Essential Dimension and Degrees of Freedom for Spatial Waveform Channels

Ram Somaraju

June 5^{th}, 2008

A thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy.

Department of Information Engineering

Research School of Information Sciences and Engineering

The Australian National University

Declaration

The contents of this thesis are the result of original research and have not been submitted for a higher degree to any other university or institution. Much of the work in this thesis is currently being prepared for submission or has been published in the following papers.

Journal Papers

1. R. Somaraju and J. Trumpf. Degrees of Freedom and Essential Dimension of Communication Channels. In preparation for submission to IEEE Transactions on Information Theory.

2. R. Somaraju and J. Trumpf. Spatial Waveform Channels. In preparation for submission to IEEE Transactions on Antennas and Propagation.

3. R. Somaraju and L. Hanlen. Uncertainty Principles for Signal Concentrations. To be submitted.

Conference Paper

1. R. Somaraju and L. Hanlen, Uncertainty principles for signal concentrations, Proceedings. 7th Australian Communications Theory Workshop, 2006. pp. 38- 42, 1-3 Feb. 2006.

2. R. Somaraju and J. Trumpf. Generalised Singular Values. The XIXth International Workshop on Operator Theory and Applications, July 22-26, 2008. Accepted for publication.

In addition, during the course of my thesis I have published the following unrelated papers, the contents of which have not been included in this thesis.

Journal Papers

1. R. Somaraju and J. Trumpf. Frequency, temperature and salinity variation of the permittivity of seawater. IEEE Transactions on Antennas and Propagation, 54(11):3441-3448, 2006.

2. Ram Somaraju and Felix Schill. A communication module and TDMA scheduling for a swarm of small submarines. Turk J Elec Engin, 15(2):283-305, 2007.

The research included in this thesis has been performed jointly with Dr. Jochen Trumpf. The majority of work is my own.

Ram Abhinav Somaraju
The Australian National University
December 3^{rd}, 2007.

Acknowledgements

I would like to express my gratitude to all the people who made it possible for me to complete this thesis. They have all supported me in their own unique way, for which I will be forever grateful.

The motivation, guidance and support provided by my supervisor Dr. Jochen Trumpf has been invaluable. Without his help I doubt that I ever would have finished this thesis. I would also like to thank Dr. Leif Hanlen, Dr. Gerard Borg, Prof Rodney Kennedy, Dr. Uwe Zimmer, Dr. Tony Pollok, Dr. Richard Webster and Felix Schill for help and assistance they provided during various stages of my thesis.

I would like to thank my family, friends and colleagues for the continued support and encouragement. Spanna, thanks for being an unbelievable sister. Amrita, you have been wonderful over the last few years. Thank you for helping me, encouraging me, supporting me and listing to my “I have a new idea” monologue over and over agian. Most importantly, I wish to thank my parents Padma and Haranath, whose love, support and willingness to sacrifice everything so that I may have a better life has been the greatest gift of all. It is to my parents that I dedicate this thesis.

Abstract

This thesis introduces Spatial Waveform Channels (SWCs) as a modelling tool to derive fundamental performance bounds for Multiple-Input Multiple-Output (MIMO) antenna systems. In practical MIMO systems, the transmitting and receiving antennas are constrained to be within finite volumes in space and the signals received on individual antennas become mutually correlated with increasing number of antennas. In SWCs one assumes that volumes in space, rather than antennas, can be used as the transmitting and receiving elements. The essential idea behind this assumption is that any rate of transmission that can be achieved by the transmitting and receiving antennas can in theory, also be achieved by using the volumes to which the antennas are constrained. However, if the bounds calculated using SWCs are to be useful, one needs to model the various physical constraints imposed on a MIMO system in the corresponding SWC. This thesis formalises the notion of an SWC and gives it an abstract mathematical definition, the structure of which enables one to impose the limitations in a MIMO system on the corresponding SWC.

