

Foundations of Field Computation

Bruce MacLennan

Copyright ©2017

Draft of

August 11, 2017

Contents

1	Introduction	3
1.1	Motivation	3
1.1.1	Truly Massive Parallelism	3
1.1.2	Field Transformation	4
1.2	History	5
1.3	Suggestions for the reader	5
2	Basic Concepts of Topology	7
2.1	Metric spaces	7
2.1.1	Metric	7
2.1.2	Examples	8
2.1.3	Metric space	10
2.1.4	Discrete metric space	11
2.1.5	Isometry	11
2.1.6	Function spaces	12
2.2	Topology	13
2.2.1	Open balls	13
2.2.2	Neighborhoods	14
2.2.3	Open and closed sets	15
2.2.4	Closure and interior of a set	16
2.2.5	Topological spaces	17
2.2.6	Bases	18
2.3	Convergence and limits	19
2.3.1	Convergence in metric spaces	19
2.3.2	Convergence in topological spaces	20
2.4	Continuity	21
2.5	Homeomorphism	22
2.6	Separability	23

2.7	Completeness	24
2.7.1	Complete metric space	24
2.7.2	Completion of metric spaces	24
2.8	Connectedness	26
2.8.1	Connection	26
2.8.2	Continua and discontinua	28
2.8.3	Path connection	28
2.8.4	Processes	29
2.9	References	30
3	Banach Spaces	31
3.1	Linear spaces	31
3.1.1	Definitions	31
3.1.2	Examples	33
3.2	Normed linear spaces	34
3.2.1	Norms	34
3.2.2	Normality	36
3.2.3	Normed linear space	36
3.3	Banach spaces	38
3.4	Differentiation in Banach spaces	38
3.4.1	Fréchet derivative	38
3.4.2	Gâteaux derivative	39
3.4.3	Properties of derivatives	39
3.5	Taylor theorem in Banach spaces	40
3.6	References	40
4	Basic Complex Analysis	41
4.1	Argand diagram	41
4.2	Geometrical Interpretations	42
4.3	Polar representation	45
4.3.1	Defined	45
4.3.2	cis function	46
4.4	Complex exponentials	47
4.4.1	Euler's Formula	47
4.4.2	Periodic Change	51
4.4.3	Complex Logarithms	58
4.5	Hyperbolic Geometry	61
4.5.1	Hyperbolic Functions	61

4.5.2	Special Relativity Theory	64
4.6	References	68
5	Hilbert Spaces	69
5.1	Inner product spaces	69
5.1.1	Inner products	69
5.1.2	Inner product norm	74
5.1.3	More on complex inner products	74
5.1.4	Orthonormality	77
5.1.5	Approximate Orthogonality	78
5.2	Hilbert spaces	80
5.2.1	Completeness	80
5.2.2	Orthonormal Bases	81
5.2.3	Generalized Fourier coefficients	82
5.2.4	Spanning vectors and orthogonal subspaces	84
5.2.5	Linear functionals	86
5.2.6	Convergence in function spaces	88
5.2.7	Adjoint operators and outer products	89
5.2.8	Commutators and the uncertainty principle	91
5.2.9	Separable Hilbert spaces and Hilbert-Schmidt operators	93
5.2.10	Tensor products	96
5.2.11	Urysohn's theorem	97
5.3	References	99
6	Gabor Representations	101
6.1	Introduction	101
6.2	The Gabor Uncertainty Principle	102
6.3	Gabor Representation of One-Dimensional Signals	110
6.4	Gabor Representation of Two-Dimensional Signals	120
6.5	Evidence for 2D Gabor Representation in Vision	125
6.6	Problems with the Gabor Representation	127
6.7	Gabor versus Wavelet Representations	128
6.8	Gabor Wavelets	132
6.9	The Orthogonality Issue	132
6.10	3D Gabor Representation of Spatiotemporal Signals	136
6.11	Conclusions	140
6.12	Appendix: Proofs	144
6.12.1	Proof of General Gabor Uncertainty Principle	144

6.12.2	Proof of Optimality of General Gabor Elementary Functions	146
7	Basic Concepts of Field Computation	149
7.1	Assumptions	149
7.1.1	Notational conventions	149
7.1.2	Fields	150
7.1.3	Field Transformations	153
7.2	Summary	153
7.3	Functional analysis preliminaries	154
7.3.1	Definitions	154
7.3.2	Orthonormal bases	154
7.4	Basic field transformations	155
7.4.1	Definition of general field product	155
7.4.2	Outer product	156
7.4.3	Kernels of linear and multilinear operators	161
8	Linear and Multilinear Operators	169
8.1	Linear Operators	169
8.1.1	Introduction	169
8.1.2	Examples	169
8.2	Multilinear Operators	184
8.2.1	Introduction	184
8.2.2	Examples	184
9	Universal Field Computation	191
9.1	Introduction	191
9.2	Approximation Based on Taylor Series	191
9.2.1	Derivatives of field transformations	191
9.2.2	Gradients of Operators	192
9.2.3	Taylor's Theorem	194
9.2.4	Summary of Taylor Series Approximation	196
9.3	General Polynomial Approximation	196
10	Local Transformations	199
10.1	Theory	199
10.2	Examples	201
10.2.1	Importance of Sigmoid Nonlinearities	201

- 10.2.2 Hyperbolic Tangent 202
- 10.2.3 Logistic Function 203
- 10.3 Replacing Local Field Polynomials by Sigmoid Functions . . . 203

- 11 The Problem of High Dimensional Gradients 207**
 - 11.1 The Problem 207
 - 11.2 Field Computation in Alternate Domains 209
 - 11.2.1 Introduction 209
 - 11.2.2 Measure Preserving Change of Domain 209
 - 11.2.3 Transformations in the Alternate Domain 210
 - 11.3 Reduction of Dimension 214

- 12 Neural Implementation 215**
 - 12.1 Discretizing Fields 215
 - 12.2 Linear Field Transformations 216
 - 12.3 Nonlinear Field Transformations 217

- A Notation 219**

Chapter 1

Introduction

1.1 Motivation

1.1.1 Truly Massive Parallelism

AI is moving into a new phase characterized by a broadened understanding of the nature of knowledge, and by the use of new computational paradigms. A sign of this transition is the growing interest in neurocomputers, optical computers, molecular computers and other massively parallel analog computers. We have argued elsewhere (MacLennan, 1987a,b, 1988a,b) that the new AI will augment the traditional deep, narrow computation with shallow, wide computation. That is, the new AI will exploit *massive parallelism*, but this means different things to different people; massive parallelism may begin with a hundred, a thousand, or a million processors. Biological evidence suggests that skillful behavior requires a very large number of processors, so many in fact that it is infeasible to treat them individually; they must be treated *en masse*. This has motivated us to propose (MacLennan, 1987b) the following definition of massive parallelism: *A computational system is **massively parallel** if the number of processing elements is so large that it may conveniently be considered a continuous quantity.* That is, a system is *massively parallel* if the processing elements can be considered a *continuous mass* rather than a *discrete ensemble* (MacLennan, 1989).

How large a number is large enough to be considered a continuous quantity? That depends on the purpose at hand. A hundred is probably never large enough; a million is probably always large enough; a thousand or ten thousand may be enough. One of the determining factors will be whether the

number is large enough to permit the application of continuous mathematics (see below).

We propose this definition of massive parallelism for a number of reasons. First, skillful behavior seems to require significant neural *mass*.¹ Second, we are interested in computers, such as optical computers and molecular computers, for which the number of processing elements *is* effectively continuous. Third, continuous mathematics is generally easier than discrete mathematics. And fourth, we want to encourage a new style of thinking about parallelism. Currently, we try to apply to parallel machines the thought habits we have acquired from thinking about sequential machines. This strategy works fairly well when the degree of parallelism is low, but it will not scale up. One cannot think individually about the 10^{20} processors of a molecular computer. Rather than postpone the inevitable, we think that it is time to develop a theoretical framework for understanding massively parallel analog computers. The principal goal of this paper is to outline such a theory.

1.1.2 Field Transformation

Our aim then is to develop a way of looking at massive parallelism that encompasses a variety of implementation technologies, including neural networks, optical computers, molecular computers and other massively parallel analog computers. What these all have in common is the ability to process in parallel amounts of data so massive as to be considered a continuous quantity. This suggests that we structure our theory around the idea of a *field*, i.e. a continuous (dense) ensemble of data. We have in mind both scalar fields (such as potential fields) and vector fields (such as gradient fields). Any operation on such a field, either to produce another field or to produce a new state of the field, can be considered massively parallel, if it operates on all the elements of the field in parallel. Indeed, it would not be feasible to serialize the processing of the field; modest degrees of parallelism cannot cope with an infinite (or nearly infinite) number of field elements.

In the remainder of this chapter we explore *field transformation computers*, that is, computers characterized by the ability to perform (in parallel) transformations on scalar and vector fields. This does not mean that field computers are unable to perform scalar calculations; in fact many field transformation computers have the scalar capabilities of conventional digital and

¹Even a bee has some 10^6 neurons (DARPA, 1988, p. 33).

analog computers. Scalars have many uses in field computation. For example, we may want to use a scalar parameter to control the rate at which a field transformation takes place (e.g., a reaction rate in a molecular computer). Similarly, we may use a scalar representing the average intensity of a field to control the contrast enhancement of that field. A scalar threshold value may be used to suppress low level noise, and so forth.

It must be stressed that there are many field computers already in existence, for example, large neurocomputers and many optical computers. What we are proposing is:

1. a name for the class of such computers
2. a theoretical framework for understanding massively parallel analog computers
3. a basis for constructing general purpose computers of this type.

1.2 History

forthcoming

1.3 Suggestions for the reader

Chapters 2 through 6 cover standard material that is a prerequisite for understanding field computation, which is the focus of the chapters that follow. I suggest that you skim the familiar material so that you can see the notational conventions that we use.²

²Briefer presentations of field computation can be found in MacLennan (2009) and MacLennan (2017).

Chapter 2

Basic Concepts of Topology

This chapter introduces basic concepts from topology, and especially metric spaces, that are a foundation for the later chapters. In particular we focus on notions of convergence and continuity that are useful in continuous computation, that is, the study of continuous information representation and processing. The concepts of separability and completeness are essential to the later presentation of Hilbert Spaces (Ch. 5).

2.1 Metric spaces

2.1.1 Metric

Definition 2.1.1 (\mathbb{R}_*) \mathbb{R}_* is the set of all nonnegative real numbers, $\mathbb{R}_* = \{x \in \mathbb{R} \mid x \geq 0\}$.

Definition 2.1.2 (Metric) A metric is a function $d : X \times X \rightarrow \mathbb{R}_*$ satisfying (for all $x, y, z \in X$):

Self-identity $d(x, y) = 0$ if and only if $x = y$.

Symmetry $d(x, y) = d(y, x)$.

Triangle inequality $d(x, y) + d(y, z) \geq d(x, z)$.

Remark 2.1.1 Although these are very natural postulates for a distance measure, they are not sacred, and in particular cases meaningful measures of similarity or dissimilarity may not satisfy them. The triangle inequality is

particularly problematic, and a measure that drops this requirement is called a semimetric. Moreover, in psychological spaces symmetry may not hold (e.g., x may seem closer to y than y seems to be to x). Finally, the Self-identity property implies that all differences are “measurable” (by the metric). A pseudo-metric is function satisfying the weaker condition $d(x, x) = 0$ instead of Self-identity.

Exercise 2.1.1 Give an example of a function that is a pseudo-metric but not a metric.

Exercise 2.1.2 For a given nondirected graph, define the distance between two vertices to be the number of edges in the minimum path between the vertices. (If there is no path between the vertices, then set the distance to ∞ or to any number greater than the number of vertices.) Is this distance a metric?

Exercise 2.1.3 For a given directed graph, define the distance from one vertex to another to be the number of edges in the minimum path from the one to the other (i.e. respecting the direction of edges). (If there is no path from one vertex to the other, then set the distance to ∞ or to any number greater than the number of vertices.) Is this distance a metric?

2.1.2 Examples

The most common metric is the L_2 or Euclidean metric on an n -dimensional vector space:

Definition 2.1.3 (Euclidean Metric) The L_2 or Euclidean metric on \mathbb{R}^n is defined

$$L_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}.$$

Exercise 2.1.4 Show that L_2 is a metric.

The L_p metrics are simple generalizations of the Euclidean metric.

Definition 2.1.4 (L_p Metrics) For any (real or integral) $p \geq 1$,

$$L_p(\mathbf{x}, \mathbf{y}) = \sqrt[p]{\sum_{k=1}^n |x_k - y_k|^p}.$$

Exercise 2.1.5 *The set of all points $\mathbf{x} \in \mathbb{R}^2$ such that $L_2(\mathbf{x}, \mathbf{y}) = r$ defines a circle of radius r around center \mathbf{y} . Thus $L_2(\mathbf{x}, 0) = 1$ defines a unit circle at the origin. Plot, by computer, the (two-dimensional) L_p -circle $L_p(\mathbf{x}, 0) = 1$, for $p = 3, 4, 5$, etc. What conclusions do you draw?*

Exercise 2.1.6 *Show that L_p is a metric for arbitrary $p > 1$. Hint: Expand the p th power by the binomial theorem.*

Consider a special case $p = 1$ of the preceding definition.

Definition 2.1.5 (L_1 metric)

$$L_1(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^n |x_k - y_k|.$$

Exercise 2.1.7 *Sketch (without use of a computer) a two-dimensional L_1 -circle.*

Exercise 2.1.8 *Prove that L_1 is a metric.*

Hamming distance is widely used in coding and information theory as well as in neural nets.

Definition 2.1.6 (Hamming distance) *The Hamming distance between two bit-strings of the same length is the number of bits that differ between them.*

Exercise 2.1.9 *Show directly that Hamming distance is a metric*

Exercise 2.1.10 *Show that Hamming distance is equivalent to the L_1 metric.*

Remark 2.1.2 *To design an error-correcting code capable of correcting errors of m or fewer bits, each legal code is surrounded by an L_1 -sphere of illegal codes (so the legal codes have a Hamming distance of $2m$). Then, if an illegal code appears in a signal, it can be replaced by the legal code that is nearest (in Hamming distance).*

The L_p definition can also be extended to $p = \infty$ by taking the limit $p \rightarrow \infty$.

Definition 2.1.7 (L_∞ metric)

$$L_\infty(\mathbf{x}, \mathbf{y}) = \lim_{p \rightarrow \infty} L_p(\mathbf{x}, \mathbf{y}) = \max_{k=1}^n |x_k - y_k|.$$

Exercise 2.1.11 Sketch (without use of a computer) a two-dimensional L_∞ -circle.

Exercise 2.1.12 Prove that L_∞ is a metric.

Sometimes the different components of a vector are not equally significant. In these cases *weighted* L_p metrics can be used.

Definition 2.1.8 (\mathbb{R}_+) \mathbb{R}_+ is the set of positive real numbers, $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$.

Definition 2.1.9 (**Weighted L_p metric**) Let $\mathbf{w} \in \mathbb{R}_+^n$ be a positive weight vector. Then, for $p \geq 1$,

$$L_p^{\mathbf{w}}(\mathbf{x}, \mathbf{y}) = \sqrt[p]{\sum_{k=1}^n w_k |x_k - y_k|^p}.$$

Exercise 2.1.13 Sketch a two-dimensional weighted L_2 circle with the weight vector $\mathbf{w} = (1, 2)$. In general, how does a weight vector affect the shape of an L_p circle?

Exercise 2.1.14 Why are the elements of \mathbf{w} required to be strictly positive (i.e. $w_k > 0$ for all k)?

Remark 2.1.3 The preceding definitions can be extended to $n = \infty$, that is, to infinite dimensional spaces, a subject that will be taken up in Ch. 3.

2.1.3 Metric space

Definition 2.1.10 (**Metric space**) A metric space is a pair (X, d) , where X is an arbitrary set (called the underlying set) and $d : X \times X \rightarrow \mathbb{R}_*$ is a metric on the underlying set.

Remark 2.1.4 Many different metric spaces can have the same underlying set. For example, (\mathbb{R}^n, L_p) , $p = 1, 2, 3, \dots$ are all different metric spaces with the same underlying set.

Exercise 2.1.15 *What can you say about the differences between the metric spaces (\mathbb{R}, L_p) for $p \geq 1$?*

Definition 2.1.11 (Euclidean space \mathcal{E}^n)

$$\mathcal{E}^n = (\mathbb{R}^n, L_2).$$

That is, the n -dimensional Euclidean metric space is the space of n -dimensional real vectors with the usual (Euclidean) metric.

Remark 2.1.5 *Sometimes we call \mathbb{R}^n a metric space, in which case we mean \mathbb{R}^n with its “usual” metric, that is, the Euclidean metric L_2 .*

2.1.4 Discrete metric space

Definition 2.1.12 (Discrete metric space) *For any set X , the discrete metric is defined:*

$$\begin{aligned} d_d(x, x) &= 0, \\ d_d(x, y) &= 1, \quad \text{if } x \neq y. \end{aligned}$$

Remark 2.1.6 *With the discrete metric there are no degrees of closeness; so far as the metric is concerned, the points are either identical or they are as different as they can be. The discrete metric is implicit in the usual treatment of alphabets in formal language theory: that symbols of the alphabet are either identical or not; they are not considered close or far on the basis of alphabetic order, shape, sound, ASCII code, or any other basis. Note, however, that the discrete metric is not limited to finite or even countable sets. For example, (\mathbb{R}, d_d) is the discrete metric space of real numbers, in which there is no sense of a real number being closer or farther to any other real numbers; it is just a set of independent points.*

Exercise 2.1.16 *Show that the discrete metric is in fact a metric.*

2.1.5 Isometry

Definition 2.1.13 (Isometry or isometric mapping) *A function $f : X \rightarrow X'$ is an isometry or isometric mapping between metric spaces (X, d) and (X', d') if and only if*

$$d(x, y) = d'[f(x), f(y)],$$

for all $x, y \in X$. That is, the function preserves distances.

Remark 2.1.7 *Many mappings between metric spaces are not isometric, yet still satisfy some interesting relation between their metrics, e.g., $d(x, y) = cd'[f(x), f(y)]$ or $d(x, y) = \log d'[f(x), f(y)]$.*

Definition 2.1.14 (Isometric spaces) *Two metric spaces are said to be isometric if there is an isometric mapping from one to the other.*

Exercise 2.1.17 *Show that isometry is an equivalence relation.*

2.1.6 Function spaces

In many applications, especially in field computation, we deal with *function spaces*, that is, spaces whose “points” are entire functions. Function spaces will be considered in detail in chapters 3 and 5; here we just consider a couple of examples. In making a metric space from a set of functions we have to decide on an appropriate notion of distance for the functions. The significance of different metrics can often be understood by thinking of the distance between two functions as being a measure of the error in the representation of one by the other. That is, if f is what we want and g is what we’ve got, then $d(f, g)$ measures the error in what we’ve got.

Definition 2.1.15 $C_\infty[a, b]$ *is the metric space comprising the set of continuous, real-valued functions on $[a, b]$ with the L_∞ (or uniform) metric:*

$$L_\infty(f, g) = \sup\{|f(x) - g(x)| \mid a \leq x \leq b\}.$$

Note that sup means, roughly, the maximum, so

$$L_\infty(f, g) = \max_{a \leq x \leq b} |f(x) - g(x)|.$$

Remark 2.1.8 *This metric space is appropriate when we are interested in the maximum difference between two functions (in the worst case, so to speak).*

Definition 2.1.16 $C_1[a, b]$ *is the metric space comprising the set of continuous, real-valued functions on $[a, b]$ with the L_1 metric:*

$$L_1(f, g) = \int_a^b |f(x) - g(x)| dx.$$

Remark 2.1.9 *This metric space is appropriate when we are interested in the total (or average) difference between two functions. In other words, we might be satisfied with an error that is, on the average, small, even if it is large in some cases.*

Definition 2.1.17 $\mathcal{C}_2[a, b]$ is the metric space comprising the set of continuous, real-valued functions on $[a, b]$ with the L_2 metric:

$$L_2(f, g) = \sqrt{\int_a^b [f(x) - g(x)]^2 dx}.$$

Remark 2.1.10 *This Euclidean function space is appropriate for many physical problems (in which energy depends quadratically on some other quantity), and is often chosen for mathematical convenience (since its derivative is linear).*

Exercise 2.1.18 *Suppose $L_2(f, g) < \epsilon$. Is there any limit on $|f(x) - g(x)|$ for a given value of x ? What does this tell you about a function approximation that is good in an L_2 sense? Suppose $L_\infty(f, g) < \epsilon$; what does this tell you about $|f(x) - g(x)|$ for particular values of x ?*

2.2 Topology

The theory of metric spaces deals with general models of nearness and distance based on a metric that quantifies distance. Topology generalizes these ideas by dealing with nearness (neighborhoods) in a context that doesn't require a metric (i.e. a quantification of distance).

2.2.1 Open balls

Definition 2.2.1 (Open Ball $B_r(c)$) *The open ball with center c and radius $r > 0$ in a metric space (X, d) is defined:*

$$B_r(c) = \{x \in X \mid d(x, c) < r\}.$$

We call $B_r(c)$ an open ball around c .

Remark 2.2.1 *Other names for the open ball $B_\epsilon(c)$ are an open sphere and an ϵ -neighborhood.*

Remark 2.2.2 (Closed balls and boundaries) *It is sometimes convenient to talk about closed balls or spheres around a point, $\{x \in X \mid d(x, c) \leq r\}$. Likewise, we can talk about the boundary of an open or closed ball: $\{x \in X \mid d(x, c) = r\}$. Note that a closed ball includes its boundary, whereas an open ball excludes it.*

Remark 2.2.3 *All of the points of an open ball are in its interior or, equivalently, none of its points are on its boundary (and hence it excludes its boundary). To see this, observe that around any point of the ball you can place another open ball that lies entirely within the first. Conversely, observe that for points on the boundary of a closed ball it is not possible to put an open ball around them that lies entirely within the closed ball. Think about this and draw pictures to make sure you understand it at an intuitive level; it will pay off later.*

Exercise 2.2.1 *Prove the preceding remarks.*

Exercise 2.2.2 *What do the open balls $B_r(c)$ look like in the metric space (\mathbb{R}, L_p) ?*

Exercise 2.2.3 *What do the closed balls look like?*

Exercise 2.2.4 *What are the boundaries of the (open or closed) balls?*

Exercise 2.2.5 *Describe the open balls in any discrete metric space.*

2.2.2 Neighborhoods

Definition 2.2.2 (Neighborhood) *A subset N of a space X is called a neighborhood of $x \in X$ if it contains some open ball around x . That is, N is a neighborhood of x if and only if for some $r > 0$, $B_r(x) \subseteq N$.*

Remark 2.2.4 *Notice that a point cannot be on the boundary of any of its neighborhoods; a neighborhood of a point must provide some “space” around that point.*

Exercise 2.2.6 *Show that each open ball around a point is a neighborhood of that point.*

Exercise 2.2.7 Show that an open ball is a neighborhood of each of its points.

Exercise 2.2.8 Give an example of a neighborhood that is not an open ball.

Exercise 2.2.9 Show that in a discrete metric space every set containing a point is a neighborhood of that point.

Exercise 2.2.10 Show that the union of two neighborhoods of a point is a neighborhood of that point. Extend your result to the union of any finite family (set) of neighborhoods of the point. Does it also extend to infinite families of neighborhoods?

Exercise 2.2.11 Show that the intersection of two neighborhoods of a point is a neighborhood of that point. Extend your result to the intersection of any finite family (set) of neighborhoods of the point. Does it also extend to infinite families of neighborhoods? (Hint: Consider the family $\{B_{1/n}(c) \mid n = 1, 2, 3, \dots\}$ of neighborhoods of c .)

2.2.3 Open and closed sets

Definition 2.2.3 (Open set) A subset of a space is called open if it is a neighborhood of each of its points.

Definition 2.2.4 (Closed set) The complement of an open set is called a closed set.

Remark 2.2.5 Intuitively, an open set excludes its boundary. Thus, open intervals (a, b) are examples of open sets in \mathcal{E} ; they exclude their boundaries $\{a, b\}$. Conversely, closed sets include their boundaries, so the closed intervals $[a, b]$ are examples of closed sets in \mathcal{E} . Sets that both include and exclude some of their boundary are neither open nor closed; examples are the half-open intervals $[a, b)$ and $(a, b]$ in \mathcal{E} .

Exercise 2.2.12 Given the foregoing remarks, it may seem that a set could not be both open and closed. However, \emptyset and \mathbb{R} are both open and closed in \mathcal{E} . Explain how this can be.

Exercise 2.2.13 Show that in a discrete metric space every subset is open and every subset is closed.

Exercise 2.2.14 Show that any union (finite or infinite) of open sets is open.

Exercise 2.2.15 Show that any finite intersection of open sets is open.

Exercise 2.2.16 Give an example of an infinite family of open sets whose intersection is not open.

Exercise 2.2.17 Prove that any (finite or infinite) intersection of closed sets is closed.

Exercise 2.2.18 Prove that any finite union of closed sets is closed.

Exercise 2.2.19 Give an example of an infinite union of closed sets that is not closed.

2.2.4 Closure and interior of a set

The following are informal definitions to help you understand open and closed sets.

Definition 2.2.5 (Interior) The interior of a set is all of its points except its boundary points.

Example 2.2.1 For example, in \mathcal{E} , (a, b) is the interior of (a, b) , $(a, b]$, $[a, b)$ and $[a, b]$.

Definition 2.2.6 (Closure) The closure of a set is all of its points as well as all of its boundary points.

Example 2.2.2 For example, in \mathcal{E} , $[a, b]$ is the closure of (a, b) , $(a, b]$, $[a, b)$ and $[a, b]$.

Answer the following questions on the basis of your intuitive understanding of these concepts (i.e., don't give formal proofs).

Exercise 2.2.20 Is the interior of a set open or closed?

Exercise 2.2.21 Is the closure of a set open or closed?

Exercise 2.2.22 What is the union of a set and its boundary?

Exercise 2.2.23 What is the (set) difference of a set and its boundary?

Exercise 2.2.24 What is the intersection of the closure of a set and the closure of its complement.

2.2.5 Topological spaces

Topological spaces are more general than metric spaces because they allow us to specify the neighborhood properties of a space directly in terms of its open sets without the need of a metric. N.B. Previously the open sets were defined in terms of a metric (by way of open balls); here the open sets will be given as part of the definition of the topology.

Definition 2.2.7 (Topological space) A topology (X, \mathcal{T}) comprises an underlying set X and a topology \mathcal{T} , which is a family of subsets of X (called the open subsets of X) satisfying:

1. The empty set is open, i.e., $\emptyset \in \mathcal{T}$.
2. The whole space is open, i.e., $X \in \mathcal{T}$.
3. The intersection of any finite number of open sets is open. That is, if $S_1, \dots, S_n \in \mathcal{T}$ then $\bigcap_{k=1}^n S_k \in \mathcal{T}$.
4. The union of any (finite or infinite) collection of open sets is open. That is, if $\mathcal{S} \subseteq \mathcal{T}$, then $\bigcup_{S \in \mathcal{S}} S \in \mathcal{T}$.

Remark 2.2.6 As before, the closed sets are those whose complements are open in the topology. Notice that a set can be both open and closed in a topology; in particular the empty set and the whole space are both open and closed in all topologies. As before, some subsets of a space may be neither open nor closed.

Remark 2.2.7 This definition of a topological space specifies the topology through its open sets. The same topology can also be specified through its closed sets, through its neighborhood system, through its interior operator, through its closure operator, and in other ways.

Remark 2.2.8 A metric space has a “natural topology” in which the open sets are just those resulting from arbitrary unions of open balls. Alternately, we can take the open sets to be all those subsets for which every point in the subset can be surrounded by an open ball in the subset. Either way, this is called the metric topology for the (metric) space.

Definition 2.2.8 (Discrete topology) The discrete topology for a set X is $\mathcal{T} = \mathcal{P}(X)$, the set of all subsets of X .

Exercise 2.2.25 Show that the metric topology of a discrete metric space is the discrete topology. Conversely, one can define a discrete metric on any discrete topological space.

Definition 2.2.9 (Subspace) A topological space (X', \mathcal{T}') is called a subspace of a topological space (X, \mathcal{T}) if $X' \subseteq X$ and

$$\mathcal{T}' = \{S \cap X' \mid S \in \mathcal{T}\}.$$

That is the \mathcal{T}' is the relativization of \mathcal{T} to the subset, and it is called the relative topology on $X' \subseteq X$.

Exercise 2.2.26 Show that the relative topology is in fact a topology.

2.2.6 Bases

Proposition 2.2.1 A subset of a topological space is open if and only if it is a union of open balls.

Exercise 2.2.27 First part of proof: Show that any union of open balls is open. (Trivial)

Exercise 2.2.28 Second part of proof: Show that an arbitrary open set U is a union of open balls. Hint: Each point $x \in U$ is surrounded by an open ball $B_\epsilon(x) \subseteq U$.

Exercise 2.2.29 Describe in words the open sets in \mathcal{E} .

Definition 2.2.10 (Base) A family of open sets is called a base for a topology if every open set is a (possibly infinite) union of sets in the base. That is, $\mathcal{B} \subseteq \mathcal{T}$ is a base for (X, \mathcal{T}) if for every $U \in \mathcal{T}$ there is a $\mathcal{F} \subseteq \mathcal{B}$ such that $U = \bigcup \mathcal{F}$.

Remark 2.2.9 A topology may have many bases, each of which is capable of generating the entire topology. (Analogously, a vector space can have many bases, and each basis is capable of generating the entire vector space. A vector space is generated from a basis by linear operation — vector addition and scalar multiplication — whereas a topology is generated from a base by the union operation.)

Proposition 2.2.2 The open balls of a metric space are a base for the metric topology.

2.3 Convergence and limits

2.3.1 Convergence in metric spaces

Definition 2.3.1 (Convergence in metric spaces) *In a metric space (X, d) a sequence of points x_1, x_2, \dots is said to converge to the limit x , written $x_n \rightarrow x$, if and only if $d(x, x_n) \rightarrow 0$, that is, $0 = \lim_{n \rightarrow \infty} d(x, x_n)$. More precisely, for any $\epsilon > 0$ there is a positive N such that $d(x, x_n) < \epsilon$ for $n > N$. That is,*

$$\forall \epsilon > 0 \exists N > 0 \forall n > N : d(x, x_n) < \epsilon.$$

Proposition 2.3.1 *In a metric space, the limit, if it exists, is unique (i.e. a sequence cannot converge to two limits). Hence, if $x_n \rightarrow x$ we can write $x = \lim_{n \rightarrow \infty} x_n$.*

Exercise 2.3.1 *Prove this proposition. Hint: Assume the sequence has two limits $x \neq x'$, let $\epsilon = \frac{1}{2}d(x, x')$ and derive a contradiction.*

Proposition 2.3.2 *In a metric space, $x_n \rightarrow x$ if and only if, for each $\epsilon > 0$, x_n is eventually in $B_\epsilon(x)$ (that is, there is a positive N such that $x_n \in B_\epsilon(x)$ for all $n > N$).*

Proposition 2.3.3 *In a metric space, $x_n \rightarrow x$ if and only if, for each neighborhood U of x , x_n is eventually in U .*

Exercise 2.3.2 *Prove these propositions.*

Proposition 2.3.4 *In a metric space, the closure of a subset is the set of all limits of sequences in the subset.*

Remark 2.3.1 *Of course, the limits of some of these sequences lie in the interior, but the important point is that we may have sequences lying in the interior with limits on (but not outside) the boundary. This is the primary significance of the boundary. For example, in $(0, 1) \subset \mathcal{E}$, we can have sequences converging to 0, 1, or any point inside the interval, but not to any points less than 0 or greater than 1.*

Remark 2.3.2 *To prove the preceding proposition, suppose $x_n \in U$ and $x_n \rightarrow x$; therefore, for any $\epsilon > 0$, x_n is eventually in $B_\epsilon(x)$. Therefore*

every such ball intersects U (since it contains some of the x_n , which are in U). Either x is in U or it's not; if it is, there is nothing to prove. If it's not, then it must be on the boundary of U , since we've shown that every ϵ -neighborhood of it intersects U .

Proposition 2.3.5 *The metric topology is determined by the convergent sequences of points.*

Remark 2.3.3 *The convergent sequences define the closures of subsets, and thus the closed subsets, and thus the topology. This observation reveals some of the intimate connections between distance, neighborhoods and convergence that is captured by the notion of a topological space. (The preceding proposition does not hold for arbitrary topological spaces, but only for those that satisfy the "first countability axiom," which will not be discussed here.)*

2.3.2 Convergence in topological spaces

The preceding propositions motivate the following definition.

Definition 2.3.2 (Convergence in topological spaces) *In a topological space a sequence is said to converge to a limit if it is eventually in every neighborhood of that limit. That is, a sequence of points x_1, x_2, \dots converges to the limit x , written $x_n \rightarrow x$, if and only if for any neighborhood U of x there is a positive N such that $x_n \in U$ for all $n > N$.*

Remark 2.3.4 *In some topological spaces it is possible for a sequence to converge to two different points, $x_n \rightarrow x$ and $x_n \rightarrow x'$, but $x \neq x'$. Since this behavior is very unintuitive, the remainder of this section will develop an example; it is not, however, necessary for the remainder of the material.*

Definition 2.3.3 (Finite complement topology) *In the finite complement topology for a set, the only open sets are those whose complements are finite. Equivalently, the finite subsets are equivalent to the closed sets.*

Exercise 2.3.3 *Show that the finite complement topology is, in fact, a topology.*

Exercise 2.3.4 *Let (X, \mathcal{T}) be a topological space in which X is infinite and \mathcal{T} is the finite complement topology. Show that in such a space every neighborhood is infinite and has a finite complement.*

Exercise 2.3.5 Suppose that in this space $x_n \rightarrow x$ and $y_n \rightarrow y$. (For example, if $X = \mathbb{R}$, we might take $1 - 1/n \rightarrow 1$ and $-1 + 1/n \rightarrow -1$.) Show that the sequence $x_1, y_1, x_2, y_2, \dots$ converges to both x and y . (For example, the sequence $\frac{1}{2}, -\frac{1}{2}, \frac{2}{3}, -\frac{2}{3}, \frac{3}{4}, -\frac{3}{4}, \dots$ converges to both $+1$ and -1 .) Hint: Pick a neighborhood U of x ; we must show x_n, y_n, \dots is eventually in U . Since $x_n \rightarrow x$ we know x_n is eventually in U , so the problem is to show that y_n is also eventually in U . Recall now, that U has a finite complement.

Remark 2.3.5 This peculiar behavior can happen because in this space the neighborhoods are “very large” and in fact have to overlap; so the neighborhoods cannot shrink down (and become nonoverlapping) in the way we expect for convergence. In topological terminology, the space is not Hausdorff, which means that any two distinct points have nonoverlapping neighborhoods. Fortunately, metric spaces and most of the other topological spaces we have to deal with are Hausdorff.

2.4 Continuity

Intuitively, a function is continuous if infinitesimal variations of its input lead to infinitesimal variations of its output (no “cliff effects”).

Definition 2.4.1 (Continuous function) A function $f : X \rightarrow X'$ between (possibly identical) metric spaces (X, d) and (X', d') is called continuous at $x \in X$ if and only if, for each positive ϵ there is a δ such that $d'[f(x), f(y)] < \epsilon$ whenever $d(x, y) < \delta$. f is called continuous when it is continuous at every $x \in X$. More precisely,

$$\forall \epsilon > 0 \exists \delta > 0 \forall y \in X : d(x, y) < \delta \implies d'[f(x), f(y)] < \epsilon.$$

Remark 2.4.1 This is, of course, the standard “ ϵ - δ definition” of continuity, extended to metric spaces. The following propositions extend it to topological spaces.

Definition 2.4.2 (Image of a set) If $f : X \rightarrow X'$ and $S \subseteq X$, then the image of S under f is defined $f[S] = \{f(x) \mid x \in S\}$.

Proposition 2.4.1 In a metric space, f is continuous at x if and only if, for each positive ϵ there is a δ such that $f[B_\delta(x)] \subset B_\epsilon(f(x))$.

Proposition 2.4.2 *In a metric space, f is continuous at x if and only if, for each neighborhood V of $f(x)$ there is a neighborhood U of x such that $f[U] \subset V$.*

Remark 2.4.2 *That is, if the variation of the input is limited to U then the variation of the output will be limited to V .*

The preceding motivate the following definition of continuity for topological spaces.

Definition 2.4.3 (Continuous topological map) *A mapping f from a topological space (X, \mathcal{T}) to a topological space (X', \mathcal{T}') is called continuous at $x \in X$ if and only if, for each neighborhood V of $f(x)$ there is a neighborhood U of x such that $f[U] \subset V$.*

Remark 2.4.3 *There are many equivalent ways of defining continuous maps on topological spaces.*

Definition 2.4.4 *A map is said to be continuous on a set if it is continuous at every point in that set.*

2.5 Homeomorphism

Remark 2.5.1 *Suppose that f is a continuous, one-to-one mapping of (X, \mathcal{T}) into (X', \mathcal{T}') . In effect, f continuously embeds X in X' , possible “deforming” it in the process, but not “tearing” it. If there is also a continuous, one-to-one function $g : X' \rightarrow X$, then we see that each space can be continuously embedded in the other. In this sense they are equivalent (ignoring the continuous deformations). These observations motivate the following definitions.*

Definition 2.5.1 (Homeomorphism) *A homeomorphism is a continuous, one-to-one, onto mapping with a continuous inverse.*

Remark 2.5.2 *N.B. A homeomorphism is not the same as a homomorphism!*

Definition 2.5.2 (Homeomorphic spaces) *Two spaces are called homeomorphic if there is a homeomorphism between them.*

Remark 2.5.3 *Homeomorphic spaces are considered topologically equivalent, because topology is most concerned with properties that are preserved by homeomorphisms (“continuous deformations”). This is the reason topology is sometimes called “rubber sheet geometry” and is the source of the old canard about topologists being unable to tell a doughnut from a coffee cup (because they are topologically equivalent). Properties that hold across homeomorphic spaces are called topological invariants.*

Remark 2.5.4 *There are many equivalent ways of defining homeomorphism and homeomorphic, which can be found in topology texts. For our purposes it is sufficient to remember that there must be a correspondence between the two sets that is one-to-one and continuous in both directions.*

2.6 Separability

Definition 2.6.1 (Dense) *A subset S of a space X is said to be dense in that space if and only if the closure of S is X*

Remark 2.6.1 *Since the closure of S includes, in addition to S , all the limits of sequences in S , we see that if every element of X can be reached by a sequence in S , then S is dense in X . Indeed, in a metric space the closure of S is exactly the set of all limits of sequences in S (Prop. 2.3.4). In other words, S is dense in X if S is sufficient to generate X either in itself or by means of convergent sequences.*

Example 2.6.1 *The set $S = (-1, 0) \cup (0, 1)$ is dense in $[-1, 1]$ since the points -1 , 0 and 1 are all limits of sequences in S .*

Example 2.6.2 *The rational numbers are dense in the reals, since every real is the limit of a sequence of rationals (e.g. its decimal approximations).*

Definition 2.6.2 (Separable) *A space is called separable if it has a countable, dense subset.*

Example 2.6.3 *The reals are separable because the rationals are a countable, dense subset.*

Remark 2.6.2 *The significance of a space being separable is that it can be generated (via limits) from a countable set, and a countable set can, in principle, be specified by finite strings over a finite alphabet. (Why?)*

2.7 Completeness

2.7.1 Complete metric space

Definition 2.7.1 (Fundamental or Cauchy sequence) A sequence x_1, x_2, \dots in a metric space (X, d) is called Cauchy or fundamental if $d(x_m, x_n) \rightarrow 0$ as $m, n \rightarrow \infty$.

Remark 2.7.1 Notice that Cauchy sequences need not have limits. For example, in the space Q of rational numbers with the usual metric ($d(q, r) = |q - r|$), a sequence of rational approximations to $\sqrt{2}$, such as $1, 1.4, 1.41, 1.414, \dots$, is Cauchy but does not have a limit in Q (since $\sqrt{2} \notin Q$).

Definition 2.7.2 (Complete) A metric space is complete if every Cauchy sequence in the space converges to a point of the space.

Remark 2.7.2 In other words, a complete space includes all its limits. Thus, for example, the real numbers with the usual metric are complete, but the rationals are not complete.

Exercise 2.7.1 Show that, with the usual metric, $[0, 1]$ is complete but $(0, 1)$ is not.

Exercise 2.7.2 Show that the integers with the usual metric are complete.

2.7.2 Completion of metric spaces

Definition 2.7.3 (Completion) A metric space (X', d') is a completion of a metric space (X, d) if X' is complete and X is isometric to a subset of X' .

Remark 2.7.3 The fact that X is isometric to a subset Y of X' means that the members of X and Y are “essentially the same”; they can be thought of as different names for the same thing. Therefore, since X' is complete, it has limits for all the sequences in Y , and these limits can be added to X in order to complete it. For example, if $f : (X, d) \rightarrow (X', d')$ is an isometric mapping and x_1, x_2, \dots is a Cauchy sequence in X , then $f(x_1), f(x_2), \dots$ will be a Cauchy sequence in X' . (Why?) Therefore $\lim f(x_n)$ exists in X' and we can adjoin this to X to provide a limit for the original sequence x_n . (This “adjoining” is a formal trick for adding elements to a set without interfering

with the elements already there; in effect they are “colored” differently so that there is no possibility of confusion.) The domain of f can then be extended to include the adjoined element, so that $f(\lim x_n) = \lim f(x_n)$.

Proposition 2.7.1 (Uniqueness) *The completion of a metric space (X, d) is “essentially unique” in that the subsets corresponding to X in completions must be isometric. That is, if X is isometric to Y in the completion (X', d') , and X is isometric to Z in the completion (X'', d'') , then Y and Z are isometric. (Why?)*

Proposition 2.7.2 *Every metric space can be completed.*

Remark 2.7.4 *Sketch of proof: Let (X, d) be a metric space and let (X^*, d^*) be the space of “Cauchy-sequence equivalence-classes” over X . That is, the elements of X^* are sets of equivalent Cauchy sequences in X , where two Cauchy sequences x_n, y_n are called equivalent if $d(x_n, y_n) \rightarrow 0$. (Show this is an equivalence relation.) The metric on X^* is defined:*

$$d^*(S, T) = \lim_{n \rightarrow \infty} d(s_n, t_n),$$

where $s \in S$ and $t \in T$. That is, the distance between two equivalence classes, S and T , of Cauchy sequences is the distance between the limits of any two representatives, s and t , of these classes. It is necessary to show that this function is well-defined, that is, that it gives the same result no matter what representatives are chosen, and that the function is a metric.

Exercise 2.7.3 *Show that d^* is well-defined and a metric.*

Remark 2.7.5 *The space (X^*, d^*) is complete. (Try to prove it!) Furthermore, it is a completion of (X, d) . To see this, notice that $x \in X$ can be embedded in X^* by identifying it with the equivalence class of the (stationary or constant) Cauchy sequence x, x, x, \dots . (Show that stationary sequences are Cauchy.) Finally, it’s necessary to show that the set of equivalence classes of stationary Cauchy sequences is dense in (X^*, d^*) . (Give it a try.)*

Remark 2.7.6 *In this way, real numbers (including irrational numbers) can be considered no more than new names for Cauchy sequences of rational numbers. This is the way mathematicians “construct” the real numbers from the rationals. Since rational numbers can be defined as pairs of integers, in this way the reals can be reduced to integers, which is the way Dedekind and other 19th mathematicians endeavored to “arithmetize geometry” (thus reducing the continuum to a discrete formal system).*

Remark 2.7.7 *In the same way, any metric space (X, d) can be completed by replacing the elements of X by equivalence classes of stationary sequences, and by adjoining to the space the equivalence classes of Cauchy sequences. Further, since all completions are essentially the same (isometric), the construction of equivalence classes of Cauchy sequences is (essentially) the only way to complete a space.*

Proposition 2.7.3 *Every metric space can be considered the dense subset of a complete metric space. (That is, by ignoring as irrelevant the isometry that embeds the original space in its completion.)*

Remark 2.7.8 *Completion, as its name suggests, is the end of the road, at least so far as limits are concerned. Since a complete space includes all the limits of Cauchy sequences, there is no way these sequences can lead to additional elements.*

2.8 Connectedness

2.8.1 Connection

Definition 2.8.1 (Connected space) *A topological space (X, \mathcal{T}) is connected if it has no disjoint, closed proper subsets $A, B \subset X$ such that $X = A \cup B$.*

Proposition 2.8.1 *Connectedness is a topological invariant.*

Definition 2.8.2 (Connected set) *A subset S of a topological space is connected if it is connected as a subspace of the topological space.*

Example 2.8.1 *Thus, \mathbb{R} and $(0, 1)$ are connected subsets of \mathcal{E} but $[0, 1] \cup [2, 3]$, $(-1, 0) \cup (0, 1)$ and $\{-1, 1\}$ are not.*

Exercise 2.8.1 *Show that a topological space is connected if and only if it has no open, disjoint proper subsets whose union is the space. Hint: Review the definition of a closed set in a topology.*

Definition 2.8.3 (Disconnected) *A space or set is called disconnected if it is not connected.*

Exercise 2.8.2 Show examples of various connected and disconnected subsets of \mathcal{E}^2 to illustrate the preceding definitions.

Definition 2.8.4 (Separated sets) Two nonempty sets in a space are said to be separated if the intersection of each with the closure of the other is empty.

Remark 2.8.1 That is, two sets are not separated if one can be reached by a convergent sequence in the other.

Exercise 2.8.3 Give example of various separated and nonseparated pairs of sets in \mathcal{E}^2 .

Proposition 2.8.2 A set is disconnected if and only if it is the union of two (or more) separated sets.

Proposition 2.8.3 Let $\mathbf{2} = \{0, 1\}$. A topological space (X, \mathcal{T}) is connected if and only if the only continuous maps from X to $\mathbf{2}$ are constant maps.

Remark 2.8.2 Sketch of proof: Since $\{0\}$ and $\{1\}$ are open in $\mathbf{2}$, their inverse images $A = f^{-1}[\{0\}]$ and $B = f^{-1}[\{1\}]$ are disjoint, open sets in X for which $X = A \cup B$. Therefore X is disconnected. Conversely, if X is the union of disjoint open sets A, B , then we can define $f(x) = 0$ for $x \in A$ and $f(x) = 1$ for $x \in B$. This function is continuous, since the inverse images of the open sets of $\mathbf{2}$ (namely, $\emptyset, \{0\}, \{1\}, \mathbf{2}$) are open.

Remark 2.8.3 This proposition says that a continuous function cannot map a connected space to two different points. It implies that discrete categories cannot be defined by a continuous function over a connected space. That is, under conditions of continuity, exact categorization is impossible.

Exercise 2.8.4 Show that any two disjoint sets in a discrete topology are separated.

Definition 2.8.5 (Connected component) A (connected) component of a set is a maximal connected subset of it. The component of a point in the space is the (unique) component that contains it.

Exercise 2.8.5 Why is the component of a point unique?

Remark 2.8.4 A connected space has one component.

Proposition 2.8.4 The components of a discrete topology are singleton sets.

2.8.2 Continua and discontinua

Different topologists define “continuum” differently, but all require a continuum to be connected. Hausdorff (*Set Theory*, 173) calls any closed, connected set a *continuum* (an open, connected set being called a *domain*). More generally (e.g. MT 264) a continuum is required to be *compact* (a concept I have avoided in this course) as well as connected. Some authors (e.g. EM 81C, MEGT 158) further require a continuum to have more than one point. For our purposes, Hausdorff’s definition is adequate:

Definition 2.8.6 (Continuum) *A continuum is a closed, connected set.*

Definition 2.8.7 (Totally disconnected) *A set is totally disconnected if its components are singleton sets.*

Remark 2.8.5 *A discrete topology is totally disconnected.*

2.8.3 Path connection

There is another, slightly more restricted definition of connection, which is often more useful.

Definition 2.8.8 (Path or arc) *A continuous function $p : [0, 1] \rightarrow X$ is called a path or arc in the space (X, \mathcal{T}) . The initial point of the path is $x_i = p(0)$ and the terminal point of the path is $x_t = p(1)$. The path is said to join or connect x_i to x_t .*

Remark 2.8.6 *Notice that a path may intersect itself any number of times.*

Remark 2.8.7 *One natural interpretation of a path is as a continuous trajectory through a continuum state space.*

Remark 2.8.8 *A path is a homeomorphic image of $[0, 1]$ with the usual metric. Indeed, this is a common definition of a path.*

Definition 2.8.9 (Path-connected) *A space is called path-connected (arc-connected, pathwise-connected, etc.) if from any point in the space there is a path to any other point in the space. A set is path-connected if it is path-connected as a subspace (i.e. in the relative topology).*

Remark 2.8.9 *Therefore, in a path-connected space, any point is reachable from any other point. Alternately, any point can be continuously transformed into any other point of the space.*

Remark 2.8.10 *Path-connected components can be defined in the same way as (connected) components.*

Remark 2.8.11 *The preceding definitions of connection are both quite intuitive, and one might wonder if they are equivalent; they are not. Consider the following set: $Y \cup C$, where Y is the y -axis between -1 and 1 , $Y = \{(0, y) \mid -1 \leq y \leq 1\}$, and C is the graph of the positive cosine-reciprocal function, $C = \{(x, y) \mid y = \cos(1/x), x > 0\}$. Notice that C oscillates more and more rapidly as it approaches the y -axis, but it never actually touches it. Therefore, $Y \cup C$ is not path connected, since there are no paths between points in Y and points in C . On the other hand, it is connected, as we can see by noting that C is approaching Y , and therefore the closure of C is not disjoint from Y .*

However,

Proposition 2.8.5 *Every path-connected set is connected.*

Exercise 2.8.6 *Prove this proposition. Hint: Suppose the set is the union of two disjoint subsets. You know there is a path from any point in one to any point in the other. Show from this that the subsets cannot be separated.*

2.8.4 Processes

We need to be able to talk about continuous processes in topological spaces, but we do not have the usual mechanisms for defining them, such as differential equations.

Definition 2.8.10 (Process) *A process on a topological space S (the state space) is a continuous function $f : S \times \mathbb{R} \rightarrow S$ satisfying the group properties:*

$$\begin{aligned} f(s, 0) &= s, \\ f[f(s, t), u] &= f(s, t + u). \end{aligned}$$

The intuitive meaning of $f(s, t)$ is the state of the system t time units after it is in state s . The two preceding equations follow from this intuitive definition.

Exercise 2.8.7 *Explain in words why you would expect each of these equations to be true.*

Remark 2.8.12 *Notice that for any $t \neq 0$ a process f defines a path from s to $f(s, t)$.*

Exercise 2.8.8 *Prove the preceding remark.*

Definition 2.8.11 (Trajectory) *If there is some $t > 0$ such that $f(s_i, t) = s_f$, then we may speak of a trajectory from s_i to s_f . Notice that by this definition of trajectory and process, a trajectory cannot intersect itself (i.e. go in two different directions from a the point of intersection).*

Exercise 2.8.9 *Explain why.*

2.9 References

Sources for this chapter include the *Encyclopedic Dictionary of Mathematics* (Mathematical Society of Japan, 1980, secs. 81A, B, C, 273C, D, J, 408B, G), McCarty (1967, chs. 3, 4, 6), Mendelson (1975, chs. 2, 3, 4), Moore (1964, chs. 2, 3, 6), and Moore (1985, ch. 3).

Chapter 3

Banach Spaces

This chapter is a sort of mid-point between the preceding chapter, on topology, and Ch. 5 on Hilbert spaces, for many of the properties commonly associated with Hilbert spaces are in fact properties of the larger class of Banach spaces. This more general framework is useful in applications of field computation in which the postulates of Hilbert spaces might not apply. We begin with a description of normed linear spaces, which are appropriate for many image-like information representations. In general, all we need are appropriate notions of adding, scaling, and measuring the “size” of things. Next we turn to Banach spaces, which have the additional property of being *complete*, that is, containing the limits of convergent sequences. This gives us very general notions of differentiation and a Banach-space version of Taylor’s theorem. Therefore we can apply these useful ideas to continuous and other non-standard information representations and develop an approximation theory that is useful in defining one notion of *universal field computation* (Ch. 9).

3.1 Linear spaces

3.1.1 Definitions

Definition 3.1.1 (Field) *Briefly, an (algebraic) field is a set on which sum and product operations are defined that satisfy the usual properties (both are commutative and associative, product distributes over sum, additive identity*

0, additive inverses, multiplicative inverses, multiplicative identity 1).¹

Example 3.1.1 *The rationals (\mathbb{Q}), reals (\mathbb{R}), and complex numbers (\mathbb{C}) are the best-known examples of algebraic fields (and the ones we will have most use of).*

Definition 3.1.2 (Linear space) *A linear space over K comprises a field K (of “scalars”), a set L (of “vectors”), and two operations, a sum $\oplus : L \times L \rightarrow L$ and a (scalar) product $\otimes : K \times L \rightarrow L$. The operators satisfy the following properties (for $a, b \in K$ and $\mathbf{x}, \mathbf{y}, \mathbf{z} \in L$):*

associative sum: $(\mathbf{x} \oplus \mathbf{y}) \oplus \mathbf{z} = \mathbf{x} \oplus (\mathbf{y} \oplus \mathbf{z})$.

zero element: *There is a zero element $\mathbf{0} \in L$ satisfying*

$$\mathbf{0} \oplus \mathbf{x} = \mathbf{x} = \mathbf{x} \oplus \mathbf{0}.$$

additive inverses: *For every $\mathbf{x} \in L$ there is a $-\mathbf{x} \in L$ such that*

$$-\mathbf{x} \oplus \mathbf{x} = \mathbf{0} = \mathbf{x} \oplus -\mathbf{x}.$$

(This is implied by the additive inverses in the field.)

commutative sum: $\mathbf{x} \oplus \mathbf{y} = \mathbf{y} \oplus \mathbf{x}$.

right distribution: $a \otimes (\mathbf{x} \oplus \mathbf{y}) = (a \otimes \mathbf{x}) \oplus (a \otimes \mathbf{y})$.

left distribution: $(a + b) \otimes \mathbf{x} = (a \otimes \mathbf{x}) \oplus (b \otimes \mathbf{x})$, *where $+$ is the addition operation of the field.*

associative product: $a \otimes (b \otimes \mathbf{x}) = (a \times b) \otimes \mathbf{x}$, *where \times is the product operation of the field.*

unit product: $1 \otimes \mathbf{x} = \mathbf{x}$, *where 1 is the unitary element (multiplicative identity) of the field. (It is easy to show $0 \otimes \mathbf{x} = \mathbf{0}$ and $-\mathbf{x} = (-1) \otimes \mathbf{x}$; do it!)*

¹In this chapter, “field” refers to an algebraic field, as opposed to the fields that are the subject of field computation as defined in Ch. 11.2.1.

