 + 2) & (3 + 2)(3 + 1) & 0 \\ 1 & (5 + 2) & (5 + 2)(5 + 1) & (5 + 2)(5 + 1)(5 - 3) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 5 & 20 & 0 \\ 1 & 7 & 42 & 84 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}
\end{aligned}$$

[71] Newton (1711), Fraser (1919), pages 9–17 ⟨Methodus differentialis: "A photographic reproduction of the original Latin text"⟩, Fraser (1919), pages 18–25 ⟨Methodus differentialis: English translation⟩, Fraser (1919), pages 1–8 ⟨historical background and notes⟩, Meijering (2002) ⟨historical background⟩, Matthews and Fink (1992), page 220

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 5 & 20 & 0 & 0 & 0 & 1 & 0 \\ 1 & 7 & 42 & 84 & 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 5 & 20 & 0 & -1 & 0 & 1 & 0 \\ 0 & 7 & 42 & 84 & -1 & 0 & 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 20 & 0 & 4 & -5 & 1 & 0 \\ 0 & 0 & 42 & 84 & 6 & -7 & 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{1}{5} & -\frac{1}{4} & \frac{1}{20} & 0 \\ 0 & 0 & 42 & 84 & 6 & -7 & 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{1}{5} & -\frac{1}{4} & \frac{1}{20} & 0 \\ 0 & 0 & 0 & 84 & 6-\frac{42}{5} & -7+\frac{42}{4} & -\frac{42}{20} & 1 \end{array}\right]$$

$$= \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{1}{5} & -\frac{1}{4} & \frac{1}{20} & 0 \\ 0 & 0 & 0 & 84 & -\frac{12}{5} & \frac{14}{4} & -\frac{42}{20} & 1 \end{array}\right]$$

$$= \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{4}{20} & -\frac{5}{20} & \frac{1}{20} & 0 \\ 0 & 0 & 0 & 84 & -\frac{24}{10} & \frac{35}{10} & -\frac{21}{10} & \frac{10}{10} \end{array}\right]$$

$$= \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{4}{20} & -\frac{5}{20} & \frac{1}{20} & 0 \\ 0 & 0 & 0 & 1 & -\frac{24}{840} & \frac{35}{840} & -\frac{21}{840} & \frac{10}{840} \end{array}\right]$$

$$\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ \frac{4}{20} & -\frac{5}{20} & \frac{1}{20} & 0 \\ -\frac{24}{840} & \frac{35}{840} & -\frac{21}{840} & \frac{10}{840} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ -\frac{9}{20} \\ \frac{79}{840} \end{bmatrix}$$

$$N_{P,3}(x) = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ -6 & -7 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \end{bmatrix}$$

$$= \left[\begin{array}{c|c|c|c} 1 & 2 & -\frac{9}{20} & \frac{79}{840} \end{array}\right] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ -6 & -7 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \end{bmatrix}$$

$$= \left[\begin{array}{c|c|c|c} 1+4-\frac{9}{10}-\frac{79}{140} & 2-\frac{27}{20}-\frac{79}{120} & -\frac{9}{20} & \frac{79}{840} \end{array}\right] \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \end{bmatrix}$$

$$= \frac{79}{840}x^3 - \frac{378}{840}x^2 - \frac{7}{840}x + \frac{2970}{840}$$

✏

APPENDIX E

C++ SOURCE CODE SUPPORT

This document seeks to conform to the principles of *Reproducible Research* as detailed at
http://reproducibleresearch.net/.

This section contains a partial C++ source code listing for ssp.exe, written by the author of this paper, which produced the TEX files used to generate the 128 or so data plot files presented in CHAPTER 3. The complete and downloadable source code for ssp.exe is to accompany any online version of this document.

There are some who might hesitate to use computer simulation to demonstrate mathematical concepts. And arguably their concern is not unfounded. There is in fact evidence to suggest that while the invention of the printing press has greatly assisted the progress of mathematical discovery, the introduction of computers has harmed it. Evidence of this hypothesis is given in Figure E.1 (page 188).[1] This graph shows the number of "notable" mathematicians alive during the last 3000 years; Here a "notable mathematician" is defined as one whose name appears in *Saint Andrew's University's Who Was There* website.[2] Note the following:

- The number of mathematicians starts to exponentially increase at about the time of Gutenberg's invention of the printing press— that is, when information of discoveries and results could be widely and economically circulated.[3]
- There is another increase after the invention of the slide rule in the early 1600s— that is, when computational power increased.
- There are huge increases around the time of the first and second industrial revolutions— that is, when there were many **applications** that called for mathematical solutions.
- After the invention of the pocket scientific calculator in 1972 and home IBM PC in 1981— machines that could often make hard-core mathmatical analysis unnecessary in real-world applications— there was a huge drop in the number of mathematicians.

E.1 Symbolic sequence routines

```
/*=========================================================================
 * Daniel J. Greenhoe
```

[1] Data for Figure E.1 (page 188) extracted from http://www-history.mcs.st-andrews.ac.uk/Timelines/WhoWasThere.html

[2] http://www-history.mcs.st-andrews.ac.uk/Timelines/WhoWasThere.html

[3] This point is also made by Resnikoff and Wells in Resnikoff and Wells (1984), page 9.

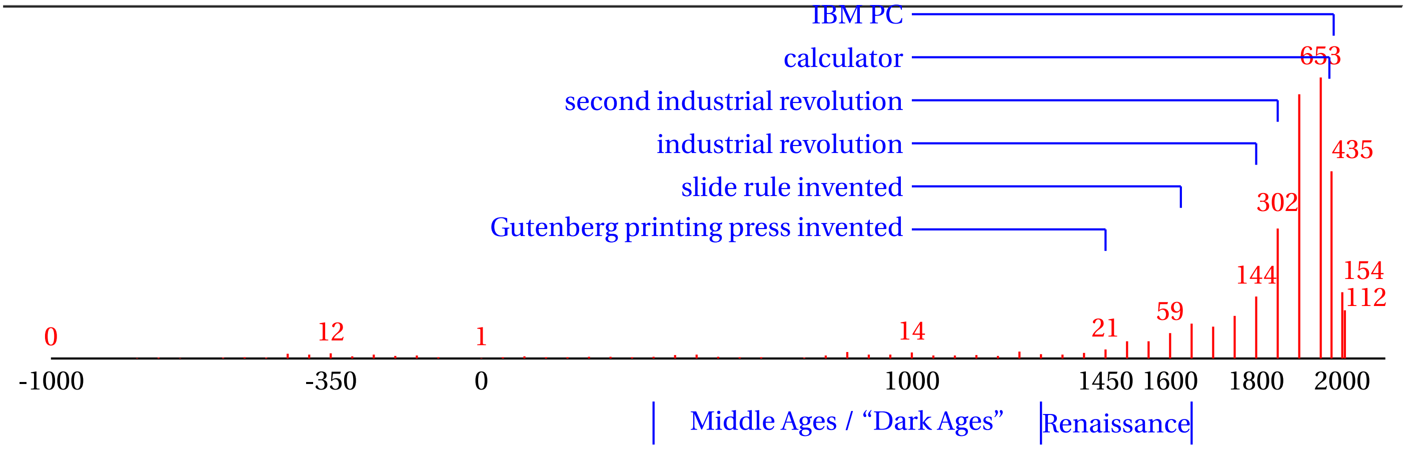


Figure E.1: Number of notable mathematicians alive over time

```
 * header file for routines for char sequence functions
 *=====================================================================*/
class symseq {
  private:
    long N;
    char *x;
  public:
    symseq(const long M);   //constructor initializing to '.'
    symseq(const long M,const unsigned seed,const char *symbols); //constructor initializing using seed
    void clear(void);               //fill sequence with the value 'A'
    char get(const long n);               //get a value from x at location n
    char get(const long n,char *symbols);  //get a value from x at location n but exit if not in
          <symbols> string
    void put(long n, char symbol);
    void put(const long start, const long end, const char symbol);
    const long getN(void){const long M=(const long)N; return M;}       //get N
    void downsample(int M, symseq *y);//downsample by a factor of <M> and write to <y>
    void  list(const long start, const long end, const char* str1, const char *str2, const int
          display, FILE *fptr);
    void  list(const long start, const long end, FILE *fptr){list(start,end, "","",1,fptr);}
    void  list(const long start, const long end, const int display, FILE *fptr){list(start,end,
          "","",1,fptr);}
    void  list(const long start, const long end, const char* str1, const char *str2, FILE *fptr){
          list(           start,           end,           str1,              str2, 1,    fptr);
          }
    void  list(long start, long end){list(start,end,"","",NULL);}  //list contents of sequence
    void  list(void){               list(0,    N-1,"","",NULL);}    //list contents of sequence
    void  list(long start){          list(start,N-1,"","",NULL);}  //list contents of sequence
    void shiftL(long n);             //shift symseq n elements to the left
    void shiftR(long n);             //shift symseq n elements to the right
    void prngseed(unsigned seed){srand(seed);}     //
    void randomize(const char *symbols);//randomize using a list of symbols
    void randomize(const unsigned seed,const char *symbols){prngseed(seed); randomize(symbols);}
    void operator=(symseq y);      //x=y
    void  operator>>=(long n){shiftR(n);} //shift symseq n elements to the right
    void  operator<<=(long n){shiftL(n);} //shift symseq n elements to the left
    int operator==(symseq y);      //test if x==y; 1 if yes, 0 if no
  };

extern long cmp(const symseq *x,const symseq *y, int showdiff, FILE *fptr);
void copy(const long start, const long end, const symseq *x, const symseq *y);
void downsample(int M, symseq *x, symseq *y);
```

```
/*=====================================================================
 * Daniel J. Greenhoe
 * routines for Real Die symseqs
 *=====================================================================*/
/*=====================================
 * headers
 *=====================================*/
#include<stdio.h>
#include<stdlib.h>
#include<string.h>
#include<math.h>
#include<symseq.h>

/*---------------------------------------------------------------------
```

```
 * constructor initializing symseq to '.'
 *----------------------------------------------------------------------------*/
symseq::symseq(long M){
  long n;
  void *memptr;
  N=M;
  memptr=malloc(N*sizeof(char));
  if(memptr==NULL){
    fprintf(stderr,"symseq::symseq memory allocation for %ld elements failed\n",M);
    exit(EXIT_FAILURE);
    }
  x = (char *)memptr;
  for(n=0; n<N; n++)x[n]='.';
  }

/*----------------------------------------------------------------------------
 * constructor initializing symseq to '.'
 *----------------------------------------------------------------------------*/
symseq::symseq(const long M,const unsigned seed,const char *symbols){
  long n;
  void *memptr;
  N=M;
  memptr=malloc(N*sizeof(char));
  if(memptr==NULL){
    fprintf(stderr,"symseq::symseq memory allocation for %ld elements failed\n",M);
    exit(EXIT_FAILURE);
    }
  x = (char *)memptr;
  randomize(seed,symbols);
  }

/*----------------------------------------------------------------------------
 * fill the symseq with the value '.'
 *----------------------------------------------------------------------------*/
void symseq::clear(void){
  long n;
  for(n=0; n<N; n++)x[n]='.';
  }

/*----------------------------------------------------------------------------
 * get a symbol from the symseq x at location n
 *----------------------------------------------------------------------------*/
char symseq::get(const long n){
  if(n<0 || n>=N){//domain check
    fprintf(stderr,"ERROR using symseq::get(n): n=%ld outside the domain [0:%ld] of the
              sequence.\n",n,N);
    exit(EXIT_FAILURE);
    }
  return x[n];
  }

/*----------------------------------------------------------------------------
 * get a symbol from the symseq x at location n
 * but exit if symbol is not in the string <range>
 * example: symbol=get(n,"ABCDEF");
 *----------------------------------------------------------------------------*/
char symseq::get(const long n,char *range){
  int M;
  int match;
  char *sptr;
  char symbol;
  if(n<0 || n>=N){//domain check
    fprintf(stderr,"\nERROR using symseq::get(n): n=%ld outside the domain [0:%ld] of the
              sequence.\n",n,N);
    exit(EXIT_FAILURE);
    }
  symbol = x[n];
  for(match=0,sptr=range;*sptr!='\0';sptr++) if(symbol==*sptr) match=1;
  if(!match){//range check
    fprintf(stderr,"\nERROR using symbol=symseq::get(%ld): symbol='%c' (0x%02X) outside the range {%s}
              of the sequence.\n",n,symbol,symbol,range);
    exit(EXIT_FAILURE);
    }
  return symbol;
  }

/*----------------------------------------------------------------------------
```

```
 * put a single value from the symseq x at location n
 *-------------------------------------------------------------------------*/
void symseq::put(long n, char symbol){
  int rval;
  if(n<0 || n>=N){//domain check
    fprintf(stderr,"ERROR using symseq::put(n): n=%ld outside the domain [0:%ld] of the sequence.\n",n,N-1);
    exit(EXIT_FAILURE);
    }
  x[n]=symbol;
  }

/*-------------------------------------------------------------------------
 * put the symbol <symbol> into the sequence <x>
 * from location <start> to location <end>
 *-------------------------------------------------------------------------*/
void symseq::put(const long start, const long end, const char symbol){
  long n;
  for(n=start;n<=end;n++) put(n,symbol);
  }

/*-------------------------------------------------------------------------
 * list contents of dieseq
 *-------------------------------------------------------------------------*/
void symseq::list(const long start, const long end, const char *str1, const char *str2, const int display, FILE *fptr){
  long n,m;
  if(strlen(str1)>0){
    if(display)printf("%s",str1);
    if(fptr!=NULL)fprintf(fptr,"%s",str1);
    }
  for(n=start,m=1; n<=end; n++,m++){
    if(display)printf("%c",get(n));
    if(m%50==0&&display)printf("\n");
    else if(m%10==0&&display)printf(" ");
    if(fptr!=NULL){
      fprintf(fptr,"%c",get(n));
      if(m%50==0)fprintf(fptr,"\n");
      else if(m%10==0)fprintf(fptr," ");
      }
    }
  if(strlen(str2)>0){
    if(display)printf("%s",str2);
    if(fptr!=NULL)fprintf(fptr,"%s",str2);
    }
  }

/*-------------------------------------------------------------------------
 * shift symseq n elements to the right inserting zeros on the left
 * example: if x = [ a b c d e f ] (N=6), then shiftR(2) results in
 *             x = [ 0 0 a b c d ] (N=6).
 *-------------------------------------------------------------------------*/
void symseq::shiftR(long n){
  long m;
  for(m=N-1;m-n>=0;m--)x[m]=x[m-n];
  for(m=0;m<n;m++)       x[m]='.';
  }

/*-------------------------------------------------------------------------
 * shift symseq n elements to the left inserting zeros on the right
 * example: if x = [ a b c d e f ] (N=6), then shiftL(2) results in
 *             x = [ c d e f 0 0 ] (N=6).
 *-------------------------------------------------------------------------*/
void symseq::shiftL(long n){
  long m;
  for(m=0;m<N-n;m++) x[m]=x[m+n];
  for(m=N-2;m<N;m++) x[m]='.';
  }

/*-------------------------------------------------------------------------
 * fill the sequence with uniformly distributed pseudo-random symbols
 * from the string <symbols>
 *-------------------------------------------------------------------------*/
void symseq::randomize(const char *symbols){
  const long N=getN();
  const int  M=strlen(symbols);
  int r,i;
```

```
  long n;
  for(n=0; n<N; n++){
    r=rand();
    i = r%M;
    put(n,symbols[i]);
    }
  }

/*=====================================
 * operators
 *=====================================*/
/*-----------------------------------------------------------------------------
 * operator symseq x = symseq y
 *-----------------------------------------------------------------------------*/
void symseq::operator=(symseq y){
  const long M=y.getN();
  long n;
  if(N!=M){
    fprintf(stderr,"ERROR using symseq x = symseq y operation: size of x (%ld) does not equal size of
              y (%ld)\n",N,M);
    exit(EXIT_FAILURE);
    }
  for(n=0;n<N;n++)x[n]=y.get(n);
  }

/*-----------------------------------------------------------------------------
 * operator symseq x == symseq y
 * compare x and y; return 1 if the same, 0 if different.
 *-----------------------------------------------------------------------------*/
int symseq::operator==(symseq y){
  const long M=y.getN();
  long n;
  int retval;
  char xsym,ysym;

  if(N!=M){
    fprintf(stderr,"ERROR using symseq x == symseq y operation: size of x (%ld) does not equal size of
              y (%ld)\n",N,M);
    exit(EXIT_FAILURE);
    }
  for(n=0,retval=1;n<N;n++){
    xsym=  get(n);
    ysym=y.get(n);
    if(xsym!=ysym) retval=0;
    }
  return retval;
  }

/*=====================================
 * external functions
 *=====================================*/
/*-----------------------------------------------------------------------------
 * compare dieseq x and dieseq y
 * return the number of locations in which the two sequences are different
 * return 0 if the same
 *-----------------------------------------------------------------------------*/
long cmp(const symseq *x, const symseq *y, int showdiff, FILE *fptr){
  const long N=x->getN();
  const long M=y->getN();
  char xsym,ysym;
  long n;
  long count;
  if(N!=M){
    fprintf(stderr,"\nERROR using cmp(symseq *x,symseq *y): size of x (%ld) != size of y
              (%ld).\n",N,M);
    exit(EXIT_FAILURE);
    }
  for(n=0,count=0;n<N;n++){
    xsym=x->get(n);
    ysym=y->get(n);
    if(xsym!=ysym){
      count++;
      if(showdiff)   fprintf(stdout,"%6ld: x[%6ld]=%c(0x%02x)
                y[%6ld]=%c(0x%02x)\n",count,n,xsym,xsym,n,ysym,ysym);
      if(fptr!=NULL) fprintf(fptr,   "%6ld: x[%6ld]=%c(0x%02x)
                y[%6ld]=%c(0x%02x)\n",count,n,xsym,xsym,n,ysym,ysym);
      }
```

```
    }
  return count;
  }

/*-------------------------------------------------------------------------
 * copy the sequence <*x> = [x_start ... x_end]
 * into the sequence <*y> = [y_0      ... y_{N-1}]
 * where N = end-start+1
 *-------------------------------------------------------------------------*/
void copy(const long start, const long end, const symseq *x, const symseq *y){
  long Nx=x->getN();
  long Ny=y->getN();
  long n,m;
  double xx;
  y->clear();
  if(end>=Nx){
    fprintf(stderr,"ERROR using copy(start,end,seqR1 *x, seqR1 *y): <end>=%ld is too large.\n",end);
    exit(EXIT_FAILURE);
    }
  if((end-start+1)!=Ny){
    fprintf(stderr,"ERROR using copy(start,seqR1 *x, seqR1 *y): length of [x_start...x_end] (%ld) !=
              length of y(%ld).\n",Nx-start,Ny);
    exit(EXIT_FAILURE);
    }
  for(n=start,m=0;n<=end;n++,m++){
    xx=x->get(n);
    y->put(m,xx);
    }
  }



/*-------------------------------------------------------------------------
 * downsample sequence by a factor of <M>
 * and write to sequence pointed to by <y>
 *-------------------------------------------------------------------------*/
void downsample(int M, symseq *x, symseq *y){
  const long Nx = x->getN();
  const long Ny = y->getN();
  char symbol;
  long n,m;
  if(M<1){//check validity of factor <M>
    fprintf(stderr,"ERROR using symseq::downsample(M,y): factor=%d must be at least 1\n",M);
    exit(EXIT_FAILURE);
    }
  if(Ny != Nx/M){//check validity of the length of output sequence <y>
    fprintf(stderr,"ERROR using symseq::downsample(M,y): length %ld of output sequence y must be N/M =
              %ld/%ld = %ld\n",Ny,Nx,M,Nx/M);
    exit(EXIT_FAILURE);
    }
  for(n=0,m=0; m<Ny; n+=M,m++){
    symbol = x->get(n);
    y->put(m,symbol);
    }
  }
```