Several properties of SWCs are examined. In particular, I study the very general concepts of *degrees of freedom at level-ϵ* and *essential dimension* for compact operators defined on normed spaces. The number of degrees of freedom at level-ϵ of an SWC can be used to determine the number of mutually uncorrelated signals in the corresponding MIMO system that has noise in the receiver proportional to ϵ. Essential dimension of the channel operator determines the number of mutually uncorrelated signals present at the receiver that is largely independent of the noise level at the receiver. Moreover, I show that the concept of degrees of freedom can be used to generalise the notion of singular values of compact operators. These generalised singular values are then used to numerically compute the degrees of freedom and essential dimension for various channels. Finally, uncertainty principles and their application in the context of SWCs are studied.

Notation and Symbols

AWGN	Additive White Gaussian Noise
ISI	Intersymbol interference
MIMO	Multiple Input Multiple Output
SWC	Spatial Waveform Channel
$B_{P,X}(\mathbf{x})$	Open ball of radius P centered at $\mathbf{x}$ in the normed space X
$\overline{B}_{P,X}(\mathbf{x})$	Closed ball of radius P centered at $\mathbf{x}$ in the normed space X
$\mathbb{C}$	Complex numbers
$\mathbb{C}^n$	n-dimensional complex space
$x \cdot y$	Dot product on $\mathbb{R}^n$
$\Im\{\cdot\}$	Imaginary part of a complex number
$\langle x, y \rangle_X$	Inner product on the space X
$\cdot \cap \cdot$	Intersection of two sets
$\mathcal{L}^2(T, \mathbb{C}^n)$	L^2 space of equivalence classes of $\mathbb{C}^n$ valued functions defined on T
$\lvert \cdot \rvert$	Modulus of an element in $\mathbb{R}^n$ or $\mathbb{C}^n$
$\lVert \cdot \rVert_X$	Norm on the space X
$\mathrm{Prob}(\cdot)$	Probability of some event
$\mathbb{R}$	Real numbers
$\Re\{\cdot\}$	Real part of a complex number
$\mathbb{R}^n$	n-dimensional Euclidian space
$\cdot \cup \cdot$	Union of two sets
$x \times y$	Vector product on $\mathbb{R}^3$
$\mathbb{Z}$	Integers
$\mathbb{Z}_0^+$	Non-negative integers
$\mathbb{Z}^+$	Positive integers

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

The advent of modern cellular systems such as mobile handsets and indoor wireless local area networks has meant that telecommunication engineers are increasingly exposed to extremely harsh environments. The classical additive white Gaussian noise (AWGN) channel [1, pp. 378], with statistically independent Gaussian noise corrupting data samples is no longer adequate to model such channels. In most practical communication systems, a signal can travel from the transmitter to the receiver over multiple paths.

Due to multi-path propagation, electromagnetic signals are subject to random fluctuations in signal amplitude, phase and/or angle of arrival. This phenomenon is known as *fading* or *scintillation* [2]. Fading can be classified into *large-scale* and *short-scale* fading. Large-scale fading refers to the average signal attenuation over large areas and is affected by prominent terrain contours. The statistics of large-scale fading can be used to estimate the mean signal attenuation as a function of distance between the transmitter and receiver. Small-scale fading refers to the rapid fluctuations in signal amplitude and phase that can be experienced as a result of small changes (as small as half a wavelength) in the position of the transmitter and/or receiver. Small-scale fading causes time-spreading of the signals and time-variance of the channel. This results in intersymbol interference (ISI) and pulse distortion at the receiver[1].

Various techniques are used to mitigate the effects of multipath propagation. These techniques include, but are not limited to, using equalizers, diversity and channel coding[2]. Equalizers compensate for the ISI introduced by the multipath channel and are generally adaptive because of the time-varying nature of the channel. One, or a combination of frequency, time, spatial or antenna polarization diversity may be used to ensure that the depth and duration of fades experienced by the receiver is reduced. Fi-

[1] See [2] and [3, ch. 4,5].
[2] See eg [3, ch. 7].

nally, channel coding can be used to correct some or all of the errors introduced by the channel. Three general types of codes: block codes, convolution codes and turbo codes are commonly used. There is a great deal of published material on the techniques that can be used to mitigate the effects of channel fading. (See [4] for a good review of mitigation techniques[3]).