Remark 3.1.1 The terms “scalar” and “vector” are meant to be suggestive, but it should not be assumed that the elements of a linear space are scalars and vectors of the familiar kind (they need not even be numbers); we will sometimes say “generalized vector” to stress this point. Indeed any set of objects may form a linear space if they can be “scaled” and “summed” in accord with the preceding definition.

Remark 3.1.2 K is called the basic field or ground field of the linear space. A linear space over the real numbers is called a real linear space, and one over the complex numbers a complex linear space.

Notation 3.1.1 For convenience, when confusion is unlikely, we write $a\mathbf{x}$ for $a \otimes \mathbf{x}$ and $\mathbf{x} + \mathbf{y}$ for $\mathbf{x} \oplus \mathbf{y}$. Further, we will write $-\mathbf{x}$ for $-1 \otimes \mathbf{x}$ and $\mathbf{x} - \mathbf{y}$ for $\mathbf{x} \oplus -\mathbf{y}$.

Definition 3.1.3 (Linear dependence and independence) A set of vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ is linearly independent if $a_1\mathbf{x}_1 + \dots + a_n\mathbf{x}_n = 0$ implies $a_1 = \dots = a_n = 0$. Conversely, they are linearly dependent if there are a_k , not all zero, such that $a_1\mathbf{x}_1 + \dots + a_n\mathbf{x}_n = 0$.

Remark 3.1.3 The significance, of course, of linear dependence is that the vectors can be written as linear combinations of the other vectors. For example, if $a_1\mathbf{x}_1 + \dots + a_n\mathbf{x}_n = 0$ with $a_1 \neq 0$, then $\mathbf{x}_1 = -a_1^{-1}(a_2\mathbf{x}_2 + \dots + a_n\mathbf{x}_n)$.

3.1.2 Examples

Example 3.1.2 The most familiar example of a linear space is the vector space of n -dimensional vectors (either real or complex).

Example 3.1.3 The preceding example can be extended to the linear space of infinite sequences of elements of a field K . (Give definitions for the sum and product operations of this linear space.)

Example 3.1.4 The set of K -valued functions is a linear space over K . (Define the sum and product operations of this linear space.)

Example 3.1.5 Linear spaces can be defined for various important subsets of the space of K -valued functions, for example continuous functions and differentiable functions.

Remark 3.1.4 *Thus the set of spatially continuous images (fields, in the sense of field computation), is a linear space. Describe the sum and product operations in intuitive terms (e.g., what is their effect on visual images).*

Exercise 3.1.1 *Show that the preceding are linear spaces. (The major problem is to show that the operations are closed, e.g., that the sum of two continuous functions is continuous.)*

Exercise 3.1.2 *Define the linear space of n -term polynomials with coefficients in K .*

Exercise 3.1.3 *Show that the space of $m \times n$ matrices forms a linear space with the usual definitions of scalar product and matrix addition.*

Definition 3.1.4 (Linear function) *A complex-valued function $f : X \rightarrow \mathbb{C}$ is called linear if for all $x, y \in X$ and $a, b \in \mathbb{C}$, $f(ax + by) = af(x) + bf(y)$.*

Exercise 3.1.4 *Show that the set of all linear functions on X forms a linear space.*

Remark 3.1.5 *There are few purely linear systems in nature. Usually some quantity cannot be increased without bound, either because some “substance” will run out, or because it will begin to behave nonlinearly — as when a high voltage arcs through a capacitor or a high current vaporizes a conductor.*

As a rule of thumb, it’s best to treat any linear model as an approximation. Further, it is unwise to depend on any properties of linear models that do not hold (at least approximately) for approximately linear models. One very useful class of approximately linear model is the saturating linear model, which is linear for moderate quantities, but approaches limits for extreme quantities. An example is a sigmoid function.

3.2 Normed linear spaces

3.2.1 Norms

3.2.1.1 Definitions

A *norm* measures the “size” or “magnitude” of an element of a linear space.

Definition 3.2.1 (Norm) An operation $N : L \rightarrow \mathbb{R}_*$ on a real or complex linear space L is a norm if and only if:

zero norm: $N(\mathbf{x}) = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

scaling: $N(a \otimes \mathbf{x}) = |a|N(\mathbf{x})$.

triangle inequality: $N(\mathbf{x} \oplus \mathbf{y}) \leq N(\mathbf{x}) + N(\mathbf{y})$.

Notation 3.2.1 $N(\mathbf{x})$ is usually written $\|\mathbf{x}\|$, possibly with subscripts to distinguish different norms.

Remark 3.2.1 The absolute value is a norm on the reals.

Exercise 3.2.1 Show that for any norm $\|\cdot\|$ on \mathbb{R} , there is a positive η such that $\|x\| = \eta|x|$ for all $x \in \mathbb{R}$. What is the value of η ?

Exercise 3.2.2 Show that a norm is a continuous mapping.

3.2.1.2 Examples

Definition 3.2.2 (L_p norms) The L_p norms of n -dimensional real or complex vectors are defined:

$$\|\mathbf{x}\|_p = \sqrt[p]{\sum_{k=1}^n |x_k|^p}.$$

More concisely, $\|\mathbf{x}\|_p^p = \sum x_k^p$. As usual, $\|\mathbf{x}\|_\infty = \max_k |x_k|$. L_2 is called the Euclidean norm.

Exercise 3.2.3 Show that that L_p norms are in fact norms.

Exercise 3.2.4 What is the relation between L_p norms and L_p metrics?

Definition 3.2.3 (L_p product norms) If L is a cartesian product of normed linear spaces L_k , that is, $L = L_1 \times \cdots \times L_n$, then an L_p norm can be defined as follows:

$$\|X\|_p = \sqrt[p]{\sum_{k=1}^n \|X_k\|_k^p},$$

where $\|X_k\|_k$ represents the norm on L_k (possibly different in each case).

Exercise 3.2.5 Show that the L_p product norm is in fact a norm.

Definition 3.2.4 $\mathcal{C}[a, b]$ is the space of all continuous real-valued functions on $[a, b]$.

Definition 3.2.5 (Uniform norm) For $f \in \mathcal{C}[a, b]$, the uniform norm is defined:

$$\|f\| = \max_{a \leq x \leq b} |f(x)|$$

Exercise 3.2.6 Show that the uniform norm is in fact a norm.

3.2.2 Normality

Definition 3.2.6 (Normalized) A (generalized) vector \mathbf{x} is normalized if its norm is 1, $\|\mathbf{x}\| = 1$. A nonzero vector can always be normalized by dividing by its norm, that is, $\mathbf{x}/\|\mathbf{x}\|$ is normalized.

Exercise 3.2.7 Use the definition of a norm to show that $\mathbf{x}/\|\mathbf{x}\|$ is normalized (for $\mathbf{x} \neq \mathbf{0}$).

Remark 3.2.2 Normalization considers (generalized) vectors independently of their absolute size, that is independently of scaling.

Exercise 3.2.8 For various values of $p \geq 1$, describe the effect of L_p -normalization on two-dimensional vectors.

Exercise 3.2.9 Describe the effect of normalization on two-dimensional visual images. Describe the effect on acoustic signals of finite duration.

3.2.3 Normed linear space

Definition 3.2.7 (Normed linear space) A normed linear space is a real or complex linear space together with a norm on that space.

Exercise 3.2.10 Show that $\mathcal{C}[a, b]$ with the uniform norm is a normed linear space.

Exercise 3.2.11 If you know some numerical analysis, then show that $\mathcal{C}[a, b]$ is separable. Hint: Apply the Weierstrass approximation theorem.

Proposition 3.2.1 *A normed linear space is a metric space under the norm-metric $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.*

Exercise 3.2.12 *Prove this proposition.*

Definition 3.2.8 (Translation and scaling invariance) *A metric is translation invariant if*

$$d(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z}) = d(\mathbf{x}, \mathbf{y}).$$

It is scaling invariant if

$$d(a\mathbf{x}, a\mathbf{y}) = |a|d(\mathbf{x}, \mathbf{y}).$$

Exercise 3.2.13 *Draw diagrams to illustrate the concepts of translation and scaling invariant metrics.*

Proposition 3.2.2 *The norm-metric is translation and scaling invariant.*

Exercise 3.2.14 *Prove this proposition.*

Definition 3.2.9 (Topologically equivalent norms) *Two norms, $\|\cdot\|$ and $\|\cdot\|'$, are topologically equivalent if there are positive real numbers ζ, η such that for all $\mathbf{x} \in L$,*

$$\zeta\|\mathbf{x}\| \leq \|\mathbf{x}\|' \leq \eta\|\mathbf{x}\|.$$

Exercise 3.2.15 *Explain why such norms should be called “topologically equivalent.” Hint: Suppose a sequence converges by one norm. What can you say about convergence under topologically equivalent norms?*

Exercise 3.2.16 *Show that in \mathcal{E}^n , the L_2 -norm is topologically equivalent to the L_∞ -norm.*

Exercise 3.2.17 *Show that in \mathcal{E}^n , all norms are topologically equivalent.*

Proposition 3.2.3 *Suppose $f : X \rightarrow Y$, where X and Y are normed linear spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. If there are positive reals ζ, η such that for all $\mathbf{x}, \mathbf{y} \in X$,*

$$\zeta\|\mathbf{x} - \mathbf{y}\|_X \leq \|f(\mathbf{x}) - f(\mathbf{y})\|_Y \leq \eta\|\mathbf{x} - \mathbf{y}\|_X,$$

then f is a homeomorphism, and X and Y are homeomorphic.

Exercise 3.2.18 *Prove this. Hint: What does the inequality say about the continuity of f and f^{-1} ?*

3.3 Banach spaces

Definition 3.3.1 (Banach space) A Banach space is a complete normed linear space.

Remark 3.3.1 Thus the significance of a Banach space is that Cauchy sequences are guaranteed to converge to limits; that is, if $\|\mathbf{x}_m - \mathbf{x}_n\| \rightarrow 0$, then we know $\lim_{n \rightarrow \infty} \mathbf{x}_n$ exists. In particular, this means that operations such as differentiation can be defined.

Exercise 3.3.1 Show that \mathcal{E}^n is a Banach space.

3.4 Differentiation in Banach spaces

3.4.1 Fréchet derivative

Definition 3.4.1 (Fréchet differentiation) Suppose X and Y are two Banach spaces and U is an open subset of X . Then $T : U \rightarrow Y$ is Fréchet differentiable at ϕ if there is a bounded linear operator $D : X \rightarrow Y$ such that the following holds. For all $\alpha \in X$ such that $\phi + \alpha \in U$, there is an $E : X \rightarrow Y$ such that

$$T(\phi + \alpha) = T(\phi) + D(\alpha) + E(\alpha)$$

and

$$\lim_{\|\alpha\| \rightarrow 0} \frac{\|E(\alpha)\|}{\|\alpha\|} = 0.$$

Under these circumstances, D is called the Fréchet derivative of T at ϕ ; it is denoted by $T'(\phi)$. The Fréchet derivative is a locally linear approximation to T ; $T'(\phi)(\alpha) = D(\alpha)$ is called the Fréchet differential of T .

Remark 3.4.1 Since a linear operator is continuous if and only if it is bounded, Fréchet derivatives are (by definition) continuous.

Proposition 3.4.1 The derivative of a linear operator is that operator: $L'(\phi) = L$.

Remark 3.4.2 Note that $T' : X \rightarrow \mathcal{L}(X, Y)$, where $\mathcal{L}(X, Y)$ is the space of all continuous (bounded) linear operators from X to Y .

Remark 3.4.3 Higher order derivatives are defined in the obvious way. Suppose $T : X \rightarrow Y$. Since $T' : X \rightarrow \mathcal{L}(X, Y)$, it is easy to see that the higher derivatives have the types:

$$\begin{aligned} T'' &: X \rightarrow \mathcal{L}(X, \mathcal{L}(X, Y)), \\ T^{(3)} &: X \rightarrow \mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X, Y))), \end{aligned}$$

and so forth. Note that each successive derivative is of “higher type” than its predecessor.

3.4.2 Gâteaux derivative

Definition 3.4.2 (Gâteaux differentiation) Suppose X and Y are Banach spaces, $U \subseteq X$ is open, and $T : U \rightarrow Y$. Then T has a Gâteaux derivative at $\phi \in U$ if, for all $\alpha \in U$ the following limit exists:

$$dT(\phi, \alpha) = \lim_{t \rightarrow 0} \frac{T(\phi + t\alpha) - T(\phi)}{t} = \frac{d}{dt} T(\phi + t\alpha)|_{t=0}.$$

We write $dT(\phi, \alpha)$ for the Gâteaux derivative of T at ϕ in the “direction” α .

Proposition 3.4.2 The Gâteaux derivative, if it exists, is unique.

3.4.3 Properties of derivatives

Proposition 3.4.3 Every Fréchet derivative is a Gâteaux derivative. Since the Gâteaux derivative is unique, the two derivatives are identical if the Fréchet exists.

Definition 3.4.3 (“Uncurried” derivatives) The spaces

$$\mathcal{L}(X, \mathcal{L}(X, \dots \mathcal{L}(X, \mathcal{L}(X, Y)) \dots))$$

[with k nestings of $\mathcal{L}(X, -)$] are isomorphic to the spaces $\mathcal{L}(X^k, Y)$, and we will often make use of this fact. We use $d^k T$ to denote that k -th order “uncurried” derivative of T :

$$\begin{aligned} d^k T &: X \rightarrow \mathcal{L}(X^k, Y), \\ d^k T(\phi)(\alpha_1, \dots, \alpha_k) &= T^{(k)}(\phi)(\alpha_1) \dots (\alpha_k). \end{aligned}$$

Proposition 3.4.4 The derivative of a composition is given by the following equation (shown in both curried and uncurried forms):

$$\begin{aligned} (T \circ U)'(\phi)(\alpha) &= T'[U(\phi)][U'(\phi)(\alpha)], \\ d(T \circ U)(\phi, \alpha) &= dT[U(\phi), dU(\phi, \alpha)]. \end{aligned}$$

3.5 Taylor theorem in Banach spaces

We consider now a generalization of the familiar Taylor theorem from real analysis. It permits the expansion of a function on a Banach space in an infinite series about a point in that space. If the Banach space is a function space, this theorem permits expanding an operator around a particular function. (Thus, in field computation, it permits expanding a field transformation around a fixed field.)

Proposition 3.5.1 (Taylor) *Suppose U is any open subset of a Banach space X and $T : X \rightarrow Y$ is a map which is C^n in U (that is, the first n derivatives of T exist). Let $\phi \in U$ and $\alpha \in X$ be such that $\phi + r\alpha \in U$ for all $r \in [0, 1]$. Then:*

$$T(\phi + \alpha) = \sum_{k=0}^{n-1} \frac{T^{(k)}(\phi)(\alpha)^k}{k!} + R_n(\phi, \alpha),$$

where

$$R_n(\phi, \alpha) = \int_0^1 \frac{(1-r)^{n-1} T^{(n)}(\phi + r\alpha)(\alpha)^n}{(n-1)!} dr.$$

Here ‘ $(\alpha)^k$ ’ denotes k occurrences of the argument α . Also note that $T^{(0)} = T$.

In uncurried form the Taylor expansion is:

$$T(\phi + \alpha) = \sum_{k=0}^{n-1} \frac{d^k T(\phi, \overbrace{\alpha, \dots, \alpha}^n)}{k!} + R_n(\phi, \alpha),$$

where

$$R_n(\phi, \alpha) = \int_0^1 \frac{(1-r)^{n-1} d^n T(\phi + r\alpha, \overbrace{\alpha, \dots, \alpha}^n)}{(n-1)!} dr,$$

and the appropriate number of α arguments (zero or more) must be supplied for $d^k T$.

3.6 References

Sources for this chapter include the *Encyclopedic Dictionary of Mathematics* (Mathematical Society of Japan, 1980, sec. 256A) and Moore (1985, chs. 5, 16).

Chapter 4

Basic Complex Analysis

Although real numbers are sufficient for most applications of field computation, complex numbers are sometimes required, as in Fourier analysis and the application of field computation is in quantum computation. Therefore the goal of this chapter is to provide an intuitive understanding of basic complex analysis, especially as it applies in Hilbert spaces; a systematic presentation of complex analysis is beyond its scope. In addition to standard material, this chapter includes a brief discussion of hyperbolic trigonometry and its applications in special relativity theory, which is intended to build intuition by stressing the analogies with ordinary (circular) trigonometry.

4.1 Argand diagram

As everyone knows, complex numbers involve $i = \sqrt{-1}$. However, it will be better at this point to forget about $\sqrt{-1}$ and understand complex numbers by means of the *Argand diagram* (Fig. 4.1). As a matter of history, mathematicians were dubious about imaginary numbers, and questioned their legitimacy, until familiarity with the Argand diagram showed that they could be thought of as ordinary two-dimensional vectors. For in the Argand diagram we simply represent the complex number $x + iy$ as a vector (x, y) . (In this sense “ i ” can be thought of as a place holder or tag to distinguish the Y-coordinate from the X-coordinate.) Then operations on complex numbers can be interpreted as operations on two-dimensional vectors, without concern for $\sqrt{-1}$. When complex numbers are represented in this way, they are said to lie in the *complex plane*. Real numbers lie along the positive and

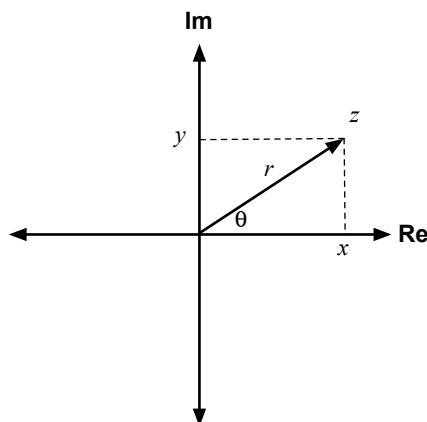


Figure 4.1: Argand diagram. $z = x + iy = re^{i\theta}$.

negative X-axis, and (pure) imaginary numbers along the positive and negative Y-axis; other points represent complex numbers with both (nonzero) real and imaginary parts. Therefore, in the complex plan the X-axis is called the *real axis* and the Y-axis is called the *imaginary axis*. (Why we should bother with complex numbers, and not simply make do with two-dimensional vectors, will become apparent as we proceed.)

Remark 4.1.1 *Notice that, unlike the real numbers, there is no natural sense in which the complex numbers can be ordered.*

Definition 4.1.1 (Cartesian components) *The $\Re : \mathbb{C} \rightarrow \mathbb{R}$ and $\Im : \mathbb{C} \rightarrow \mathbb{R}$ operators extract the Cartesian components (real and imaginary parts, respectively) of a complex number:*

$$\begin{aligned}\Re(x + iy) &= x, \\ \Im(x + iy) &= y.\end{aligned}$$

4.2 Geometrical Interpretations

The simplest use of the Argand diagram is to understand the addition and subtraction of complex numbers.

Definition 4.2.1 (Complex addition) Addition (or subtraction) of complex numbers is equivalent to vector addition in the Argand diagram:

$$(x + iy) + (x' + iy') = (x + x') + i(y + y').$$

Definition 4.2.2 (Complex multiplication) $(x + iy)(x' + iy') = (xx' - yy') + i(xy' + yx')$.

Remark 4.2.1 The definition of multiplication may seem mysterious, but it is motivated by the equation $i^2 = -1$. Thus,

$$(x + iy)(x' + iy') = xx' + iyx' + iy'x + i^2yy' = (xx' - yy') + i(xy' + yx').$$

Further, we will see that it has important implications independent of $\sqrt{-1}$.

Definition 4.2.3 (Complex conjugate) The complex conjugate \bar{z} of a complex number z is obtained by negating its imaginary part:

$$\overline{x + iy} = x - iy.$$

The notation z^* is also used for the complex conjugate.

Remark 4.2.2 The complex conjugate reflects the vector across the real (X) axis. Symmetry suggests that there ought to be an operation to reflect a complex number $x + iy$ across the imaginary (Y), yielding $-x + iy$, but it is not especially useful, so it doesn't have a name. Of course, simple negation reflects a complex number across both axes simultaneously, $-(x + iy) = -x - iy$.

Exercise 4.2.1 Prove the following:

$$\begin{aligned}\overline{\bar{x}} &= (x^*)^* = x \\ \overline{x + y} &= \bar{x} + \bar{y} \\ \overline{xy} &= \bar{x} \bar{y} \\ \overline{x/y} &= \bar{x}/\bar{y}\end{aligned}$$

Exercise 4.2.2 Show that

$$\Re z = \frac{z + \bar{z}}{2} \quad \text{and} \quad \Im z = \frac{z - \bar{z}}{2i}.$$

The Argand diagram suggests that the *magnitude* (“length”) of a complex number is a significant quantity. The length, for course, is $\sqrt{x^2 + y^2}$, but this can be expressed conveniently in terms of the complex conjugate (which is one of the reasons the complex conjugate is useful), since:

$$(x + iy)(x - iy) = x^2 + ixy - ixy - i^2y^2 = x^2 + y^2.$$

Therefore we have:

Definition 4.2.4 (Magnitude) $|z| = \sqrt{z\bar{z}}$.

Remark 4.2.3 Notice that this is consistent with the usual definition of the absolute value of a real number, since for a real r , $\bar{r} = r$; hence $|r| = \sqrt{r\bar{r}} = \sqrt{r^2}$. (Recall that, by convention, $\sqrt{}$ represents the nonnegative square root.)

Remark 4.2.4 The complex magnitude is a norm.

Remark 4.2.5 The distance between complex numbers is $|z - z'|$; it is the norm metric.

Exercise 4.2.3 What would be wrong with defining the magnitude of a complex number by $|z| = \sqrt{z^2}$. Would it be a norm? Would $|z - z'|$ be a metric?

Exercise 4.2.4 Show that $|\Re z| \leq |z|$ and $|\Im z| \leq |z|$.

Exercise 4.2.5 Show that $|z - w| \geq ||z| - |w||$.

Proposition 4.2.1 The reciprocal of a complex number is given by

$$(x + iy)^{-1} = \frac{x}{x^2 + y^2} - \frac{iy}{x^2 + y^2}. \quad (4.1)$$

Exercise 4.2.6 Derive the preceding formula by solving $wz = 1$ for w ; note that the real and complex parts of the equation can be solved separately.

Exercise 4.2.7 Write z^{-1} in terms of the complex conjugate and the magnitude. Does this simplify deriving Eq. 4.1?

Exercise 4.2.8 Derive the formula for dividing two complex numbers; you should get a formula in the form $X + iY$.

Remark 4.2.6 The complex numbers form an (algebraic) field; that is, there are both additive and multiplicative identities and inverses, both operations are commutative and associative, and multiplication distributes over addition.

4.3 Polar representation

4.3.1 Defined

Of course two-dimensional vectors can be represented in polar coordinates as well as in rectangular coordinates, but the polar representation is especially relevant to complex numbers. The radius is, of course, the magnitude of the complex number, also called the *modulus*.

Definition 4.3.1 (Magnitude or modulus) $\text{mod } z = |z| = \sqrt{x^2 + y^2}$.

The angle is measured counterclockwise from the positive X-axis and is called the *argument*, *phase*, *amplitude* or *angle* of the complex number. It can be defined as follows:

Definition 4.3.2 (Argument or phase) $\arg(x + iy) = \arctan(y/x)$.

Remark 4.3.1 *The mathematically most convenient way to measure angles is in radians, which is defined to be the area within a circle enclosed by twice the angle divided by the square radius of the circle. That is, if A is the area enclosed by the angle, then its radian measure is $\theta = 2A/r^2$. Since the circle has area πr^2 , an angle of π radians corresponds to 180° (since twice the angle includes the whole area), $\pi/2$ radians corresponds to 90° , 2π radians to 360° , etc. In general, if α is an angle in degrees, then $\theta = 2\pi(\alpha/360^\circ)$.*

Notation 4.3.1 *Because we are often interested in angles that are fractions or multiples of a complete cycle (2π radians), I have invented a kind of monogram, 2π , that I will use for 2π whenever it represents a complete cycle (360°).¹*

Remark 4.3.2 *We write $\arctan(y/x)$ so that the signs of x and y can be used to determine the quadrant of the complex plane in which the number falls. Thus $\arctan(+1/+1) = \pi/4$, $\arctan(+1/-1) = 3\pi/4$, $\arctan(-1/-1) = 5\pi/4$, and $\arctan(-1/+1) = 7\pi/4$, even though they all represent only two slopes, $+1$ and -1 .*

¹It turns out that the convention of using a single symbol for 2π goes back at least as far as H. Laurent's *Traité D'Algebra* (1889). In recent years some mathematicians have proposed using τ (standing for one *turn*) for 2π , and others have advocated for a different monogram: π . See Palais (2001) and <http://www.math.utah.edu/~palais/pi.html> (accessed 2012-05-10).

Remark 4.3.3 *It is often useful to consider \arg a multiple-valued function (like \arcsin , \arccos , etc.). Thus, for example,*

$$\arg(-1) = \pm\pi, \pm 3\pi, \pm 5\pi, \dots$$

Then, it is necessary to be clear about the range of \arg 's values; unless otherwise stated we will take it to be $[0, 2\pi)$. When the range is not important, we may write equations such as

$$\arg z = \theta \pmod{2\pi}$$

to indicate that angles are to be compared modulo 2π .

Remark 4.3.4 *If z is a complex number with magnitude r and phase θ , it's easy to see that the real part is given by $\Re z = r \cos \theta$ and the imaginary part by $\Im z = r \sin \theta$.*

Remark 4.3.5 *Notice every complex number has multiple polar representations (a property, of course, of any polar representation), since $\sin \theta = \sin(2\pi + \theta)$ and $\cos \theta = \cos(2\pi + \theta)$. In general, for any $n = 0, \pm 1, \pm 2, \dots$, $\sin \theta = \sin(2\pi n + \theta)$ and $\cos \theta = \cos(2\pi n + \theta)$. We will see that this periodicity in the phase of complex numbers makes them especially convenient for representing periodic phenomena such as waves.*

Exercise 4.3.1 *Given $z = re^{i\theta}$, show geometrically that*

$$r \cos \theta = \frac{z + \bar{z}}{2}, \quad r \sin \theta = \frac{z - \bar{z}}{2i}.$$

4.3.2 cis function

Suppose $r = |z|$ and $\theta = \arg z$; then it's easy to see:

$$z = \Re z + i\Im z = r \cos \theta + ir \sin \theta = r(\cos \theta + i \sin \theta).$$

This equation shows, in effect, how the complex number can be reconstituted from its magnitude and argument. For this purpose the “cis” ($\cos + i \sin$, pronounced “sis”) function is often used.

Definition 4.3.3 (cis function) $\text{cis } \theta = \cos \theta + i \sin \theta$.

Thus $z = r \operatorname{cis} \theta$, or, more generally,

$$z = |z| \operatorname{cis}(\arg z). \quad (4.2)$$

Proposition 4.3.1 $\operatorname{cis} \theta \operatorname{cis} \phi = \operatorname{cis}(\theta + \phi)$.

Exercise 4.3.2 *Prove this. Hint: recall from trigonometry,*

$$\begin{aligned} \sin(\theta + \phi) &= \sin \theta \cos \phi + \cos \theta \sin \phi, \\ \cos(\theta + \phi) &= \cos \theta \cos \phi - \sin \theta \sin \phi. \end{aligned}$$

Remark 4.3.6 *The equation $\operatorname{cis} \theta \operatorname{cis} \phi = \operatorname{cis}(\theta + \phi)$ suggests that cis has some similarities to the exponential function; we shall see that this is more than coincidental.*

4.4 Complex exponentials

4.4.1 Euler's Formula

4.4.1.1 IMAGINARY EXPONENTIALS

For the most part complex numbers have the same properties as real numbers, but of course it's necessary to analyse each property individually; here we will assume complex numbers are like real numbers unless stated otherwise. However, it is informative to look (informally) at the effect of taking the exponential of an imaginary number, $\exp(i\theta) = e^{i\theta}$. To do this we use the familiar power series for e^x :

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Substituting $i\theta$ for x we have:

$$\begin{aligned} e^{i\theta} &= 1 + i\theta + \frac{i^2\theta^2}{2!} + \frac{i^3\theta^3}{3!} + \frac{i^4\theta^4}{4!} + \dots \\ &= 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + \dots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots\right), \end{aligned}$$

where in the last line the real terms have been separated from the imaginary terms. As it turns out, the first parenthesized formula is the series for $\cos \theta$ and the second is that for $\sin \theta$. Thus we discover,

Proposition 4.4.1 (Euler's Formula) $e^{i\theta} = \cos \theta + i \sin \theta = \text{cis } \theta$.

It's now easy to discover the exponential of an arbitrary complex number $x + iy$:

$$e^{x+iy} = e^x e^{iy} = e^x \text{cis } y.$$

There is, however, an even more fruitful way to look at the complex exponential, since from Eq. 4.2 we see that any complex number can be written as a complex exponential:

$$z = |z|e^{i \arg z}.$$

Or, looked at another way, $re^{i\theta}$ is a complex number with magnitude (radius) r and phase angle θ .

Exercise 4.4.1 Show that $|e^{x+iy}| = |e^x|$.

Exercise 4.4.2 Use Euler's formula to prove that the following formulas are correct for real θ :

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}, \quad \sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}. \quad (4.3)$$

They also hold (by definition) for complex numbers.

Notice that -1 has a magnitude of 1 and a phase angle of π (i.e. 180°); thus we have the famous equation,

$$e^{\pi i} = -1.$$

Since a 2π (360°) rotation brings us back where we started, we also have the less famous formula,

$$e^{2\pi i} = 1$$

Obviously a phase angle that is any integral multiple of 2π will bring us back to 1. Similarly,

$$\begin{aligned} e^{\pi i/2} &= i, \\ e^{3\pi i/2} &= -i. \end{aligned}$$

Exercise 4.4.3 Show $re^{i\theta} = re^{i(2\pi m + \theta)}$ for $m = 0, 1, \dots$

Exercise 4.4.4 Show $(re^{i\theta})^* = re^{-i\theta}$.

Exercise 4.4.5 Show $re^{i\theta} = re^{-i(2\pi - \theta)}$.

4.4.1.2 MULTIPLICATION

Proposition 4.4.2 *Complex numbers can be multiplied by multiplying their magnitudes and adding their phase angles:*

$$(re^{i\theta})(se^{i\phi}) = (rs)e^{i(\theta+\phi)}.$$

Remark 4.4.1 *This proposition provides a geometric interpretation of complex multiplication, based on the proportion,*

$$1 : z :: w : wz.$$

*To multiply geometrically, construct a triangle with the sides 1 and z . Then construct a similar triangle on w , with the 1 of the first triangle corresponding to the w of the second. The result wz will be the side of the second triangle corresponding to the side z of the first. (Interestingly, this is exactly analogous to the construction Descartes used for defining the product of two real magnitudes; see ch. 4 of my book in progress, *Word and Flux*.)*

Exercise 4.4.6 *Do the construction suggested by the preceding remark, and show that the magnitudes and phase angles are correct, as given in the proposition.*

Remark 4.4.2 *Therefore a complex number, as a vector, can be rotated by multiplying by a suitable imaginary power of e . That is, to rotate z counterclockwise through an angle of θ , use $e^{i\theta}z$; for a clockwise rotation use $e^{-i\theta}z$.*

Exercise 4.4.7 *Give a rule for dividing complex numbers in terms of their magnitudes and phase angles.*

Remark 4.4.3 *We have seen that we can consider the complex number $x+iy$ as a two dimensional vector (x, y) with ordinary vector addition and a special multiplication rule. Similarly, we can consider the complex number $re^{i\theta}$ as a pair (r, θ) with a special operation that multiplies the magnitudes and adds the phase angles. In particular, whenever you have pairs of numbers for which you want to add the first components and multiply the second components, it may be worthwhile to think of them as complex numbers in polar coordinates. We will see an example shortly (Section 4.4.2).*

4.4.1.3 POWERS AND ROOTS

Proposition 4.4.3 (De Moivre's Theorem) *A complex number can be raised to the $p \geq 1$ power by raising its magnitude to the p power and by multiplying its phase by p : $(re^{i\theta})^p = (r^p)e^{ip\theta}$.*

Roots can be extracted in a similar way, but complex number bring some additional complications, as we will see by considering the “ n n -th roots of unity.” Consider first the square-root; we want to consider complex numbers z satisfying $z^2 = 1$. Writing the equation in polar form, we have

$$1 = (re^{i\theta})^2 = r^2 e^{2i\theta}.$$

To solve this, we must have $r = 1$, but we may have θ be any angle such that $2\theta = 2\pi m$ (for some $m = 0, 1, \dots$). If we restrict our attention to θ in the range $[0, 2\pi)$ (the *principal* square-roots), we see that $\theta = 0, \pi$ both solve the equation. Therefore, 1 has two square roots, $e^{0i} = 1$ and $e^{\pi i} = -1$. This is obvious enough, since $1^2 = (-1)^2 = 1$.

Now however we will apply the same method to determine the cube-roots of unity. Since $(re^{i\theta})^3 = r^3 e^{i3\theta}$, we again have $r = 1$, but now seek $\theta \in [0, 2\pi)$ such that $3\theta = 2\pi m$. Hence, $\theta = 0, (1/3)2\pi$ and $(2/3)2\pi$ are solutions.

Exercise 4.4.8 *Confirm that these θ are solutions.*

Hence, we find that 1 has three cube-roots, two of which are complex:

$$1, e^{2\pi i/3}, e^{4\pi i/3}.$$

In general we can see that 1 has n (principal) n -th roots, having phase angles satisfying $n\theta = 2\pi m$ ($m = 0, 1, \dots$), so $\theta = 2\pi m/n$. Hence,

Proposition 4.4.4 *The n principal n -th roots of unity are:*

$$1, e^{i2\pi/n}, e^{i4\pi/n}, \dots, e^{i2\pi(n-1)/n}.$$

In general, the principal values are $e^{i2\pi m/n}$, $m = 0, 1, \dots, n - 1$.

Proposition 4.4.5 *The n principal n -th roots of a complex number $z = re^{i\theta}$ are:*

$$\sqrt[n]{r}, \sqrt[n]{r} e^{i(\theta+2\pi)/n}, \sqrt[n]{r} e^{i(\theta+2\pi \cdot 2)/n}, \dots, \sqrt[n]{r} e^{i[\theta+2\pi(n-1)]/n}.$$

In general, the principal values are

$$\sqrt[n]{r} e^{i(\theta+2\pi m)/n} = \sqrt[n]{r} \operatorname{cis} \left(\frac{\theta}{n} + 2\pi \frac{m}{n} \right)$$

for $m = 0, 1, \dots, n - 1$.

Exercise 4.4.9 Prove this proposition.

4.4.2 Periodic Change

4.4.2.1 INTRODUCTION TO PERIODIC CHANGE

The polar representation of complex numbers makes them especially convenient for representing periodic processes, especially those involving sinusoidal change.

Remark 4.4.4 *It is generally convenient to measure the rate of periodic change in radians per second, its angular velocity or angular frequency. This is generally symbolized by ω , so we may write $\sin \omega t$. In some cases it is more meaningful to measure the rate of periodic change by its frequency, measured in cycles per second or Hertz. This is generally symbolized by f or ν ; since there are 2π radians per cycle, $\sin \omega t = \sin 2\pi f t = \sin 2\pi \nu t$, for example.*

Suppose we have an object rotating counter-clockwise at ω radians per second; then its motion in the plane of rotation can be written $z(t) = e^{i\omega t}$. This assumes that at time $t = 0$ the object is at an angle of zero, that is, at location $(1, 0)$, since $z(0) = e^{i\omega 0} = 1$. If instead it starts at the angle ϕ we simply write

$$z(t) = e^{i(\omega t + \phi)} = e^{i\phi} e^{i\omega t}.$$

The factor $e^{i\phi}$ simply advances the phase of the rotation by ϕ radians. Obviously, arbitrary phase shifts correspond to different imaginary exponentials.

A complex exponential representation of a periodic change may be advantageous even if it is not a circular motion in two dimensions. In some cases, two different aspects of the change correspond to the real and imaginary parts of a complex number. For example, in simple harmonic motion (such as a pendulum or oscillating spring), the position of the object is proportional to $\sin \omega t$ and its velocity is proportional to $\cos \omega t$. Therefore, the position and velocity can be combined into one complex number and written

$$\cos \omega t + i \sin \omega t = \operatorname{cis} \omega t = e^{i\omega t}.$$

(I've assumed that the position and velocity are measured in suitable units so that the motion in the complex plane is circular.)

Since the state of a simple harmonic oscillator is determined entirely by its displacement and velocity, the complex number $e^{i\omega t}$ corresponds to its state. In this case the complex plane is the (Poincaré) *phase space* of the oscillator, since it represents all its possible states. The curve $e^{i\theta} = 1$, that is, the unit circle, is this system's *trajectory* or *orbit* in phase space; it shows the possible sequence of states independent of time.

The complex exponential representation can yield additional insight into the structure of a periodic process. For example, in Newtonian mechanics the kinetic energy of a motion is proportional to the velocity squared, $K \propto \cos^2 \omega t$ in this case. With a suitable choice of units we can write $K = \cos^2 \omega t$. Also, in many simple harmonic systems the restoring force is proportional to the displacement ($F \propto -\sin \omega t$), so the potential energy, which is the integral of the force, is proportional to the square of the displacement, $U \propto \sin^2 \omega t$; with suitable units, $U = \sin^2 \omega t$. Hence the total energy in the system is

$$E = K + U = \cos^2 \omega t + \sin^2 \omega t = 1.$$

That is the total energy is conserved; the \cos^2 and \sin^2 terms reflect the fraction of the energy in the kinetic or potential form, respectively. That is, $K = E \cos^2 \omega t$ and $U = E \sin^2 \omega t$. Further, as $e^{i\omega t}$ rotates, we can see the energy shift back and forth between kinetic energy (proportional to the square of the real component, representing velocity) and potential energy (proportional to the square of the imaginary component, representing displacement).

Remark 4.4.5 *For the record, $K = mv^2/2$ and $U = kx^2/2$, where k is the force constant of an ideal spring, $F = -kx$ (Hookes Law). In this example, $v(t) = \omega \cos \omega t$ and $x(t) = \sin \omega t$, where the angular frequency is determined by $\omega^2 = k/m$.*

Even when there aren't two components corresponding to the real and imaginary parts, it may be advantageous to treat a sinusoidal motion as the real (or imaginary) part of a complex exponential, since it is often easier to manipulate exponentials than sines and cosines. That is, it may be convenient to treat a real signal $\cos \omega t$ as $\Re e^{i\omega t}$. As you probably know, periodic signals, such as sounds, can be broken down into sines and cosines (or into sines with phase shifts). Therefore, they can equally, and often more conveniently, be broken down into complex exponentials.

Finally, as you probably know, Fourier analysis involves breaking a periodic wave into sines and cosines, or into sinusoids at various phases. Therefore, it is not surprising that it can also be viewed as an analysis of a signal into complex exponentials with complex coefficients. We will take up these topics in Ch. 6, Fourier Analysis; here we mention them only to motivate the study of complex exponentials.

Exercise 4.4.10 Let $z = \rho e^{i\theta}$. Show that

$$ze^{i\omega t} + \bar{z}e^{-i\omega t} = 2\rho \cos(\theta + \omega t) = 2\rho \sin(\theta + \omega t + 2\pi/4).$$

Hint: Write $e^{\pm i\omega t}$ in cis form. This shows that a “conjugate pair of complex exponentials” is equivalent to an “amplitude and phase-shifted sinusoid.”

Exercise 4.4.11 Show that

$$a \cos \omega t + b \sin \omega t = ze^{i\omega t} + \bar{z}e^{-i\omega t}, \quad \text{where } z = \frac{a - ib}{2}.$$

Hint: Write the sine and cosine in their complex exponential forms. Thus a mixture of a sine and a cosine is equivalent to a conjugate pair of complex exponentials, which the preceding exercise shows to be equivalent to an amplitude and phase-shifted sinusoid.

Remark 4.4.6 These two exercises show the equivalence of: (1) a mixture of a sine and cosine of the same frequency (with parameters a and b), (2) an amplitude and phase-shifted sinusoid (with parameters ρ and θ), and (3) a conjugate pair of complex exponentials (with parameters $\Re z$ and $\Im z$). Therefore, a Fourier series for a signal can be equivalently viewed as a superposition of: (1) in-phase sines and cosines, (2) sinusoids of the same kind but differing phases, or (3) conjugate pairs of complex exponentials.

4.4.2.2 PHASORS

In this section I will discuss briefly a technique used in electrical engineering for analyzing circuits; it also has applications to understanding signal processing in the dendritic trees of neurons. Many passive electrical components, such as resistors, capacitors and inductors (coils) are linear. So also, to a first approximation, the passive conductance and membrane capacitance of dendrites is linear. Linearity means that if we know the behavior of a system L on complex exponentials (i.e. sines and cosines) of various frequencies

$L(e^{i\omega_k t})$, then we know its behavior on any periodic signal $s(t) = \sum_k c_k e^{i\omega_k t}$. This is because,

$$L[s(t)] = L\left(\sum_k c_k e^{i\omega_k t}\right) = \sum_k c_k L(e^{i\omega_k t}).$$

It turns out that resistors, capacitors and inductors have only two effects on sine waves: to attenuate them and to shift their phase; so also RLC (resistor-inductor-capacitor) circuits have only these two effects. Therefore, the effects of these circuits and their components are conveniently represented by complex numbers $Z = Ae^{i\theta}$, where A represents an amplitude change and θ represents a phase shift.

Remark 4.4.7 *The impedance of a R -ohm resistor is R ; that is, it does not affect the phase.*

Remark 4.4.8 *At a frequency of ω rad./sec., the impedance of a L -henry inductor is $i\omega L$.*

Remark 4.4.9 *At a frequency of ω rad./sec., the impedance of a C -farad capacitor is $1/i\omega C$.*

Exercise 4.4.12 *Write the impedance $1/i\omega C$ in rectangular form, that is, in the form $R + iX$.*

Remark 4.4.10 *When an impedance $Ae^{i\theta}$ is written in rectangular coordinates $R + iX$, the real part R is called a resistance and the imaginary part X is called a reactance. Therefore, any arbitrary RLC circuit, no matter how complicated has the effect of a resistance combined with a reactance. If the reactance is positive, it is called an inductive reactance; if it is negative, it is called a capacitive reactance. That is, an arbitrary RLC circuit behaves like a resistor combined with either an inductor (which causes phase leading) or a capacitor (which causes phase lagging)*

Electrical engineers often use the *phasor* notation $A\angle\theta$, read “ A angle θ ” for $Ae^{2\pi i\theta/360^\circ}$. The notation may be used for a circuit that causes an amplitude change A and a phase shift of θ degrees, or for a periodic signal (at

a certain frequency) of amplitude A and phase θ° . The notation is convenient because of the simple operation rules:

$$\begin{aligned} A\angle\theta \times B\angle\phi &= AB\angle(\theta + \phi), \\ A\angle\theta / B\angle\phi &= (A/B)\angle(\theta - \phi). \end{aligned}$$

With this notation, voltage, current and impedance (voltage divided by current) can all be treated as phasor quantities.

4.4.2.3 DIFFERENTIAL EQUATIONS

In section 4.4.1.1 we saw the relation between the MacLauren series for the exponential, sine and cosine functions; here we look at the relation between these functions from another perspective. These functions can also be defined in terms of simple differential equations. For example $f(x) = \sin x$ is the unique solution of $f''(x) = -f(x)$ with initial conditions $f(0) = 0$ and $f'(0) = 1$. Likewise, $f(x) = \cos x$ is the unique solution of the same equation but with initial conditions $f(0) = 1$, $f'(0) = 0$.

Exercise 4.4.13 Show that $f(t) = \sin \omega t$ is a solution to $f''(t) = -\omega^2 f(t)$ with initial conditions $f(0) = 0$ and $f'(0) = \omega$. (You are not asked to prove uniqueness.)

Exercise 4.4.14 Show that $f(t) = \cos \omega t$ is a solution to $f''(t) = -\omega^2 f(t)$ with initial conditions $f(0) = 1$ and $f'(0) = 0$. (You are not asked to prove uniqueness.)

These differential equations give us an alternate way of deriving Euler's formula from reasonable expectations about the meaning of $e^{i\theta}$. To see this, write

$$e^{i\theta} = E(\theta) + iF(\theta); \tag{4.4}$$

we will solve for E and F . Since $e^0 = 1$ we must have $E(0) = 1$ and $F(0) = 0$. Now differentiate Eq. 4.4 (assuming, or postulating, it differentiates normally), to get

$$ie^{i\theta} = E'(\theta) + iF'(\theta).$$

Substitute $\theta = 0$ and we discover (Show in detail!) that $E'(0) = 0$ and $F'(0) = 1$. Differentiating a second time yields

$$-e^{i\theta} = E''(\theta) + iF''(\theta).$$

Combining this and Eq. 4.4 shows

$$E''(\theta) + iF''(\theta) = -E(\theta) - iF(\theta).$$

Hence, $E''(\theta) = -E(\theta)$ with $E(0) = 1$ and $E'(0) = 0$, so we know $E = \cos$; similarly, $F''(\theta) = -F(\theta)$ with $F(0) = 0$ and $F'(0) = 1$, so $F = \sin$.

Exercise 4.4.15 Show that $f(t) = \text{cis } \omega t$ is a solution of $f'(t) = i\omega f(t)$ with initial condition $f(0) = 1$.

The ordinary (real) exponential function, $f(x) = e^x$ is the unique solution to the differential equation $f'(x) = f(x)$ with the initial condition $f(0) = 1$. Further, if $f(t) = ce^{\rho t}$, then $f'(t) = \rho f(t)$ and $f(0) = c$. This is the fundamental equation of exponential growth (or decay), which says that the increase (or decrease) in a quantity is proportional to the current quantity. The real number c is the initial quantity and the real number ρ is the rate of growth (for $\rho > 0$) or decay (for $\rho < 0$).

The foregoing is still true in the system of complex numbers: $f(z) = e^z$ is the unique solution of $f'(z) = f(z)$ with $f(0) = 1$. More generally, for $w \in \mathbb{C}$, ce^{wt} is the unique solution of $f'(t) = wf(t)$ with initial condition $f(0) = c$ (a complex number). The complex number c represents the initial state of the system, comprising a magnitude and phase. However, the meaning of the complex “rate” w requires some explanation.

Write w in rectangular form, $w = \rho + i\omega$. Then the exponential trajectory e^{wt} is seen to be a product of an exponential change in magnitude and a periodic cycle:

$$e^{wt}c = e^{(\rho+i\omega)t}c = e^{\rho t+i\omega t}c = e^{\rho t}e^{i\omega t}c$$

Thus $w = \rho + i\omega$ defines a rate of exponential change ρ and an angular frequency ω .

The parameter w in e^{wt} is sometimes called a *complex frequency*, since both its components are rates and its imaginary component is a rate of rotation. As we will see in Ch. ??, the “poles and zeros” of filters, which determine their behavior, are complex frequencies. Further, we will see that many systems can be reduced to a sum of complex exponentials, and are thus completely characterized by a set of complex frequencies.

The two components of a complex frequency can be termed its linear rate and its angular frequency. Therefore, we can say that many systems are a superposition of elementary systems, each determined by a linear rate and an angular frequency. In this sense, rectilinear and circular motion are the two primary motions from which almost all complex motions are composed.

Remark 4.4.11 *It is interesting that Aristotle, based on Plato's teachings, distinguished two fundamental motions: rectilinear and circular. This is precisely what we have in a complex frequency: if it is real, we have rectilinear motion; if it is imaginary, we have circular motion. Aristotle said that change in the "sublunary phenomena" (i.e. on the earth) are characterized by rectilinear motion (e.g. a dropped object), whereas the "celestial phenomena" (i.e. in the heavens) are characterized by circular motion (e.g. the motion of the stars). Newton's accomplishment was to show that a single law accounted for both kinds of motion (terrestrial and celestial).*

We may further subdivide the kinds of change based on the signs of the rates: If $w = \rho > 0$ we have an increase; if $w = \rho < 0$ we have a decrease; if $w = i\omega \neq 0$ we have a rotation (counterclockwise for $\omega > 0$, clockwise for $\omega < 0$). If $w = \rho + i\omega$, then we have a combination of rectilinear and circular motion (a spiral outward or inward).

It will be worthwhile to look at these possibilities from the perspective of the differential equation $f'(t) = wf(t)$ or, more compactly, $\dot{z} = wz$. As before, let the initial condition be $z(0) = c$, a complex number.

First suppose $w = \rho$ is real; then the differential equation is $\dot{z} = \rho z$, which means that the change in z is in the same direction as z (for $\rho > 0$), or in the opposite direction (for $\rho < 0$). (Note that $\delta z = \rho z$ is a little vector parallel, or antiparallel, to z ; when added to z it increases or decreases its length, but leaves its direction unchanged.) This sort of process causes z to move rectilinearly at an exponential rate: $z(t) = e^{\rho t}c$. Thus the initial state c grows or shrinks exponentially in time. We can see this clearly if we write the initial state in polar form, $c = ae^{i\phi}$; then $z(t) = ae^{\rho t} \times e^{i\phi}$; that is, the angle is independent of time.

Exercise 4.4.16 *Draw z , δz and $z + \delta z$ in this case.*

Next suppose $w = i\omega$ is imaginary; then the differential equation is $\dot{z} = i\omega z$. Recall that multiplication by i is equivalent to a counterclockwise rotation through 90° . Therefore, $\delta z = i\omega z$ can be thought of as a little vector perpendicular to the end of z ; it points in a counterclockwise direction for $\omega > 0$ and clockwise for $\omega < 0$. When added to z it causes it to rotate (counterclockwise or clockwise) without changing its length. This sort of process causes z to move circularly, $z(t) = e^{i\omega t}c$. Thus the initial state c rotates periodically with constant magnitude. Putting $c = ae^{i\phi}$ we have, $z(t) = ae^{i(\omega t + \phi)}$; the magnitude is constant a , but the rotation starts with a phase angle ϕ .

Exercise 4.4.17 Draw z , δz and $z + \delta z$.

In the general case $w = \rho + i\omega$, we have $\delta z = \rho z + i\omega z$, which is a composite of motion ρz parallel to z and motion $i\omega z$ perpendicular to the end of z . In this case we get a combination of exponential change and rotation, $z(t) = e^{\rho t} e^{i\omega t} c$. If we write the initial state in polar form, $c = ae^{i\phi}$, then

$$z(t) = ae^{\rho t} \times e^{i(\omega t + \phi)}.$$

We see the initial magnitude a changing exponentially by $e^{\rho t}$ and the initial phase angle ϕ rotating by ωt .

Exercise 4.4.18 Draw z , δz and $z + \delta z$.

4.4.3 Complex Logarithms

4.4.3.1 DEFINITION

Since the exponential of a complex number scales the real part exponentially to give the magnitude, and converts the imaginary part into a phase angle, we would expect the logarithm of a complex number to reverse this process, deriving the real part from the logarithm of the magnitude and the imaginary part from the phase angle. That is, since $\exp(x + iy) = e^x e^{iy}$, we expect $\ln(e^x e^{iy}) = x + iy$, or equivalently:

$$\ln(re^{i\theta}) = \ln r + i\theta.$$

This is basically correct, but there are some complications we must consider.

The basic problem is that the complex exponential is a periodic function; therefore it is not one-to-one, and so it does not have a unique inverse. In particular, we can see that

$$\ln(re^{i\theta}) = \ln r + i(\theta + m2\pi),$$

for $m = 0, \pm 1, \pm 2, \dots$. There are several ways we can deal with this.

First we may choose to restrict the angle to lie in a particular range, such as $[0, 2\pi)$ or $[-\pi, \pi)$. Thus we may talk of the *principal* value of the logarithm, as we talk of the principal value of the arcsine, arccosine, etc. (Often the principal value of the logarithm is written “Ln,” just as the principal value of the arcsine is written “Arcsin,” etc.) This convention has the disadvantage

that the identity $\ln(e^{i\theta}) = i\theta$ does not hold unless θ is restricted to the chosen range.

Second, we may simply accept that the logarithm is a *multiple-valued function*; they are not unknown in mathematics and its applications, for example we have $f(x) = \pm\sqrt{x}$ and $f(x) = \sin^{-1}x$. To use multiple-valued functions without encountering contradictions, it's necessary to restrict attention to a particular value, as determined by context, stipulation, or constraints of the application. In this case we can write $\ln(e^{i\theta}) = i\theta$, provided it's understood that the appropriate value of the logarithm must be used.

There is a third, more formal but nevertheless interesting solution, which will be discussed in Section 4.4.3.3.

Exercise 4.4.19 Show $e^{\ln z} = z$, for any of these interpretations of the complex logarithm.

Exercise 4.4.20 Show

$$\ln(zw) = \ln z + \ln w \pmod{2\pi}.$$

4.4.3.2 GEOMETRICAL INTERPRETATIONS

For any integral values of m , observe that

$$e^{x+iy} = e^{x+i(y\pm 2\pi m)}.$$

Hence the values of e^{x+iy} repeat at vertical intervals of 2π . Therefore, if we restrict attention to any infinitely wide “band” of height 2π , the logarithm will be single valued. These bands (which need not have their boundaries at multiples of π or other “reasonable” places) are called *branches* of the complex logarithm. Therefore, if we restrict attention to z in a single branch, we will have $\ln(e^z) = z$.

It will strengthen our intuitive understanding of the complex exponential and logarithm to look at how they transform various subsets of the complex plane.

First, observe that the exponential function maps a branch of the logarithm onto the entire complex plane except for the origin (since $e^z = 0$ has no solution, and so $\ln 0$ is undefined). Conversely, the logarithm maps the complex plane (minus the origin) onto its chosen branch.

4.4.3.3 RIEMANN SURFACES

forthcoming

4.4.3.4 COMPLEX POWERS

With the complex exponential and logarithm we can define arbitrary powers of complex numbers.