E.2 Die routines

```
/*=========================================================================
 * Daniel J. Greenhoe
 * header file for routines for die routines
 *    'A'--> die face value 1
 *    'B'--> die face value 2
 *    'C'--> die face value 3
 *    'D'--> die face value 4
 *    'E'--> die face value 5
 *    'F'--> die face value 6
 *=========================================================================*/
class dieseq: public symseq {
  public:
    dieseq(const long M) : symseq(M){};//constructor initializing to '.'
```

```
    dieseq(const long M,const unsigned seed) : symseq(M,seed,"ABCDEF"){};//constructor initializing
         random values
    void  randomize(void){symseq::randomize("ABCDEF");}     //
    void  randomize(unsigned seed){srand(seed); randomize();}
    int   randomize(long start, long end, int wA,int wB,int wC,int wD,int wE,int wF);
    int   randomize(unsigned seed,int wA,int wB,int wC,int wD,int wE,int wF){srand(seed);return
         randomize(0,getN()-1,wA,wB,wC,wD,wE,wF);}
    int   randomize(int wA,int wB,int wC,int wD,int wE,int wF){return
         randomize(0,getN()-1,wA,wB,wC,wD,wE,wF);}
    int   randomize(long start, long end, unsigned seed, int wA,int wB,int wC,int wD,int wE,int
         wF){srand(seed);return randomize(start,end, wA,wB,wC,wD,wE,wF);}
    char  get(long n){return symseq::get(n,"ABCDEF");}     //get a value from x at location n
    void  put(long n, char symbol){symseq::put(n,symbol);}
    seqR1 dietoR1(void);           //map die face values to R^1
    seqC1 dietoC1(void);           //map die face values to R^1
    seqR1 dietoR1pam(void);        //map die face values to R^1 using PAM scheme (symmetric about zero)
    seqR3 dietoR3(void);           //map die face values to R^3
    seqR1 histogram(const long start, const long end, int display, FILE *fptr);//compute, display, and
         write histogram
    seqR1 histogram(){return histogram(0,getN()-1,0,NULL);}//compute histogram
    seqR1 histogram(const long start, const long end){return histogram(start,end,0,NULL);}//compute
         histogram
    seqR1 histogram(int display,FILE *fptr){return histogram(0,getN()-1,1,fptr);}//print histogram to
         file
    seqR1 histogram(FILE *fptr){return histogram(0,getN()-1,0,fptr);}//print histogram to file
    void operator=(dieseq y);      //x=y
  };

extern int     die_domain(char c);//check if value is in the domain of die
extern double  die_dietoR1 (char c);
extern double  die_dietoR1pam(char c);
extern vectR3  die_dietoR3 (char c);
extern vectR6  die_dietoR6 (char c);
extern complex die_dietoC1c(char c);
```

```
/*=============================================================================
 * Daniel J. Greenhoe
 * routines for Real Die dieseqs
 *=============================================================================*/
/*=====================================
 * headers
 *=====================================*/
#include<stdio.h>
#include<stdlib.h>
#include<string.h>
#include<math.h>
#include<main.h>
#include<symseq.h>
#include<r1.h>
#include<r2.h>
#include<r3.h>
#include<r6.h>
#include<c1.h>
#include<die.h>

/*=====================================
 * prototypes
 *=====================================*/
void phistogram(seqR1 *data, const long start, const long end, FILE *ptr);

/*---------------------------------------------------------------------------
 * fill the dieseq with weighted pseudo-random die face values
 *---------------------------------------------------------------------------*/
int dieseq::randomize(long start, long end, int wA,int wB,int wC,int wD,int wE,int wF){
  int r,u;
  long n;
  char symbol;
  int sum=wA+wB+wC+wD+wE+wF;
  if(sum!=100){
    fprintf(stderr,"dieseq::randomize error: sum of weight values = %d != 100\n",sum);
    return -1;
    }
  //printf("start=%ld end=%ld  weights=(%03d %03d %03d %03d %03d %03d)\n",
           start,end,wA,wB,wC,wD,wE,wF);
  for(n=start; n<=end; n++){
    r=rand();
    u = r%100;
```

```
    if      (u<wA)              symbol='A';
    else if(u<wA+wB)            symbol='B';
    else if(u<wA+wB+wC)         symbol='C';
    else if(u<wA+wB+wC+wD)      symbol='D';
    else if(u<wA+wB+wC+wD+wE)   symbol='E';
    else                        symbol='F';
    put(n,symbol);
    }
  return 0;
  }

/*-------------------------------------------------------------------------
 * map die face values to R^1
 * A-->1  B-->2  C-->3  D-->4  E-->5  F-->6
 * all other values --> 0
 *-------------------------------------------------------------------------*/
seqR1 dieseq::dietoR1(void){
  const long N=getN();
  long n;
  char sym;
  double xR1;
  seqR1 y(N);
  for(n=0; n<N; n++){
    sym = get(n);
    xR1 = die_dietoR1(sym);
    y.put(n,xR1);
    }
  return y;
  }

/*-------------------------------------------------------------------------
 * map die face values to R^1 using PAM scheme
 * A-->-2.5  B-->-1.5  C-->-0.5  D-->0.5  E-->1.5  F-->2.5
 * all other values --> 0
 *-------------------------------------------------------------------------*/
seqR1 dieseq::dietoR1pam(void){
  const long N=getN();
  long n;
  char sym;
  double xR1;
  seqR1 y(N);
  for(n=0; n<N; n++){
    sym = get(n);
    xR1 = die_dietoR1pam(sym);
    y.put(n,xR1);
    }
  return y;
  }

/*-------------------------------------------------------------------------
 * map die face values to C^1
 *-------------------------------------------------------------------------*/
seqC1 dieseq::dietoC1(void){
  const long N=getN();
  long n;
  char sym;
  complex xC1;
  seqC1 y(N);
  for(n=0; n<N; n++){
    sym = get(n);
    xC1 = die_dietoC1c(sym);
    y.put(n,xC1);
    }
  return y;
  }

/*-------------------------------------------------------------------------
 * map die face values to R^3 sequence
 *-------------------------------------------------------------------------*/
seqR3 dieseq::dietoR3(void){
  const long N=getN();
  long n;
  char sym;
  vectR3 xR3;
  seqR3 y(N);
  for(n=0; n<N; n++){
    sym = get(n);
```

```
    xR3 = die_dietoR3(sym);
    y.put(n,xR3);
    }
  return y;
  }

/*-------------------------------------------------------------------------
 * compute histogram of dna sequence
 * return seqR1 y of length 6 where
 * y[1]-->number of dna 'A' symbols,
 * y[2]-->number of dna 'B' symbols,
 * y[3]-->number of dna 'C' symbols,
 * y[4]-->number of dna 'D' symbols,
 * y[5]-->number of dna 'D' symbols,
 * y[6]-->number of dna 'D' symbols,
 * y[0]-->number of all other values
 * y[7]-->total number of symbols y[1],y[2],...,y[6]
 *-------------------------------------------------------------------------*/
seqR1 dieseq::histogram(const long start, const long end, int display, FILE *fptr){
  seqR1 data(8);
  long n;
  long bin;
  double p;
  int i;
  char symbol;
  FILE *ptr;
  data.clear();
  for(n=start;n<=end;n++){
    symbol=get(n);
    switch(symbol){
      case 'A': bin=1; break;
      case 'B': bin=2; break;
      case 'C': bin=3; break;
      case 'D': bin=4; break;
      case 'E': bin=5; break;
      case 'F': bin=6; break;
      default : bin=0; break;
      }
    if(bin!=0) data.increment(7);
    data.increment(bin);
    }
  if(display)    phistogram(&data,start,end,stdout);
  if(fptr!=NULL) phistogram(&data,start,end,fptr  );
  return data;
  }

/*-------------------------------------------------------------------------
 * print die sequence histogram with data pointed to by <data>
 * to stream pointed to by ptr
 *-------------------------------------------------------------------------*/
void phistogram(seqR1 *data, const long start, const long end, FILE *ptr){
  const long N=end-start+1;
  long bin;
  fprintf(ptr,"\n");
  fprintf(ptr," -------------------------------------------------------------------------\n");
  fprintf(ptr,"| Histogram for sequence [x_n|n=%7ld-%7ld] (length %7ld)         |\n|",start,end,N);
  for(bin=1;bin<=6;bin++)fprintf(ptr,"         %c",'A'+(char)bin-1);
  fprintf(ptr,"     extra    |\n|");
  for(bin=1;bin<=6;bin++)fprintf(ptr,"%10.0lf",data->get(bin));
  fprintf(ptr,"%10.0lf    |\n|",data[0]);
  for(bin=1;bin<=6;bin++)fprintf(ptr," (%6.2lf%%)",data->get(bin)/(double)N*100.0);
  fprintf(ptr," (%6.2lf%%)   |\n",data->get(0)/(double)N*100.0);
  fprintf(ptr," -------------------------------------------------------------------------\n");
  }


/*=====================================
 * operators
 *=====================================*/
/*-------------------------------------------------------------------------
 * operator dieseq x = dieseq y
 *-------------------------------------------------------------------------*/
void dieseq::operator=(dieseq y){
  const long N=  getN();
  const long M=y.getN();
  long n;
  char symbol;
```

```
  if(N!=M){
    fprintf(stderr,"ERROR using dieseq x = dieseq y operation: size of x (%ld) does not equal size of
          y (%ld)\n",N,M);
    exit(EXIT_FAILURE);
    }
  for(n=0;n<N;n++){
    symbol = y.get(n);
    put(n,symbol);
    }
  }

/*=====================================
 * external operations
 *=====================================*/
/*-----------------------------------------------------------------------------
 * map die face values to R^1
 *-----------------------------------------------------------------------------*/
double die_dietoR1(char c){
  double rval;
  switch(c){
    case 'A': rval = 1.0;  break;
    case 'B': rval = 2.0;  break;
    case 'C': rval = 3.0;  break;
    case 'D': rval = 4.0;  break;
    case 'E': rval = 5.0;  break;
    case 'F': rval = 6.0;  break;
    default:
      fprintf(stderr,"ERROR using die_dietoR1(c): c=%c(0x%x) is not in the valid domain
           {A,B,C,D,E,F}\n",c,c);
      exit(EXIT_FAILURE);
    }
  return rval;
  }

/*-----------------------------------------------------------------------------
 * map die face values to R^1 PAM
 *-----------------------------------------------------------------------------*/
double die_dietoR1pam(char c){
  double rval;
  switch(c){
    case 'A': rval = -2.5;  break;
    case 'B': rval = -1.5;  break;
    case 'C': rval = -0.5;  break;
    case 'D': rval =  0.5;  break;
    case 'E': rval =  1.5;  break;
    case 'F': rval =  2.5;  break;
    default:
      fprintf(stderr,"ERROR using die_dietoR1pam(c): c=%c(0x%x) is not in the valid domain
           {A,B,C,D,E,F}\n",c,c);
      exit(EXIT_FAILURE);
    }
  return rval;
  }

/*-----------------------------------------------------------------------------
 * map die face values to complex plane C^1
 *
 *                imaginary axis
 *                      |
 *                      B=(cos90,sin90)
 *                      |
 *                      |
 * (cos150,sin150)=C    |     A=(cos30,sin30)
 *                      |
 *     -----------------|----------------- real axis
 *                      |
 * (cos210,sin210)=D    |     F=(cos330,sin330)
 *                      |
 *                      |
 *                      E=(cos270,sin270)
 *                      |
 *
 *-----------------------------------------------------------------------------*/
complex die_dietoC1c(char c){
  complex rc;
  switch(c){
    case 'A': rc = expi( 30.0/180.0*PI);  break;
```

```
    case 'B': rc = expi( 90.0/180.0*PI);  break;
    case 'C': rc = expi(150.0/180.0*PI);  break;
    case 'D': rc = expi(210.0/180.0*PI);  break;
    case 'E': rc = expi(270.0/180.0*PI);  break;
    case 'F': rc = expi(330.0/180.0*PI);  break;
    case '0': rc.put(0,0);                break;
    default:  rc.put(0,0);
    fprintf(stderr,"ERROR using dietoC1(char c): c=%c(0x%x) is not in the valid domain
              {0,A,B,C,D,E,F}. Returning (0,0).\n",c,c);
    }
  return rc;
  }

/*-----------------------------------------------------------------------------
 * map die face values to R^3
 *      |+1|      | 0|      | 0|      | 0|      | 0|      |-1|
 * A-->| 0| B-->|+1| C-->| 0| D-->| 0| E-->|-1| F-->| 0|
 *      | 0|      | 0|      |+1|      |-1|      | 0|      | 0|
 *-----------------------------------------------------------------------------*/
vectR3 die_dietoR3(char c){
  vectR3 xyz;
  switch(c){
    case 'A': xyz.put(+1, 0, 0);  break;
    case 'B': xyz.put( 0,+1, 0);  break;
    case 'C': xyz.put( 0, 0,+1);  break;
    case 'D': xyz.put( 0, 0,-1);  break;
    case 'E': xyz.put( 0,-1, 0);  break;
    case 'F': xyz.put(-1, 0, 0);  break;
    default:
      fprintf(stderr,"ERROR using die_dietoR3(c): c=%c(0x%x) is not in the valid domain
              {A,B,C,D,E,F}\n",c,c);
      exit(EXIT_FAILURE);
    }
  return xyz;
  }

/*-----------------------------------------------------------------------------
 * map die face values to R^6
 * A-->(1,0,0,0,0,0)    D-->(0,0,0,1,0,0)
 * B-->(0,1,0,0,0,0)    E-->(0,0,0,0,1,0)
 * C-->(0,0,1,0,0,0)    F-->(0,0,0,0,0,1)
 * 0-->(0,0,0,0,0,0)
 * on ERROR return (0,0,0,0,0,0)
 *-----------------------------------------------------------------------------*/
vectR6 die_dietoR6(char c){
  vectR6 rsix;
  switch(c){
    case 'A': rsix.put(1,0,0,0,0,0);  break;
    case 'B': rsix.put(0,1,0,0,0,0);  break;
    case 'C': rsix.put(0,0,1,0,0,0);  break;
    case 'D': rsix.put(0,0,0,1,0,0);  break;
    case 'E': rsix.put(0,0,0,0,1,0);  break;
    case 'F': rsix.put(0,0,0,0,0,1);  break;
    default:
      fprintf(stderr,"ERROR using dietoR6(char c): c=%c(0x%x) is not in the valid domain
              {0,A,B,C,D,E,F}.\n",c,c);
      exit(EXIT_FAILURE);
    }
  return rsix;
  }
```

E.3 Real die routines

```
/*=============================================================================
 * Daniel J. Greenhoe
 * header file for routines for real die routines
 *   'A'--> die face value 1
 *   'B'--> die face value 2
 *   'C'--> die face value 3
 *   'D'--> die face value 4
 *   'E'--> die face value 5
 *   'F'--> die face value 6
```

```
 *===========================================================================*/
class rdieseq: public dieseq {
  public:
    rdieseq(const long M) : dieseq(M){};
    rdieseq(const long M,const unsigned seed) : dieseq(M,seed){};
    void operator=(dieseq y);      //x=y
    void  operator>>=(long n){shiftR(n);} //shift rdieseq n elements to the right
    void  operator<<=(long n){shiftL(n);} //shift rdieseq n elements to the left
    int metrictbl(void);
    int Rxx(const seqR1 *Rxx, const int showcount);
    int Rxx(const seqR1 *Rxy, const int showcount,const long N, const long M, const long start, const
          long finish);
    int Rxxo(const seqR1 *rxx, const int showcount);
    double Rxx(const long m);
  };

extern rdieseq rdie_R1todie_euclid(seqR1 xyz);
extern rdieseq rdie_R3todie_larc(seqR3 xyz);
extern rdieseq rdie_R3todie0_larc(seqR3 xyz);
extern rdieseq rdie_R3todie_euclid(seqR3 xyz);
extern rdieseq rdie_R3todie0_euclid(seqR3 xyz);
extern vectR3 rdie_dietoR3(char c);
extern int     rdie_dietoR1(char c);
extern int     rdie_domain(char c);//check if value is in the domain of rdie
extern double  rdie_metric(char a, char b);
extern double  rdie_metric(rdieseq x, rdieseq y); //metric for two sequences
```

```
/*===========================================================================
 * Daniel J. Greenhoe
 * routines for Real Die rdieseqs
 *===========================================================================*/
/*=====================================
 * headers
 *=====================================*/
#include<stdio.h>
#include<stdlib.h>
#include<math.h>
#include<main.h>
#include<symseq.h>
#include<r1.h>
#include<r2.h>
#include<r3.h>
#include<r4.h>
#include<r6.h>
#include<c1.h>
#include<euclid.h>
#include<larc.h>
#include<die.h>
#include<realdie.h>

/*---------------------------------------------------------------------------
 * display real die metric table
 *---------------------------------------------------------------------------*/
int rdieseq::metrictbl(void){
  char a,b;
  for(a='A';a<='F';a++){
    for(b='A';b<='F';b++)printf("d(%c,%c)=%.1lf ",a,b,rdie_metric(a,b));
    printf("\n");
    }
  return 1;
  }

/*---------------------------------------------------------------------------
 * autocorrelation Rxx of a real die seqR1 x with 2N offset
 *---------------------------------------------------------------------------*/
int rdieseq::Rxxo(const seqR1 *rxx, const int showcount){
  const long N=getN();
  int rval;
  rval=Rxx(rxx,showcount);
  rxx->add(2*N);
  return rval;
  }

/*---------------------------------------------------------------------------
 * autocorrelation Rxx of a real die seqR1 x
 *---------------------------------------------------------------------------*/
int rdieseq::Rxx(const seqR1 *rxx, const int showcount){
```

```
  long m;
  const long N=getN();
  int rval=0;
  double rxxm;
  if(showcount)fprintf(stderr,"  Calculate %ld auto-correlation values ... n=",2*N+1);
  for(m=-N;m<=N;m++){
    if(showcount)fprintf(stderr,"%8ld",m+N);
    rxxm=Rxx(m);
    if(rxxm>0)rval=-1;
    rxx->put(m+N,rxxm);
    if(showcount)fprintf(stderr,"\b\b\b\b\b\b\b\b");
    }
  if(showcount)fprintf(stderr,"%8ld .... done.\n",m+N);
  return rval;
  }

/*-------------------------------------------------------------------------
 * autocorrelation Rxx(m)
 *-------------------------------------------------------------------------*/
double rdieseq::Rxx(const long m){
  const long mm=labs(m);
  const long N=getN();
  long n,mmm;
  double d,sum;
  char a,b;
  for(n=0,sum=0;n<(N+mm);n++){
    mmm=n-mm;
    a=(n  <0 || n  >=N)? 0.0 : get(n);
    b=(mmm<0 || mmm>=N)? 0.0 : get(mmm);
    d=(a==0 || b==0)?    1.0 : rdie_metric(a,b);
    sum+=d;
    }
  return -sum;
  }

/*=====================================
 * operators
 *=====================================*/
/*-------------------------------------------------------------------------
 * operator rdieseq x = dieseq y
 *-------------------------------------------------------------------------*/
void rdieseq::operator=(dieseq y){
  long n;
  const long N=getN();
  const long M=y.getN();
  char symbol;
  if(N!=M){
    fprintf(stderr,"ERROR using rdieseq x = rdieseq y operation: size of x (%ld) does not equal size
          of y (%ld)\n",N,M);
    exit(EXIT_FAILURE);
    }
  for(n=0;n<N;n++){
    symbol=y.get(n);
    put(n,symbol);
    }
  }

/*=====================================
 * external operations
 *=====================================*/
/*-------------------------------------------------------------------------
 * map die face values to R^1
 * A-->1 B-->2 C-->3 D-->4 E-->5 F-->6
 *-------------------------------------------------------------------------*/
int rdie_dietoR1(char c){
  int n,rval;
  char domain[6]={'A','B','C','D','E','F'};
  char element;
  for(n=0,rval=-1;n<6;n++)if(c==domain[n])rval=n;
  if(rval==-1){
    fprintf(stderr,"ERROR using rdie_dietoR1(char c): c=%c(0x%x) is not in the valid domain
          {0,A,B,C,D,E,F}\n",c,c);
    exit(EXIT_FAILURE);
    }
  return rval;
  }
```

```
/*---------------------------------------------------------------------
 * map die face values to R^3
 *      |+1|      | 0|      | 0|      | 0|      | 0|      |-1|      | 0|
 * A-->| 0| B-->|+1| C-->| 0| D-->| 0| E-->|-1| F-->| 0| 0-->| 0|
 *      | 0|      | 0|      |+1|      |-1|      | 0|      | 0|      | 0|
 *---------------------------------------------------------------------*/
vectR3 rdie_dietoR3(char c){
  vectR3 xyz;
  switch(c){
    case 'A': xyz.put(+1, 0, 0);  break;
    case 'B': xyz.put( 0,+1, 0);  break;
    case 'C': xyz.put( 0, 0,+1);  break;
    case 'D': xyz.put( 0, 0,-1);  break;
    case 'E': xyz.put( 0,-1, 0);  break;
    case 'F': xyz.put(-1, 0, 0);  break;
    default:
      fprintf(stderr,"ERROR using rdie_dietoR3(char c): c=%c(0x%x) is not in the valid domain
          {A,B,C,D,E,F}\n",c,c);
      exit(EXIT_FAILURE);
    }
  return xyz;
  }

/*---------------------------------------------------------------------
 * map R^3 values to die face values using Lagrange Arc metric
 *---------------------------------------------------------------------*/
rdieseq rdie_R3todie_larc(seqR3 xyz){
  long n;
  int m;
  long N=xyz.getN();
  double d[7];
  double smallestd;
  char closestface;
  vectR3 p,q[7];
  rdieseq rdie(N);

  //q[0].put(0,0,0);
  q[1]=rdie_dietoR3('A');
  q[2]=rdie_dietoR3('B');
  q[3]=rdie_dietoR3('C');
  q[4]=rdie_dietoR3('D');
  q[5]=rdie_dietoR3('E');
  q[6]=rdie_dietoR3('F');

  for(n=0; n<N; n++){
    p.put(xyz.getx(n),xyz.gety(n),xyz.getz(n));
    smallestd=larc_metric(p,q[1]);
    closestface='A';
    for(m=2;m<7;m++){
      d[m] = larc_metric(p,q[m]);
      if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd))){
        //---------------------------      -----------------------------------
        // bias odd samples                  bias even samples
        // towards smaller values            towards larger values
        smallestd=d[m];
        closestface='A'+m-1;
        }
      }
    rdie.put(n,closestface);
    }
  return rdie;
  }

/*---------------------------------------------------------------------
 * map R^3 values to die face values and (0,0,0) using Lagrange Arc metric
 *---------------------------------------------------------------------*/
rdieseq rdie_R3todie0_larc(seqR3 xyz){
  long n;
  int m;
  long N=xyz.getN();
  double d[7];
  double smallestd;
  char closestface;
  vectR3 p,q[7];
  rdieseq rdie(N);

  q[0].put(0,0,0);
```

```
  q[1]=rdie_dietoR3('A');
  q[2]=rdie_dietoR3('B');
  q[3]=rdie_dietoR3('C');
  q[4]=rdie_dietoR3('D');
  q[5]=rdie_dietoR3('E');
  q[6]=rdie_dietoR3('F');

  for(n=0; n<N; n++){
    p.put(xyz.getx(n),xyz.gety(n),xyz.getz(n));
    smallestd=larc_metric(p,q[0]);
    //smallestd=ae_metric(1,p,q[0]);
    closestface='0';
    for(m=1;m<7;m++){
      d[m] = larc_metric(p,q[m]);
      if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd))){
        //---------------------------------    ------------------------------------
        // bias odd samples                     bias even samples
        // towards smaller values               towards larger values
        smallestd=d[m];
        closestface='A'+m-1;
        }
      }
    rdie.put(n,closestface);
    }
  return rdie;
  }

/*---------------------------------------------------------------------------
 * map R^3 values to die face values using Euclidean metric
 *   0    A    B    C    D    E    F    A+..+F
 *---------------------------------------------------------------------------*/
rdieseq rdie_R3todie_euclid(seqR3 xyz){
  long n;
  int m;
  long N=xyz.getN();
  double d[7];
  double smallestd;
  char closestface;
  vectR3 p,q[7];
  rdieseq rdie(N);

  q[1]=rdie_dietoR3('A');
  q[2]=rdie_dietoR3('B');
  q[3]=rdie_dietoR3('C');
  q[4]=rdie_dietoR3('D');
  q[5]=rdie_dietoR3('E');
  q[6]=rdie_dietoR3('F');

  for(n=0; n<N; n++){
    p.put(xyz.getx(n),xyz.gety(n),xyz.getz(n));
    smallestd=ae_metric(1,p,q[1]);
    closestface='A';
    for(m=2;m<=6;m++){
      d[m] = ae_metric(1,p,q[m]);
      if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd))){
        //---------------------------------    ------------------------------------
        // bias towards smaller values            bias towards larger values
        smallestd=d[m];
        closestface='A'+m-1;
        }
      }
    rdie.put(n,closestface);
    }
  return rdie;
  }

/*---------------------------------------------------------------------------
 * map R^3 values to die face and (0,0,0) values using Euclidean metric
 *   0    A    B    C    D    E    F    A+..+F
 *---------------------------------------------------------------------------*/
rdieseq rdie_R3todie0_euclid(seqR3 xyz){
  long n;
  int m;
  long N=xyz.getN();
  double d[7];
  double smallestd;
  char closestface;
```

```
  vectR3 p,q[7];
  rdieseq rdie(N);

  q[0]=rdie_dietoR3('0');
  q[1]=rdie_dietoR3('A');
  q[2]=rdie_dietoR3('B');
  q[3]=rdie_dietoR3('C');
  q[4]=rdie_dietoR3('D');
  q[5]=rdie_dietoR3('E');
  q[6]=rdie_dietoR3('F');

  for(n=0; n<N; n++){
    p.put(xyz.getx(n),xyz.gety(n),xyz.getz(n));
    smallestd=ae_metric(1,p,q[0]);
    closestface='0';
    for(m=1;m<=6;m++){
      d[m] = ae_metric(1,p,q[m]);
      //if(d[m]<smallestd)
      //if(d[m]<=smallestd)
      //if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<smallestd)))
      //if(((m&0x01) && (d[m]<=smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd)))
      if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd))){
        //----------------------------     -----------------------------------
        // bias towards smaller values         bias towards larger values
        smallestd=d[m];
        closestface='A'+m-1;
        }
      //if(m&0x01){ (alternative coding)
      //  if(d[m]<smallestd){
      //    smallestd=d[m];
      //    closestface='A'+m-1;
      //    }}
      //else
      //  if(d[m]<=smallestd){
      //    smallestd=d[m];
      //    closestface='A'+m-1;
      //    }
      }
    rdie.put(n,closestface);
    }
  return rdie;
  }

/*-----------------------------------------------------------------------------
 * map R^1 values to die face values using Euclidean metric
 *-----------------------------------------------------------------------------*/
rdieseq rdie_R1todie_euclid(seqR1 xyz){
  long n;
  long N=xyz.getN();
  char closestface;
  double p;
  rdieseq rdie(N);

  for(n=0; n<N; n++){
    p = xyz.get(n);
    if(p<1.5)       closestface='A';
    else if(p>=5.5) closestface='F';
    else            closestface=(char)(p+0.5-1)+'A';
    rdie.put(n,closestface);
    }
  return rdie;
  }

/*-----------------------------------------------------------------------------
 * real die metric d(a,b)
 *   d(a,b) | 0    A    B    C    D    E     F   (b)
 *   -------|--------------------------------------
 *      a= 0| 0    1    1    1    1    1    1
 *      a= A| 1    0    1    1    1    1    2
 *      a= B| 1    1    0    1    1    2    1
 *      a= C| 1    1    1    0    2    1    1
 *      a= D| 1    1    1    2    0    1    1
 *      a= E| 1    1    2    1    1    0    1
 *      a= F| 1    2    1    1    1    1    0
 * On success return d(a,b). On error return -1.
 *-----------------------------------------------------------------------------*/
double rdie_metric(char a, char b){
```

```
  int ra=rdie_dietoR1(a);
  int rb=rdie_dietoR1(b);
  double d;

  if(ra<0)fprintf(stderr,"a=%c(0x%x) not in domain of rdie metric d(a,b)\n",a,a);
  if(rb<0)fprintf(stderr,"b=%c(0x%x) not in domain of rdie metric d(a,b)\n",b,b);

       if(ra<0)      d=-1;
  else if(rb<0)      d=-1;
  else if(ra==rb)    {d=0;  }
  else if(ra==0)     d=1;
  else if(rb==0)     d=1;
  else if(ra+rb==7)  d=2;
  else               d=1;
  return d;
  }

/*-------------------------------------------------------------------------
 * real die metric p(x,y) where x and y are rdie sequences computed as
 * p(x,y) = d(x0,y0) + d(x1,y1) + d(x2,y2) + ... + d(x{N-1},y{N-1})
 * where d(a,b) is defined above.
 * On success return d(x,y). On error return -1.
 *-------------------------------------------------------------------------*/
double rdie_metric(rdieseq x, rdieseq y){
  double rval,d;
  long n;
  long N=x.getN();
  long M=y.getN();
  long NM=(N<M)?M:N;  //NM = the larger of N and M
  for(n=0,d=0;n<NM;n++){
    rval=rdie_metric(x.get(n),y.get(n));
    if(rval<0){d+=0.0; printf("rval=%lf ",rval);}
    else d+=rval;
    }
  if(N!=M){
    fprintf(stderr,"ERROR using rdie_metric(rdieseq x,rdieseq y): size of x (%ld) does not equal the
          size of y (%ld).\n",N,M);
    exit(EXIT_FAILURE);
    }
  return d;
  }
```