The idea that the inherent diversity in multipath environments could be exploited to improve the performance of a communication system began to emerge in the 1990s. Initial results of Winters [6] showed the potential benefits of using multiple input multiple output (MIMO) antenna systems in Rayleigh fading environments. Results of Teletar [7] and Foschini *et. al.* [8] proved the theoretic potential of MIMO systems. If a system has n transmitting and m receiving antennas then assuming that the transfer matrix $\mathbf{H}$ has independent identically distributed entries, taken from a Gaussian distribution, Foschini *et. al.* [8] proved that the capacity of the channel grows linearly with $\min\{m, n\}$. Several other results have appeared in the literature with varying assumptions about channel state information at the transmitter and receiver [9–12]. Measurements on real systems corroborate the theoretical findings [13,14][4].

The underlying assumption that leads to the unbounded growth in capacity with increasing number of antennas is: *signals received on individual antennas are mutually uncorrelated*. However, in any practical communication system, both the transmitting and receiving antennas are constrained to be within some finite volume in space. Therefore, as the number of antennas increase, antenna separation reduces and the channel capacity saturates at some finite level. Several research articles address the effects of antenna separation and mutual coupling of signals on the capacity of the channel [17–24]. However these results depend on specific antenna configurations and/or scattering environments.

This motivates the following question: *Is there a fundamental limit, independent of specific antenna configurations, to the number of mutually uncorrelated signals available at the receiver?*

One approach to answering the above question consists of studying continuous spatial channels, where we assume the *volumes, to which the antennas are constrained, to be the transmitting and receiving elements*. Suppose that the transmitting and receiving antennas are constrained to be within volumes T and R respectively. We assume that a current flows in the transmitting volume T that generates an electromagnetic field in the receiving volume R that can be measured. I call such channels *Spatial Waveform Channels*[5] (SWC). We then calculate the number of *linearly independent* electromagnetic fields that can be measured by a physical receiver that measures the field only in R given that the current in the transmitter is constrained in some way. The main idea

[3] Also see [1,3,5].

[4] There is an abundance of literature on MIMO results. See [15,16] for a review of MIMO results.

[5] See definition 2.1.

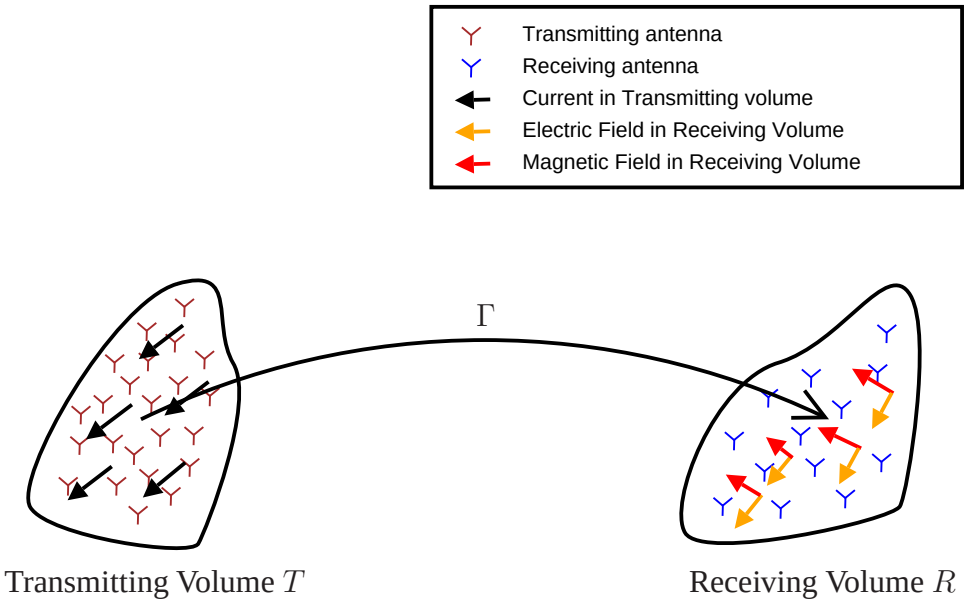

Figure 1.1: MIMO system and SWCs

being that the number of linearly independent electromagnetic fields in R gives an upper bound on the number of mutually uncorrelated signals at the receiver of a MIMO system with densely packed antennas.