Definition 4.4.1 (Complex Powers) *If z and w are complex numbers ($z \neq 0$), then z^w is defined $z^w = e^{w \ln z}$.*

Remark 4.4.12 *The complex power is multiple-valued because it is defined in terms of the complex logarithm. Therefore it's necessary to restrict attention, by context or stipulation, to a particular branch of the function (either the logarithm or the power).*

Proposition 4.4.6 *The power z^w is single valued if and only if w is an integer.*

Proposition 4.4.7 *If $w = p/q$ is a rational number in lowest terms, then z^w has exactly q values, namely the q principal q -roots of z^p .*

Proposition 4.4.8 *If w is irrational real or complex, then z^w has an infinity of values differing by $e^{2\pi m w i}$.*

Remark 4.4.13 *We have already seen (Prop. 4.4.5) that a complex number has n principal n -th roots. This is consistent with the definition of the n -th root in terms of complex powers, restricted to a branch of the logarithm:*

$$\sqrt[n]{z} = z^{1/n} = e^{(\ln z)/n}.$$

If $z = re^{i\theta}$, then, for $m = 0, 1, \dots, n - 1$,

$$\sqrt[n]{z} = \sqrt[n]{r} e^{i(\theta+2\pi m)/n} = \sqrt[n]{r} \operatorname{cis} \left(\frac{\theta}{n} + 2\pi \frac{m}{n} \right).$$

Notice that the roots have equal magnitude and angles evenly distributed around a cycle at angles $2\pi\theta/n$ and beginning at θ/n .

Exercise 4.4.21 *Plot the principal values of $\sqrt[6]{-64i}$.*

4.5 Hyperbolic Geometry

4.5.1 Hyperbolic Functions

4.5.1.1 HYPERBOLIC ANGLES

We have seen an intimate connection between the complex exponential and the (circular) trigonometric functions (sine, cosine, etc.); in this section we will explore an equally intimate connection with the *hyperbolic* trigonometric functions. First, we review the derivation of the circular functions.

Draw a circle $x^2 + y^2 = r^2$ and draw a radius at angle θ above the X-axis (Fig. 4.1, p. 42). Drop perpendiculars x and y to the X- and Y-axes. We know from trigonometry that

$$\cos \theta = x/r, \quad \sin \theta = y/r, \quad \tan \theta = y/x.$$

Finally, we have seen (Rem. 4.3.1) that the radian measure of an angle is twice the ratio of the included area to the radius squared.

Now we will undertake a similar construction, but based on the hyperbola rather than the circle; our first task is to define an appropriate measure of angles, in *hyperbolic radians*. Consider the (equilateral) hyperbola $x^2 - y^2 = r^2$, which has its arcs lying within the left and right half planes. Draw a ray from the origin at an angle of less than 45° from the X-axis, so that it intersects the right-hand half of the hyperbola. (We will deal later with angles greater than 45° .) As we did with the circular angle, we measure the hyperbolic angle by the area bounded by the curve between the radius and the X-axis; in particular the hyperbolic radian measure κ will be the ratio of twice the area to the square radius.

Proposition 4.5.1 *Suppose a radius intersects the hyperbola $x^2 - y^2 = r^2$ at the point (x, y) . Then the hyperbolic radian measure of the angle of the ray is $\ln \left(\frac{x+y}{r} \right)$.*

Exercise 4.5.1 *To determine this, first show that the area B under the curve of the hyperbola out to x is*

$$B = \frac{xy}{2} - \frac{r^2}{2} \ln \left(\frac{x+y}{r} \right).$$

Exercise 4.5.2 Show that the required area

$$A = \frac{r^2}{2} \ln \left(\frac{x+y}{r} \right)$$

by subtracting B from the area of the triangle (x, y, r) .

It then follows that the angle in hyperbolic radians is $\kappa = 2A/r^2 = \ln \left(\frac{x+y}{r} \right)$.

Now, just as for the circular functions, we will define the hyperbolic functions in terms the ratios x/r , y/r and y/x . For simplicity, use a unit circle, so $r = 1$ and $\kappa = \ln(x+y)$. Then,

$$\cosh \kappa = x, \quad \sinh \kappa = y, \quad \tanh \kappa = y/x.$$

We therefore have two equations in two unknowns:

$$1 = x^2 - y^2, \tag{4.5}$$

$$\kappa = \ln(x+y). \tag{4.6}$$

Exercise 4.5.3 Show that the solutions are

$$x = \frac{e^\kappa + e^{-\kappa}}{2}, \quad y = \frac{e^\kappa - e^{-\kappa}}{2}$$

(Thus we can “solve triangles” with the hyperbolic functions as well as with the circular, except that we don’t have protractors for measuring hyperbolic angles!) We have proved:

Proposition 4.5.2

$$\cosh \kappa = \frac{e^\kappa + e^{-\kappa}}{2}, \tag{4.7}$$

$$\sinh \kappa = \frac{e^\kappa - e^{-\kappa}}{2}, \tag{4.8}$$

$$\tanh \kappa = \frac{e^\kappa - e^{-\kappa}}{e^\kappa + e^{-\kappa}}. \tag{4.9}$$

These formulas are similar to the corresponding Eq. 4.3 for the circular sine and cosine, to which they should be carefully compared. The preceding derivation only applies to angles in the first *octant* ($0^\circ - 45^\circ$). However, by allowing x and y to be negative, it is automatically extended to all angles

within 45° of the X-axis. It is extended to angles within 45° of the Y-axis by, in effect, duplicating the above derivation with the hyperbola $y^2 - x^2 = 1$, which has its arc in the upper and lower halfplanes. That is, interchange x and y .

The equations Eq. 4.9 are true in all octants, and in fact are often stipulated as the definition of the functions. In particular, although we have justified these equations on the basis of a real-valued hyperbolic angle, we can use them to define the hyperbolic functions for any complex argument (just as can be done with the exponential formulas for the circular functions, Eq. 4.3).

4.5.1.2 HYPERBOLIC FUNCTIONS

Exercise 4.5.4 *Explore and discuss the domain and range of the hyperbolic sine, cosine and tangent over the reals; sketch their shapes (don't plot by computer; use the hyperbolic law of triangles).*

Exercise 4.5.5 *Prove the following symmetry properties: First, the hyperbolic cosine (like the circular cosine) is an even function, that is, $\cosh(-\kappa) = \cosh \kappa$. Second, the hyperbolic sine (like the circular sine) is an odd function, that is, $\sinh(-\kappa) = -\sinh \kappa$. As a consequence, the hyperbolic tangent (like the circular tangent) is also odd, $\tanh(-\kappa) = -\tanh \kappa$.*

Exercise 4.5.6 *Prove $\cosh^2 \kappa - \sinh^2 \kappa = 1$. What is the corresponding property of the circular functions?*

Exercise 4.5.7 *Prove*

$$\tanh(\kappa + \lambda) = \frac{\tanh \kappa + \tanh \lambda}{1 + \tanh \kappa \tanh \lambda}.$$

What is the corresponding property of circular functions?

Exercise 4.5.8 *Prove $\operatorname{sech}^2 \kappa = 1 - \tanh^2 \kappa$, where $\operatorname{sech} \kappa = 1/\cosh \kappa$. What is the corresponding circular property?*

Exercise 4.5.9 *Prove $\operatorname{csch}^2 \kappa = \coth^2 \kappa - 1$, where $\operatorname{csch} \kappa = 1/\sinh \kappa$ and $\coth \kappa = 1/\tanh \kappa$. What is the corresponding circular property.*

4.5.2 Special Relativity Theory

In this section we will consider briefly special relativity since (1) it illustrates the use of a mixed real/imaginary coordinate system, (2) it makes use of hyperbolic geometry and (3) it suggests ways of treating space and time together, which has relevance to wavelet processing and spatiotemporal information processing in the brain (see my report, “Gabor Representations of Spatiotemporal Visual Images”).

4.5.2.1 The Fundamental Invariance

Special relativity, which deals with the geometry of spacetime, is easier to understand by comparison with the geometry of ordinary space. First, notice that in ordinary space certain properties are dependent on the coordinate system we use, whereas others are not. For example, the x and y coordinates of a point (or vector) depend on the choice of axes, since they are projections of that point (or vector) onto the axes.

Exercise 4.5.10 *Diagram this situation.*

On the other hand, the distance between points (or the length of a vector) is independent of the coordinate system. Thus, if (x, y) and (x', y') are the coordinates of the same vector in two different coordinate systems, we can assert the *invariant* $x^2 + y^2 = x'^2 + y'^2$. We say that length is *invariant under a transformation of coordinates*.

In ordinary space there are two different ways we can measure the inclination of a line. If we measure it by slope, then the measure depends on the coordinate system, since the slope is y/x , which quantities are not invariant.

Exercise 4.5.11 *Diagram this situation.*

Further, slopes are not additive: if m and m' are the slopes of the same line in two different coordinate system, and μ is the slope of the primed system with respect to the unprimed, we might expect $m = \mu + m'$, but this is not the case.

Exercise 4.5.12 *In fact, the law of combination is:*

$$m = \frac{\mu + m'}{1 - \mu m'} \text{ or } m' = \frac{m - \mu}{1 + m\mu}.$$

Does this look familiar? Derive it by trigonometry. Notice also that $m \approx \mu + m'$ if $\mu \approx 0$; that is, if the coordinate systems deviate only slightly from each other, then slopes are approximately additive.

On the other hand, ordinary (circular) angles *are* additive, since they are invariant under coordinate transformation. Therefore, if a line has angles a and a' with respect to the X-axes of two coordinate systems, and α is the angle of the primed X'-axis to the X-axis, then $a' = \alpha + a$.

Exercise 4.5.13 *Diagram this situation.*

In relativity, events are located in four-dimensional spacetime; they have coordinates (x, y, z, t) . Now we will make several convenient assumptions. First, since relativistic effects occur in the direction of motion, and not perpendicular to the direction of motion, we will restrict our attention to a single space axis s , oriented in the direction of motion; thus spacetime coordinates will take the form (s, t) . This will make spacetime geometry easier to visualize and draw, and will simplify the mathematical notation.

Second, since time is an axis like the other three, we will measure them all in the same units, meters, which will simplify the formulas. (Imagine the needless complexity that would result from measuring north-south distances in miles and east-west distance in kilometers.) This raises the question of how to convert seconds to meters; what is the conversion factor? It turns out that it is the speed of light, $c \approx 3 \times 10^8$ m/s. We will see that this is not an arbitrary choice, but is in fact fundamental to the fabric of spacetime. If T is time in seconds and t is time in meters, then $t = cT$.

Finally, in accord with the measurement of time in meters, velocity becomes a pure number (meters/meter), for which relativity theory uses the symbol β . If $V = s/T$ is time in ordinary units, we can see that $\beta = s/t = s/(cT) = V/c$. Thus β can also be interpreted as velocity relative to the speed of light. By “natural units” I will mean the measurement of time in meters and velocity as a pure number.

Remark 4.5.1 *In our lives we range cover a vast distance along the time axis compared to our range on the spatial axes. Since there are about $\pi \times 10^7$ seconds in a year, we go about $\pi c \times 10^7 \approx 3\pi \times 10^{15} \approx 10^{16}$ meters in a year (i.e. one light-year). In our lifetimes we cover about 7×10^{17} meters on the time axis (that is, about 70 light-years, something between the distances to Aldebaràn and to Regulus). In the same amount of time the solar system*

moves about 7×10^{14} m. relative to the cosmic background radiation (since our motion with respect to it is about 3×10^5 m/s in the direction of Virgo). It is this large discrepancy, a ratio of 10^3 , between our mobility in time and space that leads to the undetectability of relativistic effects under ordinary conditions. Therefore, our average velocity, in natural units, is $\beta \approx 10^{-3}$ (i.e. speed relative to background radiation divided by speed of light).

Although we measure it in spatial units, the time axis is not just another space axis; indeed we may say that time is imaginary with respect to the spatial axes, since one consequence of the relativity postulates is that the fundamental invariant is $s^2 + (it)^2 = s^2 - t^2$.

Remark 4.5.2 *This invariant follows from the first postulate of relativity theory, which says that the velocity of light in a vacuum is the same in all reference frames. To see this suppose that a reflective object is moving to the right past us at a velocity β . When it is directly opposite us at a distance of r , suppose that it is struck by light from source distance s to our left. The distance, as measured in our reference frame, traveled by the light is $d = \sqrt{r^2 + s^2}$, so is the time, in our frame, that it took to travel it (since the speed of light = 1 in natural units): $t = \sqrt{r^2 + s^2}$. Within the reference frame of the reflective object, however, the source appears to be a distance s' to the left, so the distance the light travelled is $d' = \sqrt{r^2 + s'^2}$; likewise the time is $t' = \sqrt{r^2 + s'^2}$. Now observe:*

$$\begin{aligned} t^2 - s^2 &= (r^2 + s^2) - s^2 = r^2, \\ t'^2 - s'^2 &= (r^2 + s'^2) - s'^2 = r^2. \end{aligned}$$

We see that $t^2 - s^2 = t'^2 - s'^2$.

This quantity, $s^2 - t^2$, which is invariant under a change between reference frames in relative motion, is called the *spacetime interval* between two events; it is analogous to the Euclidean distance $x^2 + y^2$, which is invariant under change of the spatial coordinate system.

Spacetime intervals can be classified according to whether $s^2 - t^2$ is positive, negative or zero. If it is positive, the interval is called *space-like* and the *proper distance* σ is defined $\sigma^2 = s^2 - t^2$. If it is negative, the interval is called *time-like* and the *proper time* τ is defined $\tau^2 = t^2 - s^2$. If the interval is zero, it is called *light-like*. We will see that time-like intervals can be crossed by *subluminary signals* (signals travelling less than the speed of

light), and light-like intervals can be crossed only by things travelling at the speed of light. Space-like intervals could be crossed only by things travelling faster than light, which, so far as physics has been able to establish, do not exist. Therefore, space-like intervals are causally independent; causality can operate only across time-like and light-like intervals (i.e., those for which $t^2 - s^2 \geq 0$). Therefore, we will restrict our attention to this case (without loss of generality, however).

Exercise 4.5.14 *The invariant $t^2 - s^2 = \text{constant}$ should remind you of an identity that you have seen recently. What does it suggest about the formal relation of the quantities t and s ?*

4.5.2.2 MEANING OF THE HYPERBOLIC ANGLE

The invariance of spacetime interval means that, for a given pair of events, $\tau^2 = t^2 - s^2$ is constant, no matter what their distance separation s and time separation t in a given reference frame. That is, the possible s and t measurements in various reference frames is constrained by $\tau^2 = t^2 - s^2$. This means that the possible (s, t) pairs lie on an equilateral hyperbola, whose arcs lie in the positive and negative t halfplanes. From this we see, by the law of triangles for hyperbolas, that (for some hyperbolic angle κ):

$$\begin{aligned} s &= \tau \sinh \kappa, \\ t &= \tau \cosh \kappa. \end{aligned}$$

From these, the invariance of the spacetime interval follows from the properties of the hyperbolic functions:

$$t^2 - s^2 = \tau^2(\cosh^2 \kappa - \sinh^2 \kappa) = \tau^2.$$

Now we must consider the meaning of the hyperbolic angle κ . Observe that the velocity can be written in terms of hyperbolic functions:

$$\beta = \frac{s}{t} = \frac{\tau \sinh \kappa}{\tau \cosh \kappa} = \tanh \kappa.$$

Thus, $\kappa = \text{arctanh } \beta$, and so it is called the *velocity parameter*; we may say that the velocity parameter is the hyperbolic arctangent of the velocity (in natural units). The significance of the velocity parameter is that it measures the hyperbolic angle between the time axes in the two reference frames. Just

as a slope m is related to a corresponding angle θ by the circular tangent, $m = \tan \theta$, so a velocity is related to a corresponding velocity parameter by the hyperbolic tangent, $\beta = \tanh \kappa$. The appearance of the circular functions in spatial rotations is a consequence of the isotropy of x and y ; the appearance of the hyperbolic functions in spacetime transformations is a consequence of the anisotropy of s and it (i.e., time is imaginary with respect to space).

4.5.2.3 Comparison of Lorentz & Galilean Transforms

forthcoming

4.6 References

Sources for this chapter include Boas (1983, ch. 2), Bobrow (1981, §8.3), Marsden (1973, ch. 1), and Taylor & Wheeler (1966, pp. 22–6, 39–59, 92–4).

Chapter 5

Hilbert Spaces

This chapter presents the most important concepts from the theory of Hilbert spaces, which provides the principal mathematical background for field computation. Hilbert spaces extend the familiar ideas of finite-dimensional vectors and matrices to the infinite-dimensional case and, from another perspective, to continuous quantities defined over continuous spaces, that is, to fields. As for finite-dimensional spaces, the notion of an inner product is fundamental, and so we begin with the general notion of an inner product and associated concepts of orthogonality and bases. We state or prove several important results from Hilbert spaces that are useful for implementing field computations. The material in this chapter is also essential in the application of field computation to quantum computation.

5.1 Inner product spaces

5.1.1 Inner products

Definition 5.1.1 (Inner product) *An inner product is a binary operation on a real or complex linear space, $\cdot : X \times X \rightarrow \mathbb{C}$, satisfying (for $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$ and scalar $a \in \mathbb{C}$):*

(a) **positive definite:**

$$\mathbf{x} \cdot \mathbf{x} > 0, \quad \text{if } \mathbf{x} \neq \mathbf{0}, \quad (5.1)$$

$$\mathbf{x} \cdot \mathbf{x} = 0, \quad \text{if } \mathbf{x} = \mathbf{0}. \quad (5.2)$$

(b) conjugate symmetry:

$$\mathbf{x} \cdot \mathbf{y} = \overline{\mathbf{y} \cdot \mathbf{x}}. \quad (5.3)$$

(c) linearity in first argument:

$$(a\mathbf{x}) \cdot \mathbf{y} = a(\mathbf{x} \cdot \mathbf{y}), \quad (5.4)$$

$$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = (\mathbf{x} \cdot \mathbf{z}) + (\mathbf{y} \cdot \mathbf{z}). \quad (5.5)$$

Notation 5.1.1 The inner product $\mathbf{x} \cdot \mathbf{y}$ is often written $\langle \mathbf{x}, \mathbf{y} \rangle$, especially when it is defined over a function space.

Remark 5.1.1 If the linear space is real, then properties (b) and (c) are replaced by symmetry (commutativity) and bilinearity (linearity in both arguments):

(b') symmetry:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x} \quad (5.6)$$

(c') bilinearity:

$$(a\mathbf{x} + b\mathbf{y}) \cdot \mathbf{z} = a(\mathbf{x} \cdot \mathbf{z}) + b(\mathbf{y} \cdot \mathbf{z}) \quad (5.7)$$

Exercise 5.1.1 Show that for complex vector \mathbf{z} , $\langle -\mathbf{z}, \mathbf{z} \rangle = \langle \mathbf{z}, -\mathbf{z} \rangle$. (By “vector” we mean an element of any linear space.)

Exercise 5.1.2 Show $\mathbf{x} \cdot (b\mathbf{y}) = \bar{b}(\mathbf{x} \cdot \mathbf{y})$ for complex vectors. (This is called antilinearity in the second argument; see next.)

Definition 5.1.2 (Antilinear) A function f on a linear space L over the complex numbers is called antilinear if for all $\mathbf{x} \in L, z \in \mathbb{C}$, $f(z\mathbf{x}) = \bar{z}f(\mathbf{x})$. This terminology is extended in the obvious way to functions of more than one argument, such as the inner product.

Definition 5.1.3 (Sesquilinear) A binary operation is called sesquilinear if it is linear in one argument and antilinear in the other. (The prefix “sesqui-” means “one and a half.”) Therefore, the complex inner product is sesquilinear.

Exercise 5.1.3 Show that $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = (\mathbf{x} \cdot \mathbf{z}) + (\mathbf{x} \cdot \mathbf{y})$ for complex vectors.

Definition 5.1.4 (Inner-product space) An inner-product space is a real or complex linear space with an inner product.

Exercise 5.1.4 If \mathbf{x} and \mathbf{y} are n -dimensional real vectors, show that

$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^n x_k y_k$$

is an inner product. This demonstrates that \mathcal{E}^n is an inner-product space.

Exercise 5.1.5 If U and V are real column vectors, show that $U^T V$ is an inner product.

Exercise 5.1.6 If \mathbf{x} and \mathbf{y} are n -dimensional complex vectors, show that

$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^n x_k \overline{y_k}$$

is an inner product.

Definition 5.1.5 The space ℓ_2 is the inner-product space of square-summable complex sequences, that is, sequences $\mathbf{x} = (x_1, x_2, \dots)$ for which

$$\sum_{k=1}^{\infty} |x_k|^2 < \infty,$$

with the inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^{\infty} x_k \overline{y_k}.$$

Exercise 5.1.7 Show that ℓ_2 is in fact an inner-product space.

In physics, especially quantum mechanics, it is customary to use a different notation for complex inner products, which is very useful.

Definition 5.1.6 (Dirac's bracket notation)

$$\langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle = \mathbf{y} \cdot \mathbf{x} \quad (5.8)$$

Notice that the order of the arguments is reversed. As a consequence,

$$\langle \mathbf{x} | \mathbf{y} \rangle = \overline{\langle \mathbf{x}, \mathbf{y} \rangle}. \quad (5.9)$$

Exercise 5.1.8 Show that $\langle a\mathbf{x} \mid \mathbf{y} \rangle = \bar{a}\langle \mathbf{x} \mid \mathbf{y} \rangle$ and $\langle \mathbf{x} + \mathbf{y} \mid \mathbf{z} \rangle = \langle \mathbf{x} \mid \mathbf{z} \rangle + \langle \mathbf{y} \mid \mathbf{z} \rangle$.

Definition 5.1.7 (adjoint matrix) The adjoint or conjugate transpose M^\dagger of a complex matrix M is obtained by taking the complex conjugate of every element of its transpose. That is,

$$(M^\dagger)_{ij} = \overline{M_{ji}}. \quad (5.10)$$

If we let the complex conjugate apply element-wise to an entire matrix, $(\overline{M})_{ij} = \overline{M_{ij}}$, then $M^\dagger = \overline{M^T} = \overline{M}^T$. The conjugate transpose is also called the Hermitian transpose.

Remark 5.1.2 The matrix-definition of the adjoint (Def. 5.1.7) can be extended to the infinite dimensional case, i.e., to linear operators on a Hilbert space (see Sec. 5.2.7, p. 89).

Exercise 5.1.9 Show that the adjoint operation is antilinear: $(aL + bM)^\dagger = \bar{a}L^\dagger + \bar{b}M^\dagger$.

Exercise 5.1.10 If U and V are complex column vectors, show that $\langle U \mid V \rangle = U^\dagger V$ is an inner product.

Notation 5.1.2 (Dirac's bra and ket notations) In quantum mechanics, the notation $|V\rangle$ is often used to denote a column vector, and $\langle V|$ is used to represent its adjoint, $\langle V| = |V\rangle^\dagger$.¹ In effect, the $| \rangle$ frame reminds us that we are dealing with a column vector, and $\langle |$ reminds us that we are dealing with a row vector that is the adjoint of the corresponding $| \rangle$. More generally, $\langle V|$ is called the dual vector of $|V\rangle$. The notation $\langle V|$ is called a bra and is read "bra V ," and $|V\rangle$ is called a ket and read " V ket." (The reasons why will be apparent from the following exercise.)

Exercise 5.1.11 If U and V are complex column vectors, show that $\langle U \mid V \rangle = \langle U| \mid V \rangle$ (the matrix product of $\langle U|$ and $|V\rangle$).

¹More generally, $\langle V|$ represents the co-vector of $|V\rangle$ in the continuous dual space of the vector space from which $|V\rangle$ is drawn. See Sec. 5.2.5.2 (p. 86).

Exercise 5.1.12 For real-valued $f, g \in \mathcal{C}[a, b]$, show that

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

is an inner product.

Exercise 5.1.13 Show that, for complex-valued functions f and g ,

$$\langle f | g \rangle = \int_a^b \overline{f(x)}g(x)dx$$

is an inner product. (The extension of the bra-ket notation to functions is discussed in Sec. 5.2.5.2.)

Remark 5.1.3 Exercise 4.2.3 (p. 44), which you thought about earlier (Right?), reveals the reason that complex valued inner-products are defined in terms of the complex conjugate, $\mathbf{u} \cdot \mathbf{v} = \sum u_k \bar{v}_k$, $\langle u, u \rangle = \int u(t) \overline{v(t)} dt$, etc. Otherwise, the inner-product norm ($\|u\| = \sqrt{\langle u, u \rangle}$) would not always be real.

Definition 5.1.8 The notation $\langle \mathbf{x} | L | \mathbf{y} \rangle$ is defined by the following equivalent products:

$$\langle \mathbf{x} | L | \mathbf{y} \rangle = \langle \mathbf{x} | (L|\mathbf{y}) \rangle = (\langle \mathbf{x} | L) | \mathbf{y} \rangle. \quad (5.11)$$

Exercise 5.1.14 For a complex matrix M and complex column vectors \mathbf{x}, \mathbf{y} show that:

$$\langle \mathbf{x} | M | \mathbf{y} \rangle = \mathbf{x}^\dagger M \mathbf{y} = \sum_{ij} \bar{x}_i M_{ij} y_j. \quad (5.12)$$

Exercise 5.1.15 For a complex matrix M and complex vectors \mathbf{x}, \mathbf{y} show that:

$$\langle \mathbf{x} | M | \mathbf{y} \rangle^* = \langle \mathbf{x}^* | M^* | \mathbf{y}^* \rangle = \langle \mathbf{y} | M^\dagger | \mathbf{x} \rangle. \quad (5.13)$$

5.1.1.1 SCHWARZ INEQUALITY

Proposition 5.1.1 (Schwarz inequality) An inner product satisfies

$$|\mathbf{x} \cdot \mathbf{y}| \leq \sqrt{\mathbf{x} \cdot \mathbf{x}} \sqrt{\mathbf{y} \cdot \mathbf{y}}. \quad (5.14)$$

That is, $|\mathbf{x} \cdot \mathbf{y}|^2 \leq |\mathbf{x} \cdot \mathbf{x}| |\mathbf{y} \cdot \mathbf{y}|$. This is also known as the Cauchy-Schwarz-Buniakowski (C.S.B) inequality or, more briefly, as the Schwarz inequality.

Exercise 5.1.16 *Prove the Schwarz inequality Hint: For $\mathbf{x} \cdot \mathbf{y} \neq 0$, let $a = (\mathbf{x} \cdot \mathbf{y})/|\mathbf{x} \cdot \mathbf{y}|$ and let $b \in \mathbb{R}$; then apply bilinearity to get a quadratic equation with one real root. The result follows from the equations nonpositive discriminant.*

Proposition 5.1.2 (Schwarz equality) *The Schwarz inequality becomes an equality if and only if the vectors are linearly dependent:*

$$|\mathbf{x} \cdot \mathbf{y}|^2 = |\mathbf{x} \cdot \mathbf{x}| |\mathbf{y} \cdot \mathbf{y}|.$$

Exercise 5.1.17 *Prove the Schwarz equality.*

5.1.2 Inner product norm

Definition 5.1.9 (Inner-product norm) *The inner-product norm is defined $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$ or $\|\mathbf{x}\|^2 = \mathbf{x} \cdot \mathbf{x}$.*

Exercise 5.1.18 *Show that the inner-product norm is in fact a norm. Hint: For the triangle inequality, expand both sides of $\|\mathbf{x} + \mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ in terms of the inner product, and then apply the Schwarz inequality.*

Remark 5.1.4 *Thus an inner-product space is a normed linear space under the inner-product norm. Hence it is also a metric space (under the norm metric).*

Remark 5.1.5 (Schwartz inequality) *In terms of the inner-product norm, the Schwarz inequality is*

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|. \tag{5.15}$$

Remark 5.1.6 *The space ℓ_2 can be considered the infinite-dimensional analog of \mathcal{E}^n , since the infinite-dimensional vectors in ℓ_2 have finite magnitude ($\|\xi\| < \infty$ for all $\xi \in \ell_2$).*

5.1.3 More on complex inner products

In this section I will try and build some intuition for complex-valued inner-products defined over complex linear spaces. We will suppose that we have some generalized vectors \mathbf{z} and \mathbf{w} with complex-valued components z_x and

w_x . These might be finite-dimensional vectors, $\mathbf{z} = (z_1, \dots, z_n)$ and $\mathbf{w} = (w_1, \dots, w_n)$ (over the domain $\Omega = \{1, \dots, n\}$), or they might be infinite dimensional vectors or continuous functions, $z_x = z(x)$ and $w_x = w(x)$ over some domain Ω . In particular, the vectors might be two complex-valued signals evolving in time, $z_t = z(t)$ and $w_t = w(t)$. Therefore, in the following discussion I will refer to the generalized vectors \mathbf{z} and \mathbf{w} as *signals*.

Correspondingly, the inner product might be a discrete sum $\langle \mathbf{z}, \mathbf{w} \rangle = \sum_{x \in \Omega} z_x \overline{w_x}$ or an integral $\langle \mathbf{z}, \mathbf{w} \rangle = \int_{\Omega} z_x \overline{w_x} dx$. I will write the inner product as an integral, since it includes the summation as a special case (when the integral is defined by Lebesgue integration).

To understand the effect of the complex inner product, first write the signal elements in polar form: $z_x = r_x e^{i\theta_x}$, $w_x = s_x e^{i\phi_x}$. Therefore the inner product is:

$$\begin{aligned} \langle \mathbf{z}, \mathbf{w} \rangle &= \int_{\Omega} r_x e^{i\theta_x} \overline{s_x e^{i\phi_x}} dx, \\ &= \int_{\Omega} r_x e^{i\theta_x} s_x e^{-i\phi_x} dx, \\ &= \int_{\Omega} r_x s_x e^{i(\theta_x - \phi_x)} dx. \end{aligned}$$

This will be clearer if we write it in terms of the cis function:

$$\langle \mathbf{z}, \mathbf{w} \rangle = \int_{\Omega} r_x s_x \text{cis}(\theta_x - \phi_x) dx.$$

Now notice that if α is any phase angle, then $\text{cis } \theta = e^{i\alpha}$ is the corresponding *phase vector*, that is, the phase angle expressed as a vector (complex number) on the unit circle. Therefore, in the inner product the expression $\text{cis}(\theta_x - \phi_x)$ is a phase vector representing the difference in phase between the signals at x (e.g. at time x). Thus $\langle \mathbf{z}, \mathbf{w} \rangle$ is like an average phase vector, in which the phase differences at a point, $\text{cis}(\theta_x - \phi_x)$, are weighted by the joint magnitudes of the signals at that point, $r_x s_x$. Therefore phase differences are discounted where the signals are weak, but accounted more significance where they are strong.

The meaning of the inner product may become clearer by considering several examples. Consider the inner product of a signal with itself, which we know must be the square magnitude of the signal, $\langle \mathbf{z}, \mathbf{z} \rangle = \|\mathbf{z}\|^2 = \|\mathbf{r}\|^2$,

which is a real number. Observe:

$$\begin{aligned}\langle \mathbf{z}, \mathbf{z} \rangle &= \int_{\Omega} z_x^2 \operatorname{cis}(\theta_x - \theta_x) dx, \\ &= \|\mathbf{z}\|^2 \operatorname{cis} 0.\end{aligned}$$

so the average phase vector is proportional to $1 = \operatorname{cis} 0$; that is, there is no phase difference between the signals. Further, we can see that if $\mathbf{z} \approx \mathbf{w}$, with only small phase differences, then $\langle \mathbf{z}, \mathbf{w} \rangle \approx \langle \mathbf{r}, \mathbf{s} \rangle \operatorname{cis} \epsilon$, where $\operatorname{cis} \epsilon \approx 1$ is a phase vector representing little if any average phase difference.

For our second example, suppose $\mathbf{w} = -\mathbf{z}$, that is, the signals are 180° out of phase. Observe,

$$w_x = -z_x = -1 \times r_x e^{i\theta_x} = e^{i\pi} \times r_x e^{i\theta_x} = r_x e^{i(\theta_x + \pi)}.$$

Therefore the inner product is

$$\begin{aligned}\langle \mathbf{z}, \mathbf{w} \rangle &= \int_{\Omega} r_x^2 \operatorname{cis}[\theta_x - (\theta_x + \pi)] dx, \\ &= \operatorname{cis}(-\pi) \|\mathbf{r}\|^2.\end{aligned}$$

That is, the average phase vector is proportional to $\operatorname{cis}(-\pi) = -1$, representing a phase difference of $-\pi = -180^\circ$ between \mathbf{z} and \mathbf{w} . (Of course $\operatorname{cis}(-\pi) = \operatorname{cis} \pi$, that is, a phase difference of -180° is the same as a phase difference of 180° ; this is not the case for other angles.)

Exercise 5.1.19 Recall (Sec. 5.1.1) that for complex inner products commutativity is replaced by $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$. Based on the interpretation of inner product in terms of phase vectors, explain why this must be so.

Similarly, we can see that if $\mathbf{w} \approx -\mathbf{z}$, then $\langle \mathbf{z}, \mathbf{w} \rangle \approx \langle \mathbf{r}, \mathbf{s} \rangle \operatorname{cis}(-\pi \pm \epsilon)$, that is, there is an average phase shift of approximately 180° .

In summary, the inner product between two complex-valued signals is proportional to the average vector of the phase difference between the signals, weighted by their joint signal strength. Thus the inner product tends to be positive real to the extent the signals are in phase, negative real to the extent they are 180° out of phase (i.e. opposite in sign), and complex to the extent that they have a more complex average phase relation.

Remark 5.1.7 We mention in passing that the quantum mechanical wavefunction $\Psi(\mathbf{x}, t)$ does not give the probability of a particle being in a given place \mathbf{x} at a given time t ; rather, since the wavefunction is complex-valued, the probability is given by the square of the wavefunction: $|\Psi(\mathbf{x}, t)|^2 = \overline{\Psi(\mathbf{x}, t)} \Psi(\mathbf{x}, t)$, or $\|\Psi\|^2 = \langle \Psi | \Psi \rangle$.

5.1.4 Orthonormality

In the following, “vector” refers to any element of a real or complex inner-product space.

Definition 5.1.10 (Orthogonal) *Two nonzero vectors are orthogonal if and only if their inner product is zero, $\mathbf{x} \cdot \mathbf{y} = 0$.*

Definition 5.1.11 (Orthonormality) *A set of vectors is orthonormal (abbreviated ON) if they are normal and pairwise orthogonal. That is $\{\mathbf{x}_1, \mathbf{x}_2, \dots\}$ is ON if $\|\mathbf{x}_k\| = 1$ and $\mathbf{x}_j \cdot \mathbf{x}_k = 0$ for $j \neq k$.*

Definition 5.1.12 (Kronecker delta) *The Kronecker delta is defined:*

$$\begin{aligned}\delta_{jj} &= 1, \\ \delta_{jk} &= 0, \quad \text{if } j \neq k.\end{aligned}$$

Remark 5.1.8 *The Kronecker delta can be used like a conditional expression in a programming language. For example, $\delta_{jk}A + (1 - \delta_{jk})B = A$ if $j = k$ and $= B$ if $j \neq k$.*

Proposition 5.1.3 *A set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots\}$ is orthonormal if and only if $\mathbf{x}_j \cdot \mathbf{x}_k = \delta_{jk}$.*

Exercise 5.1.20 *Simplify $\sum_{k=1}^n A_j \delta_{jk} B_k$.*

Notation 5.1.3 *I will sometimes write δ_j for a (finite or infinite) sequence with a 1 in the j th position and 0 in all the rest. Thus $(\delta_j)_k = \delta_{jk}$.*

Exercise 5.1.21 *Show that the vectors $\delta_1, \dots, \delta_n \in \mathcal{E}^n$ are ON.*

Exercise 5.1.22 *Show that the sequences (infinite-dimensional vectors) $\delta_1, \delta_2, \dots \in \ell_2$ are ON.*

5.1.5 Approximate Orthogonality

As defined in mathematics, orthogonality is an exact property: two vectors either are orthogonal or they are not.² However, in practical applications of linear mathematics, especially in biology, imprecision in information representation and processing implies that exact orthogonality cannot be depended upon. In these applications, *approximate orthogonality* is more relevant, and we are more concerned that algorithms work under conditions of approximate rather than exact orthogonality. Further, we will find that in some cases approximate orthogonality is better than exact orthogonality. To illustrate this, we begin with a simple result from Hamming (1986).

Proposition 5.1.4 *If we pick any of the 2^n bipolar vectors in $\{-1, +1\}^n$, they are almost surely nearly orthogonal; specifically, as n increases, the cosine of the angle between the vectors approaches 0 almost certainly (by the weak law of large numbers), which is to say, the angle goes to 90° .*

Remark 5.1.9 *The truth of this is easy to see. Pick two random vectors $\mathbf{x}, \mathbf{y} \in \{-1, +1\}^n$ and consider the inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^n x_k y_k$. The products $x_k y_k$ will be $+1$ or -1 with equal probability. The expectation value of the sum is the mean, which is 0. The cosine of the angle between the vectors is given by*

$$\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} = \frac{\sum \pm 1}{n},$$

since $\|\mathbf{x}\| = \|\mathbf{y}\| = \sqrt{n}$ (Why?). Hence the expectation value of the cosine goes to zero with increasing n .

Next, I will extend this result to random real vectors.

Proposition 5.1.5 *Suppose two random vectors in \mathbb{R}^n are chosen with zero mean and standard deviation proportional to $1/n$ (to maintain normalization on the average). With increasing n , both the cosine of the angle and the variance of $\langle \mathbf{x}, \mathbf{y} \rangle$ approach zero, that is, the vectors are more likely approximately orthogonal.*

Remark 5.1.10 *It's easy to see that the expectation value of $\langle \mathbf{x}, \mathbf{y} \rangle$ is zero and that the cosine decreases with n , since $\mathcal{E}\{\|\mathbf{x}\|\} = \mathcal{E}\{\|\mathbf{y}\|\} = 1/n$. The*

²This section is based on MacLennan (1993, sect. 6.2).

variance is computed as follows:

$$\text{Var}\{\langle \mathbf{x}, \mathbf{y} \rangle\} = \sum_{k=1}^n \mathcal{E}\{x_k^2\} \mathcal{E}\{y_k^2\} = \sum_{k=1}^n \frac{1}{n^2} \frac{1}{n^2} = \frac{1}{n^3}.$$

Biological neurons and other analog systems are limited to low-precision calculation; therefore the difference between exact orthogonality and approximate orthogonality can have little relevance for them. Therefore, several investigators have seen the need for a notion of ϵ -orthogonality; the following discusses the Kainen's formulation (1992), who also cites related ideas.

Definition 5.1.13 (ϵ -orthogonality) *Two vectors are ϵ -orthogonal if their inner product is less than ϵ , that is, $\langle \mathbf{x}, \mathbf{y} \rangle \leq \epsilon$.*

Proposition 5.1.6 (Kainen) *The number of normalized ϵ -orthogonal n -dimensional vectors increases exponentially in n . Specifically, if $N(n)$ is the number of such vectors, then*

$$(1 - \epsilon)^{-1/2} \leq \lim_{n \rightarrow \infty} N(n)^{1/n} \leq (1 - \epsilon)^{-1/2}.$$

Remark 5.1.11 *For example, for $\epsilon = 0.1$ we have*

$$1.00504 \leq \lim_{n \rightarrow \infty} N(n)^{1/n} \leq 1.0541.$$

Hence, asymptotically,

$$1.0541^n \geq N(n) \geq 1.00504^n.$$

Therefore, if $N(5000) \geq 8.2 \times 10^{10}$ and $N(10\,000) \geq 6.7 \times 10^{21}$. Thus a 5000-dimensional space (small by brain standards, less than 0.04 sq. mm) has 82 billion 0.1-orthogonal vectors, but only 5000 exactly orthogonal vectors.

Remark 5.1.12 *The moral of this story is that in very high-dimensional spaces, such as we find in the brain, vectors that are orthogonal to within brain precision can be found by selecting random vectors. That is, random selection is a good way to develop orthogonal representations.*

5.2 Hilbert spaces

5.2.1 Completeness

Let S_r be the continuous function:

$$S_r(x) = \begin{cases} 0 & \text{if } x < 0 \\ \sin(rx) & \text{if } 0 \leq x \leq \pi/(2r) \\ 1 & \text{otherwise} \end{cases} .$$

Notice that S_r increases from 0 to 1 over a distance proportional to $1/r$. (Graph the function if you don't see this.) Next consider the sequence of functions $S_1, S_{1/2}, S_{1/3}, \dots$. Is this sequence Cauchy? (Don't prove it; just answer intuitively.) What would the limit of this sequence have to be? Notice that the limit is not continuous, so it is not a member of $\mathcal{C}[-a, a]$, which shows that $\mathcal{C}[-a, a]$ is not complete.

Definition 5.2.1 (Hilbert space) *A Hilbert space is a complete inner-product space.*

Remark 5.2.1 *Since a Hilbert space is complete, we know that every Cauchy sequence in the space has a limit in the space. Further, if an inner-product space is incomplete, it can be made into a Hilbert space by the completion process described in Sec. 2.7.2.*

Remark 5.2.2 *One of the main purposes of the theory of Hilbert spaces is to extend the familiar linear algebra of finite-dimensional vector spaces to infinite-dimensional function spaces. In general, spatially or temporally continuous signals ("fields," in the sense of field computation) are best treated as functions, and so it is convenient to treat spaces of such signals as Hilbert spaces.*

Exercise 5.2.1 *Show that n -dimensional Euclidean space \mathcal{E}^n is a Hilbert space.*

Exercise 5.2.2 *Show that the space ℓ_2 of square-summable sequences is a Hilbert space.*

We have seen that the space $\mathcal{C}[a, b]$ of continuous functions is not complete, since the limit of a sequence of continuous functions may be discontinuous. On the other hand, we have also seen that every metric space has a unique (up to isomorphism) completion. Therefore we can define a Hilbert space that is the completion of $\mathcal{C}[a, b]$.

Definition 5.2.2 *The Hilbert space $\mathcal{H}[a, b]$ is the completion of the space $\mathcal{C}[a, b]$ of continuous functions over $[a, b]$.*

Clearly $\mathcal{H}[a, b]$ contains discontinuous functions, such as the *unit step* $U = \lim_{n \rightarrow \infty} S_{1/n}$.

5.2.2 Orthonormal Bases

Definition 5.2.3 (Complete ON set) *An ON subset of a Hilbert space is called complete if each nonzero element of the space is nonorthogonal to at least one element of the set. Conversely, if $\{v_1, v_2, \dots\}$ is a complete ON set and for all k , $\langle \xi, v_k \rangle = 0$, then $\xi = \mathbf{0}$.*

Definition 5.2.4 (Orthonormal basis) *An orthonormal basis is a complete ON set.*

Remark 5.2.3 *In very rough terms, nonorthogonal vectors have some “overlap” (think of ordinary finite-dimensional vectors). Thus an ON set is complete if it has some overlap with every (nonzero) vector; it doesn’t leave any out.*

Proposition 5.2.1 *A complete ON set is a maximal ON set, in the sense that any larger set that contains it is not ON.*

Exercise 5.2.3 *Why?*

Proposition 5.2.2 *Every complete ON set in a Hilbert space has the same cardinal number.*

Definition 5.2.5 (Dimension of a Hilbert Space) *The dimension of a Hilbert space is the cardinality of a complete ON set in the space.*

Remark 5.2.4 *For a finite-dimensional space, such as \mathcal{E}^n , the dimension is a finite number (i.e. n is the dimension of \mathcal{E}^n) just as one would expect. The dimension of an infinite-dimensional space is an infinite cardinal number. For example, the dimension of ℓ_2 is \aleph_0 (read “aleph nought” or “aleph null”), the cardinality of the integers.*

Exercise 5.2.4 *Show that \aleph_0 is the cardinality of ℓ_2 .*

Proposition 5.2.3 *A Hilbert space contains an ON basis if and only if it is separable.*

Remark 5.2.5 *We will not sketch a proof of this proposition; however its truth will become clear from the following discussion.*

Proposition 5.2.4 *The set of all polynomials with rational coefficients is a countable, dense subset of $\mathcal{H}[a, b]$; hence, $\mathcal{H}[a, b]$ is separable.*

Exercise 5.2.5 *Justify this proposition by (1) explaining why the set of polynomials with rational coefficients is countable, and (2) explaining why it is dense (based on your knowledge of numerical analysis).*

5.2.3 Generalized Fourier coefficients

Proposition 5.2.5 (Generalized Fourier series) *If ξ_1, ξ_2, \dots is an ON basis, then*

$$\phi = \sum_k \langle \phi, \xi_k \rangle \xi_k = \sum_k \langle \xi_k | \phi \rangle |\xi_k\rangle = |\phi\rangle.$$

This is called a generalized Fourier series for ϕ and the $c_k = \langle \phi, \xi_k \rangle = \langle \xi_k | \phi \rangle$ are called its generalized Fourier coefficients.

Exercise 5.2.6 *Prove this. Observe that since the ξ_k are an ON basis, there exist c_k such that $\phi = \sum_k c_k \xi_k$.*

Remark 5.2.6 *Consider the generalized Fourier expansion, $|\phi\rangle = \sum_k c_k |\xi_k\rangle$, where $c_k = \langle \xi_k | \phi \rangle$. Notice that the coefficient c_k is the result of applying the linear operator $\langle \xi_k |$, which extracts the $|\xi_k\rangle$ component, to the vector $|\phi\rangle$.*

Remark 5.2.7 *The term “generalized Fourier coefficient” is used because the “ordinary” Fourier series is based on a particular basis, the sine and cosine functions.*

Definition 5.2.6 (Trigonometric Basis) *For functions over $[0, 1]$, the trigonometric basis can be defined:*

$$\begin{aligned} \xi_0(t) &= 1, \\ \xi_{2k-1}(t) &= \sqrt{2\pi} \sin(2\pi kt), \text{ for } k = 1, 2, \dots, \\ \xi_{2k}(t) &= \sqrt{2\pi} \cos(2\pi kt), \text{ for } k = 1, 2, \dots \end{aligned}$$

Proposition 5.2.6 (Parseval relation) *If ξ_1, ξ_2, \dots is an ON basis, then*

$$\langle \phi, \psi \rangle = \sum_k \langle \phi, \xi_k \rangle \overline{\langle \psi, \xi_k \rangle}.$$

That is, if $\phi = \sum c_k \xi_k$ and $\psi = \sum d_k \xi_k$, then $\langle \phi, \psi \rangle = \sum c_k \overline{d_k} = \langle \mathbf{c}, \mathbf{d} \rangle$.

Proposition 5.2.7 (Parseval equality) *If ξ_1, ξ_2, \dots is an ON basis, then*

$$\|\phi\|^2 = \sum_k |\langle \phi, \xi_k \rangle|^2.$$

That is, if $\phi = \sum c_k \xi_k$, then $\|\phi\|^2 = \sum |c_k|^2 = \|\mathbf{c}\|^2$.

Exercise 5.2.7 *Prove the preceding propositions.*

Exercise 5.2.8 *Show that if the ξ_k are an ON basis, then*

$$\langle \phi | \psi \rangle = \sum_k \langle \phi | \xi_k \rangle \langle \xi_k | \psi \rangle.$$

Proposition 5.2.8 *Any separable Hilbert space is isometric (and therefore isomorphic) to ℓ_2 .*

Remark 5.2.8 *Since a separable Hilbert space has an ON basis (ξ_1, ξ_2, \dots) , we can expand any element ϕ of the space into a generalized Fourier series. Notice that the sequence of generalized Fourier coefficients $\mathbf{c} = (\langle \phi, \xi_1 \rangle, \langle \phi, \xi_2 \rangle, \dots)$ belongs to ℓ_2 (Why?). Furthermore, \mathbf{c} and ϕ have the same norm (Why?). Thus ϕ can be isometrically mapped into ℓ_2 .*

Conversely, if $\mathbf{c} = (c_1, c_2, \dots) \in \ell_2$, then it can be isometrically mapped into $\mathcal{H}[a, b]$ as $\phi = \sum_k c_k \xi_k$. Explain why.

Proposition 5.2.9 *All Hilbert spaces with the same dimension are isometric (and therefore isomorphic).*

Remark 5.2.9 *This is because any element of the one space can be isometrically mapped to a sequence of generalized Fourier coefficients as long as the dimension of the space. This sequence can then be used as generalized Fourier coefficients to map into the second space.*

Proposition 5.2.10 *Every separable Hilbert space is either finite-dimensional, in which case it is isometric to \mathbb{C}^n , where n is its dimension, or it is infinite-dimensional, in which case it is isometric to ℓ_2 .*

Exercise 5.2.9 *Prove the preceding proposition.*

Remark 5.2.10 *Prop. 5.2.10 is fundamental for field computation, because it means that spatially continuous fields can be represented as discrete ensembles of complex coefficients, which also supports representation by discrete ensembles of neurons, as is discussed in later (Ch. 12). Separable Hilbert spaces are where the continuous meets the discrete, and so they are fundamental in understanding the relation of discrete and continuous information representation and processing.*

5.2.4 Spanning vectors and orthogonal subspaces

The following concepts are familiar from linear algebra; the theory of Hilbert spaces extends them to infinite-dimensional spaces.

Definition 5.2.7 (Span) *Suppose $\xi_1, \dots, \xi_n \in \mathcal{H}$ are linearly independent. The span of ξ_1, \dots, ξ_n is the set $S = \text{span}\{\xi_1, \dots, \xi_n\}$ of all linear combinations $\sum c_k \xi_k$ of these vectors. The set S is said to be spanned by these vectors.*

Definition 5.2.8 (Orthogonal Complement) *If $S \subset \mathcal{H}$, then the orthogonal complement of S is the set of all vectors that are orthogonal to every element of S .*

Proposition 5.2.11 *If T is the orthogonal complement of S , then the Hilbert space can be written as the direct sum of S and T , $\mathcal{H} = S \oplus T$. This means that every $\phi \in \mathcal{H}$ can be written as a sum of orthogonal $\sigma \in S$ and $\tau \in T$.*

Proposition 5.2.12 *Let $S = \text{span}\{\xi_1, \dots, \xi_n\}$ in a Hilbert space \mathcal{H} . For any $\phi \in \mathcal{H}$ there is a $\psi \in S$ that is closest to ϕ (in an L_2 sense), and $\nu = \phi - \psi$ is orthogonal to every element of S .*

Remark 5.2.11 *Thus, in the finite dimensional subspace S there is a best approximation ψ to ϕ , and the remainder ν has no “overlap” with this space. This is easy to see in the finite-dimensional case. Let $S = \{\xi_1, \xi_2\} \subset \mathcal{E}^3$. Then S is a plane in three-dimensional space. For a given ϕ , ν is a perpendicular vector from ϕ to this plane and ψ is the point in the plane intersected by this perpendicular.*

Proof: Here is a sketch of a proof of Proposition 5.2.12. Can we find a ψ such that $\nu = \phi - \psi$ is orthogonal to every element of S ? Thus, we need, for all k , $\nu \cdot \xi_k = 0$. Since ψ is in the span of ξ_1, \dots, ξ_n there are c_k such that $\psi = \sum c_k \xi_k$. We solve for ψ as follows: since $(\phi - \psi) \cdot \xi_k = 0$, we know $\phi \cdot \xi_k = \psi \cdot \xi_k$, for all k . Therefore,

$$\phi \cdot \xi_k = \left(\sum_j c_j \xi_j \right) \cdot \xi_k, \quad (5.16)$$

$$= \sum_j (\xi_k \cdot \xi_j) c_j. \quad (5.17)$$

Now let $M_{kj} = \xi_k \cdot \xi_j$ and $d_k = \phi \cdot \xi_k$. In these terms Eq. 5.17 can be written $d_k = \sum_j M_{kj} c_j$, or $\mathbf{d} = M\mathbf{c}$. Therefore we can determine ψ by solving for \mathbf{c} .

Notice that if the ξ_k are ON, then $M = I$ and $\mathbf{c} = \mathbf{d}$. However, we know only that the ξ_k are linearly independent, but in this case it can be shown that M is nonsingular. Therefore there is a unique solution $\mathbf{c} = M^{-1}\mathbf{d}$.

It remains to show that this ψ minimizes $\|\phi - \psi\|$. Therefore, let χ be any other vector in S ; we will show that $\|\phi - \chi\|$ is minimized only when $\chi = \psi$.

$$\begin{aligned} \|\phi - \chi\|^2 &= \|(\phi - \psi) - (\chi - \psi)\|^2, \\ &= \|\phi - \psi\|^2 - (\phi - \psi) \cdot (\chi - \psi) - (\chi - \psi) \cdot (\phi - \psi) + \|\chi - \psi\|^2. \end{aligned}$$

Since χ and ψ are both in S , so is $\chi - \psi$, which is therefore orthogonal to $\phi - \psi$, which is in the orthogonal complement of S ; hence $(\phi - \psi) \cdot (\chi - \psi) = 0 = (\chi - \psi) \cdot (\phi - \psi)$. Therefore,

$$\|\phi - \chi\|^2 = \|\phi - \psi\|^2 + \|\chi - \psi\|^2.$$

All the terms are nonnegative; therefore, since $\|\phi - \psi\|$ is fixed, $\|\phi - \chi\|$ will be minimized when $\chi = \psi$.

□

Remark 5.2.12 *The preceding proposition is also true if S is not a finite-dimensional subspace.*

5.2.5 Linear functionals

5.2.5.1 BOUNDED FUNCTIONALS

Definition 5.2.9 (Functional) A functional is a complex-valued function on a normed linear space.

Remark 5.2.13 (Linear Functionals) In this section we will be especially concerned with linear functionals.

Definition 5.2.10 (Bounded linear functional) A linear functional $L : X \rightarrow \mathbb{C}$ is bounded if there is a positive real number b such that $|L(\mathbf{x})| \leq b\|\mathbf{x}\|$ for all $\mathbf{x} \in X$.

Remark 5.2.14 For an example of an unbounded linear functional, consider:

$$L(\phi) = \phi'(x),$$

where x is a particular point, $a \leq x \leq b$, and ϕ' is the derivative of ϕ . This is an unbounded linear functional defined on the differentiable functions in $\mathcal{C}[a, b]$.

Proposition 5.2.13 A linear functional is bounded if and only if it is continuous.

Exercise 5.2.10 Prove the preceding proposition. Hint: First show it is continuous at $\mathbf{0}$; then show that if it's continuous at $\mathbf{0}$ it must be continuous at all $\mathbf{x} \in X$.