E.4 Spinner routines

```
/*=========================================================================
 * Daniel J. Greenhoe
 * header file for routines for spinner routines
 *   'A'--> spin face value 1
 *   'B'--> spin face value 2
 *   'C'--> spin face value 3
 *   'D'--> spin face value 4
 *   'E'--> spin face value 5
 *   'F'--> spin face value 6
 *=========================================================================*/
class spinseq: public dieseq {
  public:
    spinseq(const long M) : dieseq(M){};
    spinseq(const long M,const unsigned seed) : dieseq(M,seed){};
    seqR1   spintoR1(void);          //map spin face values to R^1
    seqR2 spintoR2(void);          //map spin face values to R^2
    void operator=(spinseq y);      //x=y
    int metrictbl(void);
    double Rxx (const long m);
    int    Rxx (const seqR1 *Rxx, const int showcount);
    int    Rxxo(const seqR1 *rxx, const int showcount);
    spinseq downsample(int factor);//downsample by a factor of <factor>
  };

extern int spin_domain(char c);//check if value is in the domain of rspin
extern spinseq spin_R1tospin_euclid(seqR1 xy);
extern spinseq spin_R2tospin_larc(seqR2 xy);
extern spinseq spin_R2tospin0_larc(seqR2 xy);
```

```
extern spinseq spin_R2tospin_euclid(seqR2 xy);
extern spinseq spin_R2tospin0_euclid(seqR2 xy);
extern vectR2  spin_spintoR2(char c);
extern double  spin_spintoR1(char c);
extern double spin_metric(char a, char b);
extern double spin_metric(spinseq x, spinseq y); //metric for two sequences
//extern seqR1 spin_correlation(spinseq x, spinseq y, int showcount);//correlation
//extern seqR1 spin_correlation(spinseq x, spinseq y){return spin_correlation(x,y,0);}//correlation
```

```
/*============================================================================
 * Daniel J. Greenhoe
 * routines for Real spin spinseqs
 *============================================================================*/
/*=====================================
 * headers
 *=====================================*/
#include<stdio.h>
#include<stdlib.h>
#include<math.h>
#include<main.h>
#include<symseq.h>
#include<r1.h>
#include<r2.h>
#include<r3.h>
#include<r4.h>
#include<r6.h>
#include<c1.h>
#include<euclid.h>
#include<larc.h>
#include<die.h>
#include<spinner.h>

/*-------------------------------------------------------------------------
 * display spinner metric table
 *-------------------------------------------------------------------------*/
int spinseq::metrictbl(void){
  char a,b;
  for(a='A';a<='F';a++){
    for(b='A';b<='F';b++)printf("d(%c,%c)=%.1lf ",a,b,spin_metric(a,b));
    printf("\n");
    }
  return 1;
  }

/*-------------------------------------------------------------------------
 * autocorrelation Rxx of a spinner seqR1 x with 2N offset
 *-------------------------------------------------------------------------*/
int spinseq::Rxxo(const seqR1 *rxx, const int showcount){
  const long N=getN();
  int rval;
  rval=Rxx(rxx,showcount);
  rxx->add(2*N);
  return rval;
  }

/*-------------------------------------------------------------------------
 * autocorrelation Rxx of a spinner seqR1 x
 *-------------------------------------------------------------------------*/
int spinseq::Rxx(const seqR1 *rxx, const int showcount){
  long m;
  const long N=getN();
  int rval=0;
  double rxxm;
  if(showcount)fprintf(stderr,"  Calculate %ld auto-correlation values ... n=",2*N+1);
  for(m=-N;m<=N;m++){
    if(showcount)fprintf(stderr,"%8ld",m+N);
    rxxm=Rxx(m);
    if(rxxm>0)rval=-1;
    rxx->put(m+N,rxxm);
    if(showcount)fprintf(stderr,"\b\b\b\b\b\b\b\b");
    }
  if(showcount)fprintf(stderr,"%8ld .... done.\n",m+N);
  return rval;
  }

/*-------------------------------------------------------------------------
 * autocorrelation Rxx(m)
```

```
 *-----------------------------------------------------------------------------*/
double spinseq::Rxx(const long m){
  const long mm=labs(m);
  const long N=getN();
  long n,nmm;
  double d,sum;
  char a,b;
  for(n=0,sum=0;n<(N+mm);n++){
    nmm=n-mm;
    a=(n   <0 || n   >=N)? 0.0 : get(n);
    b=(nmm<0 || nmm>=N)? 0.0 : get(nmm);
    d=(a==0 || b==0)?    1.0 : spin_metric(a,b);
    sum+=d;
    }
  return -sum;
  }

/*-----------------------------------------------------------------------------
 * downsample sequence by a factor of <factor>
 *-----------------------------------------------------------------------------*/
spinseq spinseq::downsample(int factor){
  const long N=getN();
  long n,m;
  long M;
  if(factor<1){
    fprintf(stderr,"ERROR using dieseq::downsample: factor=%d must be at least 1\n",factor);
    exit(EXIT_FAILURE);
    }
  M=N/factor;
  spinseq newseq(M);
  for(n=0,m=0; m<M; n+=factor,m++)newseq.put(m,get(n));
  return newseq;
  }


/*-----------------------------------------------------------------------------
 * map spin face values to R^1
 * A-->1 B-->2 C-->3 D-->4 E-->5 F-->6
 *-----------------------------------------------------------------------------*/
seqR1 spinseq::spintoR1(void){
  const long N=getN();
  long n;
  seqR1 y(N);
  for(n=0; n<N; n++)y.put(n,spin_spintoR1(get(n)));
  return y;
  }

/*-----------------------------------------------------------------------------
 * map spin face values to R^2 sequence
 *-----------------------------------------------------------------------------*/
seqR2 spinseq::spintoR2(void){
  const long N=getN();
  long n;
  seqR2 seqR2(N);
  for(n=0; n<N; n++)seqR2.put(n,spin_spintoR2(get(n)));
  return seqR2;
  }


/*=====================================
 * operators
 *=====================================*/
/*-----------------------------------------------------------------------------
 * operator spinseq x = dieseq y
 *-----------------------------------------------------------------------------*/
void spinseq::operator=(spinseq y){
  long n;
  const long N=getN();
  const long M=y.getN();
  char symbol;
  if(N!=M){
    fprintf(stderr,"\nERROR using spinseq x = spinseq y: size of x (%ld) is smaller than size of y (%ld)\n",N,M,N);
    exit(EXIT_FAILURE);
    }
  for(n=0;n<N;n++){
    symbol=y.get(n);
```

```
    put(n,symbol);
    }
  }

/*====================================
 * external operations
 *====================================*/
/*------------------------------------------------------------------------------
 * map spin face values to R^1
 *------------------------------------------------------------------------------*/
double spin_spintoR1(char c){
  double rval;
  switch(c){
    case 'A': rval = 1.0;  break;
    case 'B': rval = 2.0;  break;
    case 'C': rval = 3.0;  break;
    case 'D': rval = 4.0;  break;
    case 'E': rval = 5.0;  break;
    case 'F': rval = 6.0;  break;
    default:
      fprintf(stderr,"ERROR using spin_spintoR1(c): c=%c(0x%x) is not in the valid domain
              {A,B,C,D,E,F}\n",c,c);
      exit(EXIT_FAILURE);
    }
  return rval;
  }

/*------------------------------------------------------------------------------
 * map spinner face values to R^2
 *------------------------------------------------------------------------------*/
vectR2 spin_spintoR2(char c){
  vectR2 xy;
  switch(c){
    case 'A': xy.put(0,          -1.0);  break;
    case 'B': xy.put(+sqrt(3)/2,-0.5);  break;
    case 'C': xy.put(+sqrt(3)/2,+0.5);  break;
    case 'D': xy.put( 0,         +1.0);  break;
    case 'E': xy.put(-sqrt(3)/2,+0.5);  break;
    case 'F': xy.put(-sqrt(3)/2,-0.5);  break;
    default:
      fprintf(stderr,"ERROR: c=%c(0x%x) is not in the valid domain {0,A,B,C,D,E,F} in
              spin_spintoR2(char c)\n",c,c);
      exit(EXIT_FAILURE);
    }
  return xy;
  }

/*------------------------------------------------------------------------------
 * map R^2 values to spin face values using Lagrange Arc distance
 *------------------------------------------------------------------------------*/
spinseq spin_R2tospin_larc(seqR2 xy){
  long n;
  int m;
  long N=xy.getN();
  double d[7];
  double smallestd;
  char closestface;
  vectR2 p,q[7];
  spinseq rspin(N);

  //q[0].put(0,0,0);
  q[1]=spin_spintoR2('A');
  q[2]=spin_spintoR2('B');
  q[3]=spin_spintoR2('C');
  q[4]=spin_spintoR2('D');
  q[5]=spin_spintoR2('E');
  q[6]=spin_spintoR2('F');

  for(n=0; n<N; n++){
    p.put(xy.getx(n),xy.gety(n));
    smallestd=larc_metric(p,q[1]);
    closestface='A';
    for(m=2;m<7;m++){
      d[m] = larc_metric(p,q[m]);
      if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd))){
        //-----------------------------     ------------------------------------
        // bias odd samples                  bias even samples
```

```
        // towards smaller values            towards larger values
        smallestd=d[m];
        closestface='A'+m-1;
        }
      }
    rspin.put(n,closestface);
    }
  return rspin;
  }

/*-----------------------------------------------------------------------------
 * map R^3 values to spin face values and (0,0,0) using Lagrange Arc metric
 *-----------------------------------------------------------------------------*/
spinseq spin_R2tospin0_larc(seqR2 xy){
  long n;
  int m;
  long N=xy.getN();
  double d[7];
  double smallestd;
  char closestface;
  vectR2 p,q[7];
  spinseq rspin(N);

  q[0].put(0,0);
  q[1]=spin_spintoR2('A');
  q[2]=spin_spintoR2('B');
  q[3]=spin_spintoR2('C');
  q[4]=spin_spintoR2('D');
  q[5]=spin_spintoR2('E');
  q[6]=spin_spintoR2('F');

  for(n=0; n<N; n++){
    p.put(xy.getx(n),xy.gety(n));
    smallestd=larc_metric(p,q[0]);
    //smallestd=ae_metric(1,p,q[0]);
    closestface='0';
    for(m=1;m<7;m++){
      d[m] = larc_metric(p,q[m]);
      if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd))){
        //----------------------------      ----------------------------------
        // bias odd samples                 bias even samples
        // towards smaller values           towards larger values
        smallestd=d[m];
        closestface='A'+m-1;
        }
      }
    rspin.put(n,closestface);
    }
  return rspin;
  }

/*-----------------------------------------------------------------------------
 * map R^2 values to spin face values using Euclidean metric
 *    0    A    B    C    D    E    F    A+..+F
 *-----------------------------------------------------------------------------*/
spinseq spin_R2tospin_euclid(seqR2 xy){
  long n;
  int m;
  long N=xy.getN();
  double d[7];
  double smallestd;
  char closestface;
  vectR2 p,q[7];
  spinseq rspin(N);

  q[1]=spin_spintoR2('A');
  q[2]=spin_spintoR2('B');
  q[3]=spin_spintoR2('C');
  q[4]=spin_spintoR2('D');
  q[5]=spin_spintoR2('E');
  q[6]=spin_spintoR2('F');

  for(n=0; n<N; n++){
    p.put(xy.getx(n),xy.gety(n));
    smallestd=ae_metric(1,p,q[1]);
    closestface='A';
    for(m=2;m<=6;m++){
```

```
      d[m] = ae_metric(1,p,q[m]);
      if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd))){
        //------------------------------   -----------------------------------
        // bias towards smaller values         bias towards larger values
        smallestd=d[m];
        closestface='A'+m-1;
        }
      }
    rspin.put(n,closestface);
    }
  return rspin;
  }

/*-----------------------------------------------------------------------------
 * map R^2 values to spin face and (0,0) values using Euclidean metric
 *    0    A    B    C    D    E    F    A+..+F
 *-----------------------------------------------------------------------------*/
spinseq spin_R2tospin0_euclid(seqR3 xy){
  long n;
  int m;
  long N=xy.getN();
  double d[7];
  double smallestd;
  char closestface;
  vectR2 p,q[7];
  spinseq rspin(N);

  q[0]=spin_spintoR2('0');
  q[1]=spin_spintoR2('A');
  q[2]=spin_spintoR2('B');
  q[3]=spin_spintoR2('C');
  q[4]=spin_spintoR2('D');
  q[5]=spin_spintoR2('E');
  q[6]=spin_spintoR2('F');

  for(n=0; n<N; n++){
    p.put(xy.getx(n),xy.gety(n));
    smallestd=ae_metric(1,p,q[0]);
    closestface='0';
    for(m=1;m<=6;m++){
      d[m] = ae_metric(1,p,q[m]);
      //if(d[m]<smallestd)
      //if(d[m]<=smallestd)
      //if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<smallestd)))
      //if(((m&0x01) && (d[m]<=smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd)))
      if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd))){
        //------------------------------   -----------------------------------
        // bias towards smaller values         bias towards larger values
        smallestd=d[m];
        closestface='A'+m-1;
        }
      //if(m&0x01){ (alternative coding)
      //  if(d[m]<smallestd){
      //    smallestd=d[m];
      //    closestface='A'+m-1;
      //    }}
      //else
      //  if(d[m]<=smallestd){
      //    smallestd=d[m];
      //    closestface='A'+m-1;
      //    }
      }
    rspin.put(n,closestface);
    }
  return rspin;
  }

/*-----------------------------------------------------------------------------
 * map R^1 values to spin face values using Euclidean metric
 *-----------------------------------------------------------------------------*/
spinseq spin_R1tospin_euclid(seqR1 xy){
  long n;
  long N=xy.getN();
  char closestface;
  double p;
  spinseq rspin(N);