The aim of the thesis is to study the properties of SWCs. In particular, I study how the constraints imposed on a real MIMO system can be modeled in SWCs and how these constraints effect the properties of SWCs.

1.2 Overview

The performance of any practical MIMO system is limited by the physical constraints imposed on it. Examples of this include total power available for transmission, noise at the receiver and total volume in space that the transmitting and receiving antennas can occupy. In this section, I will discuss how the constraints considered in this thesis influence our model and properties of SWCs.

Referring to figure 1.1, assume that the transmitting and receiving antennas are constrained to volumes T and R, respectively. Also assume that a current, of density $\mathbf{J}$ is flowing in the volume T that generates a field $[\mathbf{E}\ \mathbf{H}]$ in the volume R. Here $\mathbf{E}$ is the electric field and $\mathbf{H}$ is the magnetic field. We need to impose restrictions on the volumes T and R and also the current density $\mathbf{J}$ and field $[\mathbf{E}\ \mathbf{H}]$ in order to ensure that the physical limitations imposed on the MIMO system can be captured in the model for SWCs.

The constraints on the MIMO system that are studied in this thesis are as follows:

1. The transmitting and receiving antennas are physically constrained to be within a finite volume in space.

2. The total power/energy available for transmission is finite.

3. Physical quantities, such as source current densities and electromagnetic fields must be continuous functions.

4. The final physical limitation that needs to be incorporated into the model for SWCs is more subtle. This limitation is due to the inability of any practical receiver to measure the electric field in the receiving antennas to arbitrary accuracy. There are several reasons for this including [25]

   (a) Noise in the receiver.

   (b) Dynamic range of the receiver.

   (c) Resolution of the receiver.

   Therefore, if two electric fields differ by *too little* then a practical receiver in a MIMO system will not be able to differentiate between the signals and we call such signals physically indistinguishable. We need to model this subtle concept of two electric fields differing by *too little* to complete our description of SWCs.

I give an abstract mathematical definition for an SWC in section 2.1 that enables one to use the structure of the definition to impose the constraints discussed above. An SWC consists of three parts, a normed space $\widetilde{X}_T$ of possible source current densities, a normed space $\widetilde{Y}_R$ of possible electromagnetic fields and a linear operator $\Gamma : \widetilde{X}_T \to \widetilde{Y}_R$ that determines the electromagnetic field for a given source current density. Each one of these parts is used to capture different physical limitations imposed on the MIMO system. By restricting the elements of $\widetilde{X}_T$ and $\widetilde{Y}_R$ to satisfy certain properties, we ensure that constraints 1 and 3 are satisfied. The norm on the space $\widetilde{X}_T$ is used to ensure that constraint 2 is satisfied. The norm on $\widetilde{Y}_R$ is used to ascertain if two electromagnetic fields differ by too little as required by constraint 4. This is discussed further in chapter 2. Once we have an accurate description of an SWC, we can study its properties. In this thesis, I study three properties of SWCs: *1) Degrees of freedom at level-ϵ, 2) Essential Dimension, 3) Uncertainty Principles.*

In MIMO systems with n transmitting and m receiving antennas, there will be at most $\min\{m, n\}$ linearly independent signals that a receiver can measure. However, in the case of SWCs, there are infinitely many linearly independent signals in[6]

[6]Here, $B_{1,\widetilde{X}_T}(0)$ is the unit ball centered at the origin in $\widetilde{X}_T$. Physically we can think of functions in this set as those sources that use less than 1 unit of power or energy.