5.2.5.2 RIESZ REPRESENTATION THEOREM

Definition 5.2.11 (Representer) If a bounded linear functional $L : X \rightarrow \mathbb{C}$ on an inner-product space X can be written $L(\mathbf{x}) = \langle \mathbf{x}, \mathbf{r} \rangle = \langle \mathbf{r} | \mathbf{x} \rangle$, for a fixed $\mathbf{r} \in X$ and all $\mathbf{x} \in X$, then \mathbf{r} is called a representer of L .

Proposition 5.2.14 A representer, if it exists, is unique.

Exercise 5.2.11 Prove this proposition. Hint: Assume there are two representers.

Remark 5.2.15 *The following proposition proves that there is a representer for every bounded linear functional L . It's illuminating to express this in Dirac's bra-ket notation. Then the representer will be a bra $\langle \mathbf{r} |$ so that for any vector $|\mathbf{x}\rangle$ (which we write as a ket),*

$$L|\mathbf{x}\rangle = \langle \mathbf{r} ||\mathbf{x}\rangle = \langle \mathbf{r} | \mathbf{x}\rangle.$$

That is, L and $\langle \mathbf{r} |$ are functionally identical, and so bounded linear functionals can be identified with bras. Indeed, $\langle \mathbf{r} |$ is generally defined to be the linear functional L for which $L|\mathbf{x}\rangle = \langle \mathbf{r} | \mathbf{x}\rangle$. The space of bounded linear functionals or bras $\langle \mathbf{r} |$ is the continuous dual space of the space of vectors or kets $|\mathbf{r}\rangle$.

Proposition 5.2.15 (Riesz representation theorem) *A bounded linear functional on a Hilbert space has a unique representer in that space.*

Proof: First define the null space of L ,

$$N(L) = \{x \in \mathcal{H} \mid L(x) = 0\} \quad (5.18)$$

If $N(L) = \mathcal{H}$, then let $\mathbf{r} = \mathbf{0}$ and the theorem is proved. If on the other hand $N(L) \neq \mathcal{H}$, then let M be the orthogonal complement of $N(L)$ so that $\mathcal{H} = M \oplus N(L)$. Let \mathbf{s} be any normalized element of M .

Exercise 5.2.12 *Explain why we know that such an \mathbf{s} must exist.*

We claim that $\mathbf{r} = L(\mathbf{s})\mathbf{s}$ is the representer required. To prove this we will first show that for all \mathbf{x} , $L(\mathbf{x})\mathbf{s} - L(\mathbf{s})\mathbf{x} \in N(L)$. Observe, by the linearity of L :

$$\begin{aligned} L[L(\mathbf{x})\mathbf{s} - L(\mathbf{s})\mathbf{x}] &= L[L(\mathbf{x})\mathbf{s}] - L[L(\mathbf{s})\mathbf{x}], \\ &= L(\mathbf{x})L(\mathbf{s}) - L(\mathbf{s})L(\mathbf{x}), \\ &= 0. \end{aligned}$$

Since \mathbf{s} is from the orthogonal complement of $N(L)$, which we've shown to contain $L(\mathbf{x})\mathbf{s} - L(\mathbf{s})\mathbf{x}$, we know that the latter is orthogonal to \mathbf{s} . Thus,

$$\begin{aligned} 0 &= \mathbf{s} \cdot [L(\mathbf{x})\mathbf{s} - L(\mathbf{s})\mathbf{x}], \\ &= \mathbf{s} \cdot L(\mathbf{x})\mathbf{s} - \mathbf{s} \cdot L(\mathbf{s})\mathbf{x}. \end{aligned}$$

Hence, $L(\mathbf{x})(\mathbf{s} \cdot \mathbf{s}) = L(\mathbf{s})(\mathbf{s} \cdot \mathbf{x})$; since \mathbf{s} is normalized,

$$L(\mathbf{x}) = L(\mathbf{s})\mathbf{s} \cdot \mathbf{x}.$$

□

Definition 5.2.12 *The usual norm for a bounded linear functional L on a normed linear space is:*

$$\|L\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{|L(\mathbf{x})|}{\|\mathbf{x}\|} = \sup_{\|\mathbf{x}\|=1} |L(x)|.$$

(For our purposes, $\sup = \max$.)

Proposition 5.2.16 *In a Hilbert space, the norm of a bounded linear functional is the norm of its representer. That is $\|\langle \mathbf{x} | \cdot \rangle\| = \|\mathbf{x}\|$.*

Exercise 5.2.13 *Prove this proposition.*

5.2.6 Convergence in function spaces

There are a number of useful notions of convergence in function spaces. In the following definitions, we consider the space \mathcal{F} of functions from a nonempty set D to a normed linear space X with norm $\|\cdot\|_X$. Let $\|\cdot\|_{\mathcal{F}}$ be a norm of \mathcal{F} . We can then distinguish the following notions of convergence $\phi_n \rightarrow \phi$ in the function space.

Definition 5.2.13 (Strong Convergence) $\|\phi_n - \phi\|_{\mathcal{F}} \rightarrow 0$. *This is convergence in the norm of the function space.*

Definition 5.2.14 (Pointwise Convergence) $\|\phi_n(x) - \phi(x)\|_X \rightarrow 0$ for all $x \in D$.

Definition 5.2.15 (Uniform (Pointwise) Convergence) *For all $x \in D$, $\|\phi_n(x) - \phi(x)\|_X$ converges uniformly to 0; that is, for every $\epsilon > 0$, we can select an N independent of x , such that $\|\phi_n(x) - \phi(x)\|_X < \epsilon$ for all $n > N$ and $x \in D$.*

Remark 5.2.16 *If we think of the relation between ϵ and N as representing the rate of convergence, then uniform convergence means that the rate is the same over the entire domain.*

5.2.7 Adjoint operators and outer products

The adjoint operation is extended from matrices to arbitrary linear operators on Hilbert spaces as follows:

Proposition 5.2.17 (adjoint of linear operator) *Every continuous linear operator L has a unique continuous linear adjoint operator L^\dagger satisfying $Lx \cdot y = x \cdot L^\dagger y$ for all x, y in the Hilbert space.*

Proof: Define the bounded sesquilinear³ functional $\phi(y, x) = y \cdot Lx$. Then, for each y , consider the linear functional $\psi_y(x) = \overline{\phi(y, x)}$. By the Riesz representation theorem, this functional has a unique representer r_y (depending on y) such that $\psi_y(x) = x \cdot r_y$. Therefore define $L^\dagger(y) = r_y$. It remains to show that this satisfies the conditions. Observe:

$$x \cdot L^\dagger y = r_y = \psi_y(x) = \overline{\phi(y, x)}.$$

Taking complex conjugates of both sides, we have $\phi(y, x) = L^\dagger y \cdot x$. From the definition of ϕ :

$$y \cdot Lx = L^\dagger y \cdot x. \quad (5.19)$$

Taking conjugates of both sides yields:

$$Lx \cdot y = x \cdot L^\dagger y. \quad (5.20)$$

□

Remark 5.2.17 *In Dirac's notation, the definition of the adjoint of L is the operator M satisfying*

$$\langle y | L | x \rangle = \langle y | M^* | x \rangle = \langle x | M^T | y \rangle^*. \quad (5.21)$$

Exercise 5.2.14 *For any continuous linear operator L and general vectors x, y show that:*

$$\langle x | L | y \rangle^* = \langle y | L^\dagger | x \rangle. \quad (5.22)$$

Definition 5.2.16 (Dirac outer product or dyad) *The Dirac outer product or dyad $|x\rangle\langle y|$ is the linear operator defined*

$$|x\rangle\langle y| (|z\rangle) = |x\rangle\langle y | z \rangle = |x\rangle \langle y | z \rangle \quad (5.23)$$

for all z . That is, $|x\rangle\langle y|$, when applied to $|z\rangle$, returns $|x\rangle$ weighted by $\langle y | z \rangle$. Equivalently, $|x\rangle\langle y|$ scales the vector $|x\rangle$ by the linear functional $\langle y|$.

³For “sesquilinear,” see Def. 5.1.3 (p. 70).

Notation 5.2.1 The dyad $|x\rangle\langle y|$ may be abbreviated $|x\rangle\langle y|$ and pronounced “*x-ket bra-y*” or “*x ketbra y*.”

Proposition 5.2.18 If $|\mathbf{u}\rangle$ and $|\mathbf{v}\rangle$ are finite-dimensional column vectors, then

$$|\mathbf{u}\rangle\langle\mathbf{v}| = \mathbf{u}\mathbf{v}^\dagger. \quad (5.24)$$

That is,

$$(|\mathbf{u}\rangle\langle\mathbf{v}|)_{jk} = u_j\bar{v}_k. \quad (5.25)$$

If $|\mathbf{u}\rangle$ is $m \times 1$ and $|\mathbf{v}\rangle$ is $n \times 1$, then $|\mathbf{u}\rangle\langle\mathbf{v}|$ is $m \times n$. This notation extends to the infinite dimensional case in the obvious way.

Proposition 5.2.19 (completeness relation) If $|\xi_1\rangle, |\xi_2\rangle, \dots$ is an ON basis for a Hilbert space, then

$$I = \sum_k |\xi_k\rangle\langle\xi_k|, \quad (5.26)$$

where I is the identity operation on the space.

Proof: Observe, for an arbitrary $|\phi\rangle$ in the Hilbert space,

$$\left(\sum_k |\xi_k\rangle\langle\xi_k| \right) |\phi\rangle = \sum_k (|\xi_k\rangle\langle\xi_k||\phi\rangle) = \sum_k |\xi_k\rangle\langle\xi_k|\phi\rangle,$$

which is the generalized Fourier decomposition of $|\phi\rangle$.

□

Definition 5.2.17 (normal) An operator $L : \mathcal{H} \rightarrow \mathcal{H}$ is normal if $L^\dagger L = LL^\dagger$. The same applies to square matrices.

Definition 5.2.18 (Hermitian or self-adjoint) An operator $L : \mathcal{H} \rightarrow \mathcal{H}$ is Hermitian or self-adjoint if $L^\dagger = L$. The same applies to square matrices.

Exercise 5.2.15 Show that every Hermitian operator is normal.

5.2.8 Commutators and the uncertainty principle

Definition 5.2.19 (commutator) If $L, M : \mathcal{H} \rightarrow \mathcal{H}$ are linear operators, then their commutator is defined:

$$[L, M] = LM - ML. \quad (5.27)$$

Remark 5.2.18 In effect, $[L, M]$ distills out the non-commutative part of the product of L and M . If the operators commute, then $[L, M] = \mathbf{0}$, the identically zero operator. Constant-valued operators always commute ($cL = Lc$), and so $[c, L] = \mathbf{0}$.

Definition 5.2.20 (anti-commutator) If $L, M : \mathcal{H} \rightarrow \mathcal{H}$ are linear operators, then their anti-commutator is defined:

$$\{L, M\} = LM + ML. \quad (5.28)$$

If $\{L, M\} = \mathbf{0}$, we say that L and M anti-commute, $LM = -ML$.

Exercise 5.2.16 Show that $[L, M]$ and $\{L, M\}$ are bilinear operators (linear in both of their arguments).

Exercise 5.2.17 Show that $[L, M]$ is anticommutative, i.e., $[M, L] = -[L, M]$, and that $\{L, M\}$ is commutative.

Exercise 5.2.18 Show that $LM = \frac{[L, M] + \{L, M\}}{2}$.

Definition 5.2.21 (mean of measurement) If M is a Hermitian operator representing measurement, then the mean value of the measurement of a state $|\psi\rangle$ is

$$\langle M \rangle = \langle \psi | M | \psi \rangle.$$

Definition 5.2.22 (variance and standard deviation of measurement) If M is a Hermitian operator representing measurement, then the variance in the measurement of a state $|\psi\rangle$ is

$$\text{Var}\{M\} = \langle (M - \langle M \rangle)^2 \rangle = \langle M^2 \rangle - \langle M \rangle^2.$$

As usual, the standard deviation ΔM of the measurement is defined

$$\Delta M = \sqrt{\text{Var}\{M\}}.$$

Proposition 5.2.20 *If L and M are Hermitian operators on \mathcal{H} and $|\psi\rangle \in \mathcal{H}$, then*

$$4\langle\psi | L^2 | \psi\rangle \langle\psi | M^2 | \psi\rangle \geq |\langle\psi | [L, M] | \psi\rangle|^2 + |\langle\psi | \{L, M\} | \psi\rangle|^2.$$

More briefly, in terms of average measurements,

$$4\langle L^2 \rangle \langle M^2 \rangle \geq |\langle [L, M] \rangle|^2 + |\langle \{L, M\} \rangle|^2.$$

Proof: Let $x + iy = \langle\psi | LM | \psi\rangle$. Then,

$$\begin{aligned} 2x &= \langle\psi | LM | \psi\rangle + (\langle\psi | LM | \psi\rangle)^* \\ &= \langle\psi | LM | \psi\rangle + \langle\psi | M^\dagger L^\dagger | \psi\rangle \\ &= \langle\psi | LM | \psi\rangle + \langle\psi | ML | \psi\rangle \quad \text{since } L, M \text{ are Hermitian} \\ &= \langle\psi | \{L, M\} | \psi\rangle. \end{aligned}$$

Likewise,

$$\begin{aligned} 2iy &= \langle\psi | LM | \psi\rangle - (\langle\psi | LM | \psi\rangle)^* \\ &= \langle\psi | LM | \psi\rangle - \langle\psi | ML | \psi\rangle \\ &= \langle\psi | [L, M] | \psi\rangle. \end{aligned}$$

Hence,

$$\begin{aligned} |\langle\psi | LM | \psi\rangle|^2 &= 4(x^2 + y^2) \\ &= |\langle\psi | [L, M] | \psi\rangle|^2 + |\langle\psi | \{L, M\} | \psi\rangle|^2. \end{aligned}$$

Let $|\lambda\rangle = L|\psi\rangle$ and $|\mu\rangle = M|\psi\rangle$. By the Cauchy-Schwarz inequality, $\|\lambda\| \|\mu\| \geq |\langle\lambda | \mu\rangle|$ and so $\langle\lambda | \lambda\rangle \langle\mu | \mu\rangle \geq |\langle\lambda | \mu\rangle|^2$. Hence,

$$\langle\psi | L^2 | \psi\rangle \langle\psi | M^2 | \psi\rangle \geq |\langle\psi | LM | \psi\rangle|^2.$$

The result follows. □

Proposition 5.2.21 *Prop. 5.2.20 can be weakened into a more useful form:*

$$4\langle\psi | L^2 | \psi\rangle \langle\psi | M^2 | \psi\rangle \geq |\langle\psi | [L, M] | \psi\rangle|^2,$$

or $4\langle L^2 \rangle \langle M^2 \rangle \geq |\langle [L, M] \rangle|^2$

Proposition 5.2.22 (uncertainty principle) *If Hermitian operators P and Q are measurements, then*

$$\Delta P \Delta Q \geq \frac{1}{2} |\langle \psi | [P, Q] | \psi \rangle|.$$

That is, $\Delta P \Delta Q \geq |\langle [P, Q] \rangle|/2$.

Proof: Let $L = P - \langle P \rangle$ and $M = Q - \langle Q \rangle$. By Prop. 5.2.21 we have

$$\begin{aligned} 4 \text{Var}\{P\} \text{Var}\{Q\} &= 4 \langle L^2 \rangle \langle M^2 \rangle \\ &\geq |\langle [L, M] \rangle|^2 \\ &= |\langle [P - \langle P \rangle, Q - \langle Q \rangle] \rangle|^2 \\ &= |\langle [P, Q] \rangle|^2. \end{aligned}$$

Hence,

$$2 \Delta P \Delta Q \geq |\langle [P, Q] \rangle|$$

□

5.2.9 Separable Hilbert spaces and Hilbert-Schmidt operators

5.2.9.1 SEPARABLE HILBERT SPACES

Separability is important for the use of discrete implementations, such as neural networks, for field transformations.

Proposition 5.2.23 *A bounded linear operator between separable Hilbert spaces can be represented by an infinite matrix.*

Remark 5.2.19 *We can prove this easily and in the process derive the matrix. Suppose ξ_1, ξ_2, \dots is an ON basis for the Hilbert space \mathcal{H} . Then any $|\phi\rangle \in \mathcal{H}$ can be expanded in a generalized Fourier series, $|\phi\rangle = \sum c_k |\xi_k\rangle$. Therefore, a linear operator $L : \mathcal{H} \rightarrow \mathcal{H}'$ applied to $|\phi\rangle$ can be expanded:*

$$L|\phi\rangle = L \left(\sum_k c_k |\xi_k\rangle \right) = \sum_k c_k L|\xi_k\rangle.$$

Therefore also the result $L|\phi\rangle$ of the can be written $L|\phi\rangle = \sum d_j|\zeta_j\rangle$, where ζ_1, ζ_2, \dots is an ON basis for \mathcal{H}' . Clearly, then, getting the (infinite) vector \mathbf{d} from \mathbf{c} is sufficient to get $L|\phi\rangle$ from $|\phi\rangle$. Now observe,

$$d_j = \langle \zeta_j | L | \phi \rangle = \langle \zeta_j | \sum_k c_k L | \xi_k \rangle = \sum_k c_k \langle \zeta_j | L | \xi_k \rangle.$$

Now define the infinite matrix,

$$M_{jk} = \langle \zeta_j | L | \xi_k \rangle,$$

and we see that

$$d_j = \sum M_{jk} c_k,$$

or $\mathbf{d} = M\mathbf{c}$. These infinite series converge because L is continuous (and so preserves limits).

Remark 5.2.20 *Therefore, a bounded linear operator between separable Hilbert spaces can be represented as a discrete arrangement of continuous values, namely, an infinite matrix of complex numbers.*

Remark 5.2.21 *Recall that any infinite-dimensional separable Hilbert space is isometric to ℓ_2 . As a consequence, we can treat the elements of such a space as (possibly infinite-dimensional) complex vectors, and bounded linear operators on such spaces as (possibly infinite-dimensional) complex matrices. Therefore, except where we are dealing with the physical representations of fields and field transformations, it will be convenient to ignore the difference between function spaces and infinite-dimensional vector spaces.*

Proposition 5.2.24 *For each k , $\sum_j |M_{jk}|^2 < \infty$.*

Exercise 5.2.19 *Prove this proposition.*

5.2.9.2 HILBERT-SCHMIDT OPERATORS

The following definition and propositions are essential for field computation because they determine when multilinear field transformations can be expressed as products of fields (see Sec. 7.4.1, p. 155).

Definition 5.2.23 (Hilbert-Schmidt integral operator) A linear operator $L : \mathcal{H}[a, b] \rightarrow \mathcal{H}[a, b]$ is a Hilbert-Schmidt integral operator if and only if there is a function $K : [a, b]^2 \rightarrow \mathbb{C}$ (called the kernel of L) such that for all $t \in [a, b]$,

$$(L\phi)(t) = \int_a^b K(t, s)\phi(s)ds,$$

and $\|K\|_2^2 < \infty$; that is,

$$\int_a^b \int_a^b |K(t, s)|^2 ds dt < \infty.$$

In this case, K is called a Hilbert-Schmidt kernel.

Remark 5.2.22 By analogy with matrix-vector multiplication, it is convenient to write $K\phi$ for the integral operator $\int_a^b K(t, s)\phi(s)ds$.

Remark 5.2.23 Hilbert-Schmidt operators are important for field computation, because each can be expressed as a “multiplication” by a fixed field (its kernel); see Sec. 7.4.1 (p. 155).

Proposition 5.2.25 Hilbert-Schmidt operators are bounded; in particular,

$$\|L\| < \|K\|_2.$$

Proposition 5.2.26 Each Hilbert-Schmidt operator on $\mathcal{H}[a, b]$ has a matrix representation with respect to any ON basis of $\mathcal{H}[a, b]$. Further, if M is the matrix, then $\|M\|_2 = \|K\|_2$.

Proposition 5.2.27 (spectral decomposition of normal matrices) If the matrix M is normal, then it can be unitarily diagonalized. Specifically, $M = UDU^\dagger$, where $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ is a diagonal matrix of the eigenvalues of M , and $U = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ is the matrix of corresponding orthonormal eigenvectors, which is unitary. If M is Hermitian, then the eigenvalues are real.

Remark 5.2.24 Since the eigenvectors are orthonormal,

$$M|\phi\rangle = \sum_{k=1}^n \lambda_k \langle \mathbf{e}_k | \phi \rangle \mathbf{e}_k. \quad (5.29)$$

That is, normal M scales each of the eigenvector components of $|\phi\rangle$ independently of the others.

Proposition 5.2.28 (spectral decomposition of compact Hermitian operators)

If L is a compact Hermitian operator on a Hilbert space \mathcal{H} , then there is an orthonormal basis of \mathcal{H} consisting of the eigenvectors ζ_1, ζ_2, \dots of L with corresponding real eigenvalues $\lambda_1, \lambda_2, \dots$. The set of eigenvectors is at most countably infinite with $\lambda_k \rightarrow 0$. As a consequence,

$$L|\phi\rangle = \sum_k \lambda_k \langle \zeta_k | \phi \rangle \zeta_k. \quad (5.30)$$

Remark 5.2.25 Hilbert-Schmidt integral operators are both bounded and compact.

5.2.10 Tensor products

The tensor products of functions and Hilbert spaces are important in field computation and quantum mechanics, in which they describe the state space of composite systems.

Definition 5.2.24 (tensor product of vectors) Suppose $|\phi\rangle \in \mathcal{H}$ and $|\phi'\rangle \in \mathcal{H}'$. Then the tensor product $|\phi\rangle \otimes |\phi'\rangle$ is the bilinear operator $|\phi\rangle \otimes |\phi'\rangle : \mathcal{H} \times \mathcal{H}' \rightarrow \mathbb{C}$ defined:

$$|\phi\rangle \otimes |\phi'\rangle(|\psi\rangle, |\psi'\rangle) = \langle \phi | \psi \rangle \langle \phi' | \psi' \rangle \quad (5.31)$$

for all $|\psi\rangle \in \mathcal{H}, |\psi'\rangle \in \mathcal{H}'$. The expression “ $|\phi\rangle \otimes |\phi'\rangle$ ” may be read “ ϕ tensor ϕ' .”

Definition 5.2.25 (inner product of tensor products)

$$\langle \phi \otimes \phi' | \psi \otimes \psi' \rangle = \langle \phi | \psi \rangle \langle \phi' | \psi' \rangle. \quad (5.32)$$

Proposition 5.2.29 The tensor product satisfies the following properties:

$$\begin{aligned} (c|\phi\rangle) \otimes |\psi\rangle &= c(|\phi\rangle \otimes |\psi\rangle) = |\phi\rangle \otimes (c|\psi\rangle), \\ (|\phi\rangle + |\psi\rangle) \otimes |\chi\rangle &= (|\phi\rangle \otimes |\chi\rangle) + (|\psi\rangle \otimes |\chi\rangle), \\ |\phi\rangle \otimes (|\psi\rangle + |\chi\rangle) &= (|\phi\rangle \otimes |\psi\rangle) + (|\phi\rangle \otimes |\chi\rangle). \end{aligned}$$

Notation 5.2.2 When confusion is unlikely, the tensor product $|\phi\rangle \otimes |\psi\rangle$ may be abbreviated $|\phi\rangle|\psi\rangle$ or even $|\phi, \psi\rangle$ or $|\phi\psi\rangle$.

The tensor product of separable Hilbert spaces can be defined as the space spanned by the tensor products of their respective basis vectors.

Definition 5.2.26 (tensor product of Hilbert spaces) *Suppose the vectors $|\zeta_j\rangle$ are a basis for \mathcal{H} and the vectors $|\eta_k\rangle$ are a basis for \mathcal{H}' . Consider the space \mathcal{L} of all finite linear combinations of tensor products of the basis vectors, $|\zeta_j\rangle \otimes |\eta_k\rangle$. The Hilbert space tensor product $\mathcal{H} \otimes \mathcal{H}'$ is defined to be the completion of \mathcal{L} under the inner product in Def. 5.2.25.*

Proposition 5.2.30 *Suppose the vectors $|\zeta_j\rangle$ are a basis for \mathcal{H} and the vectors $|\eta_k\rangle$ are a basis for \mathcal{H}' . An arbitrary $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}'$ can be expanded in a generalized Fourier series:*

$$|\psi\rangle = \sum_{jk} c_{jk} |\zeta_j\rangle \otimes |\eta_k\rangle, \quad (5.33)$$

where $c_{jk} = \langle \zeta_j \otimes \eta_k | \psi \rangle$.

5.2.11 Urysohn's theorem

Proposition 5.2.31 (Urysohn) *A separable metric space is homeomorphic to a subset of the Hilbert space \mathcal{E}^∞ . In this sense each separable metric space can be embedded in \mathcal{E}^∞ .*

We'll sketch the proof, which is illuminating. Suppose we have a separable metric space (X, d) . First observe that we can assume the metric is bounded by 1; that is, for all $x, y \in X$, $d(x, y) \leq 1$. If this is not the case, then we can define

$$d'(x, y) = \frac{d(x, y)}{1 + d(x, y)}, \quad (5.34)$$

which is a metric (Show this!). (This is a useful formula, worth remembering, for converting an unbounded metric to a bounded metric.) This metric induces the same topology on X since $d(x, y) \rightarrow 0$ if and only if $d'(x, y) \rightarrow 0$.

Since X is separable it has a countable, dense subset C . Since C is countable, we can enumerate its elements, $C = \{c_1, c_2, \dots\}$. The correspondence H between X and \mathcal{E}^∞ is defined as follows. For $x \in X$ define the sequence $\xi = H(x)$ by $\xi_n = d(x, c_n)/n$ for $n = 1, 2, \dots$. Notice that the c_n act like a countable set of "benchmarks" and that H replaces x by its distance to each of these benchmarks (suitably scaled to ensure the result has a finite norm).

We must show that H is a homeomorphism (i.e. a continuous one-to-one map with a continuous inverse). The remainder of the proof may be skimmed if it's not interesting.

Exercise 5.2.20 Show that H is a one-to-one function. Hint: Suppose $\xi = H(x)$ and $\zeta = H(z)$ for $x \neq z$. Show that for a suitable choice of n we are guaranteed that $\xi_n \neq \zeta_n$ and therefore that $\xi \neq \zeta$.

Next we must show that H is continuous. Suppose $\xi = H(x)$, $\zeta = H(z)$ and $d(x, z) < \epsilon$.

Exercise 5.2.21 Show that for all n , $|\xi_n - \zeta_n| < \epsilon/n$. Hint: First show that $|d(x, c_n) - d(z, c_n)| < \epsilon$.

Exercise 5.2.22 Show that

$$\|\xi - \zeta\|^2 \leq \epsilon^2 \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

The infinite series is convergent; in fact, since $\sum_{n=1}^{\infty} \frac{1}{n^2} = \pi^2/6$, we have established $\|\xi - \zeta\| \leq \epsilon\pi/\sqrt{6}$. This shows that H is uniformly continuous (and therefore continuous).

Finally we must show that H^{-1} is continuous. Let $\xi \in H[X]$ and $x = H^{-1}(\xi)$. We must show that for each $\epsilon > 0$ there is a $\delta > 0$ such that $d(x, z) < \epsilon$ whenever $z = H^{-1}(\zeta)$ and $\|\xi, \zeta\| < \delta$. Let an arbitrary ϵ be chosen. Find the minimum n such that $d(x, c_n) < \epsilon/3$. (Such an n is guaranteed to exist; why?) We claim that $\delta = \epsilon/3n$ will satisfy the requirements.

Exercise 5.2.23 Show that $|d(x, c_n) - d(z, c_n)| < \epsilon/3$. Hint: Show that $|\xi_n - \zeta_n| < \delta$. What follows from that?

Exercise 5.2.24 Show that $d(z, c_n) \leq |d(x, c_n) - d(z, c_n)| + d(x, c_n)$. Hint: Use the triangle inequality to show $|A| \leq |B - A| + |B|$.

Exercise 5.2.25 Show that $d(x, z) < \epsilon$. Apply the triangle inequality to $d(x, c_n)$ and $d(z, c_n)$.

Therefore H^{-1} is continuous (although not, in general, uniformly continuous).

Remark 5.2.26 Because of the construction of $\xi = H(x)$, we can see that the metric space is homeomorphic to a subset of a particular space Q^∞ :

Definition 5.2.27 (Fundamental Parallelopiped) In the Hilbert space \mathcal{E}^∞ , the set Q^∞ of all ξ such that $|\xi_n| \leq 1/n$ is called the fundamental parallelopiped.

5.3 References

Sources for this chapter include Berberian (1961), Hamming (1986), Kainen (1992), MacLennan (1993), Moore (1985, ch. 6), and Nemytskii & Stepanov (1989, pp. 324–326).

Chapter 6

Gabor Representations

6.1 Introduction

This is the last chapter of general background material before turning to the topic of field computation proper, which occupies the remainder of the book. The issue is the representation of continuous fields (images, signals) extended in one or more continuous dimensions, including time. We begin with a fundamental way of quantifying the information carrying capacity of a signal, which was developed by Gabor and is complementary to Shannon's better known measure. This has interesting, mathematically rigorous, connections to the Heisenberg uncertainty principle and to wave-particle duality, which are important for applications in quantum computation. Interestingly, Gabor-like representations seem to be used by the brain, especially in primary visual cortex, and so we review the evidence for this. In any case, Gabor wavelets have proved to be valuable multi-resolution representations in many practical applications. While all the mathematical essentials are here (especially in the appendices to the chapter), our principal goal is to build intuition for the material.

Dennis Gabor is best known as the father of holography, in recognition of his development of its theory in 1947. In this chapter, however, we are concerned with his theory of communication, published in 1946 (Gabor 1946), two years before Claude Shannon's more famous theory (Shannon 1948). Gabor's theory was not simply an anticipation of Shannon's (as was Hartley's, for example); rather it addresses a completely different aspect of the nature of communication. It also provides a basis for the representation and processing

of information in vision and perhaps other sensory modalities. This aspect will be our concern here.

First we review Gabor's Uncertainty Principle, which defines limits on the representation of any signal. Next we discuss the representation of signals in terms of Gabor elementary functions (Gaussian-modulated sinusoids), which is optimal in terms of the uncertainty principle and has several advantages over representations based on Fourier series and the Sampling Theorem. After reviewing John G. Daugman's research supporting the presence of Gabor's representations in mammalian vision, we discuss its pros and cons compared to wavelet-based representations. Finally we present an extension of the Gabor representation and apply it to the representation and processing of spatiotemporal patterns in the visual cortex.

6.2 The Gabor Uncertainty Principle

Gabor proved his uncertainty principle by applying to arbitrary signals the same mathematical apparatus as used in the Heisenberg-Weyl derivation of the uncertainty principle in quantum mechanics. We give our own version of this proof in the chapter appendix (Sec. 6.12.1, p. 144); here our intent is to build intuition, so we present several informal derivations.

Suppose we are trying to measure the frequency of a tone. Intuitively, the longer the sample we take, the more accurate will be our measurement (Fig. 6.1), which suggests that the error in measuring the frequency, Δf , is inversely related to the duration of the measurement, Δt . This intuition can be made a little more precise by considering a very basic kind of frequency measurement. Suppose we have a device that counts every time our signal reaches a maximum; then the number of maxima in an interval of time Δt will be the average frequency during that interval (Fig. 6.2). How long must Δt be in order to guarantee we can distinguish frequencies differing by Δf ? This will occur when the counts for f and $f + \Delta f$ are guaranteed to differ by at least one (Fig. 6.3). That is,

$$(f + \Delta f)\Delta t - f\Delta t \geq 1,$$

or,

$$\Delta f\Delta t \geq 1. \tag{6.1}$$

This is the basic *Gabor Uncertainty Principle*; it means that the product of the uncertainties in frequency and time must exceed a fixed constant, and

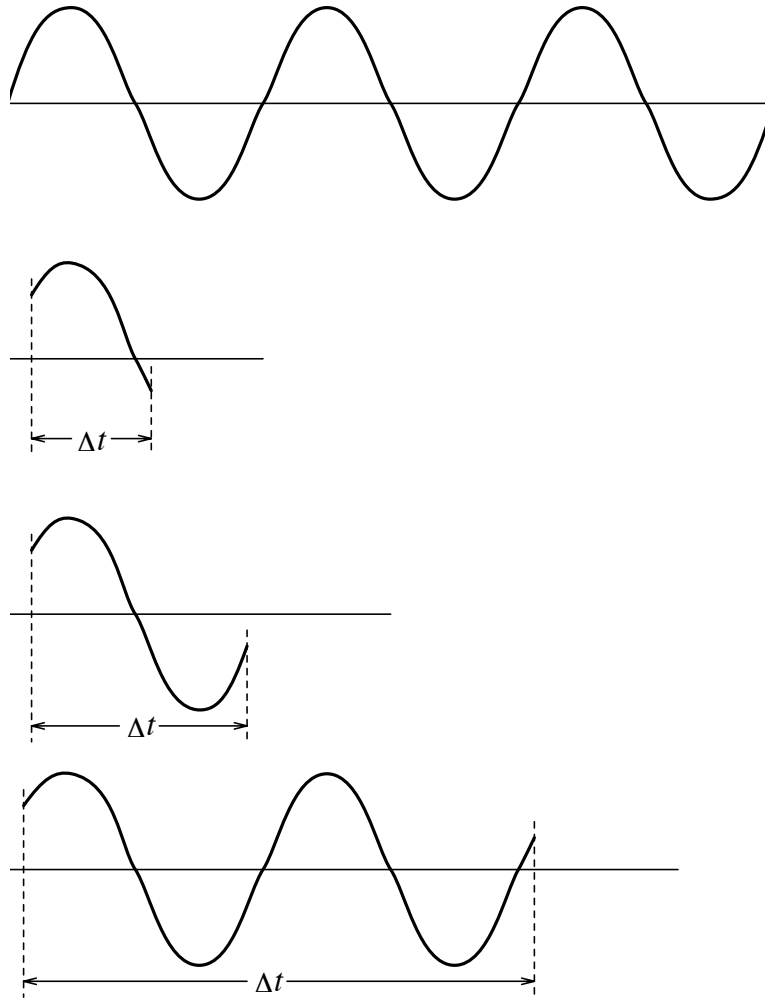


Figure 6.1: Improved frequency measurement over longer time intervals. The uncertainty in the frequency Δf decreases as the measurement interval Δt increases, and vice versa.

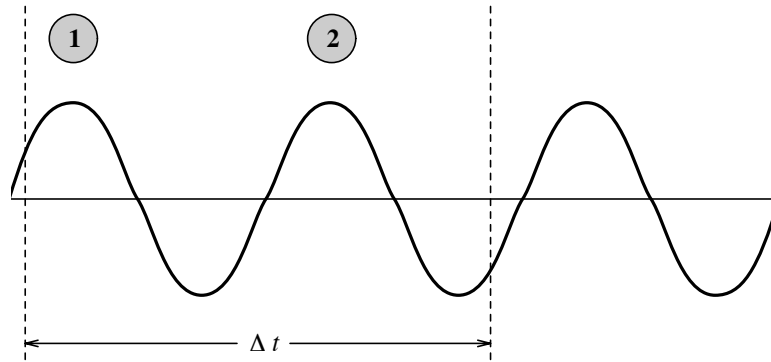


Figure 6.2: Measuring frequency by counting maxima in a given time interval. The circled numbers indicate the maxima counted during the measurement interval Δt . Since signals of other frequencies could also have the same number of maxima in that interval, there is an uncertainty Δf in the frequency.

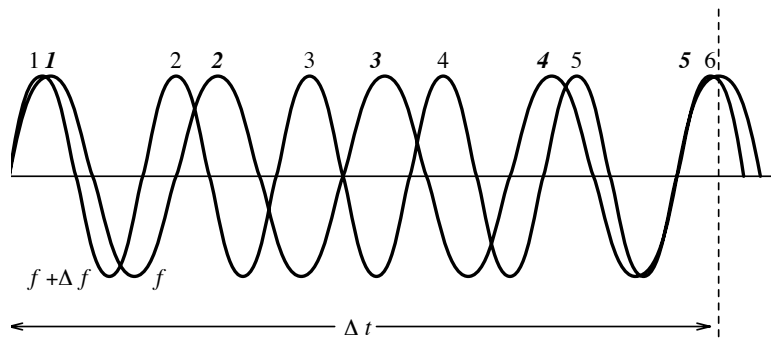


Figure 6.3: Minimum time interval Δt to detect frequency difference Δf . If two signals differ in frequency by Δf , then a measurement of duration $\Delta t \geq 1 / \Delta f$ is required to guarantee a difference in counts of maxima. Italic numbers indicate maxima of signal of frequency f ; roman numbers indicate maxima of signal of higher frequency $f + \Delta f$.

so the accuracy with which one of them can be measured limits the best possible accuracy with which the other can be measured.¹

Heisenberg's Uncertainty Principle is a simple corollary of Eq. 6.1, since according to quantum mechanics the energy of a photon is proportional to its frequency, $E = hf$. Multiplying both sides of Eq. 6.1 by h (Planck's constant) yields

$$\Delta E \Delta t \geq h,$$

which is one form of Heisenberg's principle.² Of course, Heisenberg derived his principle first; Gabor's accomplishment was to show that the same mathematical derivation applied to communication systems.

A more formal derivation of Gabor's Uncertainty Principle is based on the observation that the "spread" of a signal and its Fourier transform are inversely proportional (Fig. 6.4).³ To accomplish this we must first specify a way of measuring the spread of functions, especially when they are not *strictly local* (i.e., have noncompact support). For suppose we measure a frequency f over an interval of time Δt ; this does not imply that the frequency during that interval was always in the range $f \pm \Delta f$; it means only that the *average* frequency over that interval was in $f \pm \Delta f$. The instantaneous frequency could have varied widely, and so its spectrum might look like that in Fig. 6.5 (assumed to be centered on f). Nevertheless, we can assign a *nominal bandwidth* to the spectrum that measures its spread around the measured frequency f . Alternately we can imagine that Fig. 6.5 represents the transfer function of a band-pass filter; the nominal bandwidth is a measure of the width of the band compared with that of an ideal band-pass filter.

¹Note that we have shown a duration $\Delta t \geq 1/\Delta f$ is necessary to *discriminate* frequencies differing by Δf . On the other hand, if we *measure* a frequency f during an interval Δt , then the actual frequency could be as low as $f - 1/\Delta t$ or as high as $f + 1/\Delta t$. Therefore, the uncertainty around f is $\Delta f \geq 2/\Delta t$, giving the uncertainty principle $\Delta f \Delta t \geq 2$. Furthermore, there are other methods of measuring the frequency, such as counting sign changes (zero crossings), which would give $\Delta f \Delta t \geq 1/2$ for the discrimination case and $\Delta f \Delta t \geq 1$ for the measurement case. Thus although the exact constant depends on what we are measuring and how we are measuring it, its value doesn't much matter, since the conclusion is the same: there is a lower limit on the product of the uncertainties in the time and frequency domains. For the sake of simplicity we use the constant 1.

²Different methods of measuring ΔE and Δt yield different constants on the right-hand side, such as $\hbar = h/2\pi$ or $\hbar/2$. Again, the exact constant doesn't matter. Also, since $E = pv/2 = px/2t$, we have the other common form of the Heisenberg principle, $\Delta p \Delta x \geq 2\hbar$.

³The derivation follows Yu (1976, pp. 44–45).

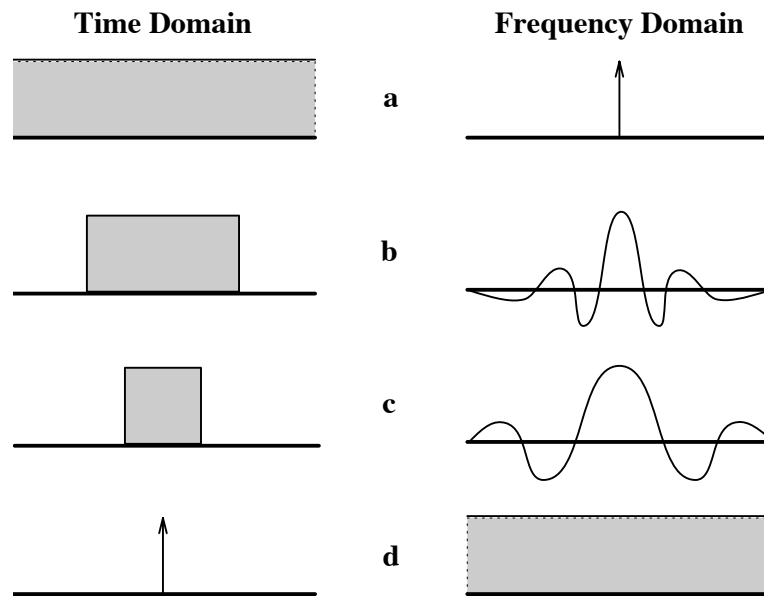


Figure 6.4: The “spread” of a signal and its Fourier transform are inversely proportional. **(a)** A constant function in the time domain corresponds to a unit impulse (Dirac delta function) in the frequency domain. **(b, c)** As the width of a pulse in the time domain decreases, its spectrum in the frequency domain spreads (spectrum shown is schematic). **(d)** A unit impulse in the time domain has a spectrum which is a constant function.

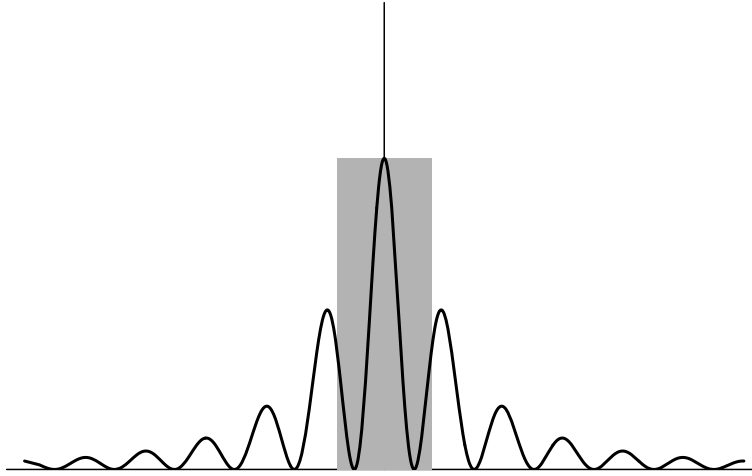


Figure 6.5: Nominal bandwidth in frequency domain of nonnegative spectrum. The nominal bandwidth is the width of a rectangular pulse (shaded) that has the same area as the continuous spectrum and has a height equal to its amplitude at the origin.

We say that nominal bandwidth measures the spectrum's *localization in the frequency domain*. Similarly, to a signal that may not actually be localized in a particular interval of time, we assign a *nominal duration* that measures its spread in time, and thus its localization in the time domain.⁴

Although there are many ways to define these measures, we define the *nominal duration* of a nonnegative signal ϕ to be the duration of a rectangular pulse of the same area and amplitude at the origin as the signal (Fig. 6.6).⁵ Thus the nominal duration Δt is defined by the equation

⁴We call a function *local* if most of its area is concentrated in a compact region; we call it *strictly local* if it has compact support (roughly, it is zero outside of a compact region). For example, the normal distribution is local, but a finite pulse is strictly local. Note that we can have a local function that is in fact more localized than a given strictly local function. We call a function *nonlocal* if its area is spread more or less uniformly over its (noncompact) domain; sine and cosine are good examples.

⁵General (possibly negative) signals are considered later. Obviously there are many ways to measure the spread of a function, for example, Gabor (1946) uses the variance, as does Hamming (1989, pp. 181–184); in Appendix 6.12.1 we use the standard deviation. The choice of measure affects only the constant on the right-hand side of the uncertainty relation.

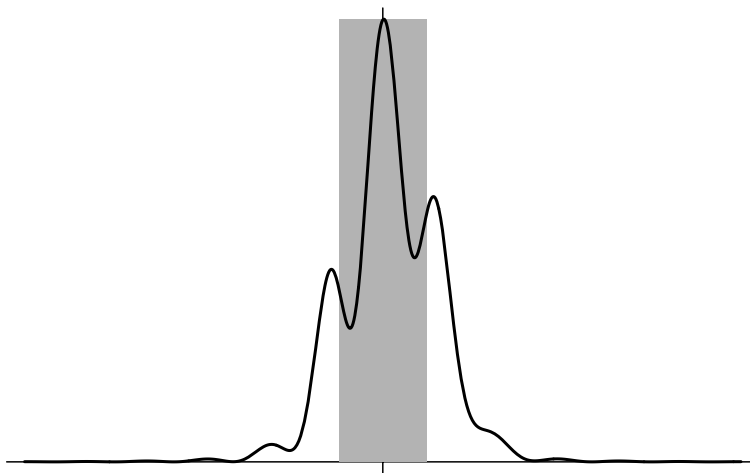


Figure 6.6: Nominal duration in time domain of nonnegative signal. The nominal duration is the duration of a rectangular pulse (shaded) that has the same area as the signal and has a height equal to its amplitude at the origin.

$$\Delta t \phi(0) = \int_{-\infty}^{\infty} \phi(t) dt \quad (\phi(t) \geq 0). \quad (6.2)$$

Similarly, the *nominal bandwidth* of the Fourier transform of ϕ , $\Phi = \mathcal{F}(\phi)$, is defined

$$\Delta f \Phi(0) = \int_{-\infty}^{\infty} \Phi(f) df \quad (\Phi(f) \geq 0). \quad (6.3)$$

Next write $\Phi(0)$ as the Fourier transform of ϕ evaluated at $f = 0$:

$$\Phi(0) = \int_{-\infty}^{\infty} \phi(t) e^{2\pi i f t} dt \Big|_{f=0} = \int_{-\infty}^{\infty} \phi(t) dt = \Delta t \phi(0).$$

Therefore,

$$\Delta t = \frac{\Phi(0)}{\phi(0)}. \quad (6.4)$$

Similarly, applying the inverse Fourier transform,

$$\phi(0) = \int_{-\infty}^{\infty} \Phi(f) e^{2\pi i f t} df \Big|_{t=0} = \int_{-\infty}^{\infty} \Phi(f) df = \Delta f \Phi(0).$$

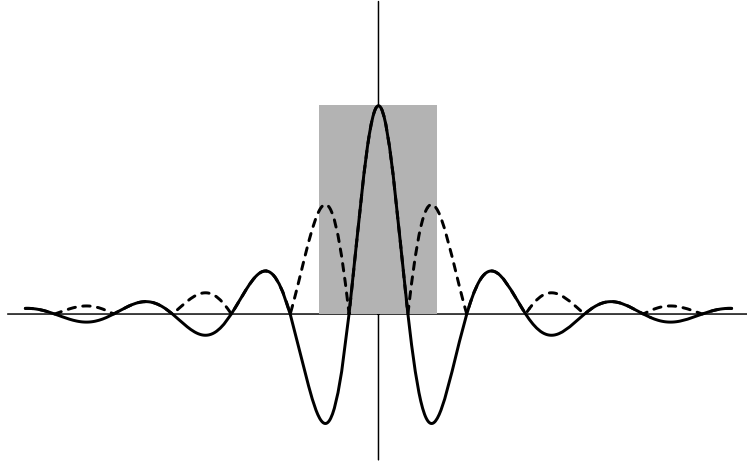


Figure 6.7: Nominal duration in time domain of arbitrary signal. Signal shown as solid line, absolute value of signal shown as dashed line. The nominal bandwidth of a spectrum is the width of a rectangular pulse (shaded) that has a height equal to the spectrum's amplitude at the origin, and that has the same area as the absolute value of the spectrum.

Therefore,

$$\Delta f = \frac{\phi(0)}{\Phi(0)}. \quad (6.5)$$

Multiplying Eq. 6.4 by Eq. 6.5 yields

$$\Delta f \Delta t = \frac{\phi(0)}{\Phi(0)} \frac{\Phi(0)}{\phi(0)} = 1. \quad (6.6)$$

Thus we see that the nominal duration and nominal bandwidth are reciprocals of each other, provided the signal and its Fourier transform are both nonnegative. In other words, there is a minimum possible simultaneous localization of the signal in the time and frequency domains.

Now we consider the general case, in which the signal and its Fourier transform may take on negative values. This is accomplished by defining the nominal spreads in terms of the absolute values of the functions (Fig. 6.7):

$$\Delta t |\phi(0)| = \int_{-\infty}^{\infty} |\phi(t)| dt, \quad (6.7)$$

$$\Delta f |\Phi(0)| = \int_{-\infty}^{\infty} |\Phi(f)| df. \quad (6.8)$$

The absolute value weakens our previous equality to an inequality:

$$\begin{aligned} \Delta t |\phi(0)| &= \int |\phi(t)| dt \geq \left| \int \phi(t) dt \right| = |\Phi(0)|, \\ \Delta f |\Phi(0)| &= \int |\Phi(f)| df \geq \left| \int \Phi(f) df \right| = |\phi(0)|. \end{aligned}$$

These equations give bounds on the nominal spreads in terms of the signal and its transform at the origin:

$$\Delta f \geq \frac{|\phi(0)|}{|\Phi(0)|}, \quad \Delta t \geq \frac{|\Phi(0)|}{|\phi(0)|}.$$

From these equations we get the general Gabor Uncertainty Principle:

$$\Delta f \Delta t \geq 1. \quad (6.9)$$

It should be noted that such an uncertainty principle applies whenever we make simultaneous measurements of a function and its Fourier transform.⁶

The implications of Gabor's principle are easier to understand by looking at it in "Fourier space," where the abscissa reflects the time domain and the ordinate the frequency domain (Fig. 6.8). Then Gabor's principle says that the spreads or uncertainties in the time and frequency measurements must define a rectangle in Fourier space whose area is at least 1. Thus we can decrease Δt , and so localize the signal better in the time domain, or decrease Δf , and so localize it better in the frequency domain, but we cannot localize it arbitrarily well in both domains simultaneously. The most we can localize signals in the Fourier domain is into rectangles of size $\Delta f \Delta t = 1$.

6.3 Gabor Representation of One-Dimensional Signals

Suppose we are transmitting information by sending signals of various frequencies of bandwidth F during an interval of time T .⁷ Suppose we sample

⁶In the language of quantum mechanics, f and t are *conjugate variables*.

⁷As Gabor notes, all real, physical signals have finite bandwidth and duration.

6.3. GABOR REPRESENTATION OF ONE-DIMENSIONAL SIGNALS 111

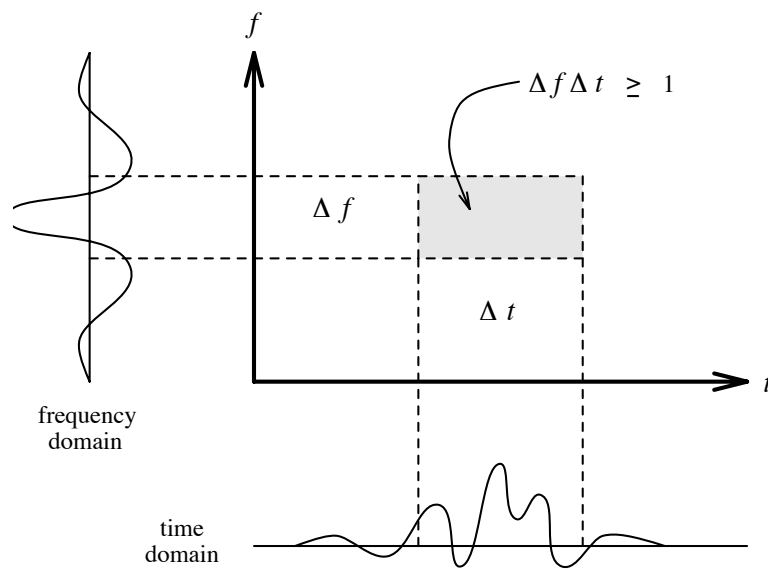


Figure 6.8: Minimum possible localization of signal in Fourier space. The product of the nominal duration Δt and nominal bandwidth Δf of a signal must be at least 1.

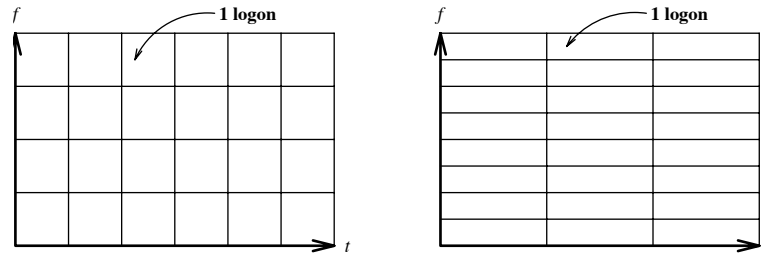


Figure 6.9: A band-limited signal is defined by a fixed number of elementary information cells. The most efficient possible (noiseless) channel divides Fourier space into information cells of size $\Delta f \Delta t = 1$. Each cell contains one *logon* of information.

the signal over intervals of length Δt to determine the signal strength at various frequencies (say, through a bank of band-pass filters). Then the closest frequencies we will be able to distinguish will be given by $\Delta f = 1/\Delta t$; that is, any frequencies differing by less than this Δf will be operationally indistinguishable. Thus our measuring apparatus divides Fourier space into *information cells* of size $\Delta f \Delta t \geq 1$ (Fig. 6.9). Since the most efficient possible (noiseless) channel will have $\Delta f \Delta t = 1$, the number of such *elementary information cells* determines the maximum amount of information that can be transmitted. No matter how we divide up Fourier space, its area gives the number of elementary information cells, and thus the number of independent quantities that can be transmitted. For example, in the simple case where $T = M\Delta t$, $F = N\Delta f$ and $\Delta f \Delta t = 1$, we are able to transmit MN independent quantities. For this reason Gabor defined a $\Delta f \Delta t = 1$ rectangle in Fourier space to be the basic *quantum* of information and called it a *logon*. Thus any device (of the given bandwidth) can transmit at most MN logons of information (in the given time interval).⁸

⁸The reader will wonder how Gabor's measure of information relates to Shannon's; in fact they are orthogonal. Gabor's measure, which may be called *structural* information, quantifies the number of possible degrees of freedom. Shannon's measure, which may be called *metrical* or *selective* information, quantifies the decrease in a priori uncertainty in a single one of these degrees of freedom. For example, in an optical device, the resolving power is equivalent to the logon content or structural information, whereas the logarithm of the number of discriminable brightness levels is equivalent to selective or Shannon information. Both notions of information are necessary for a complete theory of communication (MacKay 1969, pp. 178–180, 186–189; Cherry 1978, pp. 47–49). Re-

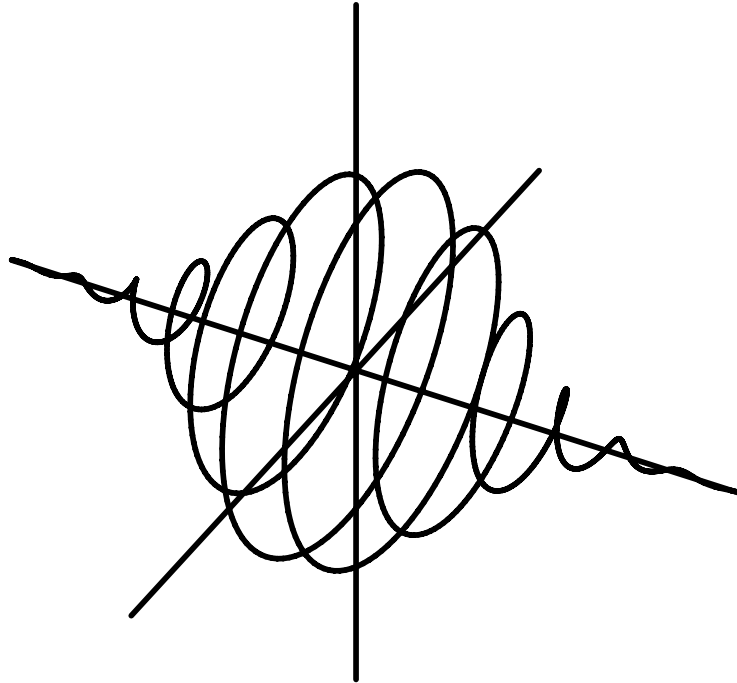


Figure 6.10: The Gaussian-modulated complex exponential, or Gabor elementary function ϕ_{jk} . The t axis goes from left to right through the center of the spiral. The imaginary axis is vertical; the real axis is horizontal, perpendicular to the other two axes. In this case $j = 0$ (no displacement from origin), $k = 1$, $\Delta f = 1$ and $\alpha^2 = 20$. The function is plotted from $t = -6$ to $t = 6$.