```

```
  for(n=0; n<N; n++){
    p = xy.get(n);
    if(p<1.5)         closestface='A';
    else  if(p>=5.5)  closestface='F';
    else              closestface=(char)(p+0.5-1)+'A';
    rspin.put(n,closestface);
    }
  return rspin;
  }

/*---------------------------------------------------------------------------
 * spinner metric d(a,b)
 *   d(a,b)  | A    B    C    D    E    F    (b)
 *   --------|-------------------------------
 *      a= A|  0    1    2    3    2    1
 *      a= B|  1    0    1    2    3    2
 *      a= C|  2    1    0    1    2    3
 *      a= D|  3    2    1    0    1    2
 *      a= E|  2    3    2    1    0    1
 *      a= F|  1    2    3    2    1    0
 * On success return d(a,b). On error return -1.
 *---------------------------------------------------------------------------*/
double spin_metric(char a, char b){
  double ra=spin_spintoR1(a);
  double rb=spin_spintoR1(b);
  double d;

  d=(double)fabs(ra-rb);
  if      (d>3.5)  d=2.0;
  else  if(d>4.5)  d=1.0;
  return d;
  }
```

E.5 DNA routines

```
/*===========================================================================
 * Daniel J. Greenhoe
 * header file for routines for DNA routines
 *===========================================================================*/
class dnaseq: public symseq {
  public:
    dnaseq(long M)  : symseq(M){};   //constructor initializing to '.'
    void    seed(unsigned seed){srand(seed);}     //
    void    randomize(void){symseq::randomize("ATCG");}     //
    void    randomize(unsigned seed){srand(seed); randomize();}
    int     randomize(long start, long end, int wA,int wT,int wC,int wG);
    int     randomize(unsigned seed,int wA,int wT,int wC,int wG){srand(seed);return
              randomize(0,getN()-1,wA,wT,wC,wG);}
    int     randomize(int wA,int wT,int wC,int wG){return randomize(0,getN()-1,wA,wT,wC,wG);}
    int     randomize(long start, long end, unsigned seed,int wA,int wT,int wC,int
              wG){srand(seed);return randomize(start,end,wA,wT,wC,wG);}
    char    get(long n){return symseq::get(n,"ATCG");}      //get a value from x at location n
    void    put(long n, char symbol){symseq::put(n,symbol);}
    void    put(dnaseq *y, const long n, const char symbol);
    void    put(dnaseq *y, long n){return put(y,n,'.');}
    void    put(const long start, const long end, char c);//put a value <c> at locations start to end
    seqR2   dnatoR2(void);          //map gsp face values to R^2
    dnaseq  downsample(int factor);//downsample by a factor of <factor>
    seqR1   dnatoR1(void);          //map dna sequence to R^1
    seqC1   dnatoC1(void);          //map dna sequence to R^1
    seqR1   dnatoR1pam(void);       //map dna sequence to R^1 using PAM scheme (symmetric about zero)
    seqR1   dnatoR1bin(void);     //map dna sequence to R^1 using AT-CG binary scheme
    seqR4   dnatoR4(void);          //map dna sequence to R^4
    double Rxx (const long m);
    int     Rxx (const seqR1 *Rxx, const int showcount);
    int     Rxxo(const seqR1 *rxx, const int showcount);

    seqR1   histogram(const long start, const long end, int display, FILE *fptr);//compute, display,
              and write histogram
    seqR1   histogram(){return histogram(0,getN()-1,0,NULL);}//compute histogram
    seqR1   histogram(const long start, const long end){return histogram(start,end,0,NULL);}//compute
              histogram
```

```
    seqR1  histogram(int display,FILE *fptr){return histogram(0,getN()-1,display,fptr);}//print histogram to file
    seqR1  histogram(FILE *fptr){return histogram(0,getN()-1,0,fptr);}//print histogram to file
    void operator=(dnaseq y);     //x=y
  };

extern long    numsym_fasta_file(const char *filename);
extern int     read_fasta_file(const char *filename, char *description, dnaseq *x);
extern int     dna_domain (char symbol);
extern vectR2  dna_dnatoR2(char symbol);
extern double  dna_dnatoR1(char symbol);
extern complex dnatoC1c    (char symbol);
extern vectR4  dnatoR4c    (char symbol);
```

```
/*============================================================================
 * Daniel J. Greenhoe
 * routines for Real gsp dnaseqs
 *============================================================================*/
/*=====================================
 * headers
 *=====================================*/
#include<stdio.h>
#include<stdlib.h>
#include<string.h>
#include<math.h>
#include<main.h>
#include<symseq.h>
#include<r1.h>
#include<r2.h>
#include<r3.h>
#include<r4.h>
#include<r6.h>
#include<c1.h>
#include<euclid.h>
#include<larc.h>
#include<dna.h>

/*=====================================
 * prototypes
 *=====================================*/
void dna_phistogram(seqR1 *data, const long start, const long end, FILE *ptr);


/*----------------------------------------------------------------------------
 * copy a dnaseq <y> into dnaseq x starting at location <n>
 * and fill any remaining locations with <c>
 *----------------------------------------------------------------------------*/
void dnaseq::put(dnaseq *y, const long n, const char symbol){
  const long N=getN();
  long i,j;
  long M=y->getN();
  if(n>=N){
    fprintf(stderr,"\nERROR using dnaseq:put(y,n,symbol): n=%ld outside sequence domain [0:%ld]\n",n,N-1);
    exit(EXIT_FAILURE);
    }
  if(!dna_domain(symbol)){
    fprintf(stderr,"\nERROR using dnaseq:put(y,n,symbol): symbol='%c'=0x%x not in sequence range {A,T,C,G}\n",symbol,symbol);
    exit(EXIT_FAILURE);
    }
  else{
    for(i=0;i<n;i++)      put(i,symbol);
    for(j=0;j<M;j++,i++)put(i,y->get(j));
    for(    ;i<N;i++)     put(i,symbol);
    }
  }

/*----------------------------------------------------------------------------
 * put the value <c> into the sequence x from location <start> to <end>
 *----------------------------------------------------------------------------*/
void dnaseq::put(const long start, const long end, const char symbol){
  const long N=getN();
  long n;
  if(start<0||end>=N||start>end){
    fprintf(stderr,"ERROR using dnaseq::put(%ld,%ld,'%c')\n",start,end,symbol);
    exit(EXIT_FAILURE);
```

```
    }
  for(n=start;n<=end;n++) put(n,symbol);
  }

/*-------------------------------------------------------------------------
 * fill the dnaseq with pseudo-random DNA values
 * using seed value <seed>
 * distributed with the weight values <wA,wT,wC,wG>
 * where each weight value wX in an integer in the closed interval [0,100]
 * and where the sum of the intervals must be 100.
 *-------------------------------------------------------------------------*/
int dnaseq::randomize(long start, long finish, int wA,int wT,int wC,int wG){
  int r,u;
  long n;
  int sum=wA+wT+wC+wG;
  char symbol;
  if(sum!=100){
    fprintf(stderr,"ERROR using dnaseq::randomize(start,finish,wA,wT,wC,wG): sum of weight values = %d != 100\n",sum);
    exit(EXIT_FAILURE);
    }
  for(n=start; n<=finish; n++){
    r=rand();
    u = r%100;
    if      (u<wA)           symbol='A';
    else if (u<wA+wT)        symbol='T';
    else if (u<wA+wT+wC)     symbol='C';
    else                     symbol='G';
    put(n,symbol);
    }
  return 0;
  }

/*-------------------------------------------------------------------------
 * map dna face values to R^1 using PAM scheme
 * A-->-1.5  T-->-0.5  C-->0.5  G-->0.5
 * all other values --> 0
 *-------------------------------------------------------------------------*/
seqR1 dnaseq::dnatoR1pam(void){
  const long N=getN();
  long n;
  char symbol;
  seqR1 seqR1(N);
  for(n=0; n<N; n++){
    symbol=get(n);
    switch(symbol){
      case 'A': seqR1.put(n,-1.5); break;
      case 'C': seqR1.put(n,-0.5); break;
      case 'T': seqR1.put(n, 0.5); break;
      case 'G': seqR1.put(n, 1.5); break;
      default:
        fprintf(stderr,"\nERROR using dnaseq:dnatoR1pam(): symbol='%c'=0x%x not in sequence range {A,T,C,G}\n",symbol,symbol);
        exit(EXIT_FAILURE);
      }
    }
  return seqR1;
  }

/*-------------------------------------------------------------------------
 * map dna face values to R^1 using AT/CG binary scheme
 * A-->1  T-->1  C-->-1  G-->-1
 * all other values --> 0
 *-------------------------------------------------------------------------*/
seqR1 dnaseq::dnatoR1bin(void){
  const long N=getN();
  long n;
  char symbol;
  seqR1 seqR1(N);
  for(n=0; n<N; n++){
    symbol=get(n);
    switch(symbol){
      case 'A': seqR1.put(n, 1); break;
      case 'C': seqR1.put(n,-1); break;
      case 'T': seqR1.put(n, 1); break;
      case 'G': seqR1.put(n,-1); break;
      default :
```

```
        fprintf(stderr,"\nERROR using dnaseq:dnatoR1bin(): symbol='%c'=0x%x not in sequence range {A,T,C,G}\n",symbol,symbol);
        exit(EXIT_FAILURE);
      }
    }
  return seqR1;
  }

/*-------------------------------------------------------------------------
 * map dna face values to C^1
 *-------------------------------------------------------------------------*/
seqC1 dnaseq::dnatoC1(void){
  const long N=getN();
  long n;
  char symbol;
  complex yy;
  seqC1 y(N);
  for(n=0; n<N; n++){
    symbol = get(n);
    yy = dnatoC1c(symbol);
    y.put(n,yy);
    }
  return y;
  }

/*-------------------------------------------------------------------------
 * map dna face values to complex plane C^1
 *
 *                    imaginary axis
 *                          |
 *                          |
 *                          |
 *  (cos135,sin135)=C       |      A=(cos45,sin45)
 *                          |
 *      ------------------|------------------ real axis
 *                          |
 *  (cos225,sin225)=T       |      G=(cos315,sin315)
 *                          |
 *                          |
 *                          |
 *
 *-------------------------------------------------------------------------*/
complex dnatoC1c(char c){
  complex rc;
  switch(c){
    case 'A': rc = expi( 45.0/180.0*PI);  break;
    case 'C': rc = expi(135.0/180.0*PI);  break;
    case 'T': rc = expi(225.0/180.0*PI);  break;
    case 'G': rc = expi(315.0/180.0*PI);  break;
    //
    //case 'A': rc = expi( 45.0/180.0*PI);  break;  //Gilleans 2007 mapping
    //case 'G': rc = expi(135.0/180.0*PI);  break;
    //case 'C': rc = expi(225.0/180.0*PI);  break;
    //case 'T': rc = expi(315.0/180.0*PI);  break;
    //
    //case 'A': rc.put(+1,+1);  break;
    //case 'G': rc.put(-1,+1);  break;
    //case 'C': rc.put(-1,-1);  break;
    //case 'T': rc.put(+1,-1);  break;
    case '0': rc.put( 0, 0);  break;
    default:  rc.put( 0, 0);
    fprintf(stderr,"ERROR using dnatoC1(char c): c=%c(0x%x) is not in the valid domain {0,A,C,T,G}. Returning (0,0).\n",c,c);
    }
  return rc;
  }

/*-------------------------------------------------------------------------
 * map dna values to R^4 sequence
 *-------------------------------------------------------------------------*/
seqR4 dnaseq::dnatoR4(void){
  const long N=getN();
  long n;
  char yc;
  seqR4 seq4(N);
  for(n=0; n<N; n++)seq4.put(n,dnatoR4c(get(n)));
  return seq4;
```

```
  }

/*-------------------------------------------------------------------------
 * map dna face values to R^4
 * A-->(1,0,0,0)      T-->(0,0,1,0)
 * C-->(0,1,0,0)      G-->(0,0,0,1)
 * 0-->(0,0,0,0)
 * on ERROR return (0,0,0,0)
 *-------------------------------------------------------------------------*/
vectR4 dnatoR4c(char c){
  vectR4 rfour;
  switch(c){
    case '0': rfour.put(0,0,0,0);  break;
    case 'A': rfour.put(1,0,0,0);  break;
    case 'C': rfour.put(0,1,0,0);  break;
    case 'T': rfour.put(0,0,1,0);  break;
    case 'G': rfour.put(0,0,0,1);  break;
    default:  rfour.put(0,0,0,0);
    fprintf(stderr,"ERROR using dnatoR4(char c): c=%c(0x%x) is not in the valid domain {A,C,T,G}.
          Returning (0,0,0,0).\n",c,c);
    }
  return rfour;
  }



/*-------------------------------------------------------------------------
 * map gsp face values to R^1
 * A-->1  T-->2  C-->3  G-->4
 *-------------------------------------------------------------------------*/
seqR1 dnaseq::dnatoR1(void){
  const long N=getN();
  long n;
  char symbol;
  seqR1 seqR1(N);
  for(n=0; n<N; n++){
    symbol = get(n);
    switch(symbol){
      case 'A': seqR1.put(n,1);  break;
      case 'T': seqR1.put(n,2);  break;
      case 'C': seqR1.put(n,3);  break;
      case 'G': seqR1.put(n,4);  break;
      default:
        fprintf(stderr,"\nERROR using dnaseq:dnatoR1(): symbol='%c'=0x%x not in sequence range
              {A,T,C,G}\n",symbol,symbol);
        exit(EXIT_FAILURE);
      }
    }
  return seqR1;
  }

/*-------------------------------------------------------------------------
 * map gsp face values to R^2 sequence
 *-------------------------------------------------------------------------*/
seqR2 dnaseq::dnatoR2(void){
  const long N=getN();
  long n;
  char symbol;
  vectR2 yy;
  seqR2 seqR2(N);
  for(n=0; n<N; n++){
    symbol=get(n);
    yy=dna_dnatoR2(symbol);
    seqR2.put(n,yy);
    }
  return seqR2;
  }

/*-------------------------------------------------------------------------
 * downsample seqR1 by a factor of <factor>
 *-------------------------------------------------------------------------*/
dnaseq dnaseq::downsample(int factor){
  const long N=getN();
  long n,m;
  long M;
  char symbol;
  if(factor<1){
```

```
    fprintf(stderr,"\nERROR using dnaseq::downsample: factor=%d must be at least 1\n",factor);
    exit(EXIT_FAILURE);
    }
  M=N/factor;
  dnaseq newseq(M);
  for(n=0,m=0; m<M; n+=factor,m++){
    symbol=get(n);
    newseq.put(m,symbol);
    }
  return newseq;
  }

/*-------------------------------------------------------------------------
 * compute histogram of dna sequence
 * return seqR1 y of length 6 where
 * y[1]-->number of dna 'A' symbols,
 * y[2]-->number of dna 'T' symbols,
 * y[3]-->number of dna 'C' symbols,
 * y[4]-->number of dna 'G' symbols,
 * y[0]-->number of all other values
 * y[5]-->total number of symbols y[1],y[2],...,y[5]
 *-------------------------------------------------------------------------*/
seqR1 dnaseq::histogram(const long start, const long end, int display, FILE *fptr){
  seqR1 data(6);
  long n;
  long bin;
  double p;
  int i;
  char symbol;
  FILE *ptr;
  data.clear();
  for(n=start;n<=end;n++){
    symbol=get(n);
    switch(symbol){
      case 'A': bin=1; break;
      case 'T': bin=2; break;
      case 'C': bin=3; break;
      case 'G': bin=4; break;
      default : bin=0; break;
      }
    if(bin!=0) data.increment(5);
    data.increment(bin);
    }
  if(display)    dna_phistogram(&data,start,end,stdout);
  if(fptr!=NULL) dna_phistogram(&data,start,end,fptr  );
  return data;
  }

/*-------------------------------------------------------------------------
 * print DNA histogram with data pointed to by <data>
 * to stream pointed to by ptr
 *-------------------------------------------------------------------------*/
void dna_phistogram(seqR1 *data, const long start, const long end, FILE *ptr){
  const long N=end-start+1;
  long bin;
  fprintf(ptr,"\n");
  fprintf(ptr," ------------------------------------------------------------------------\n");
  fprintf(ptr,"|  Histogram for dna sequence [x_n|n=%7ld-%7ld] (length %7ld)     |\n|",start,end,N);
  fprintf(ptr,"          A          T          C          G         AT         CG      other    |\n|");
  for(bin=1;bin<=4;bin++)fprintf(ptr,"%10.0lf",data->get(bin));
  fprintf(ptr,"%10.0lf%10.0lf%10.0lf
            |\n|",data->get(1)+data->get(2),data->get(3)+data->get(4),data->get(0));
  for(bin=1;bin<=4;bin++)fprintf(ptr,"  (%6.2lf%%)",data->get(bin)/(double)N*100.0);
  fprintf(ptr,"  (%6.2lf%%)  (%6.2lf%%)   (%6.2lf%%)
            |\n",(data->get(1)+data->get(2))/(double)N*100.0,(data->get(3)+data->get(4))/(double)N*100.0,data->get(0
  fprintf(ptr," ------------------------------------------------------------------------\n");
  }

/*====================================
 * operators
 *====================================*/
/*-------------------------------------------------------------------------
 * operator dnaseq x = dnaseq y
 *-------------------------------------------------------------------------*/
void dnaseq::operator=(dnaseq y){
  const long N=getN();
  const long M=y.getN();
```

```
  long n;
  char symbol;
  if(N!=M){
    fprintf(stderr,"\nERROR using dnaseq::operator=: length of x (%ld) differs from length of y
              (%ld).\n",N,M);
    exit(EXIT_FAILURE);
    }
  for(n=0;n<N;n++){
    symbol = y.get(n);
    put(n,symbol);
    }
  }

/*=====================================
 * external operations
 *=====================================*/
/*-------------------------------------------------------------------------
 * map dna symbols to R^1
 * A-->1 T-->2 C-->3 G-->4 0-->0  other-->-1
 *-------------------------------------------------------------------------*/
double dna_dnatoR1(char symbol){
  double r;
  switch(symbol){
    case 'A': r=1.0;  break;
    case 'T': r=2.0;  break;
    case 'C': r=3.0;  break;
    case 'G': r=4.0;  break;
    default :
    fprintf(stderr,"ERROR using dna_dnatoR1(symbol): symbol='%c' (0x%x) is not in the valid domain
              {A,T,C,G}\n",symbol,symbol);
    exit(EXIT_FAILURE);
    }
  return r;
  }

/*-------------------------------------------------------------------------
 * map dna symbols to R^2
 *-------------------------------------------------------------------------*/
vectR2 dna_dnatoR2(char symbol){
  vectR2 r;
  switch(symbol){
    case 'A': r.put( 1.0,      0  );  break;
    case 'T': r.put(-1.0,      0  );  break;
    case 'C': r.put( 0,      +1.0);  break;
    case 'G': r.put( 0,      -1.0);  break;
    default :
    fprintf(stderr,"ERROR using dna_dnatoR2(symbol): symbol='%c' (0x%x) is not in the valid domain
              {A,T,C,G}\n",symbol,symbol);
    exit(EXIT_FAILURE);
    }
  return r;
  }

/*-------------------------------------------------------------------------
 * map R^2 values to dna values using Euclidean metric
 *   0    A    B    C    D    E    F   A+..+F
 *-------------------------------------------------------------------------*/
dnaseq dna_R2todna_euclid(seqR2 xy){
  long n;
  int m;
  long N=xy.getN();
  double d[5];
  double smallestd;
  char closestface;
  vectR2 p,q[5];
  dnaseq rdna(N);

  //q[0].put(0,0,0);
  q[1]=dna_dnatoR2('A');
  q[2]=dna_dnatoR2('T');
  q[3]=dna_dnatoR2('C');
  q[4]=dna_dnatoR2('G');

  for(n=0; n<N; n++){
    p.put(xy.getx(n),xy.gety(n));
    smallestd=ae_metric(1,p,q[1]);
    closestface='A';
```

```
    for (m=2;m<5;m++){
      d[m]  = ae_metric(1,p,q[m]);
      if(((m&0x01) && (d[m]<smallestd)) || ((!(m&0x01)) && (d[m]<=smallestd))){
        //------------------------------      ----------------------------------
        // bias odd samples                     bias even samples
        // towards smaller values               towards larger values
        smallestd=d[m];
        switch(m){
          case 1: closestface = 'A'; break;
          case 2: closestface = 'T'; break;
          case 3: closestface = 'C'; break;
          case 4: closestface = 'G'; break;
          default: fprintf(stderr,"Error in dna_R2todna_larc(seqR2 xy)\n");
          }
        }
      }
    rdna.put(n,closestface);
    }
  return rdna;
  }

/*-----------------------------------------------------------------------------
 * map R^1 values to gsp face values using Euclidean metric
 *-----------------------------------------------------------------------------*/
dnaseq dna_R1todna_euclid(seqR1 xy){
  long n;
  long N=xy.getN();
  char closestface;
  double p;
  dnaseq rgsp(N);

  for(n=0; n<N; n++){
    p = xy.get(n);
    if(p<1.5)          closestface='A';
    else if(p>=3.5)    closestface='G';
    else if(p>=2.5)    closestface='C';
    else               closestface='T';
    rgsp.put(n,closestface);
    }
  return rgsp;
  }

/*-----------------------------------------------------------------------------
 * dna metric d(a,b)
 *     d(a,b)| 0    A    T    C    G  (b)
 *   --------|---------------------------
 *       a= 0| 0    1    1    1    1
 *       a= A| 1    0    1    1    1
 *       a= T| 1    1    0    1    1
 *       a= C| 1    1    1    0    1
 *       a= G| 1    1    1    1    0
 * On success return d(a,b). On error return -1.
 *-----------------------------------------------------------------------------*/
double dna_metric(char a, char b){
  int ra=dna_dnatoR1(a);
  int rb=dna_dnatoR1(b);
  double d;

  if(ra<0)fprintf(stderr,"a=%c(0x%x) not in domain of gsp metric d(a,b)\n",a,a);
  if(rb<0)fprintf(stderr,"b=%c(0x%x) not in domain of gsp metric d(a,b)\n",b,b);

       if(ra<0)       d=-1.0;
  else if(rb<0)       d=-1.0;
  else if(ra==rb)     d= 0.0;
  else                d= 1.0;
  return d;
  }

/*-----------------------------------------------------------------------------
 * real gsp metric p(x,y) where x and y are rgsp sequences computed as
 * p(x,y) = d(x0,y0) + d(x1,y1) + d(x2,y2) + ... + d(x{N-1},y{N-1})
 * where d(a,b) is defined above.
 * On success return d(x,y). On error return -1.
 *-----------------------------------------------------------------------------*/
double dna_metric(dnaseq x, dnaseq y){
  double rval,d;
  long n;
```

```
  long N=x.getN();
  long M=y.getN();
  long NM=(N<M)?N:M;  //NM = the smaller of N and M
  for(n=0,d=0;n<NM;n++){
    rval=dna_metric(x.get(n),y.get(n));
    if(rval<0){d+=0.0; printf("rval=%lf ",rval);}
    else d+=rval;
    }
  if(N!=M){
    fprintf(stderr,"ERROR using dna_metric(x,y): size of x (%ld) does not equal the size of y
              (%ld)\n",N,M);
    exit(EXIT_FAILURE);
    }
  return d;
  }

/*-------------------------------------------------------------------------
 * autocorrelation Rxx of a dna sequence x with 2N offset
 *-------------------------------------------------------------------------*/
int dnaseq::Rxxo(const seqR1 *rxx, const int showcount){
  const long N=getN();
  int rval;
  rval=Rxx(rxx,showcount);
  rxx->add(2*N);
  return rval;
  }

/*-------------------------------------------------------------------------
 * autocorrelation Rxx of a dna sequence x
 *-------------------------------------------------------------------------*/
int dnaseq::Rxx(const seqR1 *rxx, const int showcount){
  long m;
  const long N=getN();
  int rval=0;
  double rxxm;
  if(showcount)fprintf(stderr,"  Calculate %ld auto-correlation values ... n=",2*N+1);
  for(m=-N;m<=N;m++){
    if(showcount)fprintf(stderr,"%8ld",m+N);
    rxxm=Rxx(m);
    if(rxxm>0)rval=-1;
    rxx->put(m+N,rxxm);
    if(showcount)fprintf(stderr,"\b\b\b\b\b\b\b\b");
    }
  if(showcount)fprintf(stderr,"%8ld .... done.\n",m+N);
  return rval;
  }

/*-------------------------------------------------------------------------
 * autocorrelation Rxx(m)
 *-------------------------------------------------------------------------*/
double dnaseq::Rxx(const long m){
  const long mm=labs(m);
  const long N=getN();
  long n,nmm;
  double d,sum;
  char a,b;
  for(n=0,sum=0;n<(N+mm);n++){
    nmm=n-mm;
    a=(n   <0 || n   >=N)? 0.0 : get(n);
    b=(nmm<0 || nmm>=N)? 0.0 : get(nmm);
    d=(a==0 || b==0)?     1.0 : dna_metric(a,b);
    sum+=d;
    }
  return -sum;
  }

/*-------------------------------------------------------------------------
 * read dnan sequence from FASTA formatted file
 * and return how many symbols are in it.
 * reference: https://www.genomatix.de/online_help/help/sequence_formats.html
 *-------------------------------------------------------------------------*/
long numsym_fasta_file(const char *filename){
  FILE *fptr;
  int bufN;
  long N=0;
  char buffer[1024];