$\Gamma B_{1,\widetilde{X}_T}(0) \subset \widetilde{Y}_R$. But the finite receiver sensitivity discussed in constraint 4 above ensures that only finitely many linearly independent signals can be measured by a real receiver. The distance between any two electromagnetic fields is quantified using the norm on the receiver space of functions $\widetilde{Y}_R$ in the definition for SWCs. So, given some level ϵ, which depends on the receiver, there exists a number N and a set of N elements, $\phi_1, \ldots, \phi_N \in \widetilde{Y}_R$ such that all the signals that can be generated at the receiver differ by less than ϵ from an element in the span of $\{\phi_1, \ldots, \phi_N\}$. I call such an N the number of degrees of freedom[7] of the channel at level ϵ because a physical receiver that can measure fields to level ϵ can only measure N linearly independent signals. This in turn implies that as the number of antennas $n, m \to \infty$, the number of linearly independent signals for the MIMO system saturates at N.

Generally, in SWCs the number of degrees of freedom for a channel depends on the level ϵ. However, in several important cases [25–28] it does not change much for a large range of values of ϵ. This leads us to the concept of *essential dimension*[8] of an SWC which is the number of degrees of freedom at level ϵ for the largest range of values of ϵ. This number only depends on the channel and can be used as an upper bound for the number of mutually uncorrelated signals available at the receiver of a MIMO system with densely packed antennas, largely independent of the receiver sensitivity.

Finally, I study Uncertainty Principles for SWCs. The classical uncertainty principle constrains the amount of energy a function can simultaneously have in the frequency and time domains [29,30]. We can develop a similar principle for SWCs which tries to answer the question: *what is the maximal fraction of the total radiated energy/power that can be concentrated in the receiver?* We can therefore find a bound on the best connected mode in a MIMO system modeled by the corresponding SWC.

1.3 Summary of results

This thesis is divided into six chapters. Chapter 1 provides an overview of the subject matter studied. In the following section of this chapter, I review some of the previous results from the literature that are relevant to degrees of freedom in SWCs. In chapter 2 I discuss SWCs, explaining how the total radiated power/energy and received energy can be used to define norms on the space of transmitting and receiving functions in SWCs. In chapter 3 I show how one might define the very general concepts of degrees of freedom and essential dimension for compact operators on normed spaces and also develop sufficient machinery to compute the degrees of freedom and essential dimension for such operators. In chapter 4 I develop numerical techniques for the computation of generalised singular values of compact operators on normed spaces and show the results of numerical simulations. In chapter 5 I review the classical uncertainty principle and

[7] See definition 3.2.

[8] See definitions 3.4 and 3.5

discuss two novel principles that are more general and also give a physical interpretation in terms of SWCs for the principles. Concluding remarks are given in chapter 6.

The principal results contained in chapters 2 to 5 are summarised below.

Chapter 2: In this chapter I give a novel definition 2.1 for SWCs that is sufficiently general to encompass most[9] of the cases reviewed in section 1.4. I also prove that the total energy/power radiated and lost by a current in the transmitting volume Tdefines a norm on the vector space of all possible current densities in sections 2.3 and 2.4. Similarly, I prove that the energy stored in the electromagnetic field in the receiving volume R defines a norm on the space of all electromagnetic fields in the receiver. The definition of an SWC is such that all the constraints imposed on a MIMO system can be incorporated into it through the norms defined by the energy/power. I also prove several theorems in sections 2.3 and 2.4 which establish some useful properties of the space of transmitter and receiver functions.

Chapter 3: The concepts of degrees of freedom and essential dimensionality are very general and can be used for compact operators defined on arbitrary normed spaces. In this chapter I prove in theorem 3.1 that it makes sense to talk about degrees of freedom at level-ϵ for a compact operator defined on a normed space. I then use this theorem to give a novel definition for the number of degrees of freedom at level-ϵ for a compact operator (see definition 3.2). I prove some of the simple properties of degrees of freedom in theorem 3.2 which are useful in their own right and are also helpful in proving other theorems in later chapters.