Gabor also showed that the minimum area in Fourier space is achieved by Gaussian-modulated complex exponential functions of the form (Fig. 6.10):

$$\phi_{jk}(t) = \exp[-\pi(t - j\Delta t)^2/\alpha^2] \exp[2\pi i k \Delta f (t - j \Delta t)], \quad (6.10)$$

where $\Delta f \Delta t = 1$.⁹ Notice that the first factor leads to a Gaussian envelope

markably, by 1928 Hartley had anticipated Gabor and Shannon by suggesting that the information transmittable over a channel is proportional to $MN \log S$, where S is the number of discriminable power levels (Cherry 1978, p. 47).

⁹Although Gabor showed that these functions occupy minimum area in terms of his (variance-based) definition of nominal spread, they also do so in terms of the definitions in Eqs. 6.7 and 6.8.

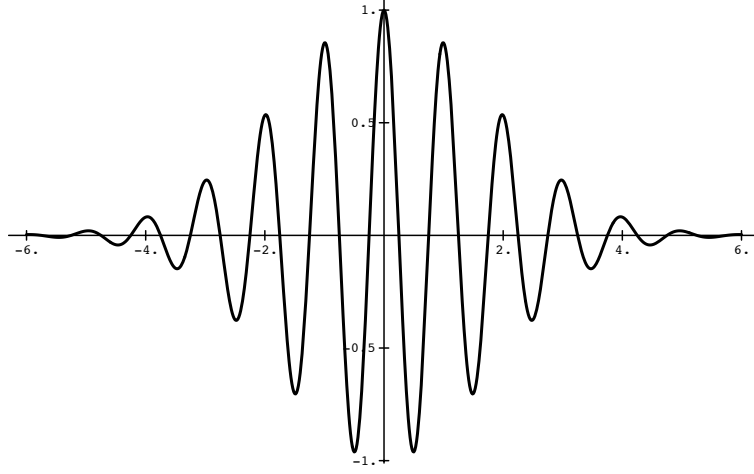


Figure 6.11: The Gaussian cosine function C_{jk} , or even-symmetric component of ϕ_{jk} . In this case $j = 0$ (no displacement from origin), $k = 1$, $\Delta f = 1$ and $\alpha^2 = 20$.

centered on $j\Delta t$, and the second factor is the conjugate exponential form of the trigonometric functions of frequency $k\Delta f$. The parameter α determines the locality (spread) of the Gaussian envelope; it is proportional to its standard deviation. So we have a periodic function modulated by a Gaussian envelope, a *coherent state* or *wave packet* in the terminology of quantum mechanics. This can be seen more clearly by using Euler's formula to rewrite Eq. 6.10 in terms of the cis (cosine + i sine) function and then in terms of the sine and cosine functions:

$$\begin{aligned}\phi_{jk}(t) &= \exp[-\pi(t - j\Delta t)^2/\alpha^2] \text{cis}[2\pi k\Delta f(t - j\Delta t)], \\ &= \exp[-\pi(t - j\Delta t)^2/\alpha^2] \cos[2\pi k\Delta f(t - j\Delta t)] + \\ &\quad i \exp[-\pi(t - j\Delta t)^2/\alpha^2] \sin[2\pi k\Delta f(t - j\Delta t)],\end{aligned}$$

Thus the Gabor elementary function is the sum of the *Gaussian cosine* and *Gaussian sine* functions (Figs. 6.11 and 6.12). If we let C_{jk} and S_{jk} represent the Gaussian cosines and sines:

$$C_{jk}(t) = \exp[-\pi(t - j\Delta t)^2/\alpha^2] \cos[2\pi k\Delta f(t - j\Delta t)], \quad (6.11)$$

$$S_{jk}(t) = \exp[-\pi(t - j\Delta t)^2/\alpha^2] \sin[2\pi k\Delta f(t - j\Delta t)], \quad (6.12)$$

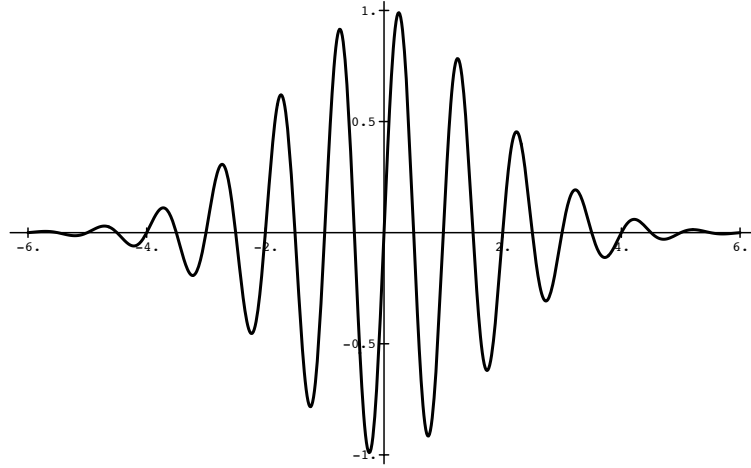


Figure 6.12: The Gaussian sine function S_{jk} , or odd-symmetric component of ϕ_{jk} . In this case $j = 0$ (no displacement from origin), $k = 1$, $\Delta f = 1$ and $\alpha^2 = 20$.

then $\phi_{jk} = C_{jk} + iS_{jk}$.

So far we have had little to say about the coefficients associated with the elementary information cells; this is the topic we now address. Suppose a rectangular region of Fourier space is divided into MN elementary information cells, and that ψ is a signal whose duration and bandwidth are confined to that region. For simplicity we assume that the cells are centered on frequencies $f = 0, \Delta f, 2\Delta f, \dots, (N-1)\Delta f$, and on times $t = 0, \Delta t, 2\Delta t, \dots, (M-1)\Delta t$ (Fig. 6.13). Gabor showed that any such (finite energy) ψ can be represented as a linear superposition of Gaussian sinusoids:

$$\psi = \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} a_{jk} C_{jk} + b_{jk} S_{jk}. \quad (6.13)$$

Each Gaussian cosine C_{jk} or sine S_{jk} is localized in the cell centered on time $j\Delta t$ and frequency $k\Delta f$; we call j the cell's *time-interval quantum number* and k its *frequency-band quantum number*. The real coefficients a_{jk} and b_{jk} show the amplitudes of Gaussian cosines and sines in each cell.

It would appear that ψ is determined by $2MN$ real coefficients, but since $S_{j0} = 0$, only $2MN - M$ of the coefficients are independent, as we can see

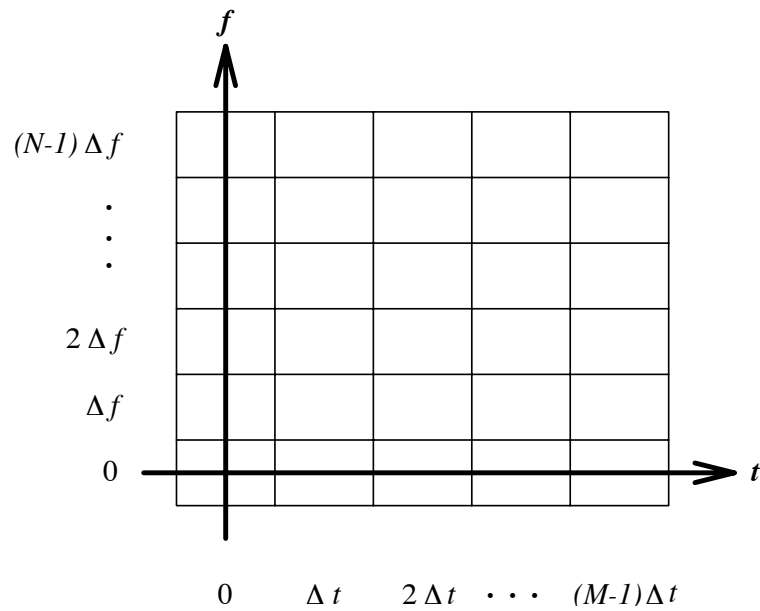


Figure 6.13: Representation of band-limited, finite-length signal by elementary information cells centered on frequencies $f = 0, \Delta f, 2\Delta f, \dots, (N-1)\Delta f$, and on times $t = 0, \Delta t, 2\Delta t, \dots, (M-1)\Delta t$. Cells are indexed by frequency-band quantum numbers $k = 0, 1, \dots, N-1$ and time-interval quantum numbers $j = 0, 1, \dots, M-1$.

by writing Eq. 6.13 in the form:

$$\psi = \sum_{j=0}^{M-1} \left\{ a_{j0} C_{j0} + \sum_{k=1}^{N-1} a_{jk} C_{jk} + b_{jk} S_{jk} \right\}.$$

The parameters a_{j0} determine the DC value of ψ in each of the M time intervals.

Just as is done in Fourier series, we can express ψ as a complex series,

$$\psi = \sum_{j=0}^{M-1} \sum_{k=-N+1}^{N-1} c_{jk} \phi_{jk}, \quad (6.14)$$

where the complex coefficients are given by:

$$\begin{aligned} c_{j0} &= a_{j0}, \\ c_{jk} &= (a_{jk} - ib_{jk})/2, \quad k > 0, \\ c_{jk} &= (a_{jk} + ib_{jk})/2, \quad k < 0. \end{aligned}$$

Notice that $c_{j,-k} = c_{j,k}^*$, the complex conjugate of $c_{j,k}$. We omit the derivation of the complex series as it can be found in any standard textbook on Fourier series.

Although there are $M(2N - 1)$ complex coefficients c_{jk} , we have seen that $M(N - 1)$ of them are complex conjugates of the others, and thus are not independent. Out of the remaining MN complex coefficients (one corresponding to each elementary information cell), the M coefficients c_{j0} are real, so once again we find that the signal is determined by $2MN - M$ real values. Thus Gabor has shown that a signal of duration T and bandwidth F has $T(2F - \Delta f)$ (real) degrees of freedom, and is thus capable of conveying $M(2N - 1)$ independent real values.¹⁰

Gabor's measure of information is consistent with the number of degrees of freedom given by the Sampling Theorem (Shannon, 1948). To see this, observe that the highest frequency elementary information cells are *centered* at frequency $(N - 1)\Delta f$; therefore their maximum frequency (as defined by their nominal spread) is $f_m = (N - 1/2)\Delta f$. The Sampling Theorem says

¹⁰Note that $FT = (M\Delta f)(N\Delta t) = MN(\Delta f\Delta t) = MN$. Another way to interpret the formula $M(2N - 1)$ is that for each time interval and frequency band we have two real parameters — an amplitude and a phase — except for the DC band, which has only an amplitude.

that to reconstruct ψ we must take equally spaced samples at a minimum of the *Nyquist frequency*, which is twice the maximum frequency. Therefore in time T the number of samples we must take is:

$$2f_m T = 2(N - 1/2)\Delta f T = (2N - 1)\Delta f \Delta t M = (2N - 1)M.$$

So Gabor's analysis and Shannon's Sampling Theorem both show that $(2N - 1)M$ real parameters determine a signal of duration T and bandwidth F .

We can also compare these results with the representation of the signal by a finite Fourier series. To do this we treat the signal ψ as periodic with period T ; then its highest frequency relative to this period is

$$H = f_m T = (N - 1/2)\Delta f (M \Delta t) = (2N - 1)M/2.$$

The signal can be represented exactly by an $H + 1$ term Fourier series:

$$\psi(t) = \sum_{n=0}^H d_n \cos(2\pi n t/T) + e_n \sin(2\pi n t/T).$$

There appear to be $2(H + 1)$ parameters d_n, e_n , but e_0 is irrelevant since the corresponding sine term is identically zero. It also can be shown that e_n is irrelevant, since (by the Sampling Theorem) the signal is determined by $2H$ points, and over these points the last sine term is linearly dependent on the other terms. Thus ψ can be represented by a Fourier series determined by $2H = (2N - 1)M$ real parameters:

$$\psi(t) = d_0 + \sum_{n=1}^{H-1} \{d_n \cos(2\pi n t/T) + e_n \sin(2\pi n t/T)\} + d_H \cos(2\pi H t/T).$$

Again, the band-limited, finite-length signal is seen to have $(2N - 1)M$ real degrees of freedom.¹¹

We have seen that any band-limited signal of finite duration can be represented by a finite superposition of Gabor elementary functions. This raises the question of whether arbitrary functions can be represented as (possibly

¹¹Some authors argue that the logon content is $MN + 1$ complex parameters or $(2N - 1)M + 1$ real parameters; it is also possible to argue an extra degree of freedom in the Fourier and Sampling representations. The practical difference is slight, since typically $MN \gg 1$, but the issue is important for information theory (MacKay 1969, pp. 185–186; Brillouin 1956, p. 97). We leave it unresolved.

6.3. GABOR REPRESENTATION OF ONE-DIMENSIONAL SIGNALS 119

infinite) superpositions of the Gabor functions. In fact it can be shown (Heil & Walnut 1989, pp. 656–657) that the set of Gabor functions ϕ_{jk} is *complete* in $L_2(\mathbb{R})$, the set of square-integrable functions. That is, any signal ψ of finite energy can be written as an infinite sum

$$\psi = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{jk} \phi_{jk}.$$

Equivalently,

$$\psi = \sum_{j,k=0}^{\infty} a_{jk} C_{jk} + b_{jk} S_{jk}.$$

(Note $S_{j0} = 0$, so the b_{j0} are irrelevant.) On the other hand, the Gabor elementary functions do not form a basis for the $L_2(\mathbb{R})$ functions, an issue addressed later (Sections 6.6 and 6.9).

There is another way to understand the relation between representations based on Gabor elementary functions, Fourier series, and the Sampling Theorem (Gabor 1946, p. 435). Notice that as $\alpha \rightarrow \infty$ the Gabor functions become

$$\begin{aligned} \phi_{jk}(t) &= \exp[2\pi i k \Delta f (t - j \Delta t)], \\ C_{jk}(t) &= \cos[2\pi k \Delta f (t - j \Delta t)], \\ S_{jk}(t) &= \sin[2\pi k \Delta f (t - j \Delta t)]. \end{aligned}$$

That is, in the $\alpha = \infty$ limit the wave packets have *no* locality, and the Gabor representation reduces to the Fourier representation, sinusoids at a spacing Δf . Conversely, as $\alpha \rightarrow 0$ the wave packets become more and more localized, and in the limit pass over into Dirac delta functions (impulses) at a spacing of Δt :

$$\begin{aligned} \phi_{jk}(t) &= \delta(t - j \Delta t) + i \delta(t - j \Delta t), \\ C_{jk}(t) &= S_{jk}(t) = \delta(t - j \Delta t). \end{aligned}$$

We see that the $\alpha = 0$ limit represents two samples (a_{jk} and b_{jk}) for each Δt interval, as required by the Sampling Theorem.

The value of the Gabor representation lies in the locality of the elementary functions. That is, although they are not *strictly* local (of compact support), their sensitivity is concentrated in a small interval of time (measured by the

nominal duration). Because of the locality property, the Gabor representation is physically more realistic than the Fourier representation, since it represents a band-limited signal of finite duration (i.e., a physically realistic signal) by a finite superposition of temporally local elementary signals. In contrast, the Fourier representation of such a signal requires an infinite superposition of nonlocal signals, and so depends on enormous cancellation in order to result in a local superposition (Strang 1989, p. 614). This and the fact that the Gabor elementary functions correspond to a quantum of information are good theoretical reasons for choosing them as representational primitives.

6.4 Gabor Representation of Two-Dimensional Signals

We have seen that any (finite energy) one-dimensional function $\psi : \mathbb{R} \rightarrow \mathbb{R}$ can be represented as a linear superposition of Gabor elementary functions, each of which represents one logon or quantum of information about the function. Although we thought of these functions as time-varying signals $\psi(t)$, it should be clear that this is not essential to the theory. $\psi(x)$ could also represent a spatial pattern, in which case the Gabor elementary functions represent information cells localized in *space* and *spatial* frequency. We must make the change to the spatial domain when we come to problems in vision, where it is necessary to consider two-dimensional functions $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$, where $\psi(x, y)$ represents the intensity at spatial location (x, y) .

It might be expected that two-dimensional signals could be represented in terms of two-dimensional analogues of Gabor elementary functions, and in the early 1980s a number of researchers suggested Gaussian-modulated sinusoids as models of the receptive fields of simple cells in visual cortex (Marčelja 1980; Daugman 1980; Watson 1982; Pribram & Carlton 1986). Our presentation is based on Daugman (1985a, 1993).

Daugman proved¹² two-dimensional analogues of Gabor's Uncertainty Principle,

$$\Delta x \Delta u \geq 1/4\pi, \quad \Delta y \Delta v \geq 1/4\pi$$

(where Δu and Δv are the uncertainties in the x and y spatial frequencies), and showed that the elementary information cells are occupied by *Gabor*

¹²Our own proofs can be found in the chapter appendices 6.12.1 and 6.12.2.

elementary functions of the form:

$$\phi_{pquv}(x, y) = \exp \left\{ -\pi \left[\frac{(x-p)^2}{\alpha^2} + \frac{(y-q)^2}{\beta^2} \right] \right\} \exp \{ 2\pi i [u(x-p) + v(y-q)] \}. \quad (6.15)$$

The first factor is a two-dimensional Gaussian distribution centered on the point (p, q) ; the second factor is the conjugate exponential form of the trigonometric functions, also centered on (p, q) . The parameters (u, v) determine the wave packet's location in the frequency domain just as (p, q) determine its location in the spatial domain. The 2D Gabor function's nominal x -spread and y -spread are α and β , and so these parameters determine its two-dimensional shape and spread.¹³

As we did for the Gabor representation of 1D signals, we will make use of Gabor elementary functions located on a regular grid in the spatial and spectral domains. In this case we index the functions by the quantum numbers j, k, l, m :

$$\begin{aligned} \phi_{jklm}(x, y) = & \exp \left\{ -\pi \left[\frac{(x-j\Delta x)^2}{\alpha^2} + \frac{(y-k\Delta y)^2}{\beta^2} \right] \right\} \times \\ & \exp \{ 2\pi i [l\Delta u(x-j\Delta x) + m\Delta v(y-k\Delta y)] \}, \end{aligned}$$

where the spacing is determined by $\Delta x \Delta u = 1$ and $\Delta y \Delta v = 1$.

The spatial frequency of the function in Eq. 6.15 is $f = \sqrt{u^2 + v^2}$ and its orientation is $\theta = \arctan(v/u)$. Conversely, $u = f \cos \theta$ and $v = f \sin \theta$. This gives an alternate form for the elementary functions:

$$\begin{aligned} \phi_{pqf\theta}(x, y) = & \exp \left\{ -\pi \left[\frac{(x-p)^2}{\alpha^2} + \frac{(y-q)^2}{\beta^2} \right] \right\} \times \\ & \exp \{ 2\pi i f [(x-p) \cos \theta + (y-q) \sin \theta] \}. \quad (6.16) \end{aligned}$$

The structure of Eq. 6.15 may be easier to understand by writing it in vector form; let $\mathbf{x} = (x, y)$ be an arbitrary point in the plane, let $\mathbf{p} = (p, q)$ be the center of the function, let $\mathbf{u} = (u, v)$ be the *wave vector* (which represents the packet's frequency along each axis). Finally, let the diagonal matrix

$$S = \begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \beta^{-1} \end{pmatrix}$$

¹³The standard deviation of the Gaussian on the x -axis is proportional to α , and on the y -axis to β ; see Sec. 6.12.2.

represent the function's shape. Then,

$$\phi_{\mathbf{p}\mathbf{u}}(\mathbf{x}) = \exp\{-\pi\|S(\mathbf{x} - \mathbf{p})\|^2\} \exp[2\pi i\mathbf{u} \cdot (\mathbf{x} - \mathbf{p})]. \quad (6.17)$$

Now it is clear that the 2D Gaussian envelope falls off with the square of the distance from \mathbf{p} scaled in accord with S . Similarly, the periodic part has its origin at \mathbf{p} . Since $\mathbf{u} \cdot (\mathbf{x} - \mathbf{p})$ projects $\mathbf{x} - \mathbf{p}$ onto \mathbf{u} , the phase of the periodic function is constant in a direction perpendicular to the wave-vector \mathbf{u} . Thus the orientation of the periodic part is given by \mathbf{u} and its frequency is given by $\|\mathbf{u}\| = f$.

The overall shape of the 2D Gabor elementary function is easiest to understand in terms of its even-symmetric (cosine) and odd-symmetric (sine) components, so we separate the periodic part of Eq. 6.17 to get $\phi_{\mathbf{p}\mathbf{u}}(\mathbf{x}) = C_{\mathbf{p}\mathbf{u}} + iS_{\mathbf{p}\mathbf{u}}$, where

$$\begin{aligned} C_{\mathbf{p}\mathbf{u}}(\mathbf{x}) &= \exp\{-\pi\|S(\mathbf{x} - \mathbf{p})\|^2\} \cos[2\pi\mathbf{u} \cdot (\mathbf{x} - \mathbf{p})], \\ S_{\mathbf{p}\mathbf{u}}(\mathbf{x}) &= \exp\{-\pi\|S(\mathbf{x} - \mathbf{p})\|^2\} \sin[2\pi\mathbf{u} \cdot (\mathbf{x} - \mathbf{p})]. \end{aligned}$$

One of these 2D Gaussian sinusoids is shown in Fig. 6.14; it can be described as an *oriented grating patch*. The 2D Gabor Uncertainty Principle can be understood by looking at Fig. 6.15. On the left we see a schematic representation of the even component of a 2D Gabor elementary function; on the right we see a schematic representation of its Fourier transform. Now consider Fig. 6.16, which shows a Gabor function like that in Fig. 6.15, but wider in the x direction. Looking at the figure we can see that its increased width will provide greater sensitivity to orientation, and this can be seen in the frequency domain, where $\Delta\theta \approx 1/(f\Delta x)$ is smaller.¹⁴ Thus there is a tradeoff between localization in the *conjugate variables* x and θ , since $\Delta x\Delta\theta \approx 1/f$. Figure 6.17 shows the effect of stretching the Gabor function in the y dimension. Just as for one-dimensional signals, the increased number of samples allows a more accurate determination of the frequency, and so decreases $\Delta f = 1/\Delta y$. Thus there is a tradeoff between localization in the conjugate variables y and f , since $\Delta y\Delta f = 1$.

¹⁴The relationship $\Delta\theta \approx 1/f\Delta x$ holds for small $\Delta\theta$.

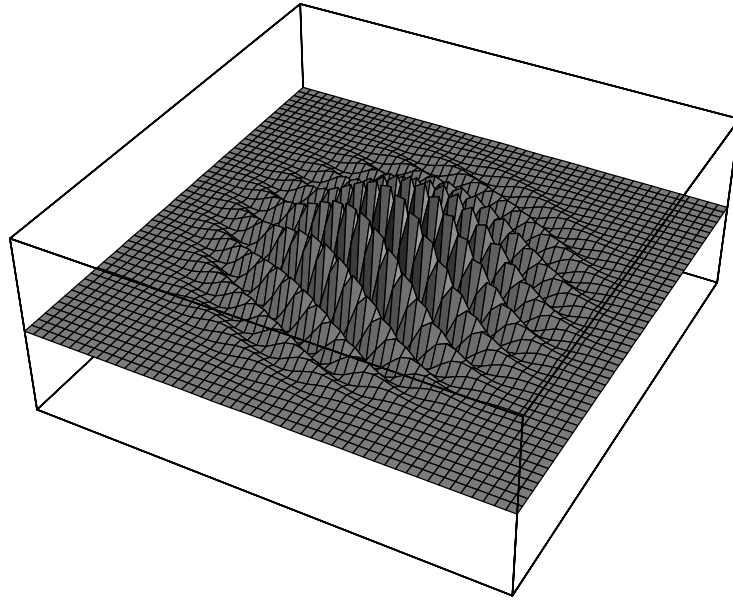


Figure 6.14: The even (cosine) component of a 2D Gabor elementary function. The function shown has $\alpha^2 = \beta^2 = 20$, $u = 1/2$, $v = 1$, and $p = q = 0$. It is plotted for all $x, y \in [-6, 6]$.

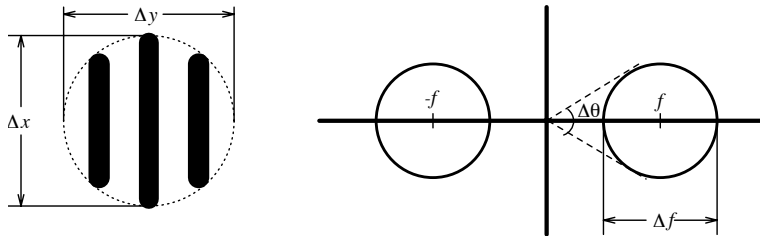


Figure 6.15: Schematic representation of even component of 2D Gabor elementary function in space and frequency domains (adapted from Daugman (1985b)).

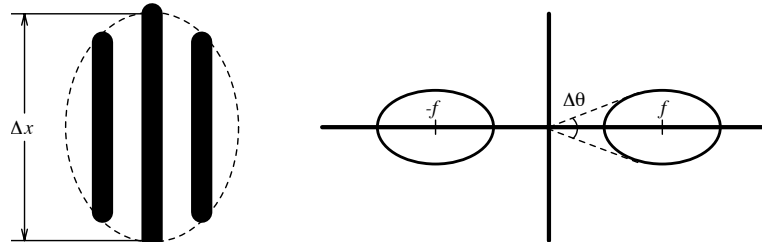


Figure 6.16: Schematic representation of even component of 2D Gabor elementary function in space and frequency domains showing Δx vs. $\Delta\theta$ tradeoff (adapted from Daugman (1985b)). Thus x and θ are conjugate variables.

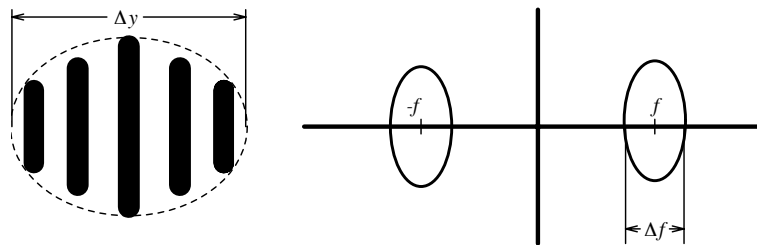


Figure 6.17: Schematic representation of even component of 2D Gabor elementary function in space and frequency domains showing Δy vs. Δf tradeoff (adapted from Daugman (1985b)). Thus y and f are conjugate variables.

6.5 Evidence for 2D Gabor Representation in Vision

Daugman (1984, 1985a, 1993) summarizes the considerable physiological evidence that 2D Gabor elementary functions are fundamental to visual processing in several mammalian species. Since proponents of alternate hypotheses, such as those based on wavelets or Laplacian edge detectors, will have to demonstrate that they can account as well for these data, we briefly review the results.

First, measurements of the receptive fields of simple cells in cat visual cortex have shown them to be like Gaussian-modulated sinusoids (Jones & Palmer 1987); Daugman has shown that 97% of them are statistically indistinguishable from the odd- or even-symmetric parts of a 2D Gabor elementary function.

Pollen and Ronner (1981) found a quadrature phase relation between pairs of simple cells in the same cortical column; that is, adjacent simple cells have grating patches that are 90° out of phase, but matched in preferred orientation and frequency. These cells could be computing the odd- and even-symmetric parts of the complex 2D Gabor function, in accord with Euler's formula:

$$e^{2\pi ifx} = \cos 2\pi fx + i \sin 2\pi fx.$$

Alternately, a simple cell could represent the sum of two Gabor elementary functions, in accord with the formulas:

$$\begin{aligned} 2 \cos 2\pi fx &= e^{-2\pi ifx} + e^{2\pi ifx}, \\ 2 \sin 2\pi fx &= ie^{-2\pi ifx} - ie^{2\pi ifx}. \end{aligned}$$

Daugman (1993) presents an additional argument in favor of the 2D Gabor elementary functions, which is based on their efficiency. An optimal image coding scheme is given by principal components analysis via the Karhunen-Loève transform. However, this method is dependent on the particular image to be encoded. To get an image-independent encoding scheme, we can make the reasonable assumption that the image statistics are locally stationary, in which case the Karhunen-Loève transform is equivalent to a windowed Fourier analysis in each of the local regions. The 2D Gabor representation is a good approximation to this scheme.

Daugman (1984) also conducted a series of *psychophysical* experiments, which allowed him to infer the tuning sensitivities of the entire visual chan-

nel (in humans). These were *masking* experiments in which a fixed grating was presented together with a grating that was variable in orientation and spatial frequency. The experiments determined how much the fixed grating interfered with the perception of the variable grating; this was determined by measuring how much the fixed grating raised the threshold at which the variable grating became visible. The underlying assumption is that the neurons in the visual channels that are involved in perceiving the fixed grating become fatigued or saturated, and so the variable grating is difficult to perceive to the extent it shares the same neural channels. Hence the degree of masking measures the response sensitivity of that neural channel to gratings of other frequencies and orientations.

The results of these psychophysical experiments were consistent with the neurophysiological data from cat visual cortex: visual channels have a frequency bandwidth of 1–2 octaves and an orientation half-bandwidth $\pm 15^\circ$ (i.e., 30° total angular bandwidth). Furthermore, the regions of sensitivity in the *spectral* domain were elliptical and twice as large in the orientation dimension as in the frequency dimension (i.e., corresponding to Fig. 6.17). Such a sensitivity profile corresponds to a width/length ratio in the *spatial* domain of $\lambda = \alpha/\beta = 1/2$, in good agreement with neurophysiological data (Jones & Palmer 1987; Movshon 1979).

For optimal (minimum uncertainty) 2D Gabor filters, a relationship can be calculated between the aspect ratio λ , the orientation half-bandwidth $\Delta\theta_{1/2}$ and the spatial frequency bandwidth r in octaves (Daugman 1985b):¹⁵

$$\Delta\theta_{1/2} = \arcsin \left(\lambda \frac{2^r - 1}{2^r + 1} \right).$$

For the observed $\lambda = 1/2$ and $r = 1.5$ octaves, the formula gives $\Delta\theta_{1/2} = 13.8^\circ$, in good agreement with the observed 15° , and supporting the hypothesis that receptive field profiles are close to Gabor elementary functions. This is reinforced by calculating the area in Fourier space of the inferred filters, which is about 2.5 times the Gabor minimum, whereas other idealized 2D filters have areas of at least 6.5 times the minimum.

¹⁵That is, $2^r = f'/f$, where f and f' are focal frequencies of two filters; we are measuring bandwidth by a ratio rather than a difference. In this case $r = 3/2$.

6.6 Problems with the Gabor Representation

One argument against the hypothesis that the vision system uses a 2D Gabor representation is that the Gabor elementary functions are not strictly local; that is, along with their Gaussian envelope, they stretch out to infinity. In mathematical terms, they have *noncompact support*. This is biologically implausible, since receptive fields are at least limited to the area of the retina, and in fact more limited than that. (Daugman notes that receptive fields die out after five or six extrema.) One answer to this argument is that the Gabor representation is intended as an idealized mathematical model, and that we shouldn't expect it to be exactly realized in the biology. In any case, the Gaussian envelope is well localized: 99.7% of its area is within 3 standard deviations of the mean, 99.994% within 4 (see Figs. 6.11, 6.12, 6.14). A receptive field that is statistically indistinguishable from a Gabor function in 97% of the cases is surely a good enough approximation to the mathematical ideal.

A second argument against the Gabor representation is that it is *nonorthogonal*.¹⁶ One result of this is that it is comparatively difficult to compute the coefficients of a 2D Gabor representation. For orthogonal representations, such as the Fourier representation and orthogonal wavelet representations (see the following section), the coefficients are computed by a simple inner product. In contrast, an algorithm for computing the coefficients of a 1D Gabor representation was not published before 1980, and Daugman uses an iterative relaxation algorithm to compute the coefficients for the 2D Gabor representation (Daugman 1993).

Further, Daugman claims that simulation studies have shown that nonorthogonal representations can lead to various artifacts, including edge echos and spurious zero-crossings (Daugman 1993). The paradox is that, as Daugman observes, nonorthogonal representations are ubiquitous in biological sensory and motor systems. Thus, whatever the disadvantages of nonorthogonality, nature seems to have found ways around them; we consider some of the possibilities in Section 6.9.

¹⁶ Indeed, the 2D Gabor elementary functions (Eq. 6.15) do not even generate a *frame*, when $\Delta x \Delta u = 1$ or $\Delta y \Delta v = 1$ (Heil & Walnut 1989, pp. 656–657; Daubechies et al. 1986, p. 1274). They do generate a frame for certain values of $\Delta u < 1/\Delta x$ and $\Delta v < 1/\Delta y$, in particular, for $\Delta u = 1/m\Delta x$, $\Delta v = 1/n\Delta y$ where $m, n = 2, 3, 4, \dots$ (Daubechies et al. 1986, p. 1275). See also Section 6.9.

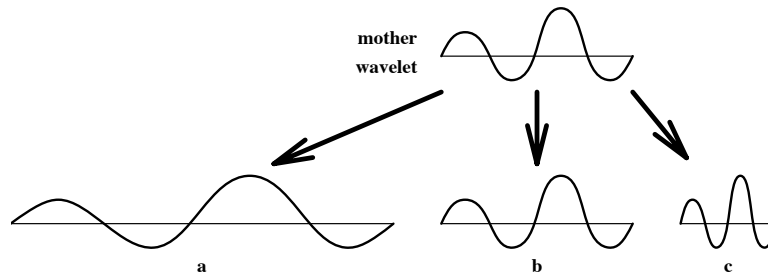


Figure 6.18: Examples of dyadic dilates of a mother wavelet. The dyadic dilates of a mother wavelet $\phi(x)$ have the form $\phi(2^k x)$ for positive and negative integers k . This figure shows the $k = -1, 0, 1$ dilates (a, b, c, respectively).

6.7 Gabor versus Wavelet Representations

Wavelets have been proposed as an alternative to Gabor elementary functions as a basis for representation in the visual system. We limit ourselves to a brief introduction.¹⁷

A family of wavelets is a complete set of functions, all generated from a *mother wavelet* by the operations of *dilation* and *translation*. Most commonly we are concerned with *dyadic* dilations and translations: A *dyadic dilate* of a function $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is a function of the form $\phi(2^k x)$, for some integer k . Thus $\phi(2^{-1}x)$ is ϕ dilated (stretched) by a factor of two (around the origin), and $\phi(2x)$ is ϕ contracted (shrunk) by a factor of two (around the origin); see Fig. 6.18.¹⁸

A *dyadic translate* of a dilated function has the form $\phi(2^k x - j)$ for some integer j . The effect of the translation is clearer if we write the function in the equivalent form $\phi[2^k(x - j/2^k)]$, since then we see that the dilate $\phi(2^k x)$ has been translated to all the dyadic points $j/2^k$ (Fig. 6.19). Thus the general form of the wavelets generated from mother wavelet ϕ by dyadic dilation and translation is:

$$\phi_{jk}(x) = \phi(2^k x - j), \quad j, k \in \mathbb{Z}. \quad (6.18)$$

Figure 6.20 shows a well-known mother wavelet, the *Haar wavelet*.

¹⁷Several good overviews of wavelets have been published, including Daubechies (1988), Strang (1989) and Heil & Walnut (1989). Our exposition is based mostly on Strang (1989).

¹⁸Some authors define the dilation by $2^{k/2}\phi(2^k x)$ so that its L_2 norm is the same as that of the mother wavelet. This is convenient if one is trying to construct an orthonormal

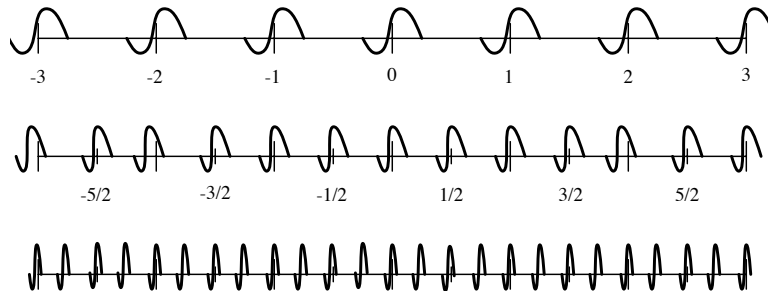


Figure 6.19: Examples of dyadic translates of a dyadic dilate of a mother wavelet. The mother wavelet $\phi(x)$ generates the family of wavelets $\phi(2^k - j)$ for all integers j and k . The first row of the figure depicts the $k = 0$ wavelets; they are centered on the integers $j = 0, \pm 1, \pm 2, \dots$. The second row shows the $k = 1$ wavelets, centered on the half-integers $j = 0, \pm 1/2, \pm 2/2, \pm 3/2, \dots$. The third row shows the $k = 2$ wavelets, centered on $j = 0, \pm 1/4, \pm 2/4, \pm 3/4, \dots$

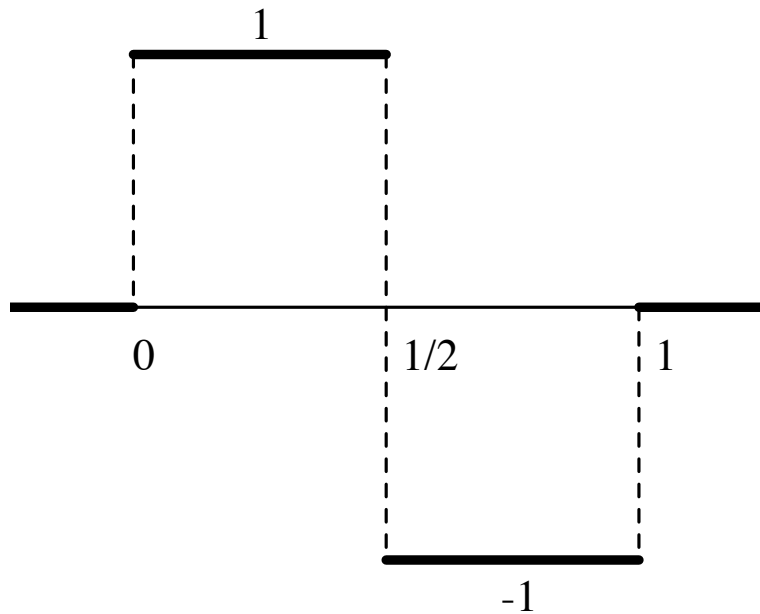


Figure 6.20: The Haar mother wavelet, which generates an orthogonal family of wavelets. A principal disadvantage of this wavelet is its discontinuity.

Next we make some observations about wavelets:

1. Although we have discussed wavelets in terms of a one-dimensional mother wavelet, it should be clear that wavelets of higher dimension can be defined in the same way:

$$\phi_{\mathbf{p}k}(\mathbf{x}) = \phi(2^k \mathbf{x} - \mathbf{p}), \quad (6.19)$$

for $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{p} \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$. Higher-dimensional wavelets are necessary to model vision.

2. Since a wavelet family is by definition *complete*, any (finite energy) function can be represented by a (possibly infinite) linear superposition of wavelets:

$$\psi = \sum_{j,k} c_{jk} \phi_{jk}.$$

This immediately raises the question of how the wavelet coefficients c_{jk} can be computed; we take it up later.

3. Families of wavelets need not be orthogonal. Although the original definition of ‘wavelet’ implied orthogonality, the term is now generally used for any complete family generated by dilation and translation. We use the term *orthogonal wavelet* for the mother wavelet of an orthogonal family. (The Haar wavelets, which are based on the mother wavelet in Fig. 6.20, are orthogonal.)
4. Although our pictures have suggested that wavelets are *strictly local* (i.e., of compact support), this is not necessarily the case. In fact, it is generally difficult to construct families of strictly local wavelets that are orthogonal, and the resulting basis functions tend to be irregular (Strang 1989, p. 615).

Why use wavelets instead of other representations, such as the Fourier or Gabor transforms? One reason is that wavelets permit functions to be represented as linear superpositions of *strictly local* elementary functions. This is especially important when the function to be represented is itself strictly local, since in this case a representation in terms of nonlocal elementary functions depends on enormous cancellation; think of the Fourier representation of a pulse.

wavelet basis, but is unnecessary for our purposes here.

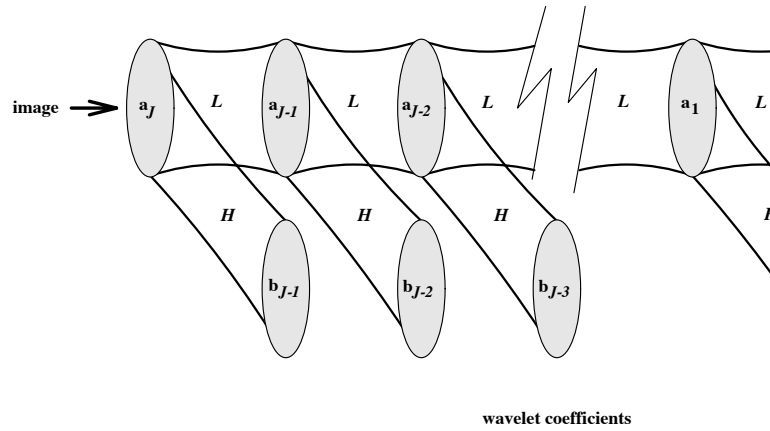


Figure 6.21: Mallat's Tree Algorithm for wavelet expansion implemented as multilayer linear neural net. Matrix H represents a linear neural network implementing a high-pass filter to extract the wavelet coefficients \mathbf{b}_j at each level of resolution. Matrix L represents a linear neural network implementing a low-pass filter that passes a "blurred" image \mathbf{a}_j on to the next stage for processing.

Representing strictly local functions in terms of strictly local functions makes sense, and wavelets are well suited to a strictly local representation. If a mother wavelet ϕ is strictly local to an interval $[-L/2, L/2]$ around the origin (i.e., its support is in this interval), then we can see that a wavelet representation of ψ is a *multiresolution*¹⁹ decomposition of the function: the wavelet coefficient c_{jk} gives information about ψ at a scale of $L/2^k$ and in the region $j/2^k$.

Coefficients are easy to compute if the wavelets are orthogonal. The simplest and most familiar way is via the inner product:

$$c_{jk} = \langle \psi, \phi_{jk} \rangle / \|\phi_{jk}\|.$$

There are also more efficient methods, such as *Mallat's Tree Algorithm* (Mallat 1989b, 1989a). Here we note only that the wavelet coefficients can be computed by a simple multilayer linear neural network (Fig. 6.21). In this algorithm, matrix H is a high-pass filter that computes the wavelet coefficients \mathbf{b}_j at resolution level 2^{-j} , and matrix L is a low-pass filter that passes

¹⁹For a review, see Daubechies (1988).

the “blurred” image on to the next stage.

6.8 Gabor Wavelets

There are some obvious similarities between wavelets and Gabor elementary functions: they are both complete families of functions, the members of which are predominantly sensitive to variations at a particular scale and at a particular location (in space or time).²⁰ Indeed, the Gabor functions can be generated from a Gaussian mother function by translation and *periodic modulation*, in the same way that wavelets are generated from a mother wavelet by translation and *dilation* (Heil & Walnut 1989).

Daugman unifies Gabor elementary functions and wavelets by defining *Gabor wavelets* (Daugman 1993). These anisotropic (oriented) wavelets are generated from a fixed Gabor elementary function (Eq. 6.17) by dilation, translation and rotation.²¹ Dilation, of course, also has the ancillary effect of changing the frequency of the Gabor function. This fits well with neurophysiological and psychophysical data indicating a *log-polar* distribution of response selectivity in cells in the visual cortex, which show an orientation half-bandwidth of $\pm 15^\circ$ and a frequency bandwidth of 1.5 octaves (Daugman 1993). That is, a space in which polar angle represents orientation and radial distance represents spatial frequency is efficiently covered by Gabor filters with aspect ratio $\lambda = 1/2$, orientation a multiple of 30° , and central frequency at radii in the ratio $2^{3/2}$ (Fig. 6.22).

6.9 The Orthogonality Issue

Of course, Gabor wavelets are *not* orthogonal, so their attractive match to the data is coupled with mathematical difficulties. But, orthogonality is a rather delicate property — functions either are or aren’t orthogonal; there are no degrees of orthogonality — and so it is probably too fragile for biology to be able to depend on it. Perhaps we should not be surprised that

²⁰Although the Gabor functions are not strictly local, their Gaussian envelope causes their greatest sensitivity to be concentrated near the center of that envelope.

²¹These are not true 2D wavelets, in the usual sense, which are generated from a mother wavelet by dilation and translation (Eq. 6.19). Thus true 2D Gabor wavelets would have the same orientation as the mother wavelet.

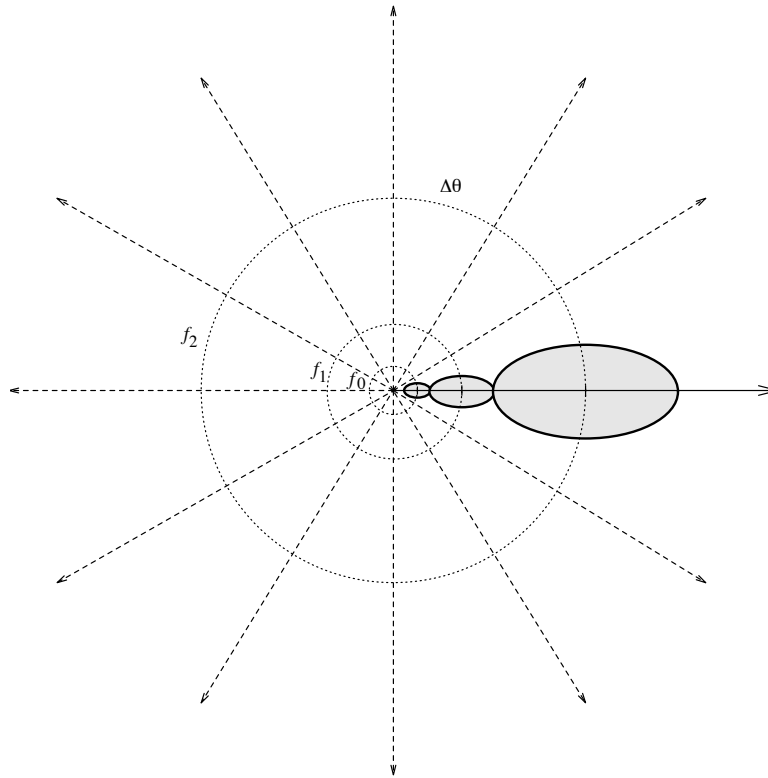


Figure 6.22: Log-polar distribution of 2D Gabor filters. Shaded ellipses represent envelopes of 2D Gabor filters with aspect ratio $\lambda = 1/2$. The filters are oriented in multiples of $\Delta\theta = 30^\circ$ and have focal frequencies f_0, f_1, f_2, \dots , where $f_k = d^k f_0$ and $d = 2^{3/2}$. Notice how effectively ellipses of these orientations, sizes and aspect ratios cover the space.

nonorthogonality is ubiquitous in biological systems; rather we should learn how nature lives with it and even exploits it.

The principal difficulty with a nonorthogonal set of elementary functions is in computing the wavelet coefficients. In other words, although we know that any finite-energy ψ can be represented by a linear superposition of the Gabor elementary functions, $\psi = \sum_{jklm} c_{jklm} \phi_{jklm}$, their nonorthogonality means that the coefficients are not defined by a simple inner product, $c_{jklm} = \langle \psi, \phi_{jklm} \rangle / \|\phi_{jklm}\|$. Daugman (1993) has described an iterative relaxation algorithm for expanding an image in terms of Gabor wavelets or other nonorthogonal codes; it operates by gradient descent in the squared L_2 error of representation:

$$\left\| \psi - \sum_{jklm} c_{jklm} \phi_{jklm} \right\|^2.$$

It is unlikely that such iterative algorithms are implemented in biological neural networks, since their speed is limited by the neuron impulse rate (say 1 msec. per impulse, with many impulses required to represent an analog quantity).²² On the other hand, iteration in local circuits in the dendritic net does not depend on impulse generation, and so could proceed much faster.

Although the Gabor coefficients cannot be computed by inner products, the evidence from receptive field studies is that the primary visual system does compute inner products, so we must question their functional role.²³ One possibility is that the inner products may be good estimates of the Gabor coefficients, and so a good place to start the relaxation process. (Daugman's algorithm does this.)

A second possibility we consider is that although the Gabor wavelets are not orthogonal, they may nevertheless be a *frame* (Heil & Walnut 1989),

²²There is at most 1 KHz. of available bandwidth since that is the maximum spike rate. Therefore, to distinguish N discrete values, we need frequencies separated by at least $\Delta f = 1000/N$. Applying the Gabor Uncertainty Principle gives $\Delta t = N/1000$ seconds, or N milliseconds to reliably transmit the value. Thus it takes at least 10 msec. to transmit an analog value with one digit of precision and at least 100 msec. to transmit it with two digits of precision.

²³Here, "primary visual system" refers to the retina, lateral geniculate nucleus and primary visual cortex (VI). Since the representational primitives of the retina + LGN system seem to be either radially symmetric differences of Gaussians or radially-symmetric Gaussian sinusoids (Pribram 1991, p. 74), the Gabor coefficients must be computed from the coefficients of these radial basis functions.

which is a generalization of a basis. If the functions ϕ_{jklm} are a frame, then there is a bounded linear operator S such that

$$\psi = \sum_{jklm} \langle \psi, S^{-1} \phi_{jklm} \rangle \phi_{jklm} = \sum_{jklm} \langle \psi, \phi_{jklm} \rangle S^{-1} \phi_{jklm}.$$

In other words, the inner products $\langle \psi, \phi_{jklm} \rangle$, which are apparently computed by the primary visual system, give the representation of ψ in terms of the dual frame $\{S^{-1} \phi_{jklm}\}$.

Now we must address the question of whether the Gabor wavelets are a frame. It has been known for some time that the 1D Gabor wavelets are *not* a frame for $\Delta f \Delta t = 1$, but Daubechies et al. (1986, p. 1275) show that they are a frame for $\Delta t = 1 / m \Delta f$ where $m = 2, 3, 4, \dots$ ²⁴

Since a 2D Gabor wavelet is the outer product of two 1D Gabor wavelets, $\phi_{jklm}(x, y) = \phi_{jl}(x) \phi_{km}(y)$, it is straight-forward to show that the 2D wavelets are a frame when $\Delta x \Delta u = 1/m$ and $\Delta y \Delta v = 1/n$, where $m, n = 2, 3, 4, \dots$. These conditions are compatible with the constraints imposed by the Gabor Uncertainty Principle. For example, $m = n = 13$ gives $\Delta x \Delta u = \Delta y \Delta v = 1/13$, which is slightly larger than the minimum $1 / 4\pi \approx 1/12.6$. Further, they are consistent with Daugman's (1984) observation that receptive fields occupy about 2.5 times the theoretical minimum area, since in the case $m = n = 8$ the functions occupy $16\pi^2/64 \approx 2.47$ times the minimum area.

Finally, we observe that there is really no *a priori* reason for the visual cortex to compute the Gabor coefficients, because there is no need for it to reconstitute the input image ψ from the coefficients:

$$\psi = \sum_{jklm} c_{jklm} \phi_{jklm}.$$

It must be remembered that this equation is only a mathematically convenient way of guaranteeing that no information is lost in computing the coefficients. Since the visual cortex harbors no homunculus, it does not need to reconstruct the image, and it may work directly in terms of the inner products.

²⁴These conditions are sufficient, but perhaps not necessary. Also note that for larger m the frame is *tighter*, which means S^{-1} is more nearly a scalar.

6.10 3D Gabor Representation of Spatiotemporal Signals

Gabor's research was motivated in part by the observation that our perception of sound is simultaneously of duration and pitch, and therefore that an analysis of sound should be in terms of elements localized in both duration and frequency (Gabor 1946, pp. 431–432). Exactly the same argument may be made for vision. In the spatial domain we see simultaneously both extent and texture (spatial frequency). Likewise, in the temporal domain we perceive simultaneously duration and motion (temporal frequency).

Thus we see that the use of 2D Gabor elementary functions to model visual image representation is unrealistic in a significant way: it ignores the *temporal structure* of images. It is as though vision were merely a succession of separate images, each independent of the next. On the other hand, if we applied to visual images the 1D Gabor functions (Section 6.3), we would capture their temporal structure, but not their spatial structure, which Daugman and others have shown to be central to understanding vision.