```

```
  if(filename==NULL) fptr=stdout;
  else               fptr=fopen(filename,"r");
  if(fptr==NULL){
    fprintf(stderr,"Unable to open file %s for reading.\n",filename);
    return -1;
    }
  while(fgets(buffer,1024,fptr)!=NULL){
    if(buffer[0]=='>'); //printf("description: %s",buffer);
    else{
      bufN = strlen(buffer);
      N += bufN-1;
      }
    }
  return N;
  }

/*-------------------------------------------------------------------------
 * read dnan sequence into <x> from FASTA formatted file
 * reference: https://www.genomatix.de/online_help/help/sequence_formats.html
 *-------------------------------------------------------------------------*/
int read_fasta_file(const char *filename, char *description, dnaseq *x){
  FILE *fptr;
  char buffer[1024];
  int bufN,i;
  long n;
  char symbol;

  if(filename==NULL) fptr=stdout;
  else               fptr=fopen(filename,"r");
  if(fptr==NULL){
    fprintf(stderr,"\nERROR using read_fasta_file(%s,...): unable to open file.\n",filename);
    exit(EXIT_FAILURE);
    }
  n=0;
  sprintf(description,"No description line found in FASTA file %s.",filename);//default description
  while(fgets(buffer,1024,fptr)!=NULL){
    if(buffer[0]=='>')strcpy(description,buffer);
    else{
      bufN = strlen(buffer);
      for(i=0;i<bufN-1;i++){
        switch(buffer[i]){
          case 'A': symbol='A'; break;
          case 'T': symbol='T'; break;
          case 'C': symbol='C'; break;
          case 'G': symbol='G'; break;
          case 'a': symbol='A'; break;
          case 't': symbol='T'; break;
          case 'c': symbol='C'; break;
          case 'g': symbol='G'; break;
          default:  symbol='x'; fprintf(stderr,"unknown character %c (%02x) in
                    dnaseq\n",buffer[i],buffer[i]);
          }
        x->put(n,symbol);
        n++;
        }
      }
    }
  fclose(fptr);
  return 0;
  }
```

E.6 Legrange arc distance routines

```
/*=========================================================================
 * Daniel J. Greenhoe
 * header file for routines for Lagrange arcs
 *=========================================================================*/
/*-------------------------------------------------------------------------
 * Lagrange arc class
 *-------------------------------------------------------------------------*/
class larcc{
  private:
```

```
    vectR2 p,q;
  public:
    larcc(vectR2 pp, vectR2 qq){p=pp; q=qq;}            //constructor
    larcc(double px, double py, double qx, double qy){p.put(px,py); q.put(qx,qy);}
    larcc(void){p.put(0,0); q.put(0,0);}
    void   setp(vectR2 pp){p=pp;}
    void   setq(vectR2 qq){q=qq;}
    void   setp(double px, double py){p.put(px,py);}
    void   setq(double qx, double qy){p.put(qx,qy);}
    vectR2  getp(void){return p;}
    vectR2  getq(void){return q;}
    double r(double theta);
    vectR2  x(double theta){return r(theta)*cos(theta);}
    vectR2  y(double theta){return r(theta)*sin(theta);}
    vectR2  xy(double theta);
    double indefint(double theta);// indefinite integral of arc length
    double arclength(void);
    double arclength(long int N);
    //double  operator|(double x, double y){double z; if(x==0)return 0; else return x/y;}  // division
              with 0/y = 0 even when y=0
    //double  operator|(double x, double y){double z; if(x==0)z=0; else z=x/y; return z;}  // division
              with 0/y = 0 even when y=0
  };

/*=====================================
 * prototypes
 *=====================================*/
extern double larc_arclength(double rp, double rq, double tdiff);
extern double larc_indefint(double rp, double rq, double thetap, double thetaq, double theta);
extern double larc_metric(const vectR2 p, const vectR2 q);
extern double larc_metric(const vectR2 p, const vectR2 q, long int N);
extern double larc_metric(const vectR3 p, const vectR3 q);
extern double larc_metric(const vectR4 p, const vectR4 q);
extern double larc_metric(const vectR6 p, const vectR6 q);
//extern vectR2 larc_findq(const vectR2 p, const double theta, const double d, const double minrq,
          const double maxrq, const double maxerror, const long N);
extern int larc_findq(const vectR2 p, const double theta, const double d, const double minrq, const
          double maxrq, const double maxerror, const long N, vectR2 *q);
extern vectR3 larc_findq(const vectR3 p, const double theta, const double phi, const double d, const
          double minrq, const double maxrq, const double maxerror, const long N);
//extern vectR3 larc_findq(vectR3 p, double theta, double phi, double d, long int N);
extern double larc_tau(const double a, const double sigma, const vectR2 p, const vectR2 q, const
          vectR2 r);
```

```
/*==========================================================================
 * Daniel J. Greenhoe
 * routines for Lagrange arcs
 * Largrange arcs are defined here in a manner analogous to
 * Lagrange polynomial interpolation.
 * Langrange polynomial interpolation is typically defined using
 * Cartesian coordinates in the R^2 plane.
 * Here, "Lagrange arcs" use basically the same idea, but are defined using
 * polar coordinates in the R^2 plane:
 *          y
 *          |    o p        Let (rp,tp) be the polar location of point p.
 *          |   /           where rp is the Euclidean distance from (0,0) to p
 *          |  /            and tp is radian measure from the x-axis to p.
 *          |/ tp           Let (rq,tq) be the polar location of point q.
 * ---------|---------- x   The "Lagrange arc" r(theta) is defined here as
 *          |\ tq                      theta -tq        theta -tp
 *          | \              r(theta) = rp ----------- + rq -----------
 *          |  o q                       tp-tq            tq-tp
 *          |
 *==========================================================================*/
/*=====================================
 * headers
 *=====================================*/
#include<stdio.h>
#include<stdlib.h>
#include<math.h>
#include<main.h>
#include<r1.h>
#include<r2.h>
#include<r3.h>
#include<r4.h>
#include<r6.h>
#include<euclid.h>
```

```
#include<larc.h>

/*-------------------------------------------------------------------------
 * path length s of Lagrange arc from a point p at polar coordinate (rp,tp)
 *                                     to point q at polar coordinate (rq,tq).
 *              __tq                  __tq
 *             |                     |                  (     dr     )^2
 *     s =     |   ds dtheta =       |     sqrt(r^2 + (--------- )     )  dtheta
 *           __|tp                 __|tp                (  dtheta    )
 *
 * reference: Paul Dawkins,
 *   http://tutorial.math.lamar.edu/Classes/CalcII/PolarArcLength.aspx
 *   https://books.google.com/books?id=b4ksCQAAQBAJ&pg=PA533
 *-------------------------------------------------------------------------*/
double larc_arclength(double rp, double rq, double tdiff){
  double y;
  const double phi=fabs(tdiff);
  const double rho=rq-rp;
  const double sp=sqrt(rp*rp*phi*phi+rho*rho);
  const double sq=sqrt(rq*rq*phi*phi+rho*rho);
  const double up=rp*rho*phi+fabs(rho)*sp;
  const double uq= rq*rho*phi+fabs(rho)*sq;
  if(rp==0)    {fprintf(stderr,"\nERROR using larc_arclength(rp,rq,tdiff): rp=%lf\n",      rp); exit(EXIT_FAILURE);}
  if(rq==0)    {fprintf(stderr,"\nERROR using larc_arclength(rp,rq,tdiff): rp=%lf\n",      rq); exit(EXIT_FAILURE);}
  if(tdiff<=0) {fprintf(stderr,"\nERROR using larc_arclength(rp,rq,tdiff): tdiff=%lf\n",   tdiff); exit(EXIT_FAILURE);}
  if(tdiff>PI) {fprintf(stderr,"\nERROR using larc_arclength(rp,rq,tdiff): tdiff=%lf>PI\n",tdiff); exit(EXIT_FAILURE);}
  //y = (larc_indefint(rp,rq,0,tdiff,tdiff)-larc_indefint(rp,rq,0,tdiff,0))/tdiff;
  //y2 = (rho>=0)? (fabs(rho)/(2*phi))*(log( rq*phi+sq)-log( rp*phi+sp))
  //             : (fabs(rho)/(2*phi))*(log(-rq*phi+sq)-log(-rp*phi+sp));
  if(fabs(rho)<=0.00000000001) y=rp*phi;
  else{
    if(up<=0){fprintf(stderr,"\nERROR using larc_arclength(rp,rq,tdiff): up=%.20lf rp=%lf rho=%lf sp=%lf\n", up,rp,rho,sp);     exit(EXIT_FAILURE);}
    if(uq<=0){fprintf(stderr,"\nERROR using larc_arclength(rp,rq,tdiff): uq=%.20lf rp=%lf rho=%lf sp=%lf\n", uq,rq,rho,sq);     exit(EXIT_FAILURE);}
    y = (rq*sq - rp*sp)/(2*rho) + fabs(rho)*(log(uq)-log(up))/(2*phi);
    }
  return y;
  }

/*-------------------------------------------------------------------------
 * indefinite integral for arc length
 * reference: http://integral-table.com/
 *            http://integral-table.com/downloads/integral-table.pdf
 *            indefinite integral (37)
 *            accessed 2015 September 19 12:29PM UTC
 * Note: This function should be viewed as DEPRECATED
 * (that is, don't use it for general computations),
 * However, this function is still useful for testing and verification of
 * larc_metric(vectR2 p, vectR2 q).
 *-------------------------------------------------------------------------*/
double larc_indefint(double rp, double rq, double thetap, double thetaq, double theta){
  double ra = (rp-rq);
  double rb = (rq*thetap-rp*thetaq);
  double a  = ra*ra;
  double b  = 2*ra*rb;
  double c  = ra*ra + rb*rb;
  double x  = theta;
  double y  = (b+2*a*x)/(4*a)*sqrt(a*x*x+b*x+c) +
            (4*a*c-b*b)/(8*a*sqrt(a))*log(2*a*x+b+2*sqrt(a*(a*x*x+b*x+c)));
//double y  = (b+2*a*x)/(4*a)*sqrt(a*x*x+b*x+c) +
          (4*a*c-b*b)/(8*a*sqrt(a))*log(fabs(2*a*x+b+2*sqrt(a*(a*x*x+b*x+c)))); // note: fabs(...) is an error in (37)
  return y;
  }

/*-------------------------------------------------------------------------
 * Lagrange arc metric from <p> to <q> in R^2
 *-------------------------------------------------------------------------*/
double larc_metric(const vectR2 p, const vectR2 q){
  const double rp=p.mag(), rq=q.mag();
  const double phi = pqtheta(p,q);
  const vectR2  pq=p-q;
```

```
  double d;
  if(rp==0 || rq==0 || phi<=0.0000001){//use Euclidean metric
    d = emetric(p,q);
    //printf("p=(%.2lf,%.2lf) q=(%.3lf,%.3lf) rq=%lf theta=%.12f PI phi=%.12f PI d=%lf
              ae\n",p.getx(),p.gety(),q.getx(),q.gety(),q.mag(),pqtheta(p,q)/PI, phi/PI,d);
    }
  else{//use Lagrange arc length
    d = larc_arclength(rp, rq, phi);
    //printf("p=(%.2lf,%.2lf) q=(%.3lf,%.3lf) rq=%lf theta=%.12f PI phi=%.12f PI d=%lf
              larc\n",p.getx(),p.gety(),q.getx(),q.gety(),q.mag(),pqtheta(p,q)/PI, phi/PI,d);
    }
  return d/PI;
  }

/*-------------------------------------------------------------------------
 * tau function for larc distance function d(p,q)
 * tau(a,sigma;p,q,r) := 2sigma[ 1/2 d^a(p,r) + 1/2 d^a(r,q) ]^(1/a)
 * reference:
 *   Daniel J. Greenhoe (2016)
 *   "Properties of distance spaces with power triangle inequalities"
 *   Carpathian Mathematical Publications, volume 8, number 1, pages 51--82
 *   doi 10.15330/cmp.8.1.51-82,
 *   http://www.journals.pu.if.ua/index.php/cmp/article/view/483
 *   https://peerj.com/preprints/2055/
 *   https://www.researchgate.net/publication/281831459
 *   section 4: Distance spaces with power triangle inequalities
 *-------------------------------------------------------------------------*/
double larc_tau(const double a, const double sigma, const vectR2 p, const vectR2 q, const vectR2 r){
  double dpr, drq;
  double tau;
  dpr = larc_metric(p,r);
  drq = larc_metric(r,q);
  tau = 2*sigma*pow((0.5*pow(dpr,a) + 0.5*pow(drq,a)),1.0/a);
  return tau;
  }

/*-------------------------------------------------------------------------
 * Lagrange metric from <p> to <q> computed numerically with resolution <N>.
 * Note: This function should be viewed as DEPRECATED
 * (that is, don't use it for general computations),
 * but instead it is strongly recommended to use larc_metric(vectR2 p, vectR2 q).
 * The function larc_metric(vectR2 p, vectR2 q) uses a closed form solution
 * (from an integral lookup table).
 * This function uses a numeric estimation
 * (by an approximated summation along the arc path).
 * However, this function is still useful for testing and verification of
 * larc_metric(vectR2 p, vectR2 q).
 *-------------------------------------------------------------------------*/
double larc_metric(const vectR2 p, const vectR2 q, const long int N){
  larcc arc(p,q);
  double d = arc.arclength(N);
  double ds=d/PI;
  return ds;
  }

/*-------------------------------------------------------------------------
 * Lagrange arc metric from <p> to <q> in R^3
 *-------------------------------------------------------------------------*/
double larc_metric(const vectR3 p, const vectR3 q){
  const double rp=p.mag(), rq=q.mag();
  const double tdiff = pqtheta(p,q);
  const vectR3  pq=p-q;
  double d;
  if(rp==0 || rq==0 || tdiff<=0) d = pq.mag();
  else if(rp==rq)                d = rp*tdiff;
  else                           d = larc_arclength(rp, rq, tdiff);
  return d/PI;
  }

/*-------------------------------------------------------------------------
 * Lagrange arc metric from <p> to <q> in R^3
 *-------------------------------------------------------------------------*/
double larc_metric(const vectR4 p, const vectR4 q){
  const double rp=p.mag(), rq=q.mag();
  const double tdiff = pqtheta(p,q);
  const vectR4  pq=p-q;
  double d;
```