The definition for degrees of freedom is a descriptive one and does not lend itself to numerical evaluation for specific channels. I therefore give a novel definition 3.3 for generalised singular values and prove in theorem 3.5 that they are generalisations of the commonly accepted notion of singular values of compact operators defined on Hilbert spaces. The advantage of defining generalised singular values is that the number of degrees of freedom of a compact operator can be characterised in terms of its singular values and this characterisation lends itself to numerical computation of degrees of freedom.

Finally, I give a novel definitions for essential dimension (3.4) and essential dimension of order-n (3.5) of a compact operator defined on a normed space. I argue that the essential dimension of an operator is different from degrees of freedom at level-ϵ in the sense that it is independent of the arbitrarily chosen constant ϵ and only depends on the operator itself. It gives a bound on the number of mutually uncorrelated signals at the receiver of a MIMO system that is largely independent of the noise at the receiver.

Chapter 4: In some cases, it is possible to write the operator in an SWC as an integral operator (see eg. [26]). Though, it may be difficult to analytically calculate the singular values of the integral operator, it is sometimes possible to calculate the singular values of another integral operator that closely approximates the original operator. I demon-

[9]The only exception is the work of Kennedy *et. al.* [31] which studies a different geometric model.

strate in this chapter that perturbation techniques can be used to show that the singular values of the approximate operator approach those of the original operator as the approximation gets better. I use these ideas to analytically calculate the singular values of the operator that describes communication using scalar waves between rectangular prisms [26]. The application of perturbation theory to the calculation of singular values of SWCs is novel.

For most situations however, analytical techniques are inadequate and I develop an important numerical technique for the computation of generalised singular values. This technique is very similar to Galerkin's method for the computation of singular values of Hilbert space operators (see eg. [32]). I prove in theorem 4.1 that if a compact operator is defined on a normed space that has a complete Schauder basis then one can use calculations on finite dimensional spaces to find approximations for the generalised singular values of the operator that are also lower bounds for the gereralised singular values. This theorem is useful because it gives a simple practical method of approximately computing the degrees of freedom. I also show results of numerical computations for several specific channels.

Chapter 5: In this chapter I review the classical uncertainty principle and use it to motivate theorem 5.7 which is a generalisation of the classical principle to arbitrary bounded operators defined on Hilbert spaces. I also give a physical interpretation of the general uncertainty principle for SWCs. I prove a second general uncertainty principle in theorem 5.9 that is particularly pertinent to SWCs.

1.4 Review of previous results from the literature

In this section I briefly review some of the previous results from the literature for degrees of freedom of spatial waveform channels[10]. One of the first papers that talks about the number of degrees of freedom of an electromagnetic field is Bucci *et. al.* [25,33]. Bucci *et. al.* [33] assume that the sources and scatterers are confined to a sphere of finite radius a and evaluate the number of degrees of freedom of an electric field generated by the source and induced current on some observation curve C. As explained in section 2.5, the channel studied by Bucci *et. al.* [25,33] can be written as an SWC. Bucci *et. al.* [25] show that the number of degrees of freedom at level-ϵ is equal to the number of singular values of the channel operator that are greater than ϵ. However, because the singular values of the channel operator show a step like behavior, for a large range of values of ϵ the number of degrees of freedom of the channel is independent of the exact value of ϵ chosen. As explained in section 3.4 my novel definition for essential dimension given in section 3.3 can be used to quantify the level-ϵ at which the singular values of Γ change most rapidly and can therefore be used to uniquely identify the step

[10] A review of previous results for uncertainty principles can be found in section 5.1.

in the singular values. Moreover, their definition of degrees of freedom is a special case of my definition of degrees of freedom given in section 3.2.

The work of Bucci *et. al.* [25] is similar to that of Miller [26] where scalar waves are studied. The main assumption in Miller [26] is that the transmitter is constrained to be in a volume $V \subset \mathbb{R}^3$ and the receiver is constrained to a volume $W \subset \mathbb{R}^3$. As shown in example 2.1, the channel studied by Miller can be written as an SWC. Miller [26] evaluates the number of modes of communication that are possible between the volumes V and W using scalar waves. Miller [26] explains how the number of modes of communication is equal to the number of significant singular values of the channel operator. Piestun and Miller [34] similarly analyse the case for vector waves and use a similar definition for the number of modes of communication. As explained in section 3.4, if we assume that a singular value of the channel operator is significant if it is greater than some constant ϵ, then the number of modes of communication is equal to the number of degrees of freedom at level-ϵ.