An obvious solution to this problem is to combine the two analyses and consider the evolution in time of two-dimensional spatial signals. Thus we will take the input to the visual cortex to be a three-dimensional signal $\psi(x, y, t)$, $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}$. Sometimes it will be more convenient to write $\psi(\mathbf{x}, t)$ where the vector $\mathbf{x} = (x, y)$ represents spatial position. Such a signal has a Fourier transform $\Psi(\zeta, \eta, \nu)$, where ζ and η are spatial frequencies and ν is a temporal frequency.

Having seen the 1D and 2D Gabor Uncertainty Principles, it is perhaps hardly surprising that there is a 3D Uncertainty Principle holding between pairs of conjugate variables:

$$\begin{aligned}\Delta x \Delta \zeta &\geq 1/4\pi, \\ \Delta y \Delta \eta &\geq 1/4\pi, \\ \Delta t \Delta \nu &\geq 1/4\pi,\end{aligned}$$

where we have defined the nominal spreads in terms of the standard deviations of the functions. For a proof of this uncertainty principle, see Sec. 6.12.1.

It is also straight-forward to show (Sec. 6.12.2) that the Gabor inequalities become equalities for the 3D Gabor elementary functions, which have the

form:

$$\begin{aligned} \phi_{pqr uvw}(x, y, t) = & \exp \left\{ -\pi \left[\frac{(x-p)^2}{\alpha^2} + \frac{(y-q)^2}{\beta^2} + \frac{(t-r)^2}{\gamma^2} \right] \right\} \times \\ & \exp \{ 2\pi i [u(x-p) + v(y-q) + w(t-r)] \}. \end{aligned} \quad (6.20)$$

The wave packet is localized around space-time coordinates (p, q, r) ; that is, it is centered at $x = p$, $y = q$ in space and $t = r$ in time. Its location in the corresponding frequency domain is given by (u, v, w) , its two spatial frequencies and one temporal frequency. This is apparent from the Fourier transform of ϕ :

$$\begin{aligned} \Phi_{pqr uvw}(\zeta, \eta, \nu) = & \exp \left\{ -\pi [(\zeta-u)^2 \alpha^2 + (\eta-v)^2 \beta^2 + (\nu-w)^2 \gamma^2] \right\} \times \\ & \exp \{ 2\pi i [x(\zeta-u) + y(\eta-v) + t(\nu-w)] \}. \end{aligned} \quad (6.21)$$

It can be shown (Sec. 6.12.2) that the standard deviations of ϕ around the x , y and t axes are proportional to α , β and γ , respectively; thus α , β and γ determine the wave packet's shape. Conversely, the standard deviations of Φ are proportional to α^{-1} , β^{-1} and γ^{-1} .

As before, the Gabor uncertainty relations permit signals to be localized in Fourier space no more accurately than a cell of size

$$\Delta x \Delta y \Delta t \Delta \zeta \Delta \eta \Delta \nu \geq 1 / 64\pi^3.$$

Indeed, the information cells can be no smaller than $1 / 4\pi$ in each pair of conjugate variables. Such cells are the information quanta for 3D signals (Fig. 6.23). The elementary information cells can be indexed by sextuples of quantum numbers, three spatiotemporal and three spectral, $\mathbf{m} = (m_1, m_2, m_3)$, $\mathbf{n} = (n_1, n_2, n_3)$, so that

$$\begin{aligned} p &= m_1 \Delta x, & u &= n_1 \Delta \zeta, \\ q &= m_2 \Delta y, & v &= n_2 \Delta \eta, \\ r &= m_3 \Delta t, & w &= n_3 \Delta \nu. \end{aligned}$$

Then the complex numbers $c_{\mathbf{mn}}$ are the Gabor coefficients of ψ if and only if

$$\psi = \sum_{\mathbf{m}} \sum_{\mathbf{n}} c_{\mathbf{mn}} \phi_{pqr uvw},$$

where the indices \mathbf{m} and \mathbf{n} have ranges appropriate to the spatiotemporal extent and bandwidth of ψ . Of course, ψ could be equally well represented by real coefficients and Gaussian sinusoids in quadrature phase.

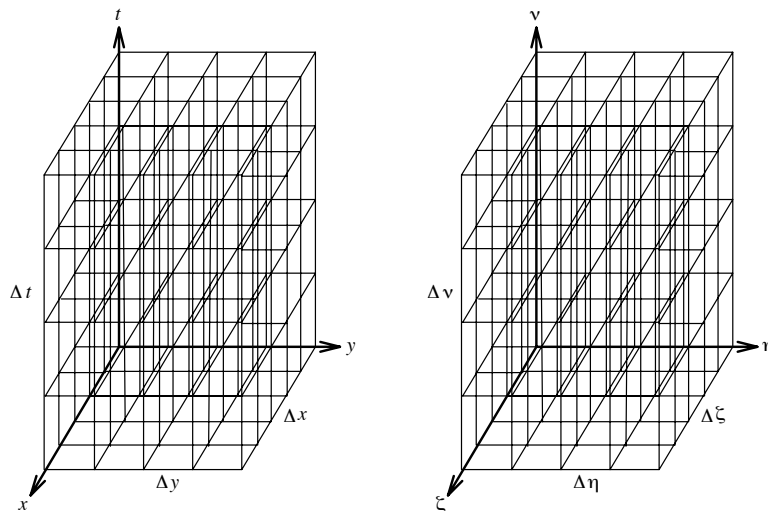


Figure 6.23: Information cells defining a 3D signal. For clarity the three spatiotemporal dimensions are shown separate from the three spectral dimensions, but it must be born in mind that each information cell is a six-dimensional rectangular space. Each such cell represents a quantum of information in Fourier space.

Next we consider the significance to vision of this quantization of Fourier space. Although it might seem natural to interpret it as a mathematical fiction, Daugman's research has already given us good reason to believe that the visual cortex is organized around a spatial Gabor representation. This suggests that we take the spatiotemporal Gabor representation quite literally and interpret the visual cortex as a bank of filters tuned to spatial frequency bands of width $\Delta\zeta$ and $\Delta\eta$, a temporal frequency band of width $\Delta\nu$, and localized to a spatial region of size $\Delta x \Delta y$. We further hypothesize that these filters accumulate information over an interval of time that is a small multiple of Δt , and produce the Gabor coefficients at the end of this interval (perhaps by relaxation during the next interval).²⁵

It is natural to identify this interval with the principal rhythm of the occipital (visual) cortex, the alpha rhythm. Slow rhythms, such as the alpha, seem to clock the generation of spike trains, just as we would expect if a set of rate-encoded Gabor coefficients were computed during each interval.²⁶ During periods of greater activity the alpha rhythm "desynchronizes" and is replaced by a higher frequency oscillation (40–60 Hz.). The results of such a decrease in Δt include greater temporal resolution, poorer temporal-frequency resolution, and less accurate computation of the Gabor coefficients — all reasonable tradeoffs in situations demanding action.

If the hypothesized correlation of Δt with the alpha rhythm is correct, then from the resting alpha frequency, 8 to 12 Hz., we can estimate the resting interval $T_\alpha \approx 100$ msec., with a range of perhaps 80 to 125 msec. Since Δt is the standard deviation of the Gabor elementary function, we can expect that T_α must be 3 or 4 times Δt (so that T_α contains 90–95% of the wave-packet). Since in Sec. 6.12.2 we show $\Delta t = \gamma / 2\sqrt{\pi}$ (Eq. 6.30), for mathematical convenience we estimate

$$T_\alpha \approx 2\sqrt{\pi}\Delta t = \gamma.$$

This implies a resting temporal frequency resolution of

$$\Delta w_\alpha = \frac{1}{4\pi\Delta t} \approx \frac{1}{4\pi(T_\alpha/2\sqrt{\pi})} = \frac{1}{2\sqrt{\pi}T_\alpha} \approx 2.8 \text{ Hz.}$$

²⁵We may compare the inhibitory wave that seems to reset cerebellar computation every 500 msec. (Pribram 1991, p. 127).

²⁶We refer here to Bland's studies of the theta rhythm in the dentate gyrus of the rabbit (Bland et al. 1978), but the same principle applies to the alpha rhythm.

The Gabor representation also sheds light on the perception of form and motion.²⁷ The parameters u , v and w determine the orientation of the elementary signal in space-time. For example, if $w = 0$ then the wave packet is perpendicular to the t -axis (Fig. 6.24) and we have the effect of a 2D (spatial) Gabor function, but localized in time. Conversely, if $w \neq 0$ then the elementary function is inclined to the time and space axes (Fig. 6.25). We can see that such a filter would respond to a grating patch moving at a fixed velocity perpendicular to the fringes. We hypothesize that 3D Gabor elementary functions of this kind explain the response characteristics of *complex* and *hypercomplex* cells in the visual cortex, which have been shown to respond to moving bars and gratings.

We can easily calculate the *phase velocity* at which the fringes move:

$$v_p = w/f,$$

where $f = \|\mathbf{u}\|$ is the spatial frequency of the grating patch. The fringes move in a direction opposite to the (spatial) wave vector \mathbf{u} (Fig. 6.25), so the velocity vector \mathbf{v} of the fringes is $-v_p$ times the *wave normal* $\mathbf{u}/\|\mathbf{u}\|$:

$$\mathbf{v} = -v_p \mathbf{u}/\|\mathbf{u}\| = -w\mathbf{u}/f^2.$$

We consider briefly the case in which the Gabor function is parallel to the time axis (Fig. 6.26), that is, $\mathbf{u} = 0$. Such a filter would respond to a uniform intensity (within its spatial receptive field) oscillating at a frequency w . We are unaware of research looking for cells with this kind of response, but it is interesting that much of the work on receptive fields has made use of *flashing* spots, and so might be consistent with the existence of such cells.

6.11 Conclusions

We have reviewed Gabor's Uncertainty Principle and Daugman's evidence for 2D Gabor filters in the visual cortex. We compared the Gabor representation with wavelet-based representations, and concluded that the Gabor representation is preferable. This is in spite of the Gabor functions not being

²⁷Of course, so would other representations in terms of signals localized in both the space-time and spectral domains, such as other 3D Fourier transforms windowed in space-time, or 3D wavelets.

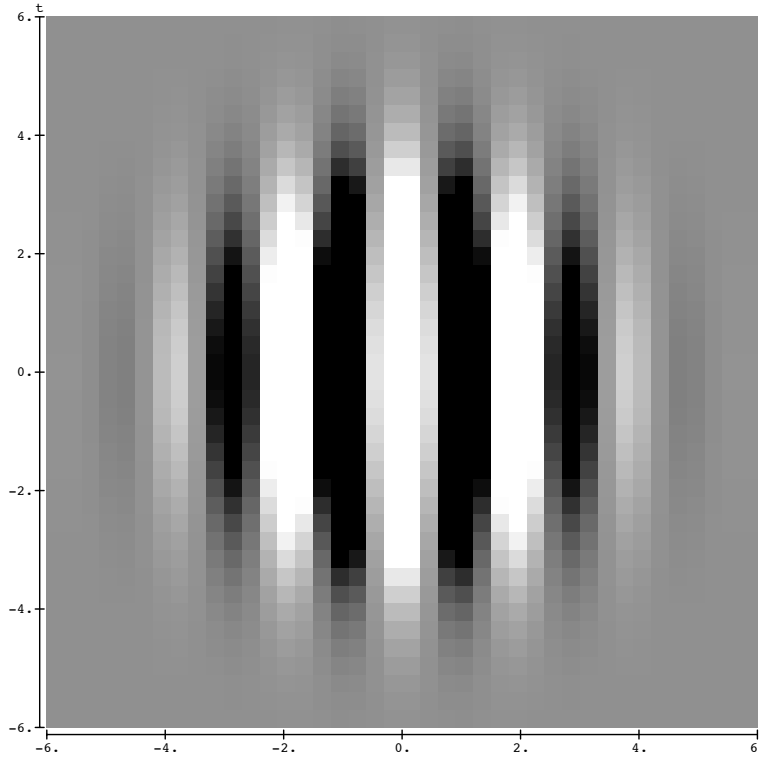


Figure 6.24: Slice through even (cosine) component of 3D Gabor elementary function oriented perpendicular ($w = 0$) to the time axis (ordinate). The abscissa is taken to be along the wave-vector \mathbf{u} , and so perpendicular to the spatial wavefronts. Such a filter is selective for stationary spatial frequency f , localized in both space and time. This function has $f = 1/2$, $\alpha^2 + \beta^2 = \gamma^2 = 20$, $u = v = 1/\sqrt{8}$, $w = 0$, $\mathbf{p} = 0$ and $r = 0$. Lighter regions are more positive, darker more negative.

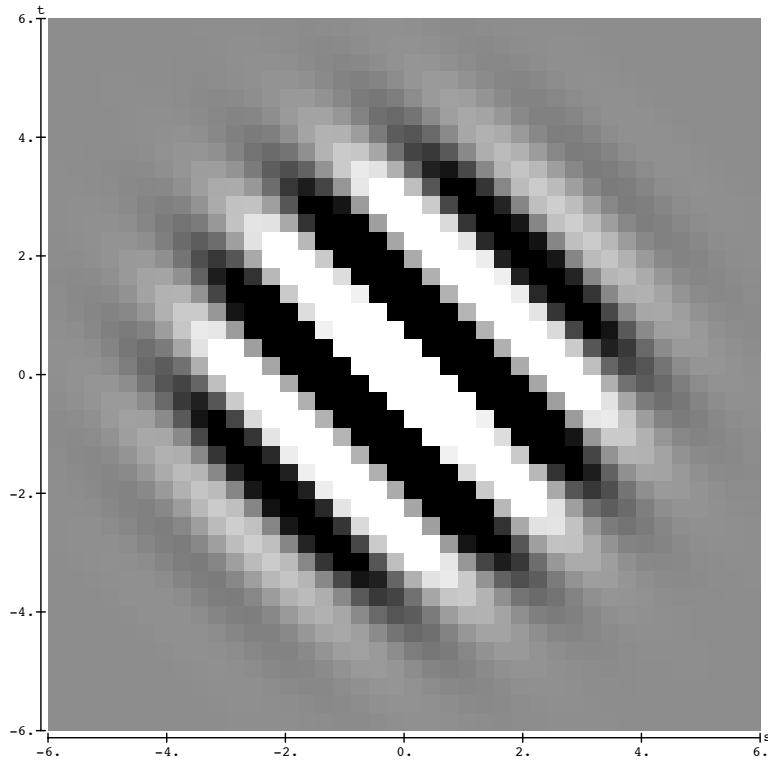


Figure 6.25: Slice through even (cosine) component of 3D Gabor elementary function inclined to the time axis ($w \neq 0$). Such a filter is selective for a spatial grating of frequency f , moving at velocity v_p , and localized in both space and time. This function has $\alpha^2 + \beta^2 = \gamma^2 = 20$, $u = v = 1/4$, $w = 1/\sqrt{8}$, $\mathbf{p} = 0$ and $r = 0$. It is selective for fringes of frequency $f = 1/\sqrt{8}$ moving at a phase velocity $v_p = w/f = 1$ to the left.

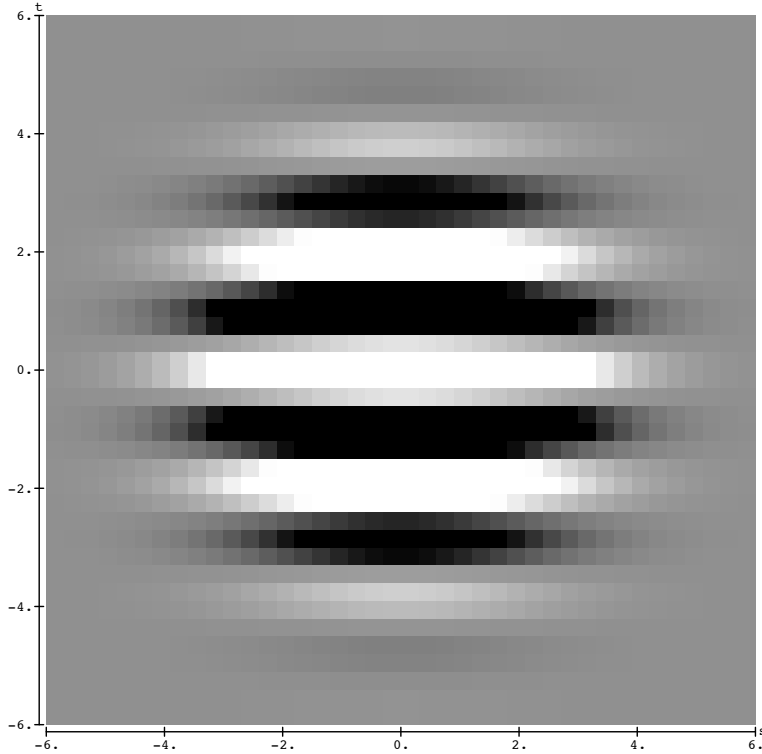


Figure 6.26: Slice through even (cosine) component of 3D Gabor elementary function oriented parallel to the time axis ($\mathbf{u} = 0$). Such a filter is selective for a spatially uniform intensity oscillating at frequency w , localized in both space and time. This function has $\alpha^2 + \beta^2 = \gamma^2 = 20$, $u = v = 0$, $w = 1/2$, $\mathbf{p} = 0$ and $r = 0$.

orthogonal. Indeed we argued that a nonorthogonal set of elementary functions might be preferable in a biological context. Finally we argued that since vision must be understood in terms of images evolving in time, the appropriate representational primitives are 3D Gabor functions. We suggested that these functions could explain the selectivity for moving edges exhibited by complex and hypercomplex cells in visual cortex, and we suggested that the alpha rhythm may correspond to the interval at which the Gabor coefficients are computed. More concrete predictions will depend on finding empirical data to constrain the parameters of the Gabor elementary functions.

6.12 Appendix: Proofs

6.12.1 Proof of General Gabor Uncertainty Principle

We prove a general Gabor Uncertainty Principle for n -dimensional functions.²⁸ Let ϕ be a function and Φ its Fourier transform; for convenience we assume $\|\phi\| = \|\Phi\| = 1$ (this is just a matter of units). We will also assume that ϕ decays to 0 at infinity; specifically we assume $s_k \phi^2(s_1, \dots, s_n) \rightarrow 0$ for all k . For this proof we will not be able to use the simple definition of nominal spread that we used for 1D signals; instead we define the nominal spread of a signal along the k th axis to be its standard deviation along that axis. Thus the spread along the k th axis is given by

$$\Delta s_k = \sqrt{\text{Var}_k\{\phi\}},$$

where Var_k is the variance along the k th axis. Similarly, in the spectral domain we define,

$$\Delta \sigma_k = \sqrt{\text{Var}_k\{\Phi\}}.$$

The variances are given by:

$$\begin{aligned} \text{Var}_k\{\phi\} &= \|s_k \phi\|^2 = \int_{\mathcal{S}} |s_k \phi|^2 d\mathbf{s} = \int_{\mathcal{S}} \phi s_k^2 \phi^* d\mathbf{s}, \\ \text{Var}_k\{\Phi\} &= \|\sigma_k \Phi\|^2 = \int_{\Sigma} |\sigma_k \Phi|^2 d\boldsymbol{\sigma} = \int_{\Sigma} \Phi \sigma_k^2 \Phi^* d\boldsymbol{\sigma}, \end{aligned}$$

²⁸The proof is a generalization of that in Hamming (1989, pp. 181–184), which is based on Gabor (1946), which is in turn based on the Heisenberg-Weyl derivation of the uncertainty principle in physics. We have already proved a more general uncertainty principle (Prop. 5.2.22, p. 93), but the following proof is more informative in the specific application of Gabor functions.

where $\mathcal{S} = \Sigma = \mathbb{R}^n$, $\mathbf{s} = (s_1, \dots, s_n) \in \mathcal{S}$ and $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n) \in \Sigma$. Our goal will be to show

$$\text{Var}_k\{\phi\} \text{Var}_k\{\Phi\} \geq \frac{1}{16\pi^2}.$$

By the Schwartz inequality we know

$$\|s_k\phi\|^2 \left\| \frac{\partial\phi}{\partial s_k} \right\|^2 \geq \left\langle s_k\phi, \frac{\partial\phi}{\partial s_k} \right\rangle^2, \quad (6.22)$$

where the bracketed expression on the right is an inner product. Since the Fourier transform is an isometry, it preserves the norm, so the norm of $\partial\phi / \partial s_k$ is the same as the norm of its Fourier transform, which is $2\pi i\sigma_k\Phi$. Therefore we can rewrite the left-hand side of Eq. 6.22 as follows:

$$\begin{aligned} \|s_k\phi\|^2 \|\partial\phi / \partial s_k\|^2 &= \|s_k\phi\|^2 \|2\pi i\sigma_k\Phi\|^2 \\ &= 4\pi^2 \|s_k\phi\|^2 \|\sigma_k\Phi\|^2 \\ &= 4\pi^2 \text{Var}_k\{\phi\} \text{Var}_k\{\Phi\}. \end{aligned} \quad (6.23)$$

Now we work on the right-hand side of Eq. 6.22:

$$\begin{aligned} \left\langle s_k\phi, \frac{\partial\phi}{\partial s_k} \right\rangle &= \int_{\mathcal{S}} s_k\phi \frac{\partial\phi^*}{\partial s_k} d\mathbf{s} \\ &= \int_{\mathcal{S}-\mathbb{R}} \int_{\mathbb{R}} s_k\phi \frac{\partial\phi^*}{\partial s_k} ds_k ds' \end{aligned} \quad (6.24)$$

where $\mathbf{s}' = (s_1, \dots, s_{k-1}, s_{k+1}, \dots, s_n) \in \mathcal{S} - \mathbb{R}$. We apply integration by parts to the innermost integral ($U = s_k\phi$, $V = \phi^*$):

$$\begin{aligned} \int s_k\phi \frac{\partial\phi^*}{\partial s_k} ds_k &= \int s_k\phi d\phi^* \\ &= s_k\phi\phi^* \Big|_{-\infty}^{\infty} - \int \phi^* d(s_k\phi) \\ &= s_k|\phi|^2 \Big|_{-\infty}^{\infty} - \int \phi^* s_k d\phi - \int \phi^* \phi ds_k. \end{aligned}$$

By our assumption that $s_k\phi^2 \rightarrow 0$ we know $s_k|\phi|^2 \Big|_{-\infty}^{\infty} = 0$, so

$$\int s_k\phi d\phi^* = - \int s_k\phi^* d\phi - \int \phi^* \phi ds_k$$

and so

$$\int s_k \phi \frac{\partial \phi^*}{\partial s_k} ds_k = -\frac{1}{2} \int \phi^* \phi ds_k.$$

Substituting this into Eq. 6.24 yields

$$\begin{aligned} \left\langle s_k \phi, \frac{\partial \phi}{\partial s_k} \right\rangle &= \int_{\mathcal{S}-\mathbb{R}} \left(-\frac{1}{2} \int \phi^* \phi ds_k \right) ds' \\ &= -\frac{1}{2} \int_{\mathcal{S}} \phi^* \phi ds \\ &= -1/2, \end{aligned} \tag{6.25}$$

since ϕ is normalized (by assumption). Therefore, combining Eqs. 6.22, 6.23 and 6.25,

$$4\pi^2 \text{Var}_k\{\phi\} \text{Var}_k\{\Phi\} \geq 1/4,$$

and so,

$$\text{Var}_k\{\phi\} \text{Var}_k\{\Phi\} \geq 1/16\pi^2.$$

We have proved the general Uncertainty Principle,

$$\Delta s_k \Delta \sigma_k \geq 1/4\pi.$$

6.12.2 Proof of Optimality of General Gabor Elementary Functions

Our task is to show that the n -dimensional Gabor elementary functions achieve the minimum area in $2n$ -dimensional Fourier space. Thus we must show

$$\frac{\|s_k \phi\|^2 \|\sigma_k \Phi\|^2}{\|\phi\|^2 \|\Phi\|^2} = \frac{1}{16\pi^2}. \tag{6.26}$$

Without loss of generality we assume that ϕ and Φ are centered at the origin (since this won't alter their variance). Notice that both functions can be written as a *cisoid* (complex sinusoid) times a product of Gaussians:

$$\phi(\mathbf{s}) = \exp(2\pi i \mathbf{s} \cdot \mathbf{u}) \prod_{j=1}^n \exp(-\pi s_j^2 / \alpha_j^2), \tag{6.27}$$

$$\Phi(\boldsymbol{\sigma}) = \exp(2\pi i \mathbf{p} \cdot \boldsymbol{\sigma}) \prod_{j=1}^n \exp(-\pi \sigma_j^2 \alpha_j^2) \tag{6.28}$$

When we compute the norms and variances of these functions the periodic parts can be ignored, since they have a constant modulus 1. Therefore,

$$\begin{aligned}\|\phi\|^2 &= \int \cdots \int \left\{ \prod_{j=1}^n \exp(-2\pi s_j^2/\alpha_j^2) \right\} ds_1 \cdots ds_n \\ &= \prod_{j=1}^n \int \exp(-2\pi s_j^2/\alpha_j^2) ds_j.\end{aligned}\quad (6.29)$$

The integrands are Gaussians, so to understand their structure better, rewrite them as normal distributions:

$$\exp(-2\pi s_j^2/\alpha_j^2) = \frac{\alpha_j}{\sqrt{2}} \left\{ \frac{1}{\sqrt{2\pi}(\alpha_j/2\sqrt{\pi})} \exp\left[-\frac{s_j^2}{2(\alpha_j^2/4\pi)}\right] \right\}.$$

The expression in curly braces is a normal distribution with mean = 0 and variance $\sigma^2 = \alpha_j^2/4\pi$. Therefore rewrite it $N_\sigma(s_j)$:

$$\exp(-2\pi s_j^2/\alpha_j^2) = \frac{\alpha_j}{\sqrt{2}} N_\sigma(s_k).$$

Since N_σ is a probability distribution, $\int N_\sigma(s_j) ds_j = 1$, so from Eq. 6.29, $\|\phi\|^2$ is the product of the normalization factors $\alpha_j/\sqrt{2}$:

$$\|\phi\|^2 = \frac{\prod_j \alpha_j}{2^{n/2}},$$

which is the first formula we need.

Next consider the variance of ϕ . It too can be rewritten, as a product of Gaussians and a quadratic factor:

$$\begin{aligned}\|s_k \phi\|^2 &= \int \cdots \int s_k^2 \exp[-2\pi(s_1^2/\alpha_1^2 + \cdots + s_n^2/\alpha_n^2)] ds_1 \cdots ds_n \\ &= \int s_k^2 \exp(-2\pi s_k^2/\alpha_k^2) ds_k \times \prod_{j \neq k} \int \exp(-2\pi s_j^2/\alpha_j^2) ds_j.\end{aligned}$$

These can be rewritten in terms of normal distributions:

$$\|s_k \phi\|^2 = \frac{\alpha_k}{\sqrt{2}} \int s_k^2 N_\sigma(s_k) ds_k \times \prod_{j \neq k} \frac{\alpha_j}{\sqrt{2}} \int N_\sigma(s_j) ds_j.$$

The first integral is the variance of the normal distribution, which we saw to be $\sigma^2 = \alpha_k^2 / 4\pi$; the remaining integrals are again 1, so

$$\begin{aligned}\|s_k\phi\|^2 &= \frac{\alpha_k}{\sqrt{2}} \frac{\alpha_k^2}{4\pi} \prod_{j \neq k} \frac{\alpha_j}{\sqrt{2}} \\ &= \frac{\alpha_k^2}{4\pi} \frac{\prod \alpha_j}{2^{n/2}}.\end{aligned}$$

Hence we see that the normalized variance of ϕ around the k th axis is

$$\frac{\|s_k\phi\|^2}{\|\phi\|^2} = \frac{\alpha_k^2 2^{-n/2} \prod \alpha_j}{4\pi 2^{-n/2} \prod \alpha_j} = \frac{\alpha_k^2}{4\pi}.$$

Thus α_k is the standard deviation along the k th axis, scaled by $2\sqrt{\pi}$:

$$\Delta s_k = \frac{\alpha_k}{2\sqrt{\pi}}. \quad (6.30)$$

Exactly the same analysis can be applied to the Fourier transform Φ , except that the variance of the normal distribution N_σ is $\sigma^2 = 1 / 4\pi\alpha_j^2$ and the normalization factor is $1 / \sqrt{2}\alpha_j$. Hence,

$$\begin{aligned}\|\Phi\|^2 &= \frac{1}{2^{n/2} \prod_j \alpha_j}, \\ \|\sigma_k\Phi\|^2 &= \frac{1}{4\pi\alpha_k^2} \frac{1}{2^{n/2} \prod_j \alpha_j}.\end{aligned}$$

Hence,

$$\frac{\|\sigma_k\Phi\|^2}{\|\Phi\|^2} = \frac{1}{4\pi\alpha_k^2},$$

and we see that the standard deviation along the k th spectral axis is α_k^{-1} , scaled by $2\sqrt{\pi}$.

Now the product of the variances in the space-time and spectral domains is easy to compute:

$$\frac{\|s_k\phi\|^2}{\|\phi\|^2} \frac{\|\sigma_k\Phi\|^2}{\|\Phi\|^2} = \frac{\alpha_k^2}{4\pi} \frac{1}{4\pi\alpha_k^2} = \frac{1}{16\pi^2}$$

and we see that the n -dimensional Gabor elementary functions achieve the minimum area given by the uncertainty principle.

Chapter 7

Basic Concepts of Field Computation

7.1 Assumptions

7.1.1 Notational conventions

To improve readability, field computation employs a number of notational conventions. They cannot be followed with absolute consistency, in part because we try to maintain consistency with established mathematical notation. Fields themselves are usually represented by Greek letters, both uppercase and lowercase. In general we use lowercase letters (ϕ, χ, ψ , etc.), but when we are dealing with fields of different dimension, we usually use uppercase letters (K, Φ, X, Ψ , etc.) for the higher dimensional fields. On the other hand Ω (possibly with subscripts) is used for the domains of sets of fields, and $\Phi(\Omega)$ is reserved for the set of fields over that domain. Lowercase italic letters (a, b, u, v, w, x, y, z , etc.) are used for scalars and the elements of field domains; we especially use r, s, t for the latter purpose. We use j, k, l, m, n for integers, but reserve i for $\sqrt{-1}$. Uppercase italic letters (F, G, T, U , etc.) are used for arbitrary field transformations, with L, M , and N generally used for linear operators. Application of a linear operator L to a field may be written $L\phi$ or $L(\phi)$. Calligraphic letters ($\mathcal{C}, \mathcal{H}, \mathcal{K}, \mathcal{L}$, etc.) are usually used for spaces. (For other notation, see Appendix A, p. 219.)

7.1.2 Fields

7.1.2.1 ASSUMED PROPERTIES OF FIELDS

Physical Realizability By *physical realizability* we mean that all fields must ultimately be represented in some physical medium (e.g., electrical potential, light intensity, chemical concentration). This places certain constraints on the fields with which we must deal. For example, fields must occupy a finite amount of space, otherwise they will not fit in our field computers. Second, the dynamic range of the fields' values are limited; concentrations, intensities, and so forth cannot be arbitrarily large. Third, a field's values vary continuously with their location in the field. Physical media will not support an infinite gradient, that is, a discontinuous change in value.

As we will see, these physical constraints translate into mathematical properties that helpfully limit the class of fields with which we must deal. On the other hand, the analysis is sometimes simplified by assuming the existence of unrealizable fields (e.g., Dirac deltas). In these cases we must be careful that the results of such analysis apply to realizable fields (see Sections 8.1.2.5 and 8.1.2.7 for examples).

Fields are Functions We treat fields as functions ϕ from a domain Ω to a codomain K , that is, $\phi : \Omega \rightarrow K$. Thus, at each point $t \in \Omega$ the field ϕ has a value $\phi(t) \in K$, which we will often write ϕ_t . The domain Ω is a metric space (Sec. 2.1). The codomain K is a subset of an algebraic field. Usually it will be some closed interval of the real numbers, but fields whose values are bounded subsets of the complex numbers are useful in some applications. In the remainder of this section we discuss constraints on the allowable domains Ω , on the codomains K , and on the functions $\phi : \Omega \rightarrow K$. Our goal will be to define $\Phi_K(\Omega)$, the space of all K -valued fields over a domain Ω .¹

Fields Belong to Linear Spaces Fields belong to *linear spaces*, which means that we can define an addition $\phi + \psi$ and scalar multiplication $a\phi$ on them that satisfy the usual properties (Sec. 3.1), which are satisfied by obvious definition of the operations:

$$\begin{aligned}(\phi + \psi)_t &= \phi_t + \psi_t, \\ (a\phi)_t &= a(\phi_t).\end{aligned}$$

¹We suppress K when it is clear from context, and write $\Phi(\Omega)$.

Notice, however, that a space of fields $\Phi(\Omega)$ cannot be closed under these operations, since that would violate our physical realizability constraints (specifically, that the dynamic range be bounded).

7.1.2.2 DOMAINS AND RANGES

Field Domains are Measure Spaces The domain Ω over which a field is defined is assumed to be a measure space. In most cases the domain will in fact be some closed and bounded subset of a Euclidean space E_n . For example, it might be a finite line segment to represent the frequency of a sound, or a closed disk to represent the light intensity over a retina. Why then have we gone to the extra generality of measure spaces?

There are two reasons. The first is that by keeping the theory general, we ensure that it applies to a wide variety of fields. At this stage it is difficult to anticipate the “shapes” we may need for our fields.

The second reason is that a finite set (when provided with a “weight” function) is a measure space. Thus, although it is our goal to encourage thinking of fields as continuously varying structures, most of the theory in fact applies to fields composed of discrete elements (such as pixel arrays and neural networks).

For our purposes, one of the most important properties of measure spaces is that we can define integration operations over them. For example, if $\phi : \Omega \rightarrow K$, then we can define the (Lebesgue) integral $\int_{\Omega} \phi_t d\mu(t)$. Here μ is the measure of the measure space. Since it is usually clear from context, we will write integrals more simply: $\int_{\Omega} \phi_t dt$. Also, note that for finite Ω and equal weighting, integration reduces to summation:

$$\int_{\Omega} \phi_t dt = \sum_{t \in \Omega} \phi_t. \quad (7.1)$$

The Domain of a Measure Space is Bounded As discussed in Section 7.1.2.1 we require that fields occupy a finite amount of space. How can this be expressed mathematically? We simply require that the “area” (volume, length) of the space be bounded, by stating:

$$\int_{\Omega} 1 dt < \infty. \quad (7.2)$$

Later (Section 7.3.1) we will see that this can be expressed in the more compact form $\|1\| < \infty$. For convenience we define the “size” of the space Ω

by:

$$\|\Omega\| = \int_{\Omega} 1 dt. \quad (7.3)$$

Hence we require that $\|\Omega\| < \infty$.

The Range of a Field is Bounded We assume that every field $\phi : \Omega \rightarrow K$ has a bound β_{ϕ} such that $|\phi_t| \leq \beta_{\phi}$ for all $t \in \Omega$. Furthermore, as mentioned above (Section 7.1.2.1), we require that a space of fields have limited dynamic range. Thus, for each space $\Phi_K(\Omega)$ there is a β such that $\beta_{\phi} \leq \beta$ for all $\phi \in \Phi(\Omega)$. Such a β is given by $\beta = \sup_{x \in K} |x| < \infty$.

7.1.2.3 BOUNDED GRADIENT

Fields are Continuous As discussed under Physical Realizability (Section 7.1.2.1) we assume that all fields are continuous (Sec. 2.4). That is, for all $t, t' \in \Omega$ and every $\epsilon > 0$ there is a $\delta > 0$ such that $|\phi_t - \phi_{t'}| < \epsilon$ whenever $d(t, t') < \delta$. Here $d(t, t')$ is the distance between the points t and t' (recall, Section 7.1.2.1, that Ω is a metric space).

Uniform Continuity As discussed in Section 7.1.2.1 we assume that there is a maximum gradient that physical fields can sustain. This implies, first of all, that fields are *uniformly continuous*, which means that for all $\epsilon > 0$ there is a δ such that $|\phi_t - \phi_{t'}| < \epsilon$ whenever $d(t, t') < \delta$. That is, for any change δ in the field's domain, there is a maximum amount ϵ that the field's value can change.

Lipschitz Condition In fact, most fields satisfy a stronger condition than uniform continuity. Specifically, there is a bounded ratio of the change of the field's value to a change in position in the field:

$$\frac{|\phi_t - \phi_{t'}|}{d(t, t')} \leq \gamma. \quad (7.4)$$

This means that such fields satisfy a *Lipschitz condition (of order 1)*.

Fields are Band Limited It is generally reasonable to assume that fields are band limited, that is that they have a frequency limit beyond which there is no information. Higher frequency variation is either physically impossible,

or represents noise that should be ignored. This means that field transformations may be assumed to have a low-pass filter on their inputs that does not affect their operation. This has important consequences for the theoretical development (Section 9.2.2.2).

7.1.3 Field Transformations

Continuity The noise that accompanies physical processes will cause slight variations in the inputs to field transformations. It is undesirable if this noise has a major effect on the output of the transformation. Therefore, we require that field transformations be *continuous*; this will ensure that small changes in the input will cause at most small changes in the output.

Using the norm that will be defined in Section 7.3.1, the continuity of field transformations can be expressed:

$$\lim_{n \rightarrow 0} \|\phi_n - \phi\| = 0 \text{ implies } \lim_{n \rightarrow 0} \|T(\phi_n) - T(\phi)\| = 0. \quad (7.5)$$

We can generally make the stronger assumption of uniform continuity: for every $\epsilon > 0$ there is a $\delta > 0$ such that $\|T(\phi) - T(\psi)\| < \epsilon$ whenever $\|\phi - \psi\| < \delta$. In effect, we are assuming that a field transformation is “uniformly insensitive” to a given quantity of noise, no matter what the input field.

Input Filters As noted above (Section 7.1.2.3), noise often manifests itself as high frequency variation in the field. Therefore, to decrease the effects of noise, field transformation units will often be constructed with a low-pass filter on their input. In other cases, the implementation will naturally suppress high frequencies, and thus behave as though the input is filtered. The foregoing permits us to assume the presence of low pass input filters on field transformations when this simplifies the mathematical analysis (see for example Sections 7.4.3.2 and 7.4.3.3).

7.2 Summary

Since fields have bounded domain (Section 7.1.2.2) and range (Section 7.1.2.2), the integral $\int_{\Omega} |\phi_t|^2 dt$ exists:

$$\int_{\Omega} |\phi_t|^2 dt \leq \beta_{\phi}^2 \int_{\Omega} 1 dt \leq \beta_{\phi}^2 \|\Omega\|. \quad (7.6)$$

We will see below (Section 7.3.1) that this is equivalent to saying that all fields have finite norms.

Since fields are also continuous (Section 7.1.2.3), we know that fields over Ω belong to the function space $L_2(\Omega)$. This is very important, because L_2 function spaces are Hilbert spaces (Sec. 5.2), which means that a large body of powerful mathematics can be brought to bear on the problems of field computation.

We cannot conclude, however, that $\Phi(\Omega) = L_2(\Omega)$, since $L_2(\Omega)$ contains many functions that do not satisfy our other constraints (e.g., Sections 7.1.2.2 and 7.1.2.3). In most cases these additional constraints will help us. Nevertheless, we must be careful that in applying properties of L_2 spaces we do not violate these other constraints.

7.3 Functional analysis preliminaries

7.3.1 Definitions

Norm of Field The norm of a field is a measure of its size. As usual, we use the inner product norm (Sec. 5.1.2):

$$\|\phi\| = \sqrt{\phi \cdot \phi}. \quad (7.7)$$

This is the usual norm for L_2 function spaces (Section 7.2).² Note that since fields have bounded domain and range (Section 7.1.2.2), we know that $\|\phi\| \leq \beta_\phi \|\Omega\|^{1/2}$ (see Section 7.2). That is, all fields are finite in size. As noted before (Section 7.2), this means that fields belong to Hilbert spaces.

7.3.2 Orthonormal bases

A *unit field* $\xi \in \Phi(\Omega)$ is a field whose norm is 1, $\|\xi\| = 1$. Since $\Phi(\Omega) \subset \mathcal{H}(\Omega)$, we know it has an orthonormal (ON) basis, ξ_0, ξ_1, \dots . We will sometimes refer to basis elements as being *higher order* or *lower order*. This refers to their location in the sequence ξ_0, ξ_1, \dots . Of course this ordering is somewhat arbitrary, and may be changed without affecting the sequence's status as an ON basis. Nevertheless, there is usually a *natural* order, such as increasing

²Although it is not the only norm upon which a theory of field computation can be based. In MacLennan (1987a,b) we used the L_1 norm. There seems to be little practical difference, but the L_2 is mathematically more convenient.

frequency for the trigonometric basis, and increasing degree for polynomial bases.

7.4 Basic field transformations

7.4.1 Definition of general field product

It is useful to define a generalization of the inner product that is continuous-dimensional analogue of products between vectors and matrices. The reason is that these products seem to be the kind of operations that we can expect general purpose field computers to compute. For example, it is well known that linear neural networks compute vector-matrix products (Rumelhart et al. 1986, Chapter 9; see also Ch. 12 below); similarly, these operations correspond to some optical phenomena (see for example Farhat et al. 1985). These considerations, together with the theoretical developments described later, lead us to base general purpose field computation on these general field products.³

Definition 7.4.1 (general field product 1) *If $\phi \in \Phi(\Omega_2)$ and $\Psi \in \Phi(\Omega_1 \times \Omega_2)$, then the product $\Psi\phi \in \Phi(\Omega_1)$ is defined:*

$$(\Psi\phi)_s = \int_{\Omega_2} \Psi_{st}\phi_t dt. \quad (7.8)$$

Remark 7.4.1 *If we let Ψ_s be the field $(\Psi_s)_t = \Psi_{st}$, and ϕ^* be the conjugate field $(\phi^*)_t = \overline{\phi_t}$ then the product can be defined:*

$$(\Psi\phi)_s = \Psi_s \cdot \phi^* = \phi \cdot \Psi_s^*. \quad (7.9)$$

(This is because is inner product must conjugate one of the fields, but the field product does not.)

Remark 7.4.2 *If the fields are complex-valued, then it is often convenient to write the product $\Psi\phi$ with Dirac's bra-ket notation (Not. 5.1.2, p. 72) as the ket $\Psi|\phi\rangle$. If we let $|\Psi_s\rangle$ be the field $|\Psi_s\rangle_t = \Psi_{st}$, then the product can be defined:*

$$(\Psi|\phi\rangle)_s = \langle\phi^* | \Psi_s\rangle = \langle\Psi_s^* | \phi\rangle. \quad (7.10)$$

³Note that the scalar, inner and outer products (see Sect. 7.4.1) can all be considered degenerate general products.

(Note that $\Psi|\phi\rangle$ represents a field product, not the application of a linear operator Ψ to $|\phi\rangle$.)

Definition 7.4.2 (general field product 2) If $\phi \in \Phi(\Omega_1)$ and $\Psi \in \Phi(\Omega_1 \times \Omega_2)$, then the product $\phi\Psi \in \Phi(\Omega_2)$ is defined:

$$(\phi\Psi)_t = \int_{\Omega_1} \phi_s \Psi_{st} ds. \tag{7.11}$$

Remark 7.4.3 For complex fields we may use Dirac's notation, $\langle\phi^*|\Psi$. If we define the transpose $|\Psi_t^\top\rangle$ so that $|\Psi_t^\top\rangle_s = \Psi_{st}$, then then we can write

$$(\langle\phi^*|\Psi)_t = \langle\phi^* | \Psi_t^\top\rangle. \tag{7.12}$$

Definition 7.4.3 (general field product 3: quadratic form) If $\phi \in \Phi(\Omega_1)$, $X \in \Phi(\Omega_1 \times \Omega_2)$, and $\psi \in \Phi(\Omega_2)$, then

$$\phi X \psi = \int_{\Omega_1} \int_{\Omega_2} \phi_s X_{st} \psi_t dt ds. \tag{7.13}$$

In Dirac's notation,

$$\phi X \psi = \langle\phi^* | X | \psi\rangle. \tag{7.14}$$

Definition 7.4.4 (general field product 4) Finally, if $\Psi \in \Phi(\Omega_1 \times \Omega_2)$ and $X \in \Phi(\Omega_2 \times \Omega_3)$, then we define the product $\Psi X \in \Phi(\Omega_1 \times \Omega_3)$ as follows:

$$(\Psi X)_{su} = \int_{\Omega_2} \Psi_{st} X_{tu} dt. \tag{7.15}$$

Any linear operator L that can be expressed in the form $L\phi = \Psi\phi$, for some field Ψ , is called an *integral operator*. Indeed, since we require $\|\Psi\| < \infty$, it is an integral operator of *Hilbert-Schmidt type* (Sec. 5.2.9). We will be interested in expressing derivatives and other linear transformations as integral operators (see Sections 7.4.3 and 9.2.2).

7.4.2 Outer product

7.4.2.1 DEFINITION

Definition 7.4.5 (field outer product) For two fields $\phi \in \Phi(\Omega_1)$, $\psi \in \Phi(\Omega_2)$ the field outer product $\phi \wedge \psi \in \Phi(\Omega_1 \times \Omega_2)$ is defined:

$$(\phi \wedge \psi)_{st} = \phi_s \psi_t. \tag{7.16}$$

Remark 7.4.4 *Since physically realizable fields are bounded (Section 7.1.2.2), their outer product always exists.*

Proposition 7.4.1 *If the fields ζ_j are a basis for $\Phi(\Omega_1)$ and the η_k are a basis for $\Phi(\Omega_2)$, then the fields $\zeta_j \wedge \eta_k$ are a basis for $\Phi(\Omega_1 \times \Omega_2)$.*

There is obviously a close connection between field outer products and tensor products (Sec. 5.2.10).

Proposition 7.4.2 *The spaces $\Phi(\Omega_1 \times \Omega_2)$ and $\Phi(\Omega_1) \otimes \Phi(\Omega_2)$ are isomorphic.*

Remark 7.4.5 *The tensor product $\phi \otimes \psi$ and the outer product $\phi \wedge \psi$ are effectively interchangeable. They satisfy*

$$\langle \phi \otimes \psi \mid \phi' \otimes \psi' \rangle = \langle \phi \wedge \psi \mid \phi' \wedge \psi' \rangle \quad (7.17)$$

Remark 7.4.6 *The Dirac outer product or dyad (Def. 5.2.16, p. 89) combines fields $|\phi\rangle$ and $|\psi\rangle$ to yield a linear operator $|\phi\rangle\langle\psi| : \Phi(\Omega_2) \rightarrow \Phi(\Omega_1)$ for which*

$$|\phi\rangle\langle\psi| |\zeta\rangle = |\phi\rangle \langle\psi \mid \zeta\rangle \quad (7.18)$$

for all $|\zeta\rangle \in \Phi(\Omega_2)$. That is, $|\phi\rangle$ is scaled by $\langle\psi \mid \zeta\rangle$. Likewise, $\langle\psi| : \Phi(\Omega) \rightarrow \mathbb{C}$ is the co-field of $|\psi\rangle \in \Phi(\Omega)$ in the continuous dual space of $\Phi(\Omega)$. Since $|\phi\rangle\langle\psi|$ is not associative, we have found it less useful in field computation than the field outer product, but the two are closely related.

Proposition 7.4.3 *The kernel of $|\phi\rangle\langle\psi|$ is $\phi \wedge \psi^\dagger$. Therefore we have the isomorphisms*

$$\phi \wedge \psi \cong |\phi\rangle\langle\psi^\dagger| \cong |\phi\rangle|\psi\rangle, \quad (7.19)$$

where the rightmost product is the tensor product $|\phi\rangle \otimes |\psi\rangle$.

7.4.2.2 USEFUL PROPERTIES

We present a few simple properties of the outer products.

Theorem 7.4.1 *If $\phi \in \Phi(\Omega_1)$ and $\psi, \zeta \in \Phi(\Omega_2)$, then*

$$(\phi \wedge \psi)\zeta = \phi(\psi \cdot \zeta^*) = \phi(\zeta \cdot \psi^*). \quad (7.20)$$

In other words, ϕ is scaled by $\psi \cdot \zeta^ = \zeta \cdot \psi^*$.*

Proof: Simply expand the product. □

Corollary 7.4.1 *If $\phi, \zeta \in \Phi(\Omega_1)$ and $\eta \in \Phi(\Omega_2)$, then*

$$\phi(\zeta \wedge \eta) = (\phi \cdot \zeta^*)\eta = (\zeta \cdot \phi^*)\eta. \quad (7.21)$$

In other words, η is scaled by $\phi \cdot \zeta^ = \zeta \cdot \phi^*$. The analogous result for the dyad or Dirac outer product is:*

$$\langle \phi | | \zeta \rangle \langle \eta | = \langle \phi | | \zeta \rangle \langle \eta |. \quad (7.22)$$

That is, $\langle \eta |$ is scaled by $\langle \phi | | \zeta \rangle$.

Theorem 7.4.2 *If $X \in \Phi(\Omega_0 \times \Omega_1)$ and $\phi, \psi \in \Phi(\Omega_2)$, then*

$$(X \wedge \phi)\psi = X(\phi \cdot \psi^*) = X(\psi \cdot \phi^*) = X\langle \phi^* | \psi \rangle = X\langle \psi^* | \phi \rangle. \quad (7.23)$$

Exercise 7.4.1 *Prove this theorem by expanding products into integrals.*

Theorem 7.4.3 *If $X \in \Phi(\Omega_0 \times \Omega_1)$, $\zeta \in \Phi(\Omega_1)$ and $\psi, \eta \in \Phi(\Omega_2)$, then*

$$(X \wedge \psi)(\zeta \wedge \eta) = X\zeta(\psi \cdot \eta^*) = X\zeta(\eta \cdot \psi^*) = X\zeta\langle \psi^* | \eta \rangle = X\zeta\langle \eta^* | \psi \rangle. \quad (7.24)$$

Proof: First expand the product as an integral over the direct product measure space:

$$\begin{aligned} [(X \wedge \psi)(\zeta \wedge \eta)]_r &= \int_{\Omega_1 \times \Omega_2} (X \wedge \psi)_{rst} (\zeta \wedge \eta)_{st} d\mu(s, t) \\ &= \int_{\Omega_1 \times \Omega_2} X_{rs} \psi_t \zeta_s \eta_t d\mu(s, t). \end{aligned}$$

Next apply Fubini's theorem, which says that the multiple integral equals the iterated integral:

$$\int_{\Omega_1 \times \Omega_2} X_{rs} \psi_t \zeta_s \eta_t d\mu(s, t) = \int_{\Omega_2} \left[\int_{\Omega_1} X_{rs} \psi_t \zeta_s \eta_t d\mu(s) \right] d\mu(t). \quad (7.25)$$

Finally, regroup and remove from under the integral sign factors not dependent on the variable of integration:

$$\begin{aligned} &= \int_{\Omega_1} X_{rs} \zeta_s d\mu(s) \int_{\Omega_2} \psi_t \eta_t d\mu(t) \\ &= (X\zeta)_r (\psi \cdot \eta^*) \\ &= [X\zeta(\psi \cdot \eta^*)]_r. \end{aligned}$$

□

Theorem 7.4.4 *If $\phi, \zeta \in \Phi(\Omega_1)$ and $\psi, \eta \in \Phi(\Omega_2)$, then*

$$(\phi \wedge \psi) \cdot (\zeta \wedge \eta) = (\phi \cdot \zeta)(\eta \cdot \psi). \quad (7.26)$$

In Dirac's notation:

$$\langle \phi \wedge \psi \mid \zeta \wedge \eta \rangle = \langle \phi \mid \zeta \rangle \langle \psi \mid \eta \rangle. \quad (7.27)$$

In both cases, note the order of the inner products, which is significant if the fields are complex-valued. Also, Dirac's notation shows that we can simply multiply $\langle \phi \mid$, the dyad $\mid \zeta \rangle \langle \psi \mid$, and $\mid \eta \rangle$.

Exercise 7.4.2 *Prove this theorem by expanding the inner and outer products.*

Corollary 7.4.2 $\|\phi \wedge \psi\| = \|\phi\| \|\psi\|$.

Proof:

$$\begin{aligned} \|\phi \wedge \psi\|^2 &= (\phi \wedge \psi) \cdot (\phi \wedge \psi) \\ &= (\phi \cdot \phi)(\psi \cdot \psi) \\ &= \|\phi\|^2 \|\psi\|^2. \end{aligned}$$

□

Corollary 7.4.3 *If $X \in \Phi(\Omega)$ and $\phi_k, \psi_k \in \Phi(\Omega_k)$, $k = 1, \dots, n$, then*

$$(X \wedge \phi_1 \wedge \dots \wedge \phi_n)(\psi_1 \wedge \dots \wedge \psi_n) = X(\phi_1 \cdot \psi_1^*) \dots (\phi_n \cdot \psi_n^*) \quad (7.28)$$

$$= X \langle \phi_1^* \mid \psi_1 \rangle \dots \langle \phi_n^* \mid \psi_n \rangle. \quad (7.29)$$

Proof: Apply Thm. 7.4.1 inductively:

$$\begin{aligned} &(X \wedge \phi_1 \wedge \dots \wedge \phi_n)(\psi_1 \wedge \dots \wedge \psi_n) \\ &= (X \wedge \phi_1 \wedge \dots \wedge \phi_{n-1})(\psi_1 \wedge \dots \wedge \psi_{n-1})(\phi_n \cdot \psi_n^*) \\ &\quad \vdots \\ &= (X \wedge \phi_1) \psi_1 (\phi_2 \cdot \psi_2^*) \dots (\phi_n \cdot \psi_n^*) \\ &= X(\phi_1 \cdot \psi_1^*)(\phi_2 \cdot \psi_2^*) \dots (\phi_n \cdot \psi_n^*). \end{aligned}$$

□

The following theorem and its corollary show the relation between outer products and iterated general products. Note that we take the general product to be left associative: $K\phi\psi = (K\phi)\psi$.