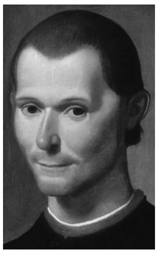
```
  if(rp==0 || rq==0 || tdiff<=0) d = pq.mag();
  else if(rp==rq)                d = rp*tdiff;
  else                           d = larc_arclength(rp, rq, tdiff);
  return d/PI;
  }

/*-------------------------------------------------------------------------
 * Lagrange arc metric from <p> to <q> in R^6
 *-------------------------------------------------------------------------*/
double larc_metric(const vectR6 p, const vectR6 q){
  const double rp=p.mag(), rq=q.mag();
  const double tdiff = pqtheta(p,q);
  const vectR6  pq=p-q;
  double d;
  if(rp==0 || rq==0 || tdiff<=0) d = pq.mag();
  else if(rp==rq)                d = rp*tdiff;
  else                           d = larc_arclength(rp, rq, tdiff);
  return d/PI;
  }

/*-------------------------------------------------------------------------
 * path length of arc computed using numeric integration
 *-------------------------------------------------------------------------*/
double larcc::arclength(long int N){
  double sum=0;
  double rp=p.mag(),     rq=q.mag();
  double tdiff=pqtheta(p,q);
  double tp=0, tq=tdiff;
  double theta=tp;
  long int n;
  vectR2 p1,p2;
  double delta=tdiff/(double)N;
  vectR2 pq=p-q;
  double d=pq.mag();//Euclidean distance(p,q)

  if(rp==0)    return d;
  if(rq==0)    return d;
  if(tdiff==0)return d;

  for (n=0; n<N; n++){
    p1 = xy(theta);
    theta += delta;
    p2 = xy(theta);
    sum += chordlength(p1,p2);
    }
  return sum;
  }

/*-------------------------------------------------------------------------
 * find the point (x(t),y(t)) on the Lagrange arc larc(p,q) at parameter <theta>
 *-------------------------------------------------------------------------*/
vectR2  larcc::xy(double theta){
  double rt=r(theta);
  vectR2 pt(rt*cos(theta),rt*sin(theta));
  return pt;
  }

/*-------------------------------------------------------------------------
 * return r(theta) for Lagrange arc(p,q)
 *-------------------------------------------------------------------------*/
double  larcc::r(double theta){
  double rp=p.mag();
  double rq=q.mag();
  double tdiff=pqtheta(p,q);
  double tp=0, tq=tdiff;
  double r =rp*(theta-tq)/(tp-tq) + rq*(theta-tp)/(tq-tp); // Lagrange polynomial of theta
  return r;
  }

/*-------------------------------------------------------------------------
 * Find a point q in R^2 orientated <phi> with respect to <p>
 * that is within a <maxerror> distance <d> from the point <p>.
 * Search for this point q using <N> search locations
 * over a radial distance from <p> of <minrq> to <maxrq>.
 * If a solution is found, place the point q at <*q> and return 1.
 * If a solution is not found and an apparent discontinuity occurred in
 * in the search, issue a warning and return 0.
```

```
 * If a solution is not found and a discontinuity apparently did NOT occur
 * in the search, issue an ERROR message and exit.
 *-------------------------------------------------------------------------*/
int larc_findq(const vectR2 p, const double theta, const double d, const double minrq, const double maxrq, const double maxerror, const long N, vectR2 *q){
  double rq,dd,ddprev,errord,bestrq,bestd,phi,smallesterror,discon1,discon2;
  vectR2 qq,bestq;
  int discontinuity=0,retval=1;

  qq.polartoxy(minrq,theta);//convert polar coor. to rectangular coordinates
  qq+=p;// search "origin" is the point p (not the R^2 origin (0,0))
  ddprev=larc_metric(p,qq);
  smallesterror=fabs(d-ddprev);

  for(rq=minrq; rq<=maxrq; rq+=(maxrq-minrq)/(double)N){
    qq.polartoxy(rq,theta);//convert polar coor. to rectangular coordinates
    qq+=p;// search "origin" is the point p (not the R^2 origin (0,0))
    dd=larc_metric(p,qq);
    if(fabs(dd-ddprev)>(maxerror*100)){
      discontinuity=1;
      retval=0;
      discon1=ddprev;
      discon2=dd;
      }
    ddprev=dd;
    errord=fabs(d-dd);
    if(errord<smallesterror){
      bestq=qq;
      bestrq=rq;
      bestd=dd;
      smallesterror=errord;
      phi = pqtheta(p,qq);
      }
    }
  if(smallesterror>maxerror){
    if(discontinuity){
      fprintf(stderr,"\nWARNING using larc_findq(vectR2 p,...): possible discontinuity,\n");
      fprintf(stderr,"  jumping from d=%lf to d=%lf.\n",discon1,discon2);
      fprintf(stderr,"  smallesterror=%lf > %lf=maxerror smallestd=%lf bestrq=%lf theta=%.12lf PI phi=%.12lf PI\n",smallesterror,maxerror,bestd,bestrq,theta/PI,phi/PI);
      }
    else{
      fprintf(stderr,"\nERROR using larc_findq(vectR2 p,...): no apparent discontinuity but...\n");
      fprintf(stderr,"  smallesterror=%lf > %lf=maxerror smallestd=%lf bestrq=%lf theta=%.12lf PI phi=%.12lf PI\n",smallesterror,maxerror,bestd,bestrq,theta/PI,phi/PI);
      exit(EXIT_FAILURE);
      }
    }
  *q = bestq;
  return retval;
  }

/*-------------------------------------------------------------------------
 * Find the polar length of a point q with radial measure tq that is a
 * distance <d> from the point <p> with polar coordinates (rp,tp)
 * using search resolution <N>
 *-------------------------------------------------------------------------*/
vectR3 larc_findq(const vectR3 p, const double theta, const double phi, const double d, const double minrq, const double maxrq, const double maxerror, const long int N){
  double rq,dd,errord,bestrq;
  vectR3 bestq(0,0,0);
  vectR3 q(0,0,0);
  double smallesterror=10000;

  for(rq=minrq; rq<=maxrq; rq+=(maxrq-minrq)/(double)N){
    q.polartoxyz(rq,theta,phi);//convert polar coor. to rectangular coordinates
    q+=p;// search "origin" is the point p (not the R^3 origin (0,0,0))
    dd=larc_metric(p,q);
    errord=fabs(d-dd);
    if(errord<smallesterror){
      bestq=q;
      bestrq=rq;
      smallesterror=errord;
      }
    }
  if(smallesterror>maxerror){
    fprintf(stderr,"\nERROR using larc_findq(vectR3 p,...):\n  smallesterror=%lf > %lf=maxerror
```

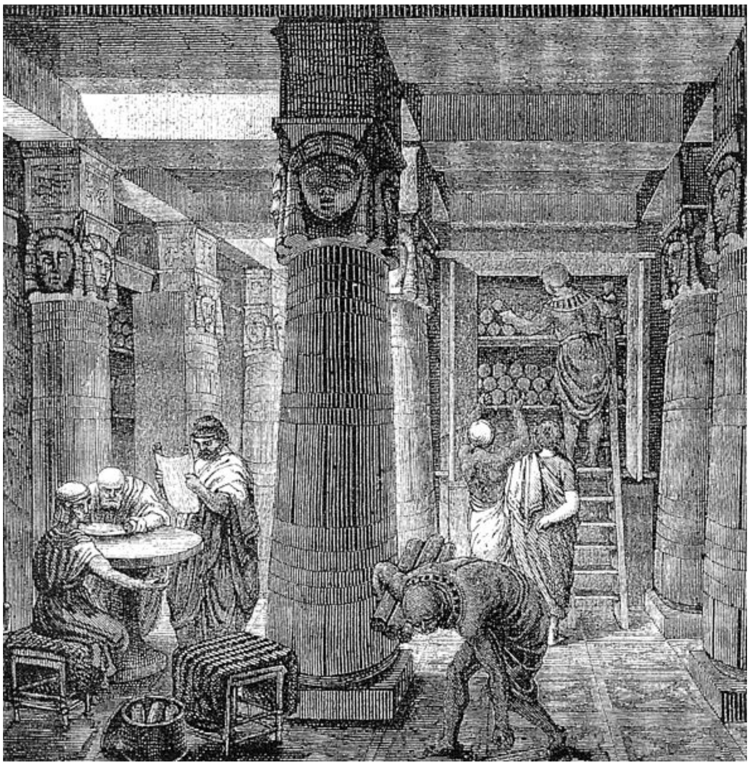
```
               bestrq=%lf  theta=%.2lf  PI\n",smallesterror,maxerror,bestrq,theta/PI);
    exit(EXIT_FAILURE);
    }
  return bestq;
  }
```

Back Matter

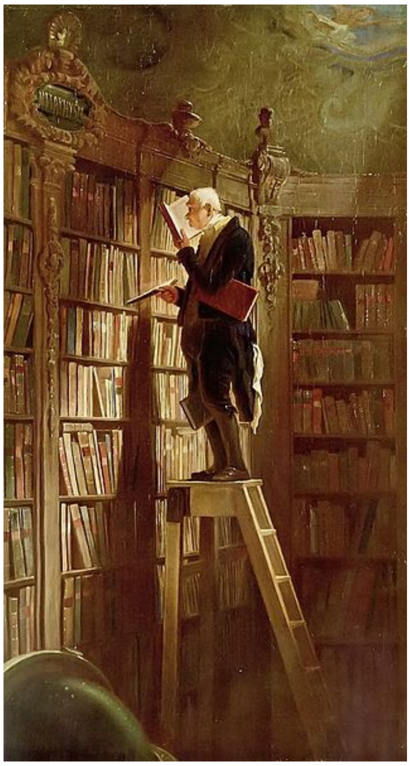
❝*It appears to me that if one wants to make progress in mathematics, one should study the masters and not the pupils.*❞
Niels Henrik Abel (1802–1829), Norwegian mathematician [4]



❝*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. [5]

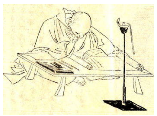


ancient library of Alexandria



The Book Worm by *Carl Spitzweg*, circa 1850

[6]



❝*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 [7]

---

[4] quote: Simmons (2007), page 187.
image: http://en.wikipedia.org/wiki/Image:Niels_Henrik_Abel.jpg, public domain

[5] quote: Machiavelli (1961), page 139?.
image: http://commons.wikimedia.org/wiki/File:Santi_di_Tito_-_Niccolo_Machiavelli%27s_portrait_headcrop.jpg, public domain

[6] http://en.wikipedia.org/wiki/File:Ancientlibraryalex.jpg, public domain http://en.wikipedia.org/wiki/File:Carl_Spitzweg_021.jpg,

[7] quote: Kenko (circa 1330)
image: http://en.wikipedia.org/wiki/Yoshida_Kenko

BIBLIOGRAPHY

Colin Conrad Adams and Robert David Franzosa. *Introduction to Topology: Pure and Applied*. Featured Titles for Topology Series. Pearson Prentice Hall, 2008. ISBN 9780131848696. URL `http://books.google.com/books?vid=ISBN0131848690`.

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`.

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`.

Tom M. Apostol. *Mathematical Analysis*. Addison-Wesley series in mathematics. Addison-Wesley, Reading, 2 edition, 1975. ISBN 986-154-103-9. URL `http://books.google.com/books?vid=ISBN0201002884`.

Aristotle. *Metaphysics*. University of Adelaide, Adelaide, 330BC? URL `http://etext.library.adelaide.edu.au/a/aristotle/metaphysics/`.

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`.

Ladislav Beran. *Orthomodular Lattices: Algebraic Approach*. Mathematics and Its Applications (East European Series). D. Reidel Publishing Company, Dordrecht, 1985. ISBN 90-277-1715-X. URL `http://books.google.com/books?vid=ISBN902771715X`.

Sterling Khazag Berberian. *Introduction to Hilbert Space*. Oxford University Press, New York, 1961. URL `http://books.google.com/books?vid=ISBN0821819127`.

Mihaly Bessenyei and Zsolt Pales. A contraction principle in semimetric spaces. *arXiv.org*, January 8 2014. URL http://arxiv.org/abs/1401.1709.

D. Besso. Teoremi elementari sui massimi i minimi. *Annuari Ist. Tech. Roma*, pages 7–24, 1879. see Bullen(2003) pages 453, 203.

M. Jales Bienaymé. Société philomatique de paris—extraits des procès-verbaux. *Scéance*, pages 67–68, June 13 1840. URL http://www.archive.org/details/extraitsdesproc46183941soci. see Bullen(2003) pages 453, 203.

Garrett Birkhoff. On the combination of subalgebras. *Mathematical Proceedings of the Cambridge Philosophical Society*, 29:441–464, October 1933. doi: 10.1017/S0305004100011464. URL http://adsabs.harvard.edu/abs/1933MPCPS..29..441B.

Garrett Birkhoff. Lattices and their applications. *Bulletin of the American Mathematical Society*, 44:1:793–800, 1938. doi: 10.1090/S0002-9904-1938-06866-8. URL http://www.ams.org/bull/1938-44-12/S0002-9904-1938-06866-8/.

Garrett Birkhoff. *Lattice Theory*. American Mathematical Society, New York, 2 edition, 1948. URL http://books.google.com/books?vid=ISBN3540120440.

R. B. Blackman and J. W. Tukey. The measurement of power spectra from the point of view of communications engineering—part ii. *The Bell System Technical Journal*, 37:485–569, March 1958. URL https://archive.org/download/bstj37-2-485.

R. B. Blackman and J. W. Tukey. *The Measurement of Power Spectra from the Point of View of Communications Engineering*. Dover Publications, New York, 1959. ISBN 486-60507-8. URL https://archive.org/download/TheMeasurementOfPowerSpectra. “unabridged and corrected republication of the work originally published in January and March, 1958, in Volume XXXVII of the *Bell System Technical Journal*”.

Leonard Mascot Blumenthal. Distance geometries: a study of the development of abstract metrics. *The University of Missouri studies. A quarterly of research*, 13(2):145, 1938.

Leonard Mascot Blumenthal. *Theory and Applications of Distance Geometry*. Oxford at the Clarendon Press, 1 edition, 1953. 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.

Umberto Bottazzini. *The Higher Calculus: A History of Real and Complex Analysis from Euler to Weierstrass*. Springer-Verlag, New York, 1986. ISBN 0-387-96302-2. URL http://books.google.com/books?vid=ISBN0387963022.

Bourbaki. *Éléments de mathématique. Prewmieère partie: Les structures fondamentales de l'analyse. Livre I: Théorie des ensembles (Fascicule des résultats)*. Hermann & Cie, Paris, 1939.

Bourbaki. *Theory of Sets*. Elements of Mathematics. Hermann, Paris, 1968. URL http://www.worldcat.org/oclc/281383.

J. L. Brenner. Limits of means for large values of the variable. *Pi Mu Epsilon Journal*, 8(3):160–163, Fall 1985. URL http://www.pme-math.org/journal/issues.html.

Jason I. Brown and Stephen Watson. The number of complements of a topology on n points is at least 2^n (except for some special cases). *Discrete Mathematics*, 154(1–3):27–39, 15 June 1996. doi: 10.1016/0012-365X(95)00004-G. URL http://dx.doi.org/10.1016/0012-365X(95)00004-G.

Andrew M. Bruckner, Judith B. Bruckner, and Brian S. Thomson. *Real Analysis*. Prentice-Hall, Upper Saddle River, N.J., 1997. ISBN 9780134588865. URL http://books.google.com/books?vid=ISBN013458886X.

Jacques Brunschwig, Geoffrey Ernest Richard Lloyd, and Pierre Pellegrin. *A Guide to Greek Thought: Major Figures and Trends*. Harvard University Press, 2003. ISBN 9780674021563. URL http://books.google.com/books?vid=ISBN0674021568.

Victor Bryant. *Metric Spaces: Iteration and Application*. Cambridge University Press, Cambridge, illustrated, reprint edition, 1985. ISBN 9780521318976. URL http://books.google.com/books?vid=ISBN0521318971.

Wlodzimierz Bryc. *The Normal Distribution: Characterizations with Applications*, volume 100 of *Lecture Notes in Statistics*. Springer Science & Business Media, 2012. ISBN 9781461225607. URL http://books.google.com/books?vid=ISBN1461225604. translated into English from Russian.

P. S. Bullen. Averages still on the move. *Mathematics Magazine*, 63(4):250–255, 1990. doi: 10.2307/2690948. URL http://www.maa.org/sites/default/files/Bullen94830627.pdf.

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.

J. C. Burkill. *The Lebesgue Integral*, volume 40 of *Cambridge Tracts in Mathematics*. Cambridge University Press, June 3 2004. ISBN 9780521604802. URL http://books.google.com/books?vid=ISBN052160480X.

Stanley Burris and Hanamantagida Pandappa Sankappanavar. *A Course in Universal Algebra*. Number 78 in Graduate texts in mathematics. Springer-Verlag, New York, 1 edition, 1981. ISBN 0-387-90578-2. URL http://books.google.com/books?vid=ISBN0387905782. 2000 edition available for free online.

Stanley Burris and Hanamantagida Pandappa Sankappanavar. A course in universal algebra. Retypeset and corrected version of the 1981 edition, 2000. URL http://www.math.uwaterloo.ca/~snburris/htdocs/ualg.html.

David Burstein, Igor Ulitsky, Tamir Tuller, and Benny Chor. Information theoretic approaches to whole genome phylogenies. In Satoru Miyano, Jill Mesirov, Simon Kasif, Sorin Istrail, Pavel A. Pevzner, and Michael Waterman, editors, *Research in Computational Molecular Biology. 9th Annual International Conference, RECOMB 2005*, volume 3500 of *Lecture Notes in Computer Science*, pages 283–295. Springer Science & Business Media, May 14–18 2005. ISBN 978-3-540-25866-7. URL http://www.researchgate.net/profile/Tamir_Tuller/publication/221530206_Information_Theoretic_Approaches_to_Whole_Genome_Phylogenies/links/0912f50b8c7a689215000000.pdf.

Herbert Busemann. *The Geometry of Geodesics*. Academic Press, 2 edition, 1955. ISBN 0486154629. URL http://books.google.com/books?vid=ISBN0486154629. a Dover 2005 edition has been published which “is an unabridged republication of the work originally published in 1955”.

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*.

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).

S. D. Chatterji. The number of topologies on n points. Technical Report N67-31144, National Aeronautics and Space Administration, July 1967. URL http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19670021815_1967021815.pdf. techreport.

Paul M. Cohn. *Basic Algebra; Groups, Rings and Fields*. Springer, December 6 2002. ISBN 1852335874. URL http://books.google.com/books?vid=isbn1852335874.

Paul M. Cohn. *Introduction to Ring Theory*. Springer Undergraduate Mathematics Series. Springer Science & Business Media, December 6 2012. ISBN 9781447104759. URL http://books.google.com/books?vid=isbn9781447104759.

Louis Comtet. Recouvrements, bases de filtre et topologies d'un ensemble fini. *Comptes rendus de l'Acade'mie des sciences*, 262(20):A1091–A1094, 1966. Recoveries, bases and filter topologies of a finite set.

Louis Comtet. *Advanced combinatorics: the art of finite and infinite*. D. Reidel Publishing Company, Dordrecht, 1974. ISBN 978-9027704412. URL http://books.google.com/books?vid=ISBN9027704414. translated and corrected version of the 1970 French edition.

Edward Thomas Copson. *Metric Spaces*. Number 57 in Cambridge tracts in mathematics and mathematical physics. Cambridge University Press, London, 1968. ISBN 978-0521047227. URL http://books.google.com/books?vid=ISBN0521047226.

Paul Corazza. Introduction to metric-preserving functions. *The American Mathematical Monthly*, 104(4):309–323, April 1999. URL http://pcorazza.lisco.com/papers/metric-preserving.pdf.

René Cori and Daniel Lascar. *Mathematical Logic: A Course with Exercises. Part II Recursion theory, Gödel's Theorems, Set Theory, Model Theory*. Oxford University Press, 2001. ISBN 0198500513. URL http://books.google.com/books?vid=ISBN0198500513.

Manuel Costa, Miguel Castro, Antony Rowstron, and Peter Key. Pic: Practical internet coordinates for distance estimation. In *Distributed Computing Systems, 2004. Proceedings. 24th International Conference on*, pages 178–187, 2004. doi: 10.1109/ICDCS.2004.1281582. URL http://research.microsoft.com/pubs/67616/tr-2003-53.pdf.

Koby Crammer, Michael Kearns, and Jennifer Wortman. Learning from multiple sources. In Bernhard Scholkopf, John Platt, and Thomas Hofmann, editors, *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*, volume 19 of *Neural information processing series*, pages 321–328. MIT Press, 2007. ISBN 9780262195683. URL http://machinelearning.wustl.edu/mlpapers/paper_files/NIPS2006_46.pdf.

Koby Crammer, Michael Kearns, and Jennifer Wortman. Learning from multiple sources. *Journal of Machine Learning Research*, 9:1757–1774, August 08 2008. URL http://repository.upenn.edu/cis_papers/407/.

Nello Cristianini and Matthew W. Hahn. *Introduction to Computational Genomics: A Case Studies Approach*. Cambridge University Press, 2007. ISBN 9781139460156. URL https://eembdersler.files.wordpress.com/2012/02/introduction-to-computational-genomics-a-case-studies-approach.pdf.

Martin D. Crossley. *Essential Topology*. Springer Undergraduate Mathematics Series. Springer Science & Business Media, 2006. ISBN 9781846281945. URL http://books.google.com/books?vid=ISBN1846281946.

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.

Stefan Czerwik. Contraction mappings in b-metric spaces. *Acta Mathematica et Informatica Universitatis Ostraviensis*, 1(1):5–11, 1993. URL http://dml.cz/dmlcz/120469.

Brian A. Davey and Hilary A. Priestley. *Introduction to Lattices and Order*. Cambridge mathematical text books. Cambridge University Press, Cambridge, 2 edition, May 6 2002. ISBN 978-0521784511. URL http://books.google.com/books?vid=ISBN0521784514.

Philip J. Davis and Reuben Hersh. *The Mathematical Experience*. Houghton Mifflin Books, Boston, 1999. ISBN 0395929687. URL http://books.google.com/books?vid=ISBN0395929687.

Sheldon W. Davis. *Topology*. McGraw Hill, Boston, 2005. ISBN 007-124339-9. URL http://www.worldcat.org/isbn/0071243399.

Richard Dedekind. Was sind und was sollen die zahlen? In Robert Fricke, Emmy Noether, and /:Oystein Ore, editors, *Gesammelte mathematische Werke*, pages 335–391. Druck and Verlag von Friedr. Vieweg and Sohn Akt.-Ges., Braunschweig, 1888a. URL http://resolver.sub.uni-goettingen.de/purl?PPN23569441X. What are and what should be numbers?

Richard Dedekind. The nature and meaning of numbers. In *Essays on the Theory of Numbers*, pages 14–58. The Open Court Publishing Company, Chicago, 1888b. URL http://www.gutenberg.org/etext/21016. 1901 English translation of *Was sind und was sollen die Zahlen?*

Richard Dedekind. Ueber die von drei moduln erzeugte dualgruppe. *Mathematische Annalen*, 53:371–403, January 8 1900. URL http://resolver.sub.uni-goettingen.de/purl/?GDZPPN002257947. Regarding the Dual Group Generated by Three Modules.