Miller [26] also analytically evaluates the singular values of the channel if the volumes V and W are rectangular prisms that are aligned along one of the Cartesian coordinate axis. Miller [26] uses the paraxial approximation of the channel's Green's function to calculate the singular values. I use perturbation theory to find bounds on the approximation error as explained in section 4.1. Miller shows that if V and W are rectangular prisms, then the singular values of the channel operator show a step like behavior and claims that the number of modes of communication is essentially independent of how one defines a singular value to be significant. As explained in section 3.4, the definition of essential dimension given in section 3.3 can be used to calculate the number of modes of communication if the singular values of the channel operator show a step like behavior.

Hanlen *et. al.* build on the results of Miller [26] for scalar waves to include the effect of scatterers [35,36]. They assume that reflective scatterers are present and evaluate a new Green's function to account for the scatterers. Then the singular values of the channel operator are calculated numerically. Simulations are used to calculate the singular values for different scatterer positions. Xu *et. al.* [37] build on the results of Piestun *et. al.* [34] for vector waves to include the effect of scatterers. They numerically evaluate the Green's function for vector waves in the presence of reflective scatterers in two dimensions using the finite moments method (FMM). Xu *et. al.* [37] then numerically calculate the singular values of the operator defined using the Green's function. As explained in section 2.5, the channels studied by Xu *et. al.* [37] and Hanlen *et. al.* [36] can be written as SWCs. Xu *et. al.* [37] give a novel definition for the number of degrees of freedom for the channel in terms of the singular value of the channel operator, the power available for transmission and noise in the receiver. As explained in section 3.4 their definition is equivalent to a special case of my definition of degrees of freedom at level-ϵ. Simulation results in both [36] and [37] indicate that the singular values of the

channel operator decrease gradually with no sudden knee like behavior.

This is quite unlike the behavior predicted by Poon *et. al.* [28] who consider communication between finite volume transmitters and receivers in the presence of scatterers. But, instead of considering individual scatterers, they consider a cluster of scatterers that are constrained to be within some finite volume. Poon *et. al.* [28] use a heuristic argument to show that the singular values σ_n of the channel operator show a step like behavior with eigenvalues close to 1 for $n < N_{dof}$ and close to 0 otherwise. Here N_{dof} depends on the channel operator and the finite transmitting and receiving volumes. As explained in section 3.4, my definition of essential dimension can be used to uniquely identify N_{dof}. It should be noted that this step like behavior is observed in most analytical calculations of singular values where some assumptions are made regarding the scatterer and/or antenna configurations (e.g. cluster of scatterers within a finite volume). In contrast, if numerical simulations are used to calculate the singular values of channels, wherein the scatterers are placed at random locations, the singular values do not generally show a step-like behavior[11]. My numerical simulations show no step-like behavior for singular values (Note that the physical model I simulated is considerably different from that of Poon *et. al.* [28]. Therefore one cannot directly compare these results.)

A step like behavior of eigenvalues is also found in [31,38,39]. The approach of Kennedy *et. al.* [31] is substantially different from those discussed earlier in this section wherein the transmitting volume was assumed to be constrained. Kennedy *et. al.* [31] consider the possible wavefields within a volume constrained to a radius R provided all sources are in the far field (outside a ball of radius R_S). The channel model used in these papers is considerably different from that considered in this thesis. The physical interpretation of these channels is discussed further in section 3.4. Kennedy *et. al.* [31] use the error in approximating the field in the receiving volume by a finite dimensional set of functions to evaluate the degrees of freedom of the wavefields in the receiving volume. A similar approach can also be found in Dickins *et. al.* [40]. However, as explained in section 3.4, these definitions cannot be directly compared with degrees of freedom at level-ϵ or with essential dimension of the channel.