Theorem 7.4.5 *Suppose $K \in \Phi(\Omega \times \Omega_2 \times \Omega_1)$, $\phi \in \Omega_1$ and $\psi \in \Omega_2$. Then:*

$$K\phi\psi = K(\psi \wedge \phi). \tag{7.30}$$

Proof: Simply expand the general products as integrals:

$$\begin{aligned} (K\phi\psi)_r &= \int_{\Omega_2} (K\phi)_{rs}\psi_s ds \\ &= \int_{\Omega_2} \int_{\Omega_1} K_{rst}\phi_t dt \psi_s ds \\ &= \int_{\Omega_2} \int_{\Omega_1} K_{rst}\psi_s\phi_t ds dt \\ &= \int_{\Omega_2} \int_{\Omega_1} K_{rst}(\psi \wedge \phi)_{st} ds dt. \end{aligned}$$

By Fubini's theorem (see proof of Theorem 7.4.2.2) the iterated integral may be replaced by the multiple integral over the direct product space:

$$\begin{aligned} &= \int_{\Omega_2 \times \Omega_1} K_{rst}(\psi \wedge \phi)_{st} d(s, t) \\ &= [K(\psi \wedge \phi)]_r. \end{aligned}$$

□

Corollary 7.4.4 *Suppose $K \in \Phi(\Omega \times \Omega_n \times \cdots \times \Omega_1)$ and $\phi_k \in \Phi(\Omega_k)$, for $1 \leq k \leq n$. Then:*

$$K\phi_1\phi_2 \cdots \phi_n = K(\phi_n \wedge \cdots \wedge \phi_2 \wedge \phi_1). \tag{7.31}$$

Proof: An inductive application of the theorem.

□

7.4.3 Kernels of linear and multilinear operators

7.4.3.1 NEED FOR KERNELS

We explore the conditions under which a linear field transformation $L : \Phi(\Omega_1) \rightarrow \Phi(\Omega_2)$ can be expressed as a general field product, $L(\phi) = K\phi$, for some $K \in \Phi(\Omega_2 \times \Omega_1)$. This condition is equivalent to saying that L is an *integral operator (of Hilbert-Schmidt type) with kernel K* (recall Sec. 5.2.9). It is an important condition, because field products can be computed by neural networks and other massively parallel devices. The kernel is the analog of the matrix representing a linear transformation in the finite-dimensional case.

After the linear case (Section 7.4.3.2), we explore the multilinear case (i.e., the case for multi-argument operators that are linear in each of their arguments; see Section 7.4.3.3).

7.4.3.2 KERNELS OF LINEAR OPERATORS

Theorem 7.4.6 *Suppose ξ_0, ξ_1, \dots is a basis for the Hilbert space. Let*

$$K = \sum_{k=0}^{\infty} L(\xi_k) \wedge \xi_k^*. \quad (7.32)$$

If this field exists, then it is the kernel of L , $L(\phi) = K\phi$. In Dirac's notation:

$$K = \sum_{k=0}^{\infty} L(|\xi_k\rangle)\langle\xi_k|. \quad (7.33)$$

Proof: To see this, expand the Fourier series for ϕ and make use of linearity of L :

$$\begin{aligned} L(\phi) &= L \left[\sum_{k=0}^{\infty} |\xi_k\rangle \langle \xi_k | \phi \rangle \right] \\ &= \sum_k L(|\xi_k\rangle \langle \xi_k | \phi \rangle) \\ &= \sum_k L(|\xi_k\rangle) \langle \xi_k | \phi \rangle \\ &= \sum_k [L(|\xi_k\rangle) \langle \xi_k |] | \phi \rangle \end{aligned}$$

$$\begin{aligned}
 &= \left[\sum_k L(|\xi_k\rangle)\langle\xi_k| \right] |\phi\rangle \\
 &= K\phi.
 \end{aligned}$$

This completes the proof. □

The preceding result assumes that the kernel exists. Sufficient conditions are established next. It will be useful to have the following definitions:

Definition 7.4.6 *A linear transformation $F : \Phi(\Omega) \rightarrow \Phi(\Omega)$ is called a filter if the basis elements e_k are eigenvectors of the transformation. That is, there are λ_k such that $F(e_k) = \lambda_k e_k$. The sequence $\lambda_0, \lambda_1, \dots$ is called the transfer function of the filter.*

Definition 7.4.7 *A sequence λ_k is called absolutely summable if*

$$\sum_{k=0}^{\infty} |\lambda_k| < \infty. \tag{7.34}$$

Another way of saying this is that the sequence belongs to the space l_1 .

Theorem 7.4.7 *Suppose L can be written in the form $L = G \circ F$, where F is a filter whose transfer function is absolutely summable, and G is a continuous linear transformation. Then the kernel of L exists and is*

$$K = \sum_{k=0}^{\infty} L(\xi_k) \wedge \xi_k^*. \tag{7.35}$$

Proof: We show that the norm of the kernel is finite.

$$\begin{aligned}
 \left\| \sum L(\xi_k) \wedge \xi_k^* \right\| &\leq \sum \|L(\xi_k) \wedge \xi_k^*\| \\
 &= \sum \|L(\xi_k)\| \|\xi_k^*\|, \text{ by Cor. 7.4.2} \\
 &= \sum \|L(\xi_k)\|, \text{ since the } \xi_k \text{ are normalized} \\
 &= \sum \|G[F(\xi_k)]\|.
 \end{aligned}$$

Since F is a filter we know $F(\xi_k) = \lambda_k \xi_k$. Hence,

$$\begin{aligned}
&= \sum \|G(\lambda_k \xi_k)\| \\
&= \sum |\lambda_k| \|G(\xi_k)\|, \text{ since } G \text{ is linear} \\
&= \sum |\lambda_k| \beta \|\xi_k\|, \text{ since } G \text{ is continuous and hence bounded} \\
&= \beta \sum |\lambda_k| \\
&< \infty, \text{ since the } \lambda_k \text{ are absolutely summable.}
\end{aligned}$$

□

Hence, if the filter sufficiently suppresses the higher-order components of its argument, the kernel will exist. This is certainly the case when F has a sharp cutoff: $\lambda_k = 0$ for all k greater than some N .

In practice the conditions on the preceding theorem are not a problem. Since most fields are band-limited (Section 7.1.2.3) they can be written $\phi = F(\phi)$ for a filter F with a sharp cutoff. Also, since we generally want our field transformations to be insensitive to higher-order noise, it is useful to express them in the form $G \circ F$, where G is the “ideal” or “goal” transformation and F is an appropriate filter.

Finally, we show that if L can be written as a field product, then the sum-of-outer-products series converges.

Theorem 7.4.8 *Suppose that $L : \Phi(\Omega_1) \rightarrow \Phi(\Omega_2)$ satisfies $L(\phi) = K\phi$ and $\|K\| < \infty$. Then $\|\sum_k L(\xi_k) \wedge \xi_k^*\| < \infty$.*

Proof: The hypothesis amounts to supposing that L is an integral operator with Hilbert-Schmidt kernel K . From Prop. 5.2.23 (p. 93) we know that any such operator on a separable Hilbert space (such as a space of fields) can be represented by an infinite matrix

$$M_{mn} = \langle \beta_m | L\xi_n \rangle, \quad (7.36)$$

where the β_m are an ON basis for $\Phi(\Omega_2)$. In addition, this matrix satisfies

$$\sum_m \sum_n M_{mn}^2 < \infty. \quad (7.37)$$

Now compute a bound on the sum-of-outer-products field:

$$\begin{aligned}
 \left\| \sum_m L(\xi_m) \wedge \xi_m^* \right\|^2 &= \sum_m \sum_n \langle L(\xi_m) \wedge \xi_m^* \mid L(\xi_n) \wedge \xi_n^* \rangle \\
 &= \sum_m \sum_n \langle L\xi_m \mid L\xi_n \rangle \langle \xi_m^* \mid \xi_n^* \rangle \\
 &= \sum_n \|L\xi_n\|^2, \text{ since the } \xi_k \text{ are ON.}
 \end{aligned}$$

Next expand $L\xi_n$ in a Fourier series and apply Parseval's equality (Prop. 5.2.7, p. 83):

$$\begin{aligned}
 &= \sum_n \left\| \sum_m \langle \beta_m \mid L\xi_n \rangle \beta_m \right\|^2 \\
 &= \sum_n \sum_m |\langle \beta_m \mid L\xi_n \rangle|^2 \\
 &= \sum_m \sum_n M_{mn}^2 \\
 &< \infty.
 \end{aligned}$$

□

7.4.3.3 KERNELS OF MULTILINEAR OPERATORS

The results in Section 7.4.3.2 are easily extended to multilinear operators. An operator

$$M : \Phi(\Omega_1) \times \cdots \times \Phi(\Omega_n) \rightarrow \Phi(\Omega) \quad (7.38)$$

is *multilinear* if it is linear in each of its arguments. Our goal is to find a field

$$K \in \Phi(\Omega \times \Omega_n \times \cdots \times \Omega_1). \quad (7.39)$$

such that

$$M(\phi_1, \dots, \phi_n) = K\phi_1 \cdots \phi_n = K(\phi_n \wedge \cdots \wedge \phi_1). \quad (7.40)$$

That the two products above are equivalent is established in Section 7.4.2.1.

Theorem 7.4.9 *If M is a multilinear operator, then its kernel is*

$$K = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} M(\xi_{k_1}, \dots, \xi_{k_n}) \wedge \xi_{k_n}^* \wedge \cdots \wedge \xi_{k_1}^*. \quad (7.41)$$

if the sum exists.

Proof: Suppose $\phi_j \in \Phi(\Omega_j)$, $j = 1, \dots, n$, and $c_{jk} = \langle \xi_k | \phi_j \rangle$, $k = 0, \dots$. That is, c_{jk} is the k -th generalized Fourier coefficient of ϕ_j . Then, since M is multilinear we can expand:

$$\begin{aligned} M(\phi_1, \dots, \phi_n) &= M\left(\sum c_{1k}\xi_k, \dots, \sum c_{nk}\xi_k\right) \\ &= \sum_{k_1} c_{1k_1} M\left(\xi_{k_1}, \dots, \sum c_{nk}\xi_k\right) \\ &= \sum_{k_1} \cdots \sum_{k_n} c_{1k_1} \cdots c_{nk_n} M(\xi_{k_1}, \dots, \xi_{k_n}). \end{aligned}$$

At this point it will be convenient to work from the other side:

$$\begin{aligned} &K(\phi_n \wedge \cdots \wedge \phi_1) \\ &= \left[\sum_{k_1} \cdots \sum_{k_n} M(\xi_{k_1}, \dots, \xi_{k_n}) \wedge \xi_{k_n}^* \wedge \cdots \wedge \xi_{k_1}^* \right] (\phi_n \wedge \cdots \wedge \phi_1) \\ &= \sum_{k_1} \cdots \sum_{k_n} [M(\xi_{k_1}, \dots, \xi_{k_n}) \wedge (\xi_{k_n}^* \wedge \cdots \wedge \xi_{k_1}^*)] (\phi_n \wedge \cdots \wedge \phi_1) \\ &= \sum_{k_1} \cdots \sum_{k_n} [M(\xi_{k_1}, \dots, \xi_{k_n}) \langle \xi_{k_n} | \phi_n \rangle \cdots \langle \xi_{k_1} | \phi_1 \rangle], \text{ by Cor. 7.4.3} \\ &= \sum_{k_1} \cdots \sum_{k_n} M(\xi_{k_1}, \dots, \xi_{k_n}) c_{nk_n} \cdots c_{1k_1}. \end{aligned}$$

The two expansions can be seen to be equal. □

Theorem 7.4.10 *If M is a continuous multilinear operator whose inputs are filtered:*

$$M(\phi_1, \dots, \phi_n) = G(F_1\phi_1, \dots, F_n\phi_n) \quad (7.42)$$

then

$$K = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} M(\xi_{k_1}, \dots, \xi_{k_n}) \wedge \xi_{k_n}^* \wedge \cdots \wedge \xi_{k_1}^*. \quad (7.43)$$

exists and is the kernel of M , provided the transfer functions of the filters F_k are absolutely summable.

Proof: Let $\lambda^{(i)}$ be the transfer function of the i -th filter F_i ; hence $\lambda^{(i)}_1, \lambda^{(i)}_2, \dots$ are its eigenvalues. As in Theorem 7.4.7 we compute a bound on $\|K\|$.

$$\begin{aligned} \|K\| &= \left\| \sum_{k_1} \cdots \sum_{k_n} M(\xi_{k_1}, \dots, \xi_{k_n}) \wedge \xi_{k_n}^* \wedge \cdots \wedge \xi_{k_1}^* \right\| \\ &\leq \sum_{k_1} \cdots \sum_{k_n} \|M(\xi_{k_1}, \dots, \xi_{k_n}) \wedge \xi_{k_n}^* \wedge \cdots \wedge \xi_{k_1}^*\| \\ &= \sum_{k_1} \cdots \sum_{k_n} \|M(\xi_{k_1}, \dots, \xi_{k_n})\|. \end{aligned}$$

Now note that

$$\begin{aligned} M(\xi_{k_1}, \dots, \xi_{k_n}) &= G(F_1 \xi_{k_1}, \dots, F_n \xi_{k_n}) \\ &= G(\lambda^{(1)}_{k_1} \xi_{k_1}, \dots, \lambda^{(n)}_{k_n} \xi_{k_n}) \\ &= \lambda^{(1)}_{k_1} \cdots \lambda^{(n)}_{k_n} G(\xi_{k_1}, \dots, \xi_{k_n}). \end{aligned}$$

Continuing the derivation of the bound:

$$\begin{aligned} \|K\| &\leq \sum_{k_1} \cdots \sum_{k_n} |\lambda^{(1)}_{k_1} \cdots \lambda^{(n)}_{k_n}| \|G(\xi_{k_1}, \dots, \xi_{k_n})\| \\ &\leq \beta \sum_{k_1} \cdots \sum_{k_n} |\lambda^{(1)}_{k_1} \cdots \lambda^{(n)}_{k_n}|. \end{aligned}$$

The last step follows from the assumption that G is bounded (continuous) with bound β . Continuing:

$$\begin{aligned} &= \beta \sum_{k_1} \cdots \sum_{k_n} |\lambda^{(1)}_{k_1}| \cdots |\lambda^{(n)}_{k_n}| \\ &= \beta \prod_{i=1}^n \sum_k |\lambda^{(i)}_k|. \end{aligned}$$

This will be finite if each of the $\sum_k |\lambda^{(i)}_k|$ are finite.

□

7.4.3.4 KERNELS IN TERMS OF GENERALIZED FUNCTIONS

In this section we derive an alternative formula for the kernel that is often easier to use than those given in Sections 7.4.3.2 and 7.4.3.3. The new formula is easily understood through the analogy with finite-dimensional spaces. Recall that a linear transformation $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$ can be represented by a matrix-vector product $L(x) = Mx$ in which the k th column of M is $L(\delta_k)$. That is, $(M^T)_k = L(\delta_k)$. Here the δ_k represent the basis vectors defined by the *Kronecker delta functions*:

$$(\delta_k)_j = \begin{cases} 1, & \text{if } k = j \\ 0, & \text{if } k \neq j \end{cases} . \quad (7.44)$$

Now consider the infinite dimensional case and suppose L is a Hilbert-Schmidt operator, $L(\phi) = K\phi$, where K is given in Section 7.4.3.2. We claim that the t th “column” of K is $L(\delta_t)$, that is, $(K^T)_t = L(\delta_t)$. Here the δ_t are the *Dirac delta functions*:

$$\delta_t(s) = \begin{cases} \infty, & \text{if } s = t \\ 0, & \text{if } s \neq t \end{cases} . \quad (7.45)$$

Although “generalized functions” such as the Dirac delta are not physically realizable, their use often simplifies the derivation of physically realizable fields. See Section 8.1.2.5 for a further discussion.

To establish our claim we need to make use of the well-known “sifting property” of the Dirac delta:

$$\delta_t \cdot \phi = \int_{\Omega} \delta_t(s) \phi_s^* ds = \phi_t^* . \quad (7.46)$$

With this in hand it is easy to show that $(K^T)_t = L(\delta_t)$:

$$\begin{aligned} (K^T)_t &= \left[\sum_k L(\xi_k) \wedge \xi_k^* \right]_t^T \\ &= \left[\sum_k \xi_k^* \wedge L(\xi_k) \right]_t \\ &= \sum_k \xi_k^*(t) L(\xi_k) \end{aligned}$$

$$\begin{aligned}
&= L \left[\sum_k \xi_k^*(t) \xi_k \right] \\
&= L \left[\sum_k (\delta_t \cdot \xi_k) \xi_k \right].
\end{aligned}$$

The expression in brackets is the generalized Fourier series for δ_t , so we conclude $(\mathbf{K}^T)_t = L(\delta_t)$. We state the formula for \mathbf{K} in two ways, also using an alternate notation for δ_t (explained in Section 8.1.4).

$$(\mathbf{K}^T)_t = L(\delta_t) = L(\Delta_t^1), \quad (7.47)$$

$$\mathbf{K}_{st} = [L(\delta_t)]_s = [L(\Delta_t^1)]_s. \quad (7.48)$$

We state without proof the analogous formulas for the kernel of a multilinear operator:

$$\mathbf{K}_{st_n \dots t_1} = [M(\delta_{t_1}, \dots, \delta_{t_n})]_s, \quad (7.49)$$

$$(\mathbf{K}^T)_{t_n \dots t_1} = M(\delta_{t_1}, \dots, \delta_{t_n}). \quad (7.50)$$

In the second equation the transpose must be interpreted as being around the first dimension.

Chapter 8

Linear and Multilinear Operators

8.1 Linear Operators

8.1.1 Introduction

Many of the most important field transformations are linear and multilinear operators. However, in their ideal forms they often do not satisfy the conditions in Theorems 7.4.7 and 7.4.10 for the existence of their kernels. In these cases we have to consider the approximation of the ideal operation by a field product.

8.1.2 Examples

8.1.2.1 COMPLEX CONJUGATE

Definition 8.1.1 (complex conjugate) *The complex conjugate operation ϕ^* computes the element-wise complex conjugate of a field:*

$$(\phi^*)_t = \overline{\phi_t}. \quad (8.1)$$

Formula 8.1.1 (complex conjugate)

$$\phi^* = K\phi, \quad \text{where } K = \sum_k \xi_k^* \wedge \xi_k^*. \quad (8.2)$$

Proof: Simply expand the product:

$$\begin{aligned}
(\mathbf{K}\phi)_s &= \int_{\Omega} \mathbf{K}_{st} \phi_t dt \\
&= \int_{\Omega} \left[\sum_k \xi_k^* \wedge \xi_k^* \right]_{st} \phi_t dt \\
&= \int_{\Omega} \sum_k \xi_k^*(s) \xi_k^*(t) \phi_t dt \\
&= \sum_k \xi_k^*(s) \int_{\Omega} \xi_k^*(t) \phi_t dt \\
&= \sum_k \xi_k^*(s) (\phi \cdot \xi_k) \\
&= \left[\sum_k \xi_k(s) (\phi \cdot \xi_k) \right]^* \\
&= (\phi_s)^*,
\end{aligned}$$

since the bracketed expression in the second-to-last line is the Fourier series for ϕ evaluated at s .

□

8.1.2.2 DEFINITE INTEGRAL

Definition 8.1.2 (defint) *The definite integral operator* $\text{defint} : \Phi(\Omega) \rightarrow \mathbb{R}$ *simply computes the total value of the field:*

$$\text{defint } \phi = \int_{\Omega} \phi_t dt. \quad (8.3)$$

If we let $\mathbf{1}_t = 1$ be the constant 1 field, then we can define the definite integral:

$$\text{defint } \phi = \phi \cdot \mathbf{1}. \quad (8.4)$$

The definite integral is often useful, especially for computing the mean of a field. For example, by subtracting from a field its mean, we may get maximum use of the dynamic range of a field storage unit. The definite integral is also useful for automatic gain control.

Formula 8.1.2 (defint)

$$\text{defint } \phi = \langle \mathbf{1} \mid \phi \rangle. \quad (8.5)$$

Since $\mathbf{1}$ is a physically realizable field, no approximation is involved in this field computation of the definite integral.

Proof: Simply observe:

$$\begin{aligned} \langle \mathbf{1} \mid \phi \rangle &= \int_{\Omega} \mathbf{1}_t \phi_t dt \\ &= \int_{\Omega} \phi_t dt \\ &= \text{defint } \phi. \end{aligned}$$

Alternately, we compute the kernel according to Equation 7.32:

$$\begin{aligned} K &= \sum_{k=0}^{\infty} (\text{defint } \xi_k) \wedge \xi_k^* \\ &= \sum_{k=0}^{\infty} \langle \mathbf{1} \mid \xi_k \rangle \wedge \xi_k^*. \end{aligned}$$

Now note that $\langle \mathbf{1} \mid \xi_k \rangle$ is a scalar, and that an outer product with a scalar is the same as a scalar product. That is, if a is a scalar, then:

$$a \wedge \phi = \phi \wedge a = a\phi. \quad (8.6)$$

Therefore, the formula for the kernel is:

$$K = \sum_{k=0}^{\infty} \langle \mathbf{1} \mid \xi_k \rangle \xi_k^* = \left(\sum_{k=0}^{\infty} \langle \xi_k \mid \mathbf{1} \rangle \xi_k \right)^*. \quad (8.7)$$

But this is just the Fourier expansion (Sec. 5.2.3) of $\mathbf{1}^* = \mathbf{1}$, so $K = \mathbf{1}$.

□

8.1.2.3 INDEFINITE INTEGRAL

Definition 8.1.3 (*f*) For illustrative purposes we take $\Omega = [0, 1]$. The indefinite integral $\int \phi$ of a field ϕ is then defined:

$$\left(\int \phi \right)_s = \int_0^s \phi_t dt. \quad (8.8)$$

Formula 8.1.3 (*f*)

$$\int \phi = \Delta^0 \phi. \quad (8.9)$$

where

$$\Delta^0_{st} = \begin{cases} 1, & \text{if } s \geq t \\ 0, & \text{if } s < t \end{cases}. \quad (8.10)$$

The *unit step field* (or *Heaviside field*) Δ^0 can be visualized as follows: it is 1 above the $s = t$ diagonal and zero below it. Although Δ^0 is discontinuous (and thus violates our physical realizability constraints, Section 7.1.2.1), it can be approximated arbitrarily closely by continuous functions.

The point of the Δ^0 notation will become apparent later (Formulas 8.1.4 and 8.1.5). An alternative formula for the indefinite integral will be derived in Section 8.2.2.4.

Proof: To see that Δ^0 is the kernel, observe:

$$\begin{aligned} \left(\int \phi \right)_s &= \int_0^s \phi_t dt \\ &= \int_0^1 \Delta^0_{st} \phi_t dt \\ &= (\Delta^0 \phi)_s. \end{aligned}$$

To see that Δ^0 exists, observe:

$$\begin{aligned} \|\Delta^0\|^2 &= \int_0^1 \int_0^1 (\Delta^0_{st})^2 dt ds \\ &= \int_0^1 \int_0^s 1 dt ds \\ &= \int_0^1 s ds \\ &= 1/2 \\ &< \infty. \end{aligned}$$

□

8.1.2.4 PRODUCT MASK

Definition 8.1.4 (\times) *The product mask $\mu \times \phi$ computes a point-wise product between the given field ϕ and a fixed field μ :*

$$(\mu \times \phi)_t = \mu_t \phi_t. \quad (8.11)$$

This is obviously a linear operator. In addition to its obvious use for masking out part of a field, it may also be used with defint to compute weighted averages of fields.

Formula 8.1.4 (\times)

$$\mu \times \phi = K\phi, \quad (8.12)$$

where

$$K_{st} = \mu_s \delta(s - t). \quad (8.13)$$

Here we have made use of the *Dirac delta function* (or *unit impulse function*). This “generalized function” has the value $+\infty$ at the origin, and the value 0 everywhere else. An alternative notation for K is:

$$K_{st} = \mu_s \Delta^1_{st}, \quad (8.14)$$

where:

Definition 8.1.5 (Δ^1)

$$\Delta^1_{st} = \delta(s - t). \quad (8.15)$$

We prefer this notation because Δ^1 is the derivative along the second coordinate of Δ^0 , which was defined in Formula 8.1.3.

$$\Delta^1_{st} = d\Delta^0_{st}/dt. \quad (8.16)$$

Since the kernel K is defined in terms of the *unit impulse field* Δ^1 , it is not physically realizable, and we must use an approximation; see p. 174.

Proof: This formula for the kernel can be established by Equation 7.48:

$$K_{st} = (\mu \times \Delta^1_t)_s = \mu_s \Delta^1_{ts} = \mu_s \Delta^1_{st} \quad (8.17)$$

(since Δ^1 is symmetric).

□

Notice that K is not physically realizable; indeed, it is not even a Hilbert-Schmidt kernel (since $\|K\| = \infty$). We turn to this problem next.

8.1.2.5 DELTA FUNCTIONS AND PHYSICAL REALIZABILITY

The field K , defined as it is in terms of the Dirac delta function, violates several of our physical realizability conditions (Section 7.1.2.1), since it is infinite valued at the origin and discontinuous. However, we can approximate it by various realizable functions. For example, we can approximate it by a square wave, $\delta(x) \approx S_\epsilon(x)$, where:

$$\begin{aligned} S_\epsilon(x) &= 1/\epsilon, \text{ if } -\epsilon/2 < x < +\epsilon/2, \\ S_\epsilon(x) &= 0, \text{ otherwise.} \end{aligned}$$

Clearly, $\delta = \lim_{\epsilon \rightarrow 0} S_\epsilon$. Similarly, we could approximate δ by a triangular wave, or a Gaussian distribution, or any number of other standard functions.

All of these approximations have the effect of “smearing out” the product $\mu \times \phi$. For example, with S_ϵ :

$$\begin{aligned} (K\phi)_s &= \int_{\Omega} \mu_s S_\epsilon(t-s) \phi_t dt \\ &= \mu_s \int_{s-\epsilon/2}^{s+\epsilon/2} \phi_t dt. \end{aligned}$$

Notice that μ_s is multiplied by the average of the values of ϕ in an interval of width ϵ centered on s .

From time to time we will make use of singularity functions such as the Dirac delta function. Although they are convenient for the theoretical development, keep in mind that physical realizability requires them to be approximated. This is not very different from the familiar situation of numerical approximation on digital computers. On the other hand, delta functions are significant in the theoretical development, since they show us when an operator is *local*, that is, the value of the output field at a point depends on the value of the input field at only one or a few points. This is important, because local operators can be implemented with very sparse interconnections between layers in a neural network (see ? for more information).

8.1.2.6 DERIVATIVE

Definition 8.1.6 (D) *There are of course many derivative and derivative-like operators that can be defined on spaces of fields. In this case we take $\Omega = [a, b]$ and consider the derivative operator on this closed interval: $D\phi = \phi'$; that is, $(D\phi)_t = d\phi_t/dt$.*

Formula 8.1.5 (D)

$$D\phi = \Delta^2\phi, \quad (8.18)$$

where

$$\Delta^2_{st} = -\delta'(s-t). \quad (8.19)$$

Here we make use of the *doublet* δ' , which is the derivative of the *Dirac delta function* or *unit impulse function* δ (see Frm. 8.1.4). The doublet is a very unusual function. Its value is zero everywhere, except “just to the left” of 0, where its value is $+\infty$, and “just to the right” of 0, where its value is $-\infty$. Therefore the *doublet field* Δ^2 has the value $-\infty$ “just below” the $s = t$ line, and the value $+\infty$ “just above” it. Since the field Δ^2 is defined in terms of this “generalized function,” it is not physically realizable; this issue is discussed later (Section 8.1.2.7).

Note that Δ^2 is the derivative of Δ^1 along its second coordinate:

$$\Delta^2_s = D\Delta^1_s. \quad (8.20)$$

See Section 8.2.2.4 for an alternative approach to the field computation of the derivative.

Proof: The simplest proof of this result uses the following “sifting” property of the doublet:

$$-\phi'_s = \int_{\Omega} \delta'(s-t)\phi_t dt. \quad (8.21)$$

Hence, letting $\Delta^2_{st} = -\delta'(s-t)$, it is immediate that

$$\phi'_s = \langle \Delta^2_s | \phi \rangle. \quad (8.22)$$

Alternately, use the formula for the kernel given in Section 7.4.3.4:

$$K_{ts} = (D\Delta^1_s)_t = \Delta^2_{st}. \quad (8.23)$$

Note however that Δ^2 is not a Hilbert-Schmidt kernel, since $\|K\| = \infty$.

□

It may be instructive to consider the special case of the trigonometric basis. Suppose $a = -\pi$, $b = \pi$ and:

$$\begin{aligned} \xi_0(t) &= 1/\sqrt{2\pi} \\ \xi_{2n-1}(t) &= \cos nt/\sqrt{2\pi}, \quad n = 1, 2, \dots \\ \xi_{2n}(t) &= \sin nt/\sqrt{2\pi}, \quad n = 1, 2, \dots \end{aligned}$$

First apply the formula for the kernel (Eq. 7.32):

$$\begin{aligned} K &= \sum_k D\xi_k \wedge \xi_k^* \\ &= \sum_k \xi'_k \wedge \xi_k^*. \end{aligned}$$

For the trigonometric basis, observe that:

$$\begin{aligned} \xi'_0(t) &= 0 \\ \xi'_{2n-1}(t) &= -n \sin nt / \sqrt{2\pi} = -n\xi_{2n}(t) \\ \xi'_{2n}(t) &= n \cos nt / \sqrt{2\pi} = n\xi_{2n-1}(t). \end{aligned}$$

Hence,

$$\begin{aligned} \xi'_0 &= \mathbf{0} \\ \xi'_{2n-1} &= -n\xi_{2n} \\ \xi'_{2n} &= n\xi_{2n-1}. \end{aligned}$$

Hence, the kernel is

$$\begin{aligned} K &= \sum_{n=1}^{\infty} (\xi'_{2n-1} \wedge \xi_{2n-1}^* + \xi'_{2n} \wedge \xi_{2n}^*) \\ &= \sum_{n=1}^{\infty} n(\xi_{2n-1} \wedge \xi_{2n}^* - \xi_{2n} \wedge \xi_{2n-1}^*). \end{aligned}$$

Now note that

$$\begin{aligned} (\xi_{2n-1} \wedge \xi_{2n}^*)_{st} &= \xi_{2n-1}(s)\xi_{2n}^*(t) \\ &= (\cos ns / \sqrt{2\pi})(\sin nt / \sqrt{2\pi})^* \\ &= \sin nt \cos ns / 2\pi, \text{ since the sin and cos are real.} \end{aligned}$$

Similarly,

$$(\xi_{2n} \wedge \xi_{2n-1}^*)_{st} = \cos nt \sin ns / 2\pi. \quad (8.24)$$

Therefore,

$$\begin{aligned} (\xi_{2n-1} \wedge \xi_{2n}^* - \xi_{2n} \wedge \xi_{2n-1}^*)_{st} &= (\sin nt \cos ns - \cos nt \sin ns) / 2\pi \\ &= \sin(nt - ns) / 2\pi \\ &= \sin n(t - s) / 2\pi. \end{aligned}$$

Therefore the kernel is defined by:

$$K_{st} = \frac{1}{2\pi} \sum_{n=1}^{\infty} n \sin n(t-s). \quad (8.25)$$

This is the Fourier series for $-\delta'(t-s)$ as can be seen by noting that the sines “pile up” to $-\infty$ on the negative side of 0, and to $+\infty$ on the positive side of 0. Clearly K is not physically realizable; see Section 8.1.2.5 as well as the following section.

8.1.2.7 DOUBLET AND PHYSICAL REALIZABILITY

The field Δ^2 , defined as it is in terms of the doublet, violates several of our physical realizability conditions (Section 7.1.2.1), since it is infinite valued at the origin and discontinuous. However, just as we did for Δ^1 (Frm. 8.1.2.5), we can approximate Δ^2 by various realizable functions. For example, we can truncate the Fourier series given above. More directly, we can approximate it by a square wave, $\Delta^2_{st} \approx S_\epsilon(s, t)$, where:

$$\begin{aligned} S_\epsilon(s, t) &= -2/\epsilon^2, & \text{if } -\epsilon/2 < s-t < 0 \\ S_\epsilon(s, t) &= +2/\epsilon^2, & \text{if } 0 < s-t < +\epsilon/2 \\ S_\epsilon(s, t) &= 0, & \text{otherwise.} \end{aligned}$$

Clearly, $\Delta^2 = \lim_{\epsilon \rightarrow 0} S_\epsilon$. By using S_ϵ we are approximating the derivative by the “difference”:

$$\phi'_t \approx \frac{\int_0^{\epsilon/2} \phi_t dt - \int_{-\epsilon/2}^0 \phi_t dt}{\epsilon}. \quad (8.26)$$

Notice that in field computation the difference is computed between two “average” values, whereas in digital computation it is computed between the values at two points.

8.1.2.8 DISCRETE FOURIER TRANSFORM

Definition 8.1.7 Let $\beta_0, \beta_1, \beta_2, \dots$ be an orthogonal basis for $\Phi(\Omega_2)$ and suppose $\Omega_1 = \{0, 1, 2, \dots, N\}$. The result of the discrete Fourier transform $F_N \phi$ is an $N+1$ element discrete field σ , such that σ_n is the n th Fourier coefficient of ϕ with respect to the β_n . That is,

$$\sigma_n = (F_N \phi)_n = \phi \cdot \beta_n. \quad (8.27)$$

Note that we extract only the first $N + 1$ coefficients. There are two reasons for this. First, extracting all the coefficients would require Ω_1 to be infinite (i.e. the natural numbers), which violates our condition that the domains of fields be bounded (Section 7.1.2.2). The second reason is that higher order coefficients frequently represent noise, and so should be suppressed. Nevertheless, later in this section we show a way to capture the full spectrum in a physically realizable field.

Formula 8.1.6

$$F_N \phi = K \phi. \quad (8.28)$$

where

$$K = \sum_{n=0}^N \delta_n \wedge \beta_n^*. \quad (8.29)$$

Here δ_n is the *Kronecker delta function* defined by:

$$\begin{aligned} \delta_n(n) &= 1, \\ \delta_n(m) &= 0, n \neq m. \end{aligned}$$

Alternately we can define the Kronecker delta in a way analogous to the Dirac delta (Section 8.1.4). Thus $\delta(k)$ is 1 if $k = 0$, but 0 if $k \neq 0$. Then K is defined:

$$K_{mt} = \sum_{n=0}^N \delta(m - n) \beta_n(t). \quad (8.30)$$

Observe that K is in effect an “array” of the basis functions $\beta_0, \beta_1, \dots, \beta_N$; $K_m = \beta_m$. Alternately, we may say that the kernel *is* the (truncated) series of basis functions.

Proof: Although it is straight-forward to show the correctness of the formula for K by expanding the product $K\phi$, we will show that the formula can be derived directly from the definition, $(F_N \phi)_n = \phi \cdot \beta_n$. First note that any $N + 1$ -element “array” (discrete field) A satisfies

$$A = \sum_{n=0}^N \delta_n A_n. \quad (8.31)$$

since

$$\begin{aligned}
 A_k &= \left(\sum_n \delta_n A_n \right)_k \\
 &= \sum_n \delta_n(k) A_n \\
 &= \delta_k(k) A_k \\
 &= A_k.
 \end{aligned}$$

Therefore, the Fourier transform can be represented:

$$\begin{aligned}
 \mathbf{F}_N \phi &= \sum_{n=0}^N \delta_n (\mathbf{F}_N \phi)_n \\
 &= \sum_{n=0}^N \delta_n (\phi \cdot \beta_n) \\
 &= \sum_{n=0}^N (\delta_n \wedge \beta_n^*) \phi \\
 &= \left(\sum_{n=0}^N \delta_n \wedge \beta_n^* \right) \phi \\
 &= \mathbf{K} \phi.
 \end{aligned}$$

□

8.1.2.9 REPRESENTING THE FULL SPECTRUM

To represent the full spectrum, we must find some way of fitting the infinite set of natural numbers into finite space. Although there are a number of ways of doing this, we choose a representation that may have some practical applications. We define the following family of “regressive pulse functions”:

$$\begin{aligned}
 r_n(x) &= 1, \text{ if } 2^{-n} < x < 2^{-n-1} \\
 r_n(x) &= 0, \text{ otherwise.}
 \end{aligned}$$

In other words, r_0, r_1, \dots are a series of exponentially narrower contiguous pulses of unit amplitude. We will use the nonzero portion of each pulse to

represent a Fourier coefficient; therefore higher order coefficients will occupy exponentially less space. It is easy to show that this transformation is the integral operator:

$$F_\infty \phi = K\phi = \left(\sum_{n=0}^{\infty} r_n \wedge \beta_n \right) \phi, \quad (8.32)$$

provided that the infinite sum exists. To show that it does, we assume that the β_n are orthonormal (vice merely orthogonal) and observe that $\|r_n\| = 2^{-n-1}$. Then, derive:

$$\begin{aligned} \|K\| &\leq \sum_n \|r_n \wedge \beta_n^*\| \\ &= \sum_n \|r_n\| \|\beta_n^*\| \\ &= \sum_n 2^{-n-1} \\ &= 1 < \infty. \end{aligned}$$

8.1.2.10 CONTINUOUS FOURIER TRANSFORM

Definition 8.1.8 *In this case we replace the discrete series of basis functions β_0, β_1, \dots by a family of functions that depends continuously on a parameter ω . For example, we may replace the complex exponential basis $\beta_n(t) = e^{-in\omega t}$ by the family:*

$$\beta_\omega(t) = e^{-i\omega t}. \quad (8.33)$$

If we are concerned with the spectrum in only a finite interval $\Omega = [a, b]$, then there is no difficulty extending the discrete case to the continuous case. Here we define:

$$(F_{[a,b]}\phi)_\omega = \phi \cdot \beta_\omega. \quad (8.34)$$

If we want the full spectrum, then we have a representation problem, since the domain of the resulting field would be $[0, \infty)$, which is not physically realizable. In this case we choose some continuous monotonic mapping ρ from $[0, \infty)$ into a bounded domain. An example ρ is

$$\rho(\omega) = 1/(\omega + 1). \quad (8.35)$$

The resulting definition of the transform function is:

$$(F_\rho \phi)_{\rho(\omega)} = \phi \cdot \beta_\omega. \quad (8.36)$$

The field returned by $F_\rho\phi$ is physically realizable, since the image of $[0, \infty)$ under ρ is $[0, 1]$.

Formula 8.1.7 In the band-limited case we use the analogous formula to that in Section 8.1.6:

$$F_{[a,b]}\phi = K\phi, \quad (8.37)$$

where

$$K_{\omega t} = \beta_\omega^*(t). \quad (8.38)$$

To compute the full spectrum, a different kernel is required:

$$K_{st} = \beta_{\rho^{-1}(s)}^*(t). \quad (8.39)$$

Note that ρ^{-1} exists, since ρ is monotonic. Also note that, as before, the kernel is the orthonormal basis (possibly reindexed to ensure a bounded domain).

Proof: First consider the band-limited case. Derive:

$$\begin{aligned} (K\phi)_\omega &= K_\omega \cdot \phi^* \\ &= \phi \cdot K_\omega^* \\ &= \phi \cdot \beta_\omega \\ &= (F_{[a,b]}\phi)_\omega. \end{aligned}$$

Further, the field K exists, since f_ω depends continuously on ω .

For the full-spectrum case derive:

$$\begin{aligned} (K\phi)_s &= K_s \cdot \phi^* \\ &= \phi \cdot K_s^* \\ &= \phi \cdot \beta_{\rho^{-1}(s)}^* \\ &= (F_\rho\phi)_s. \end{aligned}$$

Hence $(K\phi)_{\rho(\omega)} = \phi \cdot \beta_\omega^*$. It remains to show that K exists. Assume an orthonormal basis, $\|\beta_\omega\| = 1$. Then:

$$\begin{aligned} \|K\|^2 &= \int_{\Omega_1 \times \Omega_2} K_{st}^2 d(s, t) \\ &= \int_{\Omega_1} \int_{\Omega_2} |\beta_{\rho^{-1}(s)}^*(t)|^2 dt ds \end{aligned}$$

$$\begin{aligned}
&= \int_{\Omega_1} \|\beta_{\rho^{-1}(s)}^*\|^2 ds \\
&= \int_{\Omega_1} 1 ds \\
&= |\Omega_1| < \infty.
\end{aligned}$$

□

8.1.2.11 INVERSE FOURIER TRANSFORM

Definition 8.1.9 We consider the case of the finite-spectrum discrete Fourier transform, although it will be apparent that the solution can be simply extended to the full-spectrum or continuous cases. Our goal is as follows. Suppose that the generalized coefficients of ϕ beyond the N th are zero, then we want F_N^{-1} such that:

$$F_N^{-1}(F_N\phi) = \phi \quad (8.40)$$

We are given a discrete field c such that c_n is the n th generalized Fourier coefficient, $n = 0, 1, \dots, N$. Then F_N^{-1} takes the simple form:

$$F_N^{-1}c = \sum_{n=0}^N c_n \beta_n. \quad (8.41)$$

Formula 8.1.8

$$F_N^{-1}c = K^\dagger c. \quad (8.42)$$

where K is the kernel of the discrete Fourier transform (Frm. 8.1.6), which is simply the “array” of conjugate basis functions. If we use F_N for this kernel, then the inverse transform is expressed more obviously by:

$$F_N^{-1}c = F_N^\dagger c. \quad (8.43)$$

Proof: Recall (Frm. 8.1.6) that the kernel of the discrete Fourier transform is defined:

$$K_{nt} = \beta_n^*(t) = \beta_{nt}^*. \quad (8.44)$$

That is, $K_{tn}^\dagger = \beta_n(t)$. Therefore,

$$\begin{aligned}
(F_N^{-1}c)_t &= \sum_n c_n \beta_n(t) \\
&= \sum_n c_n K_{tn}^\dagger \\
&= (K^\dagger c)_t.
\end{aligned}$$

Hence $\mathbf{F}_N^{-1}c = \mathbf{K}^\dagger c$.

□

8.1.2.12 USING FOURIER METHODS TO COMPUTE LINEAR OPERATORS

The Fourier transform permits a way computing linear operators that is particularly suitable for neural implementation ?. To see this, observe that for $L : \Phi(\Omega) \rightarrow \Phi(\Omega')$ and $\phi \in \Phi(\Omega)$:

$$\begin{aligned} L\phi &= L\left(\sum_n c_n \beta_n\right) \\ &= \sum_n c_n L(\beta_n) \\ &= \sum_n c_n \lambda_n. \end{aligned}$$

Here the c_n are the Fourier coefficients of ϕ , $c = \mathbf{F}_N \phi$, and the $\lambda_n \in \Phi(\Omega')$ are the values of L on the basis functions (i.e., $\lambda_0, \lambda_1, \dots$ is the *transfer function* of L). Then

$$L(\phi) = c\Lambda = (\mathbf{F}_N \phi)\Lambda = \Lambda^\mathbf{T}(\mathbf{F}_N \phi), \quad (8.45)$$

where $\Lambda \in \Phi(\{1, \dots, N\} \times \Omega')$ is defined

$$\Lambda_n = L(\beta_n). \quad (8.46)$$

In other words, any linear operator can be computed by extracting the Fourier coefficients of its argument and using these to weight the values of the operator on the basis fields. The advantage of this for neural implementation is that if the input is band limited, then there are only a finite number of coefficients. These can be represented by the “hidden units” between the neural layers that compute \mathbf{F}_N and $\Lambda^\mathbf{T}$. The kernel of the linear operator is of course $\Lambda^\mathbf{T}\mathbf{F}_N$.

The foregoing suggests a generalization based on the continuous Fourier transform. We may compute any linear operator by:

1. Taking its Fourier transform (discrete or continuous).
2. Multiplying the result by a (discrete or continuous) product mask representing the operator’s transfer function.

3. Taking the corresponding inverse Fourier transform.

This suggests that a general purpose field computer could be structured around Fourier transforms.

8.2 Multilinear Operators

8.2.1 Introduction

Like the linear operators discussed in Section 8.1, many of the multilinear operators discussed here are not implementable by physically realizable field products. They can, however, be approximated in a straight-forward way.

8.2.2 Examples

8.2.2.1 LOCAL PRODUCT

Definition 8.2.1 *The local product $\phi \times \psi$ of two fields over the same domain is defined:*

$$(\phi \times \psi)_s = \phi_s \psi_s. \quad (8.47)$$

Formula 8.2.1

$$\phi \times \psi = M\phi\psi = M(\psi \wedge \phi), \quad (8.48)$$

where

$$M_s = \Delta^1_s \wedge \Delta^1_s. \quad (8.49)$$

That is, $M_{stu} = \Delta^1_{st} \Delta^1_{su}$.

The unit impulse field Δ^1 is defined in Def. 8.1.5 (p. 173). Since Δ^1 is not physically realizable, M will have to be approximated; see Sec. 8.1.2.5 (p. 174). Notice that the field product $M\phi\psi$ is a very inefficient way to compute the local product $\phi \times \psi$. The field product brings together all the possible combinations (ψ_s, ϕ_t) , but the field M ignores all except those for which $s = t$. It is wasteful to use the power of global computation where only local computation is required. This suggests that most general-purpose field computers will have the local product operation built in.

Proof: It is easy to check that the formula for M is correct:

$$\begin{aligned}
[M(\psi \wedge \phi)]_s &= (\psi \wedge \phi) \cdot M_s^* \\
&= (\psi \wedge \phi) \cdot (\Delta_s^1 \wedge \Delta_s^1)^* \\
&= (\psi \cdot \overline{\Delta_s^1})(\overline{\Delta_s^1} \cdot \phi) \\
&= \psi_s \phi_s.
\end{aligned}$$

The last two steps follow from Thm. 7.4.4 (Section 7.4.2.1, p. 159) and the sifting property of Δ^1 (Eq. 7.46, p. 167).

It is also easy to derive the formula for M directly from the formula for the kernel of a multilinear operator (Thm. 7.4.9, p. 164):

$$M = \sum_k \sum_l (\xi_k \times \xi_l) \wedge \xi_l^* \wedge \xi_k^*. \quad (8.50)$$

Hence,

$$\begin{aligned}
M_{stu} &= \sum_k \sum_l (\xi_k \times \xi_l)_s \xi_l^*(t) \xi_k^*(u) \\
&= \sum_k \sum_l \xi_l(s) \xi_l(t) e \xi_l^*(t) \xi_k^*(u) \\
&= \left[\sum_l \xi_l(s) \xi_l^*(t) \right] \left[\sum_k \xi_k(s) \xi_k^*(u) \right] \\
&= \left[\sum_l \xi_l^*(s) \xi_l(t) \right]^* \left[\sum_k \xi_k^*(s) \xi_k(u) \right]^* \\
&= \left[\sum_l (\Delta_s^1 \cdot \xi_l) \xi_l(t) \right]^* \left[\sum_k (\Delta_s^1 \cdot \xi_k) \xi_k(u) \right]^*.
\end{aligned}$$

The last step is by the sifting property, $\Delta_s^1 \cdot \xi_l = \xi_l^*$. Now observe that in the brackets we have the Fourier series for Δ^1 , hence:

$$\begin{aligned}
M_{stu} &= \overline{\Delta_s^1(t)} \overline{\Delta_s^1(u)} \\
&= \Delta_s^1(t) \Delta_s^1(u) \\
&= (\Delta_s^1 \wedge \Delta_s^1)_{tu}.
\end{aligned}$$

□

8.2.2.2 LINEAR MEASURE SPACES

In order to define the convolution and correlation of arbitrary fields we need a subtraction operation on the domains of fields. That is, for any $\phi \in \Phi(\Omega)$ for which we might want a convolution or correlation, we need $s - t \in \Omega$ to be defined for all $s, t \in \Omega$.

To this end, we define a *linear measure space* to be a measure space that is also a linear space. Therefore, it has addition, subtraction and scalar multiplication operations satisfying the usual properties.

There is one difficulty with this definition. Since a linear space must be closed with respect to its operations (addition, subtraction and scalar multiplication), it cannot be bounded, since it must contain at for every real number a and every $t \in \Omega$. Thus linear measure spaces violate one of our physical realizability constraints, namely, that the domains of fields be bounded (Section 7.1.2.2). This is a problem that is commonly faced in the analysis of linear, shift-invariant systems, since most implementable systems are bounded, and hence not completely shift-invariant. We shall take the same pragmatic approach here that is commonly applied in that analysis: apply the theory based on linear spaces, but be careful of “edge effects.”

8.2.2.3 CONVOLUTION

Definition 8.2.2 *If Ω is a linear measure space (Section 8.2.2.2), then the convolution $\phi * \psi$ of two fields $\phi, \psi \in \Phi(\Omega)$ is defined:*

$$(\phi * \psi)_s = \int_{\Omega} \phi_{s-t} \psi_t dt. \quad (8.51)$$

Formula 8.2.2

$$\phi * \psi = K\phi\psi = K(\psi \wedge \phi), \quad (8.52)$$

where

$$K_{stu} = \delta(s - t - u). \quad (8.53)$$

Since the field K is defined in terms of the unit impulse field $\delta = \Delta^1$ (Section 8.1.2.5), it must in practice be approximated.

Proof: It is easiest to establish this result by direct expansion of the product:

$$(K\phi\psi)_s = \int_{\Omega} (K\phi)_{st} \psi_t dt$$

$$\begin{aligned}
&= \int_{\Omega} \int_{\Omega} K_{stu} \phi_u du \psi_t dt \\
&= \int_{\Omega} \int_{\Omega} \delta(s - t - u) \phi_u du \psi_t dt \\
&= \int_{\Omega} \phi_{s-t} \psi_t dt \\
&= (\phi * \psi)_s.
\end{aligned}$$

The second to last step follows from the sifting property of δ .

□

Remark 8.2.1 Notice that if $\phi, \psi \in \Phi(\Omega)$, then $K \in \Phi(\Omega \times \Omega \times \Omega)$, which is of much higher dimension than the fields being convolved. We will discuss alternative approaches to convolution later (Sec. ??).

8.2.2.4 USE OF CONVOLUTION TO IMPLEMENT LINEAR, SHIFT-INVARIANT OPERATORS

Many linear, shift-invariant operators can be implemented more efficiently by convolution than by general field product. For example, the derivative operator (Defn. 8.1.6, p. 174) can be implemented by:

$$D\phi = -\delta' * \phi, \quad (8.54)$$

where δ' is the unit doublet (Frm. 8.1.5, p. 175). To see this, observe

$$(-\delta' * \phi)_s = - \int \delta'_{s-t} \phi_t dt = \phi'_s \quad (8.55)$$

by the sifting property of the doublet (p. 175). This formula should be compared with that in Frm. 8.1.5 (p. 175):

$$D\phi = \Delta^2 \phi. \quad (8.56)$$

Although neither δ' nor Δ^2 is physically realizable, δ' has the advantage that it is of lower dimension: $\delta' \in \Phi(\Omega)$, but $\Delta^2 \in \Phi(\Omega \times \Omega)$. Therefore, δ' will generally be easier to represent in field computers.

For another example of the use of convolution, consider the indefinite integral (Defn. 8.1.3, p. 172), which is also shift invariant. Define the *Heaviside field* v to be a slice through Δ^0 :

$$v_s = \Delta_{s0}^0. \quad (8.57)$$

It has the property:

$$v_s = \begin{cases} 1, & \text{if } s \geq 0 \\ 0, & \text{if } s < 0 \end{cases} . \quad (8.58)$$

This field can be convolved with an arbitrary field to compute its indefinite integral:

$$\int \phi = v * \phi. \quad (8.59)$$

To see this, observe:

$$\begin{aligned} (v * \phi)_s &= \int_{\Omega} v_{s-t} \phi_t dt \\ &= \int_{\Omega} \Delta_{s-t,0}^0 \phi_t dt \\ &= \int_{\Omega} \Delta_{st}^0 \phi_t dt \\ &= \left(\int \phi \right)_s . \end{aligned}$$

The foregoing examples suggest that convolution is a useful operation to include in general purpose field computers.

8.2.2.5 CORRELATION

Definition 8.2.3 *If Ω is a linear measure space (Section 8.2.2.2), then the correlation $\phi \star \psi$ of two fields $\phi, \psi \in \Phi(\Omega)$ is defined:*

$$(\phi \star \psi)_s = \int_{\Omega} \phi_{t-s}^* \psi_t dt. \quad (8.60)$$

Remark 8.2.2 *Notice that for real-valued fields correlation differs from convolution only in having ‘ $t-s$ ’ where the latter has ‘ $s-t$ ’. (For complex-valued fields, correlation also has a complex conjugate, as you see here.)*

Remark 8.2.3 *The correlation of two different fields is usually called cross-correlation, whereas the correlation of a field with itself is called autocorrelation.*

Formula 8.2.3

$$\phi \star \psi = K\phi\psi = K(\psi \wedge \phi), \quad (8.61)$$

where

$$K = \sum_{k,l} (\xi_k \star \xi_l) \wedge \xi_l^* \wedge \xi_k^*. \quad (8.62)$$

Proof: This is simply an application of Thm. 7.4.3.3 (p. 164).

□

Chapter 9

Universal Field Computation

9.1 Introduction

In Sections ?? and ?? we discussed the utility of general purpose field computers, and claimed that there exists a *universal set* of field operations that permits the approximation of any field transformation in a large and useful class. In this chapter we present one such universal set and give several examples of the resulting approximations.

9.2 Approximation Based on Taylor Series

9.2.1 Derivatives of field transformations

Our goal is to find ways of approximating field transformations. The notion of a derivative is important for understanding approximations of real functions (for instance, in understanding the Taylor series). For many of the same reasons, it is necessary to investigate the derivatives of field transformations. However, since fields are functions (Sec. 7.1.2.1), we need the derivative of an *operator* on a function space. There are two kinds of derivatives that may be defined, the Fréchet (Sec. 3.4.1) and Gâteaux (Sec. 3.4.2). For fields they happen to be equivalent (Sec. 3.4.3). It will be noted that these derivatives are defined on functions between *Banach spaces* (complete normed linear spaces). Since Hilbert spaces are Banach spaces, we can apply these results (Sec. 5).

9.2.2 Gradients of Operators

9.2.2.1 DEFINITION OF GRADIENT

Based on the analogy with finite-dimensional spaces (MacLennan, 1987b) we define the *gradient* of a field transformation $T : \Phi(\Omega_1) \rightarrow \Phi(\Omega_2)$ at a point $\phi \in \Phi(\Omega_1)$ to be the field $K \in \Phi(\Omega_2 \times \Omega_1)$ satisfying the following property:

$$T'(\phi)(\alpha) = K\alpha, \text{ for all } \alpha \in \Phi(\Omega_1). \quad (9.1)$$

In other words, the derivative T' is an integral operator with (Hilbert-Schmidt) kernel K . Sufficient conditions for the existence of a gradient are discussed below (p. 192).