René Descartes. *Discours de la méthode pour bien conduire sa raison, et chercher la verite' dans les sciences*. Jan Maire, Leiden, 1637a. URL http://www.gutenberg.org/etext/13846.

René Descartes. *Discourse on the Method of Rightly Conducting the Reason in the Search for Truth in the Sciences*. 1637b. URL http://www.gutenberg.org/etext/59.

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.

Elena Deza and Michel-Marie Deza. *Dictionary of Distances*. Elsevier Science, Amsterdam, 2006. ISBN 0444520872. URL http://books.google.com/books?vid=ISBN0444520872.

Michel-Marie Deza and Elena Deza. *Encyclopedia of Distances*. Springer, 2009. ISBN 3642002331. URL http://www.uco.es/users/ma1fegan/Comunes/asignaturas/vision/Encyclopedia-of-distances-2009.pdf.

Michel-Marie Deza and Elena Deza. *Encyclopedia of Distances*. Springer, Bücher, 3 edition, 2014. ISBN 3662443422. URL http://books.google.com/books?vid=ISBN3662443422.

Emmanuele DiBenedetto. *Real Analysis*. Birkhäuser Advanced Texts. Birkhäuser, Boston, 2002. ISBN 0817642315. URL http://books.google.com/books?vid=ISBN0817642315.

Jean Alexandre Dieudonné. *Foundations of Modern Analysis*. Academic Press, New York, 1969. ISBN 1406727911. URL http://books.google.com/books?vid=ISBN1406727911.

Jozef Doboš. *Metric Preserving Functions*. Štroffek, 1998. ISBN 9788088896302. URL http://prof.jozef.xn--dobo-j6a.eu/files/2012/03/mpf1.pdf.

Tadeusz Dobrowolski and Jerzy Mogilski. Regular retractions onto finite dimensional convex sets. In Krzysztof Jarosz, editor, *Function Spaces: The Second Conference*, volume 172 of *lecture notes in pure and applied mathematics*, pages 85–99, proceedings of the conference at Edwardsville, 1995. Marcel Dekker Inc., New York.

J. Alejandro Dominguez-Torres. The origin and history of convolution i: continuous and discrete convolution operations. online, November 02 2010. URL http://www.slideshare.net/Alexdfar/origin-adn-history-of-convolution.

J. Alejandro Dominguez-Torres. A history of the convolution operation. *IEEE Pulse, Retroscope*, January/February 2015. URL http://pulse.embs.org/january-2015/history-convolution-operation/.

John R. Durbin. *Modern Algebra; An Introduction*. John Wiley & Sons, Inc., 4 edition, 2000. ISBN 0-471-32147-8. URL http://www.worldcat.org/isbn/0471321478.

W.D. Duthie. Segments of ordered sets. *Transactions of the American Mathematical Society*, 51(1): 1–14, January 1942. doi: 10.2307/1989978. URL http://www.jstor.org/stable/1989978.

Euclid. *Elements*. circa 300BC. URL http://farside.ph.utexas.edu/euclid.html.

L. Euler. De eximio usu methodi interpolationum in serierum doctrina. *Opuscula Analytica Petropoli*, 1:157–210, 1783. URL http://www.e-rara.ch/zut/content/structure/1197147. English translation of title: "An examination of the use of the method of interpolation in the doctrine of series".

J.W. Evans, Frank Harary, and M.S. Lynn. On the computer enumeration of finite topologies. *Communications of the ACM — Association for Computing Machinery*, 10:295–297, 1967. ISSN 0001-0782. URL http://portal.acm.org/citation.cfm?id=363282.363311.

David Ewen. *The Book of Modern Composers*. Alfred A. Knopf, New York, 1950. URL http://books.google.com/books?id=yHw4AAAAIAAJ.

David Ewen. *The New Book of Modern Composers*. Alfred A. Knopf, New York, 3 edition, 1961. URL http://books.google.com/books?id=bZIaAAAAMAAJ.

Ronald Fagin, Ravi Kumar, and D. Sivakumar. Comparing top k lists. *SIAM Journal on Discrete Mathematics*, 17(1):134–160, 2003a. doi: 10.1137/S0895480102412856. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.86.3234&rep=rep1&type=pdf.

Ronald Fagin, Ravi Kumar, and D. Sivakumar. Comparing top k lists. In *In Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, pages 28–36. Society for Industrial and Applied Mathematics, 2003b. doi: 10.1137/S0895480102412856. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.119.6597.

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.

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.

Abraham Fraenkel. Zu den grundlagen der cantor-zermeloschen mengenlehre. *Mathematische Annalen*, 86:230–237, 1922. URL http://dz-srv1.sub.uni-goettingen.de/cache/toc/D36920.html.

Duncan C. Fraser. *Newton's Interpolation Formulas*. C. & E. Layton, London, 1919. URL https://archive.org/details/newtonsinterpola00frasrich.

Maurice René Fréchet. Sur quelques points du calcul fonctionnel (on some points of functional calculation). *Rendiconti del Circolo Matematico di Palermo*, 22:1–74, 1906. Rendiconti del Circolo Matematico di Palermo (Statements of the Mathematical Circle of Palermo).

Maurice René Fréchet. *Les Espaces abstraits et leur théorie considérée comme introduction a l'analyse générale*. Borel series. Gauthier-Villars, Paris, 1928. URL http://books.google.com/books?id=9czoHQAACAAJ. Abstract spaces and their theory regarded as an introduction to general analysis.

Avner Friedman. *Foundations of Modern Analysis*. Holt, Rinehart and Winston, Inc. New York, 1970. ISBN 9780486640624. URL http://books.google.com/books?vid=ISBN0486640620. A 1982 "unabridged and corrected" Dover "republication" has been published.

Paul Abraham Fuhrmann. *A Polynomial Approach to Linear Algebra*. Springer Science+Business Media, LLC, 2 edition, 2012. ISBN 978-1461403371. URL http://books.google.com/books?vid=ISBN1461403375.

Lorenzo Galleani and Roberto Garello. The minimum entropy mapping spectrum of a dna sequence. *IEEE Transactions on Information Theory*, 56(2):771–783, February 2010. ISSN 0018-9448. doi: 10.1109/TIT.2009.2037041. URL http://dx.doi.org/10.1109/TIT.2009.2037041.

Fred Galvin and Samuel David Shore. Completeness in semimetric spaces. *Pacific Journal Of Mathematics*, 113(1):67–75, March 1984. doi: 10.2140/pjm.1984.113.67. URL http://msp.org/pjm/1984/113-1/pjm-v113-n1-p04-s.pdf.

C. F. Gauss. Theoria interpolationis methodo nova tractata. In *Werke*, volume III, pages 265–327. Königlichen Gesellschaft der Wissenschaften, Göttingen, Germany, 1866.

Michael C. Gemignani. *Elementary Topology*. Addison-Wesley Series in Mathematics. Addison-Wesley Publishing Company, Reading, Massachusetts, 2 edition, 1972. ISBN 9780486665221. URL http://books.google.com/books?vid=ISBN0486665224. A 1990 Dover “unabridged and corrected” edition has been published.

GenBank. *GenBank*. NCBI: National Center for Biotechnology Information, Rockville Pike Bethesda MD 20894 USA, 2014. URL http://www.ncbi.nlm.nih.gov.

GenBank-AF086833.2. *Ebola virus—Mayinga, Zaire, 1976, complete genome*. NCBI: National Center for Biotechnology Information, Rockville Pike Bethesda MD 20894 USA, February 13 2013. URL http://www.ncbi.nlm.nih.gov/nuccore/AF086833.2. accession AF086833.2.

GenBank-DS982815.1. *Carica papaya supercontig_1446 genomic scaffold*. NCBI: National Center for Biotechnology Information, Rockville Pike Bethesda MD 20894 USA, March 11 2015. URL http://www.ncbi.nlm.nih.gov/nuccore/DS982815.1. accession DS982815.1.

GenBank-NC_004718.3. *SARS coronavirus, complete genome*. NCBI: National Center for Biotechnology Information, Rockville Pike Bethesda MD 20894 USA, September 02 2011. URL http://www.ncbi.nlm.nih.gov/nuccore/30271926. accession NC_004718.3.

GenBank-NZ_CM003360.1. *Melissococcus plutonius strain 49.3 plasmid pMP19*. NCBI: National Center for Biotechnology Information, Rockville Pike Bethesda MD 20894 USA, August 21 2015. URL http://www.ncbi.nlm.nih.gov/nuccore/NZ_CM003360.1. accession NZ_CM003360.1.

Jean Dickinson Gibbons, Ingram Olkin, and Milton Sobel. *Selecting and Ordering Populations: A New Statistical Methodology*. John Wiley & Sons, New York, 1977. ISBN 9781611971101. URL http://books.google.com/books?vid=ISBN1611971101. A 1999 unabridged and corrected republication has been made available as an “SIAM Classics edition”, ISBN 9781611971101.

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.

Mariusz Grabiec, Yeol Je Cho, and Viorel Radu. *On Nonsymmetric Topological and Probabilistic Structures*. Nova Publishers, 2006. ISBN 9781594549175. URL http://books.google.com/books?vid=ISBN1594549176.

I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Elsevier, Amsterdam, 7 edition, February 23 2007. ISBN 9780080471112. URL http://books.google.com/books?vid=ISBN0080471110. translated from Russian.

Ivor Grattan-Guinness. *Convolutions in French mathematics, 1800-1840: from the calculus and mechanics to mathematical analysis and mathematical physics. Volume I: The Settings*, volume 2 of *Science networks. Historical Studies*. Birkhäuser Verlag, Basel, 1 edition, 1990. ISBN 9783764322373. URL http://books.google.com/books?vid=ISBN3764322373.

George A. Grätzer. *General Lattice Theory*. Birkhäuser Verlag, Basel, 2 edition, January 17 2003. ISBN 3-7643-6996-5. URL http://books.google.com/books?vid=ISBN3764369965.

Daniel J. Greenhoe. *Wavelet Structure and Design*, volume 3 of *Mathematical Structure and Design series*. Abstract Space Publishing, August 2013. ISBN 9780983801139. URL http://books.google.com/books?vid=ISBN0983801134. revised online version available at https://www.researchgate.net/publication/312529555.

Daniel J. Greenhoe. Order and metric geometry compatible stochastic processing. *PeerJ PrePrints*, February 19 2015. doi: 10.7287/peerj.preprints.844v1. URL https://peerj.com/preprints/844/. This paper has been submitted to the journal *Stochastic Systems* (http://www.i-journals.org/ssy/) on 2015 February 18 and as of 2017 January 18 (23 months later) no editor decision has yet been received.

Daniel J. Greenhoe. An extension to the spherical metric using polar linear interpolation. *PeerJ Preprints*, 4(e2467v1), 2016a. doi: 10.7287/peerj.preprints.2467v1. URL https://doi.org/10.7287/peerj.preprints.2467v1.

Daniel J. Greenhoe. Properties of distance spaces with power triangle inequalities. *Carpathian Mathematical Publications*, 8(1):51–82, 2016b. ISSN 2313-0210. doi: 10.15330/cmp.8.1.51-82. URL http://www.journals.pu.if.ua/index.php/cmp/article/view/483. preprint versions are available at http://www.researchgate.net/publication/281831459 and https://peerj.com/preprints/2055.

Daniel J. Greenhoe. Order and metric compatible symbolic sequence processing. *PeerJ Preprints*, May 18 2016c. doi: 10.7287/peerj.preprints.2052v1/supp-1. URL https://peerj.com/preprints/2052v1/.

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. Halmos. *Finite Dimensional Vector Spaces*. Princeton University Press, Princeton, 1 edition, 1948. ISBN 0691090955. URL http://books.google.com/books?vid=isbn0691090955.

Paul R. Halmos. *Measure Theory*. The University series in higher mathematics. D. Van Nostrand Company, New York, 1950. URL http://www.amazon.com/dp/0387900888. 1976 reprint edition available from Springer with ISBN 9780387900889.

Paul Richard Halmos. *Naive Set Theory*. The University Series in Undergraduate Mathematics. D. Van Nostrand Company, Inc., Princeton, New Jersey, 1960. ISBN 0387900926. URL http://books.google.com/books?vid=isbn0387900926.

Frank Haray. *Graph Theory*. Addison-Wesley Publishing Company, Reading-Massachusetts, 1969. URL http://www.dtic.mil/dtic/tr/fulltext/u2/705364.pdf.

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.

Felix Hausdorff. *Grundzüge der Mengenlehre*. Von Veit, Leipzig, 1914. URL http://books.google.com/books?id=KTs4AAAAMAAJ. Properties of Set Theory.

Felix Hausdorff. *Set Theory*. Chelsea Publishing Company, New York, 3 edition, 1937. ISBN 0828401195. URL http://books.google.com/books?vid=ISBN0828401195. 1957 translation of the 1937 German *Grundzüge der Mengenlehre*.

Robert W. Heath. A regular semi-metric space for which there is no semi-metric under which all spheres are open. *Proceedings of the American Mathematical Society*, 12: 810–811, 1961. ISSN 1088-6826. URL http://www.ams.org/journals/proc/1961-012-05/S0002-9939-1961-0125562-9/.

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.

Juha Heinonen. *Lectures on Analysis on Metric Spaces*. Universitext Series. Springer Science & Business Media, January 1 2001. ISBN 9780387951041. URL http://books.google.com/books?vid=ISBN0387951040.

Omar Hijab. *Introduction to Calculus and Classical Analysis*. Undergraduate Texts in Mathematics. Springer, 4 edition, February 9 2016. ISBN 9783319284002. URL http://books.google.com/books?vid=ISBN3319284002.

John G. Hocking and Gail S. Young. *Topology*. Addison-Wesley Publishing Company, Reading Massachusettes USA and London England, 1961. URL https://archive.org/details/Topology_972.

Larry Hoehn and Ivan Niven. Averages on the move. *Mathematics Magazine*, 58(3):151–156, May 1985. doi: 10.2307/2689911. URL http://www.jstor.org/stable/2689911.

Alfred Edward Housman. *More Poems*. Alfred A. Knopf, 1936. URL http://books.google.com/books?id=rTMiAAAAMAAJ.

T. E. Hull and A. R. Dobell. Random number generators. *S.I.A.M. Review*, 4(3):230–253, July 1962. doi: 10.1137/1004061. URL http://epubs.siam.org/doi/abs/10.1137/1004061.

Chris J. Isham. *Modern Differential Geometry for Physicists*. World Scientific Publishing, New Jersey, 2 edition, 1999. ISBN 9810235623. URL http://books.google.com/books?vid=ISBN9810235623.

C.J. Isham. Quantum topology and quantisation on the lattice of topologies. *Classical and Quantum Gravity*, 6:1509–1534, November 1989. doi: 10.1088/0264-9381/6/11/007. URL http://www.iop.org/EJ/abstract/0264-9381/6/11/007.

Alan Jeffrey. *Handbook of Mathematical Formulas and Integrals*. Academic Press, 1 edition, May 19 1995. ISBN 0123825806. URL http://books.google.com/books?vid=ISBN1483295141. "Many of the entries are based on the updated fifth edition of Gradshteyn and Ryzhik's "Tables of Integrals, Series, and Products," though during the preparation of the the book, results were also taken from various other reference works." — Preface, page xix.

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.

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/.

R. Jiménez and J. E. Yukich. *Statistical distances based on Euclidean graphs*, pages 223–240. Springer Science & Business Media, 2006. ISBN 9780387233949. URL http://books.google.com/books?vid=ISBN0387233946.

Kapli D. Joshi. *Introduction To General Topology*. New Age International, 1 edition, 1983. ISBN 9780852264447. URL http://books.google.com/books?vid=ISBN0852264445.

William J. Kennedy Jr. and James E. Gentle. *Statistical Computing*, volume 33 of *Statistics: A Series of Textbooks and Monographs*. CRC Press, 1980. ISBN 0824768981. URL http://books.google.com/books?vid=isbn0824768981.

M. Jůza. A note on complete metric spaces. *Matematicko-Fyzikálny Časopis*, 6:143–148, 1956.

John Leroy Kelley. *General Topology*. University Series in Higher Mathematics. Van Nostrand, New York, 1955. ISBN 0387901256. URL http://books.google.com/books?vid=ISBN0387901256. Republished by Springer-Verlag, New York, 1975.

J. C. Kelly. Bitopological spaces. *Proceedings of the London Mathematical Society*, s3-13(1):71–89, 1963. doi: 10.1112/plms/s3-13.1.71. URL http://plms.oxfordjournals.org/content/s3-13/1/71.extract.

Yoshida Kenko. *The Tsuredzure Gusa of Yoshida No Kaneyoshi. Being the meditations of a recluse in the 14th Century (Essays in Idleness)*. circa 1330. URL http://www.humanistictexts.org/kenko.htm. 1911 translation of circa 1330 text.

Mohamed A. Khamsi and W.A. Kirk. *An Introduction to Metric Spaces and Fixed Point Theory*. John Wiley, New York, 2001. ISBN 978-0471418252. URL http://books.google.com/books?vid=isbn0471418250.

William Kirk and Naseer Shahzad. *Fixed Point Theory in Distance Spaces*. SpringerLink: Bücher. Springer, October 23 2014. ISBN 9783319109275. URL http://books.google.com/books?vid=isbn3319109278.

John R. Klauder. *A Modern Approach to Functional Integration*. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, 2010. ISBN 0817647902. URL http://books.google.com/books?vid=isbn0817647902. john.klauder@gmail.com.

Andrei Nikolaevich Kolmogorov. *Foundations of the theory of probability*. Chelsea Publishing Company, New Yourk, 2 edition, 1933. 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”.

A. Korselt. Bemerkung zur algebra der logik. *Mathematische Annalen*, 44(1):156–157, March 1894. ISSN 0025-5831. doi: 10.1007/BF01446978. URL http://www.springerlink.com/content/v681m56871273j73/. referenced by Birkhoff(1948)p.133.

V. Krishnamurthy. On the number of topologies on a finite set. *The American Mathematical Monthly*, 73(2):154–157, February 1966. URL http://www.jstor.org/stable/2313548.

Carlos S. Kubrusly. *The Elements of Operator Theory*. Springer, 1 edition, 2001. ISBN 9780817641740. URL http://books.google.com/books?vid=ISBN0817641742.

Carlos S. Kubrusly. *The Elements of Operator Theory*. Springer, 2 edition, 2011. ISBN 9780817649975. URL http://books.google.com/books?vid=ISBN0817649972.

Kazimierz Kuratowski. Sur la notion d'ordre dans la theorie des ensembles (on the concept of order in set theory). *Fundamenta Mathematicae*, 2:161–171, 1921. URL http://matwbn.icm.edu.pl/ksiazki/fm/fm2/fm2122.pdf.

Kazimierz Kuratowski. *Introduction to Set Theory and Topology*, volume 13 of *International Series on Monographs on Pure and Applied Mathematics*. Pergamon Press, New York, 1961. URL http://www.worldcat.org/oclc/1023273.

Lagrange. Leçons élémentaires sur les mathématiques données à l'école normale en 1795, séances des ecoles normales, 1794–1795. In *Œuvres complètes*, volume 7, pages 183–288. 1795. URL http://math-doc.ujf-grenoble.fr/cgi-bin/oeitem?id=OE_LAGRANGE__7_183_0.

J. L. Lagrange. Leçons élémentaires sur les mathématiques données à l'école normale en 1795. In *Œuvres complètes*, volume 7, pages 187–288. Journal de l'École polytechnique, 1877. URL http://sites.mathdoc.fr/cgi-bin/oeitem?id=OE_LAGRANGE__7_183_0.

Edmund Landau. *Foundations of analysis: the arithmetic of whole, rational, irrational, and complex numbers. A supplement to textbooks on the differential and integral calculus*. AMS Chelsea Publishing, Providence, Rhode Island, USA, 3 edition, 1966. URL http://books.google.com/books?vid=ISBN082182693X. 1966 English translation of the German text *Grundlagen der Analysis*.

Nicolas K. Laos. *Topics in Mathematical Analysis and Differential Geometry*, volume 24 of *Series in pure mathematics*. World Scientific, 1998. ISBN 9789810231804. URL http://books.google.com/books?vid=ISBN9810231806.

J. G. Leathem. *Volume and surface integrals used in physics*, volume 1 of *Cambridge Tracts in Mathematics and Mathematical Physics*. Cambridge University Press, 1 edition, 1905. URL http://archive.org/details/volumesurfaceint01leatuoft.

Dekang Lin. An information-theoretic definition of similarity. In *Proceedings of the Internatioinal Conference on Machine Learning*, 1998. URL http://webdocs.cs.ualberta.ca/~lindek/papers/sim.pdf.

M. Loève. *Probability Theory I*, volume 45 of *Graduate Texts in Mathematics*. 4 edition, 1977. ISBN 9780387902104. URL http://books.google.com/books?vid=ISBN0387902104.

Proclus Lycaeus. *Proclus: A Commentary on the First Book of Euclid's Elements*. Princeton University Press, circa 450. ISBN 0691020906. URL http://books.google.com/books?vid=OCLC03902909. translation published in 1992. reprint edition, 1992.

Niccolò Machiavelli. *The Literary Works of Machiavelli: Mandragola, Clizia, A Dialogue on Language, and Belfagor, with Selections from the Private Correspondence*. Oxford University Press, 1961. ISBN 0313212481. URL http://www.worldcat.org/isbn/0313212481.

Saunders MacLane and Garrett Birkhoff. *Algebra*. AMS Chelsea Publishing, Providence, 3 edition, 1999. ISBN 0821816462. URL http://books.google.com/books?vid=isbn0821816462.

I. J. Maddox. *Elements of Functional Analysis*. Cambridge University Press, Cambridge, 2, revised edition, 1989. ISBN 9780521358682. URL books.google.com/books?vid=ISBN052135868X.

Roger Duncan Maddux. *Relation Algebras*. Elsevier Science, 1 edition, July 27 2006. ISBN 0444520139. URL http://books.google.com/books?vid=ISBN0444520139.

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.

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.

Lech Maligranda and W. Orlicz. On some properties of functions of generalized variation. *Monatshefte für Mathematik*, 104(1):53–65, March 01 1987. ISSN 1436-5081. doi: 10.1007/BF01540525. URL http://gdz.sub.uni-goettingen.de/download/PPN362162050_0104/LOG_0008.pdf.

John H. Matthews and Kurtis D. Fink. *Numerical Methods Using MatLab 3rd edition*. Prentice-Hall, Inc., Upper River, NJ, 1992. ISBN 0-13-013164-4.

George McCarty. *Topology: An Introduction With Application to Topological Groups*. International Series in Pure and Applied Mathematics. McGraw-Hill Book Company, New York, 1967. URL http://www.amazon.com/dp/0486656330. 1988 Dover edition available.

Erik Meijering. A chronology of interpolation: From ancient astronomy to modern signal and image processing. *Proceedings of the IEEE*, 90(3):319–342, March 2002. URL http://bigwww.epfl.ch/publications/meijering0201.pdf.

Gregor Mendel. *Experiments In Plant Hybridization*. 1853. URL http://old.esp.org/foundations/genetics/classical/gm-65.pdf.

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).

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.

Dragoslav S. Mitrinović, J. E. Pečarić, and Arlington M. Fink. *Classical and New Inequalities in Analysis*, volume 61 of *Mathematics and its Applications (East European Series)*. Kluwer Academic Publishers, Dordrecht, Boston, London, 2010. ISBN 978-90-481-4225-5. URL http://www.amazon.com/dp/0792320646.

Ilya S. Molchanov. *Theory of Random Sets*. Probability and Its Applications. Springer, 2005. ISBN 185233892X. URL http://books.google.com/books?vid=ISBN185233892X.

H. P. Mulholland. On generalizations of minkowski's inequality in the form of a triangle inequality. *Proceedings of the London Mathematical Society*, s2-51:294–307, 1950. URL http://plms.oxfordjournals.org/content/s2-51/1/294.extract. received 1946 October 10, read 1947 June 19.

Markus Müller-Olm. 2. complete boolean lattices. In *Modular Compiler Verification: A Refinement-Algebraic Approach Advocating Stepwise Abstraction*, volume 1283 of *Lecture Notes in Computer Science*, chapter 2, pages 9–14. Springer, September 12 1997. ISBN 978-3-540-69539-4. URL http://link.springer.com/chapter/10.1007/BFb0027455. Chapter 2.

James R. Munkres. *Topology*. Prentice Hall, Upper Saddle River, NJ, 2 edition, 2000. ISBN 0131816292. URL http://www.amazon.com/dp/0131816292.

Mangesh G Murdeshwar. *General Topology*. New Age International, 2 edition, 1990. ISBN 9788122402469. URL http://books.google.com/books?vid=isbn8122402461.

Isaac Newton. Methodus differentialis. In *Analysis Per Quantitatum Series, Fluxiones ac Differentias : cum Enumeratio Linearum Tertii Ordinis*. 1711. "A photographic reproduction of the original Latin text ...and a translation are given" in Fraser (1919), pages 9–17 (Latin) and pages 18–25 (English).

Timothy S. Norfolk. When does a metric generate convex balls? not sure about the year, 1991. URL http://www.math.uakron.edu/~norfolk/.

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.

Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing*. Prentice Hall, 2 edition, 1999. ISBN 9780137549207. URL http://www.amazon.com/dp/0137549202.

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.

J/ orgen Bang-Jensen and Gregory Gutin. *Digraphs: Theory, Algorithms and Applications*. Springer-Verlag, Berlin, corrected version of first edition edition, August 15 2007. URL http://www.cs.rhul.ac.uk/books/dbook/.

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, New York, 3 edition, 1991. ISBN 0070484775. URL http://books.google.com/books?vid=ISBN0070484775.

C. W. Patty. Bitopological spaces. *Duke Mathematical Journal*, 34(3):387–391, 1967. doi: 10.1215/S0012-7094-67-03442-4. URL http://projecteuclid.org/euclid.dmj/1077377140.

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.

Giuseppe Peano. *Árithmetices principa, nova methodo exposita*. Fratres Bocca, 1889a. URL https://archive.org/details/arithmeticespri00peangoog. The principles of arithmetic presented by a new method.

Giuseppe Peano. The principles of arithmetic, presented by a new method. In Jean Van Heijenoort, editor, *From Frege to G odel : A Source Book in Mathematical Logic, 1879-1931*, pages 85–97. Harvard University Press (1967), Cambridge, Massachusetts, 1889b. ISBN 0674324498. URL http://www.amazon.com/dp/0674324498. translation of *Árithmetices principa, nova methodo exposita*.

Charles Sanders Peirce. Note b: the logic of relatives. In *Studies in Logic by Members of the Johns Hopkins University*, pages 187–203. Little, Brown, and Co., Boston, 1883. URL http://www.archive.org/details/studiesinlogic00peiruoft.

C.S. Peirce. On the algebra of logic. *American Journal of Mathematics*, 3(1):15–57, March 1880. URL http://www.jstor.org/stable/2369442.

J. E. Pečarić, Frank Proschan, and Yung Liang Tong. *Convex Functions, Partial Orderings, and Statistical Applications*, volume 187 of *Mathematics in Science and Engineering*. Academic Press, San Diego, California, 1992. ISBN 978-0125492508. URL http://books.google.com/books?vid=ISBN0125492502.

Henri Poincaré. *La Science et l'hypothèse*. 1902a. URL http://fr.wikisource.org/wiki/La_Science_et_l%27hypoth%C3%A8se. (Science and Hypothesis).

Henri Poincaré. *Science and Hypothesis*. Dover Publications (1952), New York, 1902b. ISBN 0486602214. URL http://books.google.com/books?vid=isbn0486602214. translation of La Science et l'hypothèse.

Jeffrey C. Pommerville. *Fundamentals of Microbiology*. Jones & Bartlett Publishers, January 1 2013. ISBN 9781449647964. URL http://books.google.com/books?vid=isbn1449647960.

S. Ponnusamy. *Foundations of Functional Analysis*. CRC Press, 2002. ISBN 9780849317170. URL http://books.google.com/books?vid=ISBN0849317177.

K. M. M. Prabhu. *Window Functions and Their Applications in Signal Processing*. CRC Press, 2013. ISBN 9781466515840. URL https://books.google.com/books?vid=ISBN1466515848.

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.

John Ratcliffe. *Foundations of Hyperbolic Manifolds*, volume 149 of *Graduate Texts in Mathematics*. Springer Science & Business Media, 2013. ISBN 9781475740134. URL http://books.google.com/books?vid=ISBN1475740131.

Howard L. Resnikoff and Raymond O'Neil Wells. *Mathematics in Civilization*. Dover Publications, New York, 1984. ISBN 0486246744. URL http://books.google.com/books?vid=ISBN0486246744.

Hugo Ribeiro. Sur les espaces à métrique faible. *Portugaliae mathematica*, 4(1):21–40, 1943. ISSN 0032-5155. URL https://eudml.org/doc/114615.

Frigyes Riesz. Stetigkeitsbegriff und abstrakte mengenlehre. In Guido Castelnuovo, editor, *Atti del IV Congresso Internazionale dei Matematici*, volume II, pages 18–24, Rome, 1909. Tipografia della R. Accademia dei Lincei. URL http://www.mathunion.org/ICM/ICM1908.2/Main/icm1908.2.0018.0024.ocr.pdf. 1908 April 6–11.

Stefan Rolewicz. *Metric Linear Spaces*. Mathematics and Its Applications. D. Reidel Publishing Company, 1 edition, 1985. ISBN 9789027714800. URL http://books.google.com/books?vid=ISBN9027714800.

Maxwell Rosenlicht. *Introduction to Analysis*. Dover Publications, New York, 1968. ISBN 0-486-65038-3. URL http://books.google.com/books?vid=ISBN0486650383.

Gian-Carlo Rota. The number of partitions of a set. *The American Mathematical Monthly*, 71(5): 498–504, May 1964. URL http://www.jstor.org/stable/2312585.

Gian-Carlo Rota. The many lives of lattice theory. *Notices of the American Mathematical Society*, 44 (11):1440–1445, December 1997. URL http://www.ams.org/notices/199711/comm-rota.pdf.

Sergiu Rudeanu. *Lattice Functions and Equations*. Discrete Mathematics and Theoretical Computer Science. Springer Science & Business Media, July 30 2001. ISBN 9781852332662. URL http://books.google.com/books?vid=ISBN1852332662.

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.

et. al Runtao He. Analysis of multimerization of the sars coronavirus nucleocapsid protein. *Biochemical and Biophysical Research Communications*, 316(2):476–483, April 2 2004. URL http://www.sciencedirect.com/science/article/pii/S0006291X04003250.

et. al. S. G. Gregory. The dna sequence and biological annotation of human chromosome 1. *Nature: International Weekly Journal of Science*, 441:315–321, May 18 2006. doi: 10.1038/nature04727. URL http://www.nature.com/nature/journal/v441/n7091/abs/nature04727.html.

H. Salzmann, T. Grundhöfer, H. Hähl, and R. Löwen. *The Classical Fields: Structural Features of the Real and Rational Numbers*, volume 112 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2007. ISBN 9780521865166. URL http://books.google.com/books?vid=ISBN0521865166.

Eric Schechter. *Handbook of Analysis and Its Foundations*. Academic Press, 1996. ISBN 9780080532998. URL http://books.google.com/books?vid=ISBN0080532993.

Gert Schubring. *Conflicts Between Generalization, Rigor, and Intuition: Number Concepts Underlying the Development of Analysis in 17th–19th Century France and Germany*. Sources and studies in the history of mathematics and physical sciences. Springer, New York, 1 edition, June 2005. ISBN 0387228365. URL http://books.google.com/books?vid=ISBN0387228365.

Berthold Schweizer and Abe Sklar. *Probabilistic Metric Spaces*. Elsevier Science Publishing Co., 1983. ISBN 9780444006660. URL http://books.google.com/books?vid=ISBN0486143759. A 2005 Dover edition (ISBN 9780486143750) has been published which is “an unabridged republication of the work first published by Elsevier Science Publishing Co., Inc., in 1983.”.

Mícheál Ó Searcóid. *Metric Spaces*. Springer Undergraduate Mathematics Series. Springer Science & Business Media, 2006. ISBN 9781846286278. URL http://books.google.com/books?vid=ISBN1846286271.

Frank L. Severance. *System Modeling and Simulation: An Introduction*. John Wiley & Sons, August 16 2001. ISBN 0471496944. URL http://books.google.com/books?vid=ISBN0471496944.

Alexander Shen and Nikolai Konstantinovich Vereshchagin. *Basic Set Theory*, volume 17 of *Student mathematical library*. American Mathematical Society, Providence, July 9 2002. ISBN 0821827316. URL http://books.google.com/books?vid=ISBN0821827316. translated from Russian.

A. N. Sherstnev. Random normed spaces. questions of completeness. *Kazan. Gos. Univ. Uchen. Zap.*, 122(4):pages 3–20, 1962. URL http://mi.mathnet.ru/uzku138.

Judith Silver and Erik Stokes. A great circle metric. In *Proceedings of the West Virginia Academy of Science*, volume 78, pages 8–13, 2007. URL http://www.marshall.edu/wvas/public.html.

David Simmons. Sequence. Technical report, University of North Texas, Department of Mathematics, 1155 Union Circle #311430, Denton, TX 76203, January 21 15:38 2016. URL https://en.wikipedia.org/w/index.php?title=Sequence&oldid=700944642.

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.

Rajeev Singh, Ramon E. Vasquez, and Reena Singh. Comparison of daubechies, coiflet, and symlet for edge detection. In Stephen K. Park and Richard D. Juday, editors, *SPIE Proceedings*, volume 3074 of *Visual Information Processing VI*, July 22 1997. doi: 10.1117/12.280616. URL http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=926491.

Neil J. A. Sloane. On-line encyclopedia of integer sequences. World Wide Web, 2014. URL http://oeis.org/.

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.

Lynn Arthur Steen and J. Arthur Seebach. *Counterexamples in Topology*. Springer-Verlag, 2, revised edition, 1978. URL http://books.google.com/books?vid=ISBN0486319296. A 1995 "unabridged and unaltered republication" Dover edition is available.

Anne K. Steiner. The lattice of topologies: Structure and complementation. *Transactions of the American Mathematical Society*, 122(2):379–398, April 1966. URL http://www.jstor.org/stable/1994555.

James Stewart. *Single Variable Essential Calculus*. Cenage Learning, 2 edition, February 10 2012. ISBN 9781133112761. URL http://books.google.com/books?vid=ISBN1133112765.

Ronald A. Stoltenberg. On quasi-metric spaces. *Duke Mathematical Journal*, 36(1):65–71, 1969. doi: 10.1215/S0012-7094-69-03610-2. URL http://projecteuclid.org/euclid.dmj/1077378129.

Jeffrey Stopple. *A Primer of Analytic Number Theory: From Pythagoras to Riemann*. Cambridge University Press, Cambridge, June 2003. ISBN 0521012538. URL http://books.google.com/books?vid=ISBN0521012538.

Patrick Suppes. *Axiomatic Set Theory*. Dover Publications, New York, 1972. ISBN 0486616304. URL http://books.google.com/books?vid=ISBN0486616304.

Wilson Alexander Sutherland. *Introduction to Metric and Topological Spaces*. Oxford University Press, illustrated, reprint edition, 1975. ISBN 9780198531616. URL http://books.google.com/books?vid=ISBN0198531613.

Jenö Szirmai. The densest geodesic ball packing by a type of nil lattices. *Contributions to Algebra and Geometry*, 48(2):383–397, 2007. URL http://www.emis.ams.org/journals/BAG/vol.48/no.2/b48h2szi.pdf.

Brian S. Thomson, Andrew M. Bruckner, and Judith B. Bruckner. *Elementary Real Analysis*. www.classicalrealanalysis.com, 2 edition, 2008. ISBN 9781434843678. URL http://classicalrealanalysis.info/com/Elementary-Real-Analysis.php.

Wolfgang J. Thron. *Topological structures*. Holt, Rinehart and Winston, New York, 1966. URL http://books.google.com/books?id=JRM_AAAAIAAJ.

Hugh Ansfrid Thurston. *The Number-System*. Blackie, London, 1956. URL http://books.google.com/books?vid=ISBN0486458067. 2007 Dover republication also available.

Heinrich Franz Friedrich Tietze. Beiträge zur allgemeinen topologie i. *Mathematische Annalen*, 88 (3–4):290–312, 1923. URL http://link.springer.com/article/10.1007/BF01579182.

Elmer Tolsted. An elementary derivation of cauchy, hölder, and minkowski inequalities from young's inequality. *Mathematics Magazine*, 37:2–12, 1964. URL http://mathdl.maa.org/mathDL/22/?pa=content&sa=viewDocument&nodeId=3036.

Donald M. Topkis. Minimizing a submodular function on a lattice. *Operations Research*, 26(2): 305–321, April 1 1978. URL http://dx.doi.org/10.1287/opre.26.2.305.

Robert W. Vallin. Continuity and differentiability aspects of metric preserving functions. *Real Analysis Exchange*, 25(2):849–868, 1999. URL projecteuclid.org/euclid.rae/1230995419.

M.L.J. van de Vel. *Theory of Convex Structures*, volume 50 of *North-Holland Mathematical Library*. North-Holland, Amsterdam, 1993. ISBN 978-0444815057. URL http://books.google.com/books?vid=ISBN0444815058.

Remco C. Veltkamp. Shape matching: Similarity measures and algorithms. *Shape Modeling and Applications, SMI 2001 International Conference on*, pages 188–197, May 2001. doi: 10.1109/SMA.2001.923389. URL http://dspace.library.uu.nl/bitstream/handle/1874/2540/2001-03.pdf.

Remco C. Veltkamp and Michiel Hagedoorn. Shape similarity measures, properties and constructions. In Robert Laurini, editor, *Advances in Visual Information Systems 4th International Conference, VISUAL 2000 Lyon*, volume 1929 of *Lecture Notes in Computer Science*, pages 467–476, France, November 2–4 2000. Springer. ISBN 978-3-540-40053-0. URL http://link.springer.com/chapter/10.1007/3-540-40053-2_41.

Paul M. B. Vitányi. Information distance in multiples. *IEEE Transactions on Information Theory*, 57(4), April 2011. URL http://arxiv.org/pdf/0905.3347.pdf.

Richard F. Voss. Evolution of long-range fractal correlations and 1/f noise in dna base sequences. *Physical Review Letters*, 68(25):3805–3808, June 22 1992. doi: 10.1103/PhysRevLett.68.3805. URL http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.68.3805.

Charles Walmsley. *An Introductory Course Of Mathematical Analysis*. Cambridge University Press, 1920. URL https://archive.org/details/introductorycour032788mbp.

Edward Waring. Problems concerning interpolations. *Philosophical Transactions of the Royal Society of London*, 69:59–67, 1779. URL http://www.jstor.org/stable/106408.

J. D. Watson and F. H. C. Crick. Molecular structures of nucleic acids: A struture for deoxyribose nucleic acid. *Nature*, pages 737–738, April 25 1953a. doi: 10.1038/171737a0. URL http://www.nature.com/nature/journal/v171/n4356/pdf/171737a0.pdf.

J. D. Watson and F. H. C. Crick. Genetical implications of the structure of deoxyribonucleic acid. *Nature*, pages 964–967, May 30 1953b. doi: 10.1038/171964b0. URL http://www.nature.com/nature/journal/v171/n4361/pdf/171964b0.pdf.

Hermann Weyl. Emmy noether. *Scripta Mathematica...*, 3:201–220, April 26 1935. URL http://link.springer.com/chapter/10.1007/978-1-4684-0535-4_6. memorial address at Bryn Mawr College.

Norbert Wiener. A simplification of the logic of relations. *Proceedings of the Cambridge Philosophical Society*, 17:387–390, 1914. URL http://books.google.com/books?vid=ISBN0674324498&pg=PA224. reprinted in Jean Van Heijenoort's *From Frege to G odel : A Source Book in Mathematical Logic, 1879-1931* page 224–227.

Stephen Willard. *General Topology*. Addison-Wesley Series in Mathematics. Addison-Wesley, 1970. ISBN 9780486434797. URL http://books.google.com/books?vid=ISBN0486434796. a 2004 Dover edition has been published which "is an unabridged republication of the work originally published by the Addison-Wesley Publishing Company ...1970".

W. A. Wilson. On quasi-metric spaces. *American Journal of Mathematics*, 53(3):675–684, July 1931a. URL http://www.jstor.org/stable/2371174.

Wallace Alvin Wilson. On semi-metric spaces. *American Journal of Mathematics*, 53(2):361–373, April 1931b. URL http://www.jstor.org/stable/2370790.

Robert S. Wolf. *Proof, Logic, and Conjecture; The Mathematician's Toolbox*. W.H. Freeman and Company, 1998. ISBN 0716730502. URL http://www.worldcat.org/isbn/0716730502.

Qinglan Xia. The geodesic problem in quasimetric spaces. *Journal of Geometric Analysis*, 19(2):452–479, April 2009. doi: 10.1007/s12220-008-9065-4. URL http://link.springer.com/article/10.1007/s12220-008-9065-4.

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.

Ernst Zermelo. Untersuchungen über die grundlagen der mengenlehre i. *Mathematische Annalen*, 65:261–281, 1908a. URL http://dz-srv1.sub.uni-goettingen.de/sub/digbib/loader?ht=VIEW&did=D38200.

Ernst Zermelo. Investigations in the foundations of set theory. In Jean Van Heijenoort, editor, *From Frege to G odel : A Source Book in Mathematical Logic, 1879-1931*, pages 199–215. Harvard University Press, 1967, Cambridge, Massechusettes, 1908b. URL http://www.amazon.com/dp/0674324498.

Vladimir A. Zorich. *Mathematical Analysis I*. Universitext Series. Springer Science & Business Media, 2004. ISBN 9783540403869. URL http://books.google.com/books?vid=ISBN3540403868.

REFERENCE INDEX

SUBJECT INDEX

License

…last page …please stop reading …