The notation $\nabla T(\phi)$ denotes the gradient of T at ϕ . Thus,

$$T'(\phi)(\alpha) = [\nabla T(\phi)]\alpha. \quad (9.2)$$

It will also be convenient to use the notation $\nabla_\alpha T(\phi)$ for the *directional derivative* of T in the “direction” α :

$$\nabla_\alpha T(\phi) = \nabla T(\phi)\alpha = T'(\phi)(\alpha). \quad (9.3)$$

This permits ∇_α to be treated as an operator; operator techniques are exploited in MacLennan (1987a,b).

It will be useful to consider the form taken by higher order directional derivatives. Supposing that all the gradients exist, observe:

$$\begin{aligned} T^{(k)}(\phi)(\alpha_1) \cdots (\alpha_k) &= \nabla^k T(\phi)\alpha_1 \cdots \alpha_k \\ &= \nabla^k T(\phi)(\alpha_k \wedge \cdots \wedge \alpha_1) \\ &= \nabla_{\alpha_k} \cdots \nabla_{\alpha_1} T(\phi). \end{aligned}$$

Thus, a k -th order differential can be expressed as a k -fold product with the k -th order gradient (if it exists), or as a product between the gradient and the k -fold outer product. These relationships depend on the properties of outer products discussed in Sec. 7.4.2.2.

9.2.2.2 WHEN CAN A DERIVATIVE BE EXPRESSED AS A PRODUCT?

Another way of asking this is “When do gradients exist?” Yet another way of asking it is, “When are derivatives integral operators (of Hilbert-Schmidt type)?”

In this section we present practical sufficient conditions for the existence of gradients. Since derivatives are linear (or multilinear) these conditions are direct applications of the results in Sec. 7.4.3.

For practical applications it seems reasonable to assume that field transformations have filtered inputs (see Sec. 7.4.3.2 for the definition of a filter). Typically there are limitations on the gradients sustainable in the media used to represent the input fields to a transformation (Sec. 7.1.2.3). Also, higher order components typically represent noise, and we do not want our transformations to be excessively sensitive to noise (Sec. 7.1.3).

Therefore, we consider the form of derivatives of transformations that can be written $T = G \circ F$, where F is a filter (Theorem 7.4.7) and G is the “ideal” operator to be computed. By the formula for the derivative of a composition (Prop. 3.4.4, p. 39) we have:

$$\begin{aligned} T'(\phi)(\alpha) &= (G \circ F)'(\phi)(\alpha) \\ &= G'[F(\phi)][F'(\phi)(\alpha)] \\ &= (G' \circ F)(\phi)[F'(\phi)(\alpha)]. \end{aligned}$$

Notice that the input to G' is also filtered by F : $(G' \circ F)(\phi)$. Also note that since F is a filter it is linear, and so by Prop. 3.4.1 (p. 38) $F'(\phi) = F$. Hence,

$$T'(\phi)(\alpha) = (G' \circ F)(\phi)[F(\alpha)]. \quad (9.4)$$

That is,

$$T'(\phi) = (G' \circ F)(\phi) \circ F. \quad (9.5)$$

Hence, $T'(\phi)$ can be written in the form $H \circ F$, and so the gradient $\nabla T(\phi)$ exists. Since the gradient has the appropriate form $(H \circ F)$, the higher order gradients also exist.

The existence of gradients of multilinear operators exist under similar circumstances.

Summarizing, gradients exist for operators that are sufficiently insensitive to higher-order (typically, higher frequency) components of their inputs. In particular, if the operators are band-limited (insensitive to all components beyond a certain order) then the gradients exist. However, they also exist if there response rolls off as an absolutely summable sequence.

9.2.3 Taylor's Theorem

9.2.3.1 TAYLOR'S EXPANSION IN DERIVATIVES

The standard Taylor theorem from functional analysis, like the more familiar Taylor theorem from real analysis, permits the expansion of a function in an infinite series about a point. The difference is that in the present case the function is a field transformation, and the point is a field. The theorem was stated in Sec. 3.5, but we repeat it here for convenience.

Theorem 9.2.1 (Taylor) *Suppose U is any open subset of $\Phi(\Omega_1)$ and $T : \Phi(\Omega_1) \rightarrow \Phi(\Omega_2)$ is a map which is C^n in U (that is, the first n derivatives of T exist). Let $\phi \in U$ and $\alpha \in \Phi(\Omega_1)$ be such that $\phi + \theta\alpha \in U$ for all $\theta \in [0, 1]$. Then:*

$$T(\phi + \alpha) = \sum_{k=0}^{n-1} \frac{T^{(k)}(\phi)(\alpha)^k}{k!} + R_n(\phi, \alpha), \quad (9.6)$$

where

$$R_n(\phi, \alpha) = \int_0^1 \frac{(1-\theta)^{n-1} T^{(n)}(\phi + \theta\alpha)(\alpha)^n}{(n-1)!} d\theta. \quad (9.7)$$

Here ' $(\alpha)^k$ ' denotes k occurrences of the argument α . Also note that $T^{(0)} = T$.

In uncurried form the Taylor expansion is:

$$T(\phi + \alpha) = \sum_{k=0}^{n-1} \frac{d^k T(\phi, \overbrace{\alpha, \dots, \alpha}^n)}{k!} + R_n(\phi, \alpha), \quad (9.8)$$

where

$$R_n(\phi, \alpha) = \int_0^1 \frac{(1-\theta)^{n-1} d^n T(\phi + \theta\alpha, \overbrace{\alpha, \dots, \alpha}^n)}{(n-1)!} d\theta. \quad (9.9)$$

and the appropriate number of α arguments (zero or more) must be supplied for $d^k T$.

9.2.3.2 TAYLOR'S EXPANSION IN GRADIENTS

If the first n gradients of T are defined (Section 9.2.2.2), then its Taylor expansion is:

$$T(\phi + \alpha) = T(\phi) + \sum_{k=1}^n \frac{\nabla_{\alpha}^k T(\phi)}{k!} + R_n(\phi, \alpha). \quad (9.10)$$

As shown in Section 9.2.2.1, the k th term can be written in any of the following forms:

$$\begin{aligned}\frac{1}{k!}\nabla_{\alpha}^k T(\phi) &= \frac{1}{k!}\nabla^k T(\phi)\overbrace{\alpha\cdots\alpha}^k \\ &= \frac{1}{k!}\nabla^k T(\phi)\alpha^{(k)}.\end{aligned}$$

where by $\alpha^{(k)}$ we mean the k -fold outer product $\overbrace{\alpha\wedge\alpha\wedge\cdots\wedge\alpha}^k$. The two forms on the right are especially useful, since they separate the part of the term which is fixed by the point of expansion, $\nabla^k T(\phi)$, from the part which is variable, α .

The remainder term is:

$$R_n(\phi, \alpha) = \int_0^1 \frac{(1-\theta)^{n-1}\nabla_{\alpha}^n(\phi+\theta\alpha)}{(n-1)!}d\theta. \quad (9.11)$$

9.2.3.3 HORNER'S RULE EXPANSION

As is done for conventional polynomials, we can eliminate the need to compute higher (outer product) powers of α by using a form of ‘‘Horner’s Rule.’’ Consider the 3-term Taylor expansion:

$$T(\phi+\alpha) \approx T(\phi) + \nabla T(\phi)\alpha + \frac{1}{2}\nabla^2 T(\phi)\alpha^{(2)}. \quad (9.12)$$

This can be written

$$T(\phi+\alpha) \approx T(\phi) + \left[\nabla T(\phi) + \frac{1}{2}\nabla^2 T(\phi)\alpha\right]\alpha. \quad (9.13)$$

In general, define

$$Q_k(\phi, \alpha) = \nabla^k T(\phi) + \frac{1}{k+1}Q_{k+1}(\phi, \alpha)\alpha. \quad (9.14)$$

for $k \geq 0$ where $\nabla^0 T = T$. Then the infinite Taylor expansion is given by

$$T(\phi+\alpha) = Q_0(\phi, \alpha). \quad (9.15)$$

The Horner’s Rule expansion has direct relevance to implementation of field transformations by neural networks incorporating conjunctive synapses (sigma-pi units); see ?.

9.2.4 Summary of Taylor Series Approximation

We have already seen (Ch. 8) that any reasonable linear or multilinear field transformation can be approximated by a general field product. The results of this section show that reasonable (i.e. sufficiently differentiable) nonlinear transformations can be approximated by a *field polynomials*, that is, by a local sum of general products. We conclude that the following constitute a *universal set* of operators:

1. Local sum: $(\phi + \psi)_t = \phi_t + \psi_t$,
2. General product: $(\Psi X)_{su} = \int_{\Omega} \Psi_{st} X_{tu} dt$.

(Note that the scalar and outer products, which also appear in field polynomials, are degenerate cases of the general product, and so are not strictly necessary.) Of course, practical general purpose field computers will implement a larger set of primitive operations.

9.3 General Polynomial Approximation

The Taylor series approximation of field transformations suffers from a limitation similar to that of the Taylor series approximation of real functions, namely, the approximation is only *locally good*. That is, since the Taylor series extrapolates from a fixed point, the accuracy tends to fall off rapidly with the distance from that point. This is not an important limitation if our only purpose is to establish a universal set of operations. If, however, we are interested in practical field computation, then the limitation is significant. In this case we require polynomials that satisfy some *global* criterion of goodness.

The problem of the polynomial approximation of field transformations can be put as follows. Consider the n th degree polynomial:

$$P_n(\phi) = K_0 + K_1\phi + K_2\phi^{(2)} + \cdots + K_n\phi^{(n)}. \quad (9.16)$$

How can we choose the fields K_0, \dots, K_n so as to minimize the “distance” between P_n and the desired transformation T ? The difficulty is to define an appropriate distance between field transformations. The usual development of an approximation theory presumes an inner product norm and a basis. Unfortunately, we have not found a suitable way to define an inner product on field transformations.

One way to compare field transformations is to compare their values on a finite set of input fields:

$$\delta(T, U) = \sum_{k=1}^m \|T(\phi_k) - U(\phi_k)\|. \quad (9.17)$$

Although this measure is only a pseudo-metric,¹ it is nevertheless useful for a number of purposes.

¹For a pseudo-metric $\delta(x, y) = 0$ need not imply $x = y$.

Chapter 10

Local Transformations

10.1 Theory

In this section we consider the special case of (nonlinear) *local* transformations. These are transformations in which each point of the output field is a function of the corresponding point of the input field:

$$[T(\phi)]_t = F_t(\phi_t). \quad (10.1)$$

We write \bar{F} for the local transformation that applies F_t at each point t ; thus:

$$[\bar{F}(\phi)]_t = F_t(\phi_t). \quad (10.2)$$

Suppose that for all $t \in \Omega$, $F_t : [a, b] \rightarrow K$. Then the type of F is

$$F : \Omega \rightarrow [a, b] \rightarrow K, \quad (10.3)$$

and hence the type of \bar{F} is:

$$\bar{F} : \Phi_{[a,b]}(\Omega) \rightarrow \Phi_K(\Omega). \quad (10.4)$$

Although Taylor's theorem can be used to derive the power series of a local transformation (MacLennan, 1987a), a more general result is just as easy to obtain:

Theorem 10.1.1 *Suppose that $\bar{F} : \Phi_{[a,b]}(\Omega) \rightarrow \Phi_K(\Omega)$ and that the series*

$$F_t(x) = \sum_{k=0}^{\infty} \alpha_{kt} x^k \quad (10.5)$$

converges uniformly with respect to t and x . Then \bar{F} is given by the following (L_2) convergent series:

$$\bar{F}(\phi) = \sum_{k=0}^{\infty} \alpha_k \times \phi^k. \quad (10.6)$$

Here ϕ^k denotes the k -fold local product:

$$\begin{aligned} \phi^0 &= \mathbf{1} \\ \phi^{k+1} &= \phi \times \phi^k, \quad k \geq 0. \end{aligned}$$

We call transformations such as Eq. 10.6, *local field polynomials* or *local power series*, since the powers are computed by local products.

Proof: Let $\epsilon > 0$ be chosen; we must show that there is an N such that

$$\left\| \bar{F}(\phi) - \sum_{k=0}^n \alpha_k \times \phi^k \right\| < \epsilon \quad (10.7)$$

whenever $n \geq N$. Let $\zeta = \epsilon|\Omega|^{-1/2}$. Since the series for $F_t(x)$ converges uniformly, we know that there is an N , independent of t and x , such that

$$\left| F_t(x) - \sum_{k=0}^n \alpha_{kt} x^k \right| < \zeta \quad (10.8)$$

whenever $n > N$. Now consider:

$$\begin{aligned} \left\| \bar{F}(\phi) - \sum_{k=0}^n \alpha_k \times \phi^k \right\|^2 &= \int_{\Omega} \left[\bar{F}(\phi) - \sum_{k=0}^n \alpha_k \times \phi^k \right]_t^2 dt \\ &= \int_{\Omega} \left[F_t(\phi_t) - \sum_{k=0}^n \alpha_{kt} \phi_t^k \right]^2 dt \\ &\leq \int_{\Omega} \zeta^2 dt \\ &= \zeta^2 |\Omega| \\ &= \epsilon^2. \end{aligned}$$

Hence,

$$\left\| \bar{F}(\phi) - \sum_{k=0}^n \alpha_k \times \phi^k \right\| < \epsilon. \quad (10.9)$$

□

The common case where F is a constant function, that is, $F_t = f$ for all t , is especially useful:

Corollary 10.1.1 *If $\bar{f} : \Phi_I(\Omega) \rightarrow \Phi_K(\Omega)$ and I is in the interval of convergence of*

$$f(x) = \sum_{k=0}^{\infty} a_k x^k, \quad (10.10)$$

then this series converges:

$$\bar{f}(\phi) = \sum_{k=0}^{\infty} a_k \phi^k. \quad (10.11)$$

Here $a_k \phi^k$ denotes a scaling of ϕ^k by $a_k \in \mathbb{C}$. If this is not a primitive operation, then it can be accomplished by a local product with the constant field $a_k \mathbf{1}$.

Proof: It is well-known that if I is in the interval of convergence of $\sum_{k=0}^{\infty} a_k x^k$ then the series converges uniformly in I . Hence the theorem applies.

□

10.2 Examples

10.2.1 Importance of Sigmoid Nonlinearities

We begin by exploring the field computation of two local transformations with a “sigmoid” shape. Such transformations are important because of their applications in neural networks. The functional behavior of the most common artificial neurons are defined by the equation

$$y_i = \sigma \left(\sum_{k=1}^N W_{ij} x_j \right). \quad (10.12)$$

This defines the activity level y_i of an output neuron i ($1 \leq i \leq M$) in terms of the activities x_j of the input neurons j ($1 \leq j \leq N$). The “weights” W_{ij} reflect the strength and polarity (excitatory or inhibitory) of the synapses

between neurons j and i . The sigmoid function σ is a nonlinear function whose effect is to “sharpen up” the value computed by the summation; it acts as a “soft threshold.” Notice that if we think of x , y and W as finite fields, then the preceding equation can be expressed as a field computation:

$$y = \bar{\sigma}(Wx). \quad (10.13)$$

We are of course most interested in the case where x , y and W are continuous fields, but the mathematics is the same. In the following subsections we discuss the field computation of two common sigmoid transformations.¹

10.2.2 Hyperbolic Tangent

Definition 10.2.1 *The hyperbolic tangent sigmoid transformation is*

$$\overline{\tanh} : \Phi_{[a,b]}(\Omega) \rightarrow \Phi_{[-1,1]}(\Omega), \quad (10.14)$$

where $-\pi/2 < a < b < \pi/2$. Its effect as a soft threshold can be seen in its continuous variation between these values:

$$\begin{aligned} \overline{\tanh}(-\infty) &= -1, \\ \overline{\tanh}(0) &= 0, \\ \overline{\tanh}(+\infty) &= +1. \end{aligned}$$

The $\overline{\tanh}$ function is most useful when we want the nonlinearity to preserve the sign of the input. It also arises naturally in analog VLSI implementations of neural networks (Mead, 1989, p. 69 and passim).

Formula 10.2.1

$$\overline{\tanh}(\phi) = \phi - \frac{1}{3}\phi^3 + \frac{2}{15}\phi^5 - \frac{17}{315}\phi^7 + \dots + \frac{2^{2n}(2^{2n} - 1)B_{2n}}{(2n)!}\phi^{2n-1} + \dots, \quad (10.15)$$

where B_{2n} is the $2n$ -th Bernoulli number.

Proof: This follows directly from Cor. 10.1.1 and the Maclaurin series for $\overline{\tanh}$ (National Bureau of Standards, 1965), whose interval of convergence is $(-\pi/2, \pi/2)$. □

¹For neural networks there is little practical difference between different sigmoids, since the effect of scaling can be accomplished by modifying the weights, and the effect of translation by applying a constant bias to the neuron; see ?.

10.2.3 Logistic Function

Definition 10.2.2 *The sigmoid function most commonly used in neural networks is the logistic function, given by*

$$\text{lgst}(x) = \frac{1}{1 + e^{-x/T}}. \quad (10.16)$$

It can also be defined as a translated, scaled hyperbolic tangent:

$$\text{lgst}(x) = \frac{1}{2} \tanh\left(\frac{x}{2T}\right) + \frac{1}{2}. \quad (10.17)$$

It is a soft threshold, as can be seen from its behavior:

$$\begin{aligned} \text{lgst}(-\infty) &= 0, \\ \text{lgst}(0) &= 1/2, \\ \text{lgst}(+\infty) &= 1. \end{aligned}$$

This sigmoid is most useful when the output is to be interpreted as a probability. The parameter T , commonly called “computational temperature,” adjusts the slope of the sigmoid at the origin (which is $1/T$). At $T = 0$ it becomes a step function (threshold). We consider here the corresponding local transformation $\overline{\text{lgst}} : \Phi_{[a,b]}(\Omega) \rightarrow \Phi_{[0,1]}(\Omega)$, where $-\pi T < a < b < \pi T$.

Formula 10.2.2

$$\overline{\text{lgst}}(\phi) = \frac{1}{2} + \frac{1}{4T}\phi - \frac{1}{48T^3}\phi^3 + \frac{1}{480T^5}\phi^5 - \frac{17}{80640T^7}\phi^7 + \dots \quad (10.18)$$

Proof: The series can be derived by direct differentiation, or from the series for \tanh by the relation $\text{lgst}(x) = \frac{1}{2} \tanh(x/2T) + \frac{1}{2}$. Since the radius of convergence for the \tanh series is $\pi/2$, the radius for lgst will be πT .

□

10.3 Replacing Local Field Polynomials by Sigmoid Functions

Sections 10.2.2 and 10.2.3 showed how a local sigmoid transformation could be computed by a local polynomial. On the other hand, artificial neurons often have a built in sigmoid function, and we can expect many field computers

will provide a local sigmoid transformation. For this reason it is important to ask the converse question: When can a field polynomial be computed by a local sigmoid transformation? The following equations characterize a sigmoid centered at the origin, but with a bias of b , a slope m at the origin, and asymptotic values $b \pm \mu$:

$$\sigma(0) = b, \quad (10.19)$$

$$\sigma(x) - b = b - \sigma(-x), \quad (10.20)$$

$$\sigma(+\infty) = b + \mu, \quad (10.21)$$

$$\sigma'(+\infty) = 0, \quad (10.22)$$

$$\sigma'(0) = m, \quad (10.23)$$

$$\sigma'(x) \geq 0, \quad (10.24)$$

$$\sigma'(x) < m, \quad x \neq 0. \quad (10.25)$$

Now consider a power series for σ :

$$\sigma(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots \quad (10.26)$$

Equation 10.19 tells us $a_0 = b$. Also, Eq. 10.20 tells us $\sigma(x) + \sigma(-x) = 2b$. But,

$$\sigma(-x) = a_0 - a_1x + a_2x^2 - a_3x^3 + \dots \quad (10.27)$$

Therefore,

$$2b = \sigma(x) + \sigma(-x) = 2b + 2a_2x^2 + 2a_4x^4 + \dots, \quad (10.28)$$

and we conclude that the power series contains only odd powers. Furthermore, by Eq. 10.23 and

$$\sigma'(x) = a_1 + 3a_3x^2 + \dots \quad (10.29)$$

we know $a_1 = m$. Thus the power series for a sigmoid must look like this:

$$\sigma(x) = b + mx + a_3x^3 + \dots \quad (10.30)$$

This suggests that we ask when a cubic of the form

$$f(x) = b + mx + a_3x^3 \quad (10.31)$$

can be approximated by a sigmoid function. This cubic can be easily seen to satisfy Eqs. 10.19, 10.20 and 10.23. Since $f'(x) = m + 3a_3x^2$, it will satisfy Eq. 10.25 only if $a_3 < 0$. Thus we rewrite it

$$f(x) = b + mx - ax^3, \quad (10.32)$$

where $a > 0$. The cubic differs from the sigmoid in its asymptotic properties (Eqs. 10.21 and 10.22), so we must restrict our attention to the portion of the cubic between its extrema. To determine their location, set $f'(r) = 0$ and then since $m - 3ar^2 = 0$ we find

$$r = \pm\sqrt{m/3a}. \quad (10.33)$$

Therefore, we will be able to approximate $f(x)$ by $\sigma(x)$ only if $|x| \leq r$, the *radius* of the cubic sigmoid. Equation 10.24 is satisfied within this radius.

For a specific example, suppose we want to approximate Eq. 10.32 by a scaled, translated hyperbolic tangent:

$$\sigma(x) = b + \mu \tanh(\nu x). \quad (10.34)$$

First, observe that

$$\sigma(0) = b, \quad (10.35)$$

$$\sigma'(0) = \mu\nu. \quad (10.36)$$

Therefore, if we let $m = \mu\nu$, then f and σ will agree at 0 in their value and derivative. We need an additional condition to determine the values of μ and ν . Since the argument of x is restricted to $[-r, r]$ we require $f(r) = \sigma(r)$. Substituting into the power series for \tanh yields

$$b + mr - ar^3 = b + \mu[\nu r - (\nu r)^3/3 + \dots]. \quad (10.37)$$

Therefore,

$$mr - ar^3 = \mu\nu r - (\mu\nu^3/3)r^3 + \dots. \quad (10.38)$$

If we neglect higher order terms, and let $m = \mu\nu$, then we have

$$mr - ar^3 = mr - (m\nu^2/3)r^3. \quad (10.39)$$

Hence, $a = m\nu^2/3$ and so $\nu = \sqrt{3a/m} = r^{-1}$. Thus, the sigmoid approximates the cubic if

$$\nu \approx r^{-1}, \quad (10.40)$$

$$\mu \approx mr. \quad (10.41)$$

That is, the cubic

$$f(x) = b + mx - ax^3, \quad |x| \leq r = \sqrt{m/3a} \quad (10.42)$$

may be approximated by the sigmoid

$$\sigma(x) = b + mr \tanh(x/r). \quad (10.43)$$

Therefore, if we must compute the local field polynomial

$$T(\phi) = b\mathbf{1} + mL\phi - a(L\phi)^3, \quad (10.44)$$

then this can be accomplished by a composition of a local sigmoid and a linear transformation:

$$T(\phi) \approx \bar{\sigma}(L\phi). \quad (10.45)$$

This is especially important for neural networks, which generally have the effect of a linear transformation followed by a sigmoid nonlinearity. In this case the input weight matrix is $r^{-1}L$ and the output is scaled by mr and biased by b .²

²Of course this is not the only way to approximate a local cubic by a sigmoid. We could for example pick μ and ν to minimize the L_2 error: $\|\sigma - f\|_2$.

Chapter 11

The Problem of High Dimensional Gradients

11.1 The Problem

Consider the k th term in the Taylor expansion:

$$\frac{1}{k!} \nabla^k T(\phi) \alpha^{(k)}. \quad (11.1)$$

If $\alpha \in \Phi[a, b]$ is a field defined over a closed segment of the real line, then $\alpha^{(k)}$ is a field defined over a k -dimensional hypercube. Unfortunately, fields of dimension higher than 3 are, seemingly, unrealizable.¹ In other words, there seems to be no way to build a field computer that can store fields like $\alpha^{(k)}$, for $k > 3$. The immediate problem of raising α to a high power can be eliminated by alternate forms of the Taylor expansion, such as that given by Horner's Rule (Sec. 9.2.3.3, p. 195), but we are left with the problem that the gradients are of high dimension. In particular, if $T : \Phi(\Omega_1) \rightarrow \Phi(\Omega_2)$, then

$$\nabla^k T(\phi) \in \Phi(\Omega_2 \times \Omega_1^k). \quad (11.2)$$

In general, for any field polynomial $\sum_{k=0}^N K_k \phi^{(k)}$ the coefficient fields are of successively higher dimension, $K_k \in \Phi(\Omega_2 \times \Omega_1^k)$.

There are several ways to avoid higher dimensional fields. One, which is discussed in detail in ?, is to discretize the field. A second solution is to

¹See Caulfield (1987) for an exception, however.

represent one dimension by time. For example, if the field Ψ_{st} is represented by a time-varying field $(\Psi_s)_t$, then in the field product

$$(\Psi\phi)_s = \int_{\Omega} \Psi_{st}\phi_t dt \quad (11.3)$$

the integral can be computed by accumulating the product $\Psi_{st}\phi_t$ over an interval of time $\Omega = [t_o, t_f]$. Of course this approach buys only one additional dimension. A third solution is applicable when a higher dimensional field is *sparse*, that is, zero over most of its domain. For example, if $\Psi \in \Phi(\Omega \times \Omega)$ satisfies

$$\Psi_{st} = 0, \quad \text{if } |s - t| > \epsilon, \quad (11.4)$$

then it may be approximately represented by the lower dimensional $\psi \in \Phi(\Omega)$:

$$\phi_s = \int_{s-\epsilon}^{s+\epsilon} \Psi_{st} dt. \quad (11.5)$$

We can replace a higher dimensional general product $\Psi\phi$ by a lower dimensional local product $2\epsilon(\psi \times \phi)$ as shown by this derivation:

$$\begin{aligned} (\Psi\phi)_s &= \int_{\Omega} \Psi_{st}\phi_t dt \\ &= \int_{s-\epsilon}^{s+\epsilon} \Psi_{st}\phi_t dt \\ &\approx \int_{s-\epsilon}^{s+\epsilon} \psi_s \phi_t dt \\ &= \psi_s \int_{s-\epsilon}^{s+\epsilon} \phi_t dt \\ &\approx \psi_s \int_{s-\epsilon}^{s+\epsilon} \phi_s dt \\ &= 2\epsilon\psi_s\phi_s. \end{aligned}$$

There are also various mixed strategies, and no doubt other solutions that may be useful in various situations.

11.2 Field Computation in Alternate Domains

11.2.1 Introduction

Another approach to the problem of higher dimensional gradients is to map the gradient into a lower dimensional space, and do the corresponding computations in this lower dimensional space. For example, suppose $\Omega \subset E^2$ is an appropriate subset of Euclidean two-space. Further suppose that $\Psi, X \in \Phi(\Omega^2)$ are two higher dimensional (in fact four-dimensional) fields. It seems that there ought to be a continuous function $R : \Phi(\Omega^2) \rightarrow \Phi(\Omega')$, with $\Omega' \subset E^2$, that maps four-dimensional spaces into two-dimensional spaces in such a way that we can find an operation $IP : \Phi(\Omega') \times \Phi(\Omega') \rightarrow \mathbb{R}$ that does a four-dimensional inner product on the two-dimensional surrogates of the fields:

$$F \cdot G = IP[R(\Psi), R(X)]. \quad (11.6)$$

In fact, such continuous maps R exist; they are based on space-filling curves, such as Peano curves (see Section 11.3). There remains the problem of whether lower dimensional correspondents of the product operation (IP in this example) exist. In this section we show that they do.

11.2.2 Measure Preserving Change of Domain

Suppose we have a field $\Psi \in \Phi(\Omega)$ and we want to represent it by a field $\psi \in \Phi(\Omega')$. A typical motivation for this would be that Ω' is of lower dimension than Ω . Thus we want $R : \Phi(\Omega) \rightarrow \Phi(\Omega')$ such that $\psi = R(\Psi)$. The change of domain is accomplished by letting

$$R(\Psi) = \Psi \circ C, \quad (11.7)$$

where $C : \Omega' \rightarrow \Omega$ is bijective.² The transformation R has the correct domain and range, and loses no information (since C is bijective). It is also easy to establish that R is linear:

$$\begin{aligned} [R(a\Psi + bX)]_x &= [(a\Psi + bX) \circ C]_x \\ &= (a\Psi + bX)_{C_x} \\ &= a\Psi_{C_x} + bX_{C_x} \\ &= a(\Psi \circ C)_x + b(X \circ C)_x \\ &= a[R(\Psi)]_x + b[R(X)]_x. \end{aligned}$$

²We will impose additional constraints shortly.

Hence $R(a\Psi + bX) = aR(\Psi) + bR(X)$.

We now impose an additional condition on C : we require it to be *measure-preserving*. That is, for $S \subseteq \Omega$ and $S' = C^{-1}[S]$ (i.e., S' is the inverse image of S under C) we have

$$\int_S f_t dt = \int_{S'} f(C_s) ds. \quad (11.8)$$

That is,

$$\int_S \Phi_t dt = \int_{S'} [R(\Phi)]_s ds. \quad (11.9)$$

Under this assumption we find that R is an *isometry* (isometric transformation), since

$$\begin{aligned} \|R(\Psi) - R(X)\|^2 &= \|R(\Psi - X)\|^2 \\ &= \|(\Psi - X) \circ C\|^2 \\ &= \int_{\Omega'} (\Psi - X)_{C_s}^2 ds \\ &= \int_{\Omega} (\Psi - X)_t^2 dt \\ &= \|\Psi - X\|^2. \end{aligned}$$

An isometry is a homeomorphism, so we've shown that under these assumptions the spaces $\Phi(\Omega)$ and $\Phi(\Omega')$ are homeomorphic (and in fact isometric). Further, since an isometry is necessarily continuous, we've shown that the transformation R is continuous.

11.2.3 Transformations in the Alternate Domain

To complete the replacement of a domain Ω by a more convenient domain Ω' , we must also replace transformations T on $\Phi(\Omega)$ by corresponding transformations T' on $\Phi(\Omega')$ so that

$$T(\Psi) = T'[R(\Psi)]. \quad (11.10)$$

(In some cases it is more convenient to have $R[T(\Psi)] = T'[R(\Psi)]$.)

11.2.3.1 INNER PRODUCT

We consider some examples, of which the simplest is the inner product. We want $\text{IP} : \Phi(\Omega') \times \Phi(\Omega') \rightarrow \mathbb{R}$ such that

$$\Psi \cdot X = \text{IP}[R(\Psi), R(X)]. \quad (11.11)$$

We proceed to derive IP:

$$\begin{aligned} \Psi \cdot X &= \int_{\Omega} \Psi_t X_t dt \\ &= \int_{\Omega} (\Psi \circ C \circ C^{-1})_t (X \circ C \circ C^{-1})_t dt \\ &= \int_{\Omega} [R(\Psi)]_{C_t^{-1}} [R(X)]_{C_t^{-1}} dt \\ &= \int_{\Omega'} [R(\Psi)]_s [R(X)]_s ds \\ &= R(\Psi) \cdot R(X). \end{aligned}$$

Hence, let

$$\text{IP}(\psi, \chi) = \psi \cdot \chi, \quad (11.12)$$

and then $\Psi \cdot X = \text{IP}[R(\Psi), R(X)]$, as we might expect.

11.2.3.2 LOCAL TRANSFORMATIONS

By an analogous derivation it is easy to see that if $\bar{f}(\Psi, X)$ is any local binary operation, then

$$R[\bar{f}(\Psi, X)] = \bar{f}[R(\Psi), R(X)]. \quad (11.13)$$

Simply observe that

$$\begin{aligned} \{R[\bar{f}(\Psi, X)]\}_s &= [\bar{f}(\Psi, X)]_{C_s} \\ &= f(\Psi_{C_s}, X_{C_s}) \\ &= f\{[R(\Psi)]_s, [R(X)]_s\} \\ &= \{\bar{f}[R(\Psi), R(X)]\}_s. \end{aligned}$$

That is R commutes with local transformations.

11.2.3.3 GENERAL PRODUCT

Now we consider a more useful example, the general product. Suppose

$$\begin{aligned}\Psi &\in \Psi(\Omega'' \times \Omega), \\ X &\in \Psi(\Omega).\end{aligned}$$

We want to replace the product ΨX by a product $\psi \chi$ in which

$$\begin{aligned}\psi &\in \Phi(\Omega'' \times \Omega'), \\ \chi &\in \Phi(\Omega').\end{aligned}$$

Derive:

$$\begin{aligned}(\Psi X)_s &= \int_{\Omega} \Psi_{st} X_t dt \\ &= \int_{\Omega} [\Psi \circ (\mathbf{I} \times C) \circ (\mathbf{I} \times C)^{-1}]_{st} (X \circ C \circ C^{-1})_t dt,\end{aligned}$$

where the direct product $(\mathbf{I} \times C)_{st} = (s, C_t)$. Notice that $(\mathbf{I} \times C)$ is an isomorphism, and let

$$S(\Psi) = \Psi \circ (\mathbf{I} \times C). \quad (11.14)$$

Then continue the derivation:

$$\begin{aligned}(\Psi X)_s &= \int_{\Omega} [S(\Psi)]_{s, C_t^{-1}} [R(X)]_{C_t^{-1}} dt \\ &= \int_{\Omega'} [S(\Psi)]_{su} [R(X)]_u du \\ &= [S(\Psi)R(X)]_s.\end{aligned}$$

Hence,

$$\Psi X = S(\Psi)R(X). \quad (11.15)$$

11.2.3.4 OUTER PRODUCT

Next we consider the computation of an altered outer product:

$$\text{OP}(\phi, \psi) = R(\phi \wedge \psi). \quad (11.16)$$

Our goal is to compute $\text{OP}(\phi, \psi)$ without generating the higher-dimensional field $\phi \wedge \psi$. The outer product is a bilinear operator, but can be expressed as a local product of two linear operators, as follows:

$$\phi \wedge \psi = (\phi \wedge \mathbf{1}) \times (\mathbf{1} \wedge \psi). \quad (11.17)$$

This is convenient, since R commutes with local transformations:

$$R(\phi \wedge \psi) = R(\phi \wedge \mathbf{1}) \times R(\mathbf{1} \wedge \psi). \quad (11.18)$$

Each of the two factors are linear operators, so they have kernels which we compute as follows (Eq. 7.48):

$$\mathbf{K}_{st} = [(\delta_t \wedge \mathbf{1}) \circ C]_s = (\delta_t \wedge \mathbf{1})_{C_s}, \quad (11.19)$$

$$\mathbf{K}'_{st} = [(\mathbf{1} \wedge \delta_t) \circ C]_s = (\mathbf{1} \wedge \delta_t)_{C_s}. \quad (11.20)$$

Then,

$$\text{OP}(\phi, \psi) = \mathbf{K}\phi \times \mathbf{K}'\psi. \quad (11.21)$$

This formula involves no higher-dimensional fields (except \mathbf{K} and \mathbf{K}' , which can instead be provided as primitive operators).

11.2.3.5 OUTER PRODUCT POWER

Finally, we define an altered outer-product power:

$$P_n(\phi) = R[\phi^{(n)}]. \quad (11.22)$$

To accomplish this it will be useful to have two domain bijections:

$$C : \Omega' \rightarrow \Omega \times \Omega', \quad (11.23)$$

$$C_0 : \Omega' \rightarrow \Omega, \quad (11.24)$$

and the corresponding representation transformations:

$$R(\Psi) = \Psi \circ C, \quad (11.25)$$

$$R_0(\Psi) = \Psi \circ C_0. \quad (11.26)$$

Note that $R : \Phi(\Omega \times \Omega') \rightarrow \Phi(\Omega')$ and $R_0 : \Phi(\Omega) \rightarrow \Phi(\Omega')$. Then we define the powers recursively:

$$P_1(\phi) = R_0(\phi), \quad (11.27)$$

$$P_{n+1}(\phi) = \text{OP}[\phi, P_n(\phi)], \quad n \geq 1, \quad (11.28)$$

where $\text{OP}(\phi, \psi) = R(\phi \wedge \psi)$. Note that the computation of $P_n(\phi)$ makes use of no fields of dimension greater than $\Omega \times \Omega'$.

11.2.3.6 FIELD POLYNOMIALS

The preceding derivations show that an arbitrary field polynomial can be implemented by products over an alternative domain as follows:

$$\sum_{n=0}^N K_n \phi^{(n)} = K_0 + \sum_{n=1}^N S(K_n) P_n(\phi). \quad (11.29)$$

Notice that all the fields that constitute this polynomial belong to $\Phi(\Omega)$, $\Phi(\Omega')$ or $\Phi(\Omega \times \Omega')$. In particular:

$$K_0 \in \Phi(\Omega), \quad (11.30)$$

$$S(K_n) \in \Phi(\Omega \times \Omega'), \quad n \geq 1, \quad (11.31)$$

$$P_n(\phi) \in \Phi(\Omega'). \quad (11.32)$$

11.3 Reduction of Dimension

We now return to the problem of the reduction of dimension. We are given Ω and Ω' , where Ω' is assumed to be of lower dimension than Ω . To apply the preceding results we need a measure-preserving bijection $C : \Omega' \rightarrow \Omega$. Fortunately, such functions exist, although they are a little unusual. For example, suppose that $\Omega = [0, 1]^2$ and $\Omega' = [0, 1]$. Then there are various space-filling curves (such as the Peano curve) $C : [0, 1] \rightarrow [0, 1]^2$. These functions are continuous, because they are curves, and bijections, because they are space-filling. (They are not, however, homeomorphisms, because their inverses are not continuous, a result due to Brouwer (Hausdorff, 1957, Section 36, pp. 228–236).

We also require that the space-filling curve C be measure-preserving. This presents no difficulty if we take a curve C that is the limit of a sequence C_n of functions $C_n : \Omega \rightarrow \Omega_n$ that divide Ω into narrower and narrower “strips” and preserve the measure. For example, we may have $\Omega_n = [0, 1/2^n] \times [0, 2^n]$. In practice, we are not much concerned about what happens in the limit, since we have to use one of the finite approximations C_n anyway.

The foregoing discussion suggests that space-filling curves and, more generally, fractal curves may be important in the representation of higher dimensional fields. Thus it is especially intriguing that some structures in the brain seem to have a fractal geometry.

Chapter 12

Neural Implementation

12.1 Discretizing Fields

The goal of *discretizing* a field is to replace the continuous ensemble $\phi_t, t \in \Omega$, by a finite set of values $u_i, i = 1, 2, \dots, n$. There are many ways this can be done. For example, we can use *sampling*, selecting at random n points t_i from Ω , and letting $u_i = \phi_{t_i}$. Other discretization methods depend on partitioning Ω :

$$\Omega = \Omega_1 \cup \Omega_2 \cup \dots \cup \Omega_n, \quad (12.1)$$

where $\Omega_i \cap \Omega_j = \emptyset$. Given such a partition, we can select $t_i \in \Omega_i$ and let $u_i = \phi_{t_i}$. This t_i may be selected randomly or by some regular process (e.g., each t_i may be the “center” of the corresponding Ω_i). The foregoing methods all suffer from *sampling error*, i.e., the values at the chosen points t_i might not be representative of the values at all the points in Ω_i .

Sampling error can be avoided by averaging the field’s value over each partition. Thus we let

$$u_i = \frac{\int_{\Omega_i} \phi_t dt}{|\Omega_i|}. \quad (12.2)$$

This process leads to its own kind of error – *averaging error* – but it is more tractable in many situations. In the remainder of Ch. 12 we assume that field are discretized by averaging.¹

¹It should be observed that much of traditional numerical analysis is concerned with the discretization of continuous functions, and hence is relevant here.

12.2 Linear Field Transformations

Consider first a product between a fixed field L and a variable input field ϕ :²

$$\psi = L\phi. \quad (12.3)$$

Suppose that the input field's domain has been partitioned:

$$\Omega = \Omega_1 + \Omega_2 + \cdots + \Omega_n, \quad (12.4)$$

and that the output field's domain has also been partitioned:

$$\Omega' = \Omega'_1 + \Omega'_2 + \cdots + \Omega'_m. \quad (12.5)$$

Let \mathbf{u} be the discretization of the input field:

$$u_j = \omega_j \int_{\Omega_j} \phi_t dt, \quad (12.6)$$

where ω_j normalizes for the size of Ω_j :

$$\omega_j = |\Omega_j|^{-1}. \quad (12.7)$$

Let \mathbf{v} be the discretization of the output field:

$$v_i = \omega'_i \int_{\Omega'_i} \psi_s ds, \quad (12.8)$$

where

$$\omega'_i = |\Omega'_i|^{-1}. \quad (12.9)$$

Next derive a formula for \mathbf{v} in terms of \mathbf{u} :

$$\begin{aligned} v_i &= \omega'_i \int_{\Omega'_i} \psi_s ds \\ &= \omega'_i \int_{\Omega'_i} (L\phi)_s ds \\ &= \omega'_i \int_{\Omega'_i} \int_{\Omega} L_{st} \phi_t dt ds \\ &= \omega'_i \int_{\Omega'_i} \sum_j \int_{\Omega_j} L_{st} \phi_t dt ds. \end{aligned}$$

²We use '+' to emphasize that the Ω_i are disjoint.

Since the average value of ϕ_t over Ω_j is u_j , we substitute the latter in the equation to get:

$$\begin{aligned} v_i &\approx \omega'_i \int_{\Omega'_i} \sum_j \int_{\Omega_j} L_{st} u_j dt ds \\ &= \sum_j u_j \omega'_i \int_{\Omega'_i} \int_{\Omega_j} L_{st} dt ds. \end{aligned}$$

Now let

$$M_{ij} = \omega'_i \int_{\Omega'_i} \int_{\Omega_j} L_{st} dt ds, \quad (12.10)$$

and we have

$$v_i \approx \sum_j u_j M_{ij}. \quad (12.11)$$

Hence, as expected, \mathbf{v} can be approximated by a matrix-vector product:

$$\mathbf{v} \approx M\mathbf{u}. \quad (12.12)$$

Clearly the approximation will be good to the extent that the fields L and ϕ are smooth, and to the extent that the partition is fine in the regions of highest gradient.

We relate the foregoing to neural networks. The product $L\phi$ can be approximately implemented by a single layer neural network with linear neurons as follows. Let the activity level u_j of the j th input neuron be the average value of ϕ over Ω_j . Let the interconnection weights (synaptic strengths) be given by Eq. 12.10. Then the activity v_i of output neuron i approximates the average value of ψ over Ω'_i .

12.3 Nonlinear Field Transformations

forthcoming

Appendix A

Notation

Notation	Description	Definition
\mathbb{Z}	integers	
\mathbb{Q}	rationals	
\mathbb{C}	complex numbers	
\mathbb{R}	real numbers	
\mathbb{R}_+	positive reals	Def. 2.1.8 (p. 10)
\mathbb{R}_*	non-negative reals	Sec. 2.1.1 (p. 7)
\mathcal{E}^n	n -dimensional Euclidean space	Def. 2.1.11 (p. 11)
$d(\cdot, \cdot)$	arbitrary metric	Sec. 2.1.1 (p. 7)
d_d	discrete metric	Sec. 2.1.4 (p. 11)
L_p	L_p metric	Sec. 2.1.2 (p. 8)
$L_p^{\mathbf{w}}$	weighted L_p metric	Def. 2.1.9 (p. 10)
\mathcal{C}_1		Def. 2.1.16 (p. 12)
\mathcal{C}_2		Def. 2.1.17 (p. 13)
\mathcal{C}_∞		Def. 2.1.15 (p. 12)
$\mathcal{C}[a, b]$	continuous real functions	Def. 3.2.4 (p. 36)
\mathcal{H}	Hilbert space	Def. 5.2.2 (p. 81)
$\mathcal{L}(X, Y)$	bounded linear operators	Prop. 3.4.1 (p. 38)
$B_r(c)$	open ball	Def. 2.2.1 (p. 13)
$f[S]$	image of a set	Def. 2.4.2 (p. 21)
$\mathbf{2}$	the set $\{0, 1\}$	Prop. 2.4.2 (p. 21)
$\mathbf{0}$	zero vector or field	Def. 3.1.2 (p. 32)
\oplus	addition on linear space	Def. 3.1.2 (p. 32)
\oplus	direct sum	Prop. 5.2.11 (p. 84)
\otimes	scaling on linear space	Def. 3.1.2 (p. 32)

Notation	Description	Definition
$N(\cdot)$	arbitrary norm	Def. 3.2.1 (p. 35)
$\ \cdot\ $	arbitrary norm	Not. 3.2.1 (p. 35)
$\ \cdot\ _p$	L_p -norm	Def. 3.2.2 (p. 35)
$dT(\phi, \alpha)$	Gâteaux derivative	Sec. 3.4.2 (p. 39)
$F \circ G$	function composition	
\bar{z}, z^*	complex conjugate	Def. 4.2.3 (p. 43)
\Re	real part of complex number	Ex. 4.2.2 (p. 43)
\Im	imaginary part of complex number	Ex. 4.2.2 (p. 43)
$ \cdot $	absolute value or modulus	Def. 4.2.4 (p. 44), Def. 4.3.1 (p. 45)
\arg	argument or phase	Def. 4.3.2 (p. 45)
2π	$2 \times \pi$ representing one cycle	Not. 4.3.1 (p. 45)
cis	$\cos + i \sin$ function	Sec. 4.3.2 (p. 46)
$A \angle \theta$	phasor notation	p. 54
$\mathbf{x} \cdot \mathbf{y}$	inner product	Def. 5.1.1 (p. 69)
$\langle \mathbf{x}, \mathbf{y} \rangle$	inner product	Not. 5.1.1 (p. 70)
$\langle \phi \psi \rangle$	Dirac inner product	Def. 5.1.6 (p. 71)
$\langle \phi $	Dirac bra notation	Not. 5.1.2 (p. 72)
$ \psi \rangle$	Dirac ket notation	Def. 5.1.2 (p. 72)
$\langle \phi L \psi \rangle$	Dirac quadratic form	Def. 5.1.8 (p. 73)
$ \phi \rangle \langle \psi $	Dirac outer product	Def. 5.2.16 (p. 89), Def. ?? (p. ??)
ℓ_1	absolutely-summable sequences	Def. 7.4.3.2 (p. 162)
ℓ_2	square-summable sequences	Def. 5.1.5 (p. 71)
T^\dagger	adjoint (conjugate or Hermetian transpose)	Def. 5.1.7 (p. 72), Prop. 5.2.17 (p. 89)
$\mathcal{E}\{\cdot\}$	expectation value	
$\text{Var}\{\cdot\}$	variance	
$\Phi_K(\Omega)$	field space	Sec. 7.1.2 (p. 150)
$\Psi\phi, \phi\Psi, \phi X\psi, \Psi X$	field products	Sec. 7.4.1 (p. 155)
\wedge	field outer product	Def. 7.4.5 (p. 156)
δ_t	Dirac unit impulse	Eq. 7.45 (p. 167)

Bibliography

- Berberian, S. (1961). *Introduction to Hilbert Space*. New York: Oxford University Press.
- Bland, B. H., Anderson, P., Ganes, T., & Sveen, O. (1978). Automated analysis of rhythmicity of physiologically identified hippocampal formation neurons. *Experimental Brain Research*, 38, 205–219.
- Boas, M. L. (1983). *Mathematical Methods in the Physical Sciences*. New York, NY: John Wiley & Sons, second edition.
- Bobrow, L. S. (1981). *Elementary Linear Circuit Analysis*. New York, NY: Holt, Rinehart & Winston.
- Brillouin, L. (1956). *Science and Information Theory*. New York, NY: Academic Press.
- Caulfield, H. J. (1987). Parallel N^4 weighted optical interconnections. *Applied Optics*, 26(19), 4039–4040.
- Cherry, C. (1978). *On Human Communication*. Cambridge, MA: MIT Press.
- DARPA (1988). *DARPA Neural Network Study*. AFCEA International Press.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Comm. Pure Applied Math.*, 41, 909–996.
- Daubechies, I., Grossman, A., & Meyer, Y. (1986). Painless non-orthogonal expansions. *Journal of Mathematical Physics*, 27, 1271–83.
- Daugman, J. (1993). An information-theoretic view of analog representation in striate cortex. In E. Schwartz (Ed.), *Computational Neuroscience* (pp. 403–423). Cambridge: MIT Press.

- Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, 20, 847–856.
- Daugman, J. G. (1984). Spatial visual channels in the fourier plane. *Vision Research*, 24, 891–910.
- Daugman, J. G. (1985a). Representational issues and local filter models of two-dimensional spatial visual encoding. In D. Rose & V. G. Dobson (Eds.), *Models of the Visual Cortex* (pp. 96–107). New York: John Wiley & Sons.
- Daugman, J. G. (1985b). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, 2(7), 1160–1169.
- Farhat, N. H., Psaltis, D., Prata, A., & Paek, E. (1985). Optical implementation of the hopfield model. *Applied Optics*, 24, 1469–1475.
- Gabor, D. (1946). Theory of communication. *Journal of the Institution of Electrical Engineers*, 93, Part III, 429–57.
- Hamming, R. W. (1986). *Coding and Information Theory*. Englewood-Cliffs, NJ: Prentice-Hall.
- Hamming, R. W. (1989). *Digital Filters*. Englewood-Cliffs, NJ: Prentice-Hall, third edition.
- Hausdorff, F. (1957). *Set Theory*. New York, NY: Chelsea. Transl. John R. Aumann et al.
- Heil, C. & Walnut, D. (1989). Continuous and discrete wavelet transforms. *SIAM Review*, 31(4), 628–66.
- Jones, J. & Palmer, L. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58(6), 1233–1258.
- Kainen, P. C. (1992). *Orthogonal Dimension and Tolerance*. unpublished, Industrial Math, Washington, DC.
- MacKay, D. M. (1969). *Information, Mechanism and Meaning*. Cambridge, MA: MIT Press.

- MacLennan, B. (1987a). *Field Computation and Nonpropositional Knowledge*. Technical Report NPS52-87-040, Naval Postgraduate School.
- MacLennan, B. (1988a). Field computation: A model of massively parallel computation in electronic, optical, molecular and biological systems. In *Parallel Models of Intelligence: How Can Slow Components Think So Fast?*, AAAI Spring Symposium (pp. 180–183).
- MacLennan, B. (1989). *Continuous Computation: Taking Massive Parallelism Seriously*. Technical Report CS-89-83, University of Tennessee, Knoxville, Department of Computer Science.
- MacLennan, B. J. (1987b). Technology-independent design of neurocomputers: The universal field computer. In M. Caudill & C. Butler (Eds.), *Proceedings of the IEEE First International Conference on Neural Networks*, volume 3 (pp. 39–49).: IEEE Press.
- MacLennan, B. J. (1988b). Logic for the new AI. In J. H. Fetzer (Ed.), *Aspects of Artificial Intelligence* (pp. 163–192). Dordrecht: Kluwer Academic Publishers.
- MacLennan, B. J. (1993). Information processing in the dendritic net. In K. H. Pribram (Ed.), *Rethinking Neural Networks: Quantum Fields and Biological Data* (pp. 161–197). Hillsdale, NJ: Lawrence Erlbaum. Also available from web.eecs.utk.edu/~mclennan.
- MacLennan, B. J. (2009). Field computation in natural and artificial intelligence. In R. Meyers et al. (Ed.), *Encyclopedia of Complexity and System Science* chapter 6, entry 199, (pp. 3334–3360). Springer.
- MacLennan, B. J. (2017). Field computation: A framework for quantum-inspired computing. In S. Bhattacharyya, U. Maulik, & P. Dutta (Eds.), *Quantum Inspired Computational Intelligence: Research and Applications* chapter 3, (pp. 85–110). Cambridge MA: Morgan Kaufmann / Elsevier.
- Mallat, S. (1989a). Multiresolution approximations and wavelet orthonormal bases of $L^2(r)$. *Transactions American Mathematical Society*, 315(1), 69–87.

- Mallat, S. (1989b). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 674–693.
- Marsden, J. E. (1973). *Basic Complex Analysis*. San Francisco, CA: W.H. Freeman & Co.
- Marčelja, S. (1980). Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America*, 70, 1297–1300.
- Mathematical Society of Japan (1980). *Encyclopedic Dictionary of Mathematics*. Cambridge: MIT Press. S. Iyanaga and Y. Kawada (eds.).
- McCarty, G. (1967). *Topology*. New York, NY: McGraw-Hill.
- Mead, C. (1989). *Analog VLSI and Neural Systems*. Addison-Wesley.
- Mendelson, B. (1975). *Introduction to Topology*. Allyn & Bacon.
- Moore, R. E. (1985). *Computational Functional Analysis*. Chichester: Ellis Horwood.
- Moore, T. O. (1964). *Elementary General Topology*. Englewood Cliffs: Prentice-Hall.
- Movshon, J. (1979). Two-dimensional spatial frequency tuning of cat striate cortical neurons. *Society for Neuroscience Abstracts*, 9, 799.
- National Bureau of Standards (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. U.S. Government Printing Office.
- Nemytskii, V. V. & Stepanov, V. V. (1989). *Qualitative Differential Equations*. New York, NY: Dover, reprint of 1960 Princeton Univ. Press edition.
- Palais, B. (2001). π is wrong! *The Mathematical Intelligencer*, 23(3), 7–8.
- Pollen, D. & Ronner, S. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science*, 212, 1409–1411.
- Pribram, K. (1991). *Brain and Perception: Holonomy and Structure in Figural Processing*. Hillsdale, NJ: Lawrence Erlbaum.

- Pribram, K. H. & Carlton, E. H. (1986). Holonomic brain theory in imaging and object perception. *Acta Psychologica*, 63, 175–210.
- Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423, 623–656. Reprinted in Shannon & Weaver (1963), *The Mathematical Theory of Communication*, Urbana, IL: University of Illinois Press.
- Strang, G. (1989). Wavelets and dilation equations: A brief introduction. *SIAM Review*, 31(4), 614–627.
- Taylor, E. F. & Wheeler, J. A. (1966). *Spacetime Physics*. San Francisco, CA: W.H. Freeman & Co.
- Watson, A. B. (1982). Summation of grating patches indicate many types of detector at one retinal location. *Vision Research*, 22, 17–25.
- Yu, F. T. S. (1976). *Optics and Information Theory*. New York: John Wiley & Sons.