Newsam and Barakat [41] approach the problem of finding the degrees of freedom of an operator differently. The primary interest of [41] is in the inverse problem: Suppose $g(t) = \int k(s-t)f(t)dt$. Here $g(t), f(t) \in \mathcal{L}^2(\mathbb{R}, \mathbb{C})$. Then given g can we determine f? As explained in section 3.4 their definition is equivalent to a special case of the definition of degrees of freedom at level-ϵ.

All the channels discussed so far in this section constrain the total current in the transmitting volume in some sense. Jensen and Wallace [42,43] also study communication between finite transmitting and receiving volumes using electromagnetic waves. They assume that the source current is constrained to be within a finite volume T and

[11] c.f. simulation results in [34,36,37]

impose a restriction on the total power available for transmission. Wallace *et. al.* [43] claim that if only the total power radiated is bounded then there are infinitely many channels available between the transmitting and receiving volumes. They further claim that if the super- directivity ratio is also bounded for a source current then only finitely many channels transfer sufficient power from the transmitter to the receiver. No proofs are provided for these claims in [42,43]. Further, unlike the SWCs described in chapter 2, Wallace *et. al.* [43] do not constrain the total power lost as heat in the transmitting volume.

Chapter 2

Spatial Waveform Channels

As explained in the previous chapter, SWCs may be used to find bounds on the performance of MIMO systems with an arbitrarily large number of transmitting and/or receiving antennas that are constrained to a finite volume. A MIMO system's performance is limited by several constraints imposed on it. So, in order to obtain useful bounds on MIMO systems one needs to model the constraints imposed on it in the corresponding SWC. The approach taken in this thesis to achieve the above is to give a definition for SWCs that enables one to impose the required constraints by using the structure of the definition. This definition, given in the following section is novel and is sufficiently general to encompass most[1] of the cases discussed in section 1.4.

There are two distinct cases considered here. Firstly, we assume that the underlying MIMO system, which is constrained to be within some finite volume, can only radiate for a finite duration of time. In this case we also assume that a finite amount of energy is available for transmission. I refer to this case as the *finite energy* case. Any real MIMO system is subject to the above restriction and this is the most important case considered here. In the second case I assume that the MIMO system can radiate energy for all time $t \in (-\infty, \infty)$. In this case I assume that the total power available for transmission is finite. I refer to this case as the *finite power* case. In both the finite energy and finite power cases I assume a monochromatic time dependence of $e^{j\omega t}$. Here, ω is the angular frequency of oscillation.

Finite energy case: The important constraints considered in this thesis and the approach taken to model them in SWCs is as follows:

1. **Constraint:** The transmitting and receiving antennas are physically constrained to be within a finite volume in space.

[1] The only exception being that studied by Kennedy *et. al.* [31] where the physical situation considered is much different.

Approach taken to model the constraint: This constraint can easily be incorporated into the SWC model by requiring that the transmitting volume T and receiving volume R be compact subsets of $\mathbb{R}^3$. I then impose the restriction that the source current density must be zero outside T and the electromagnetic field can only be measured within the volume R.

2. **Constraint:** The total energy available for transmission is finite.
   Approach taken to model the constraint: Suppose X_T is the space of Lebesgue measurable $\mathbb{C}^3$-valued functions defined on T that are square integrable. It turns out[2] that the total energy required to set up a current density $\mathbf{J} \in X_T$ in T defines a semi-norm $\rho_{X_T}(\cdot)$ on X_T. Let $\widetilde{X}_T$ be the set of equivalence classes of functions in X_T such that if $\widetilde{x} \in \widetilde{X}_T$ and if $x_1, x_2 \in \widetilde{x}$ then $\rho_{X_T}(x_1 - x_2) = 0$. The semi-norm ρ_{X_T} induces a norm $\|\cdot\|_{\widetilde{X}_T}$ on the space $\widetilde{X}_T$. I assume that the source current density is modeled by an equivalence class of functions which have norm less than the energy available for transmission.

   Now, suppose that Y_R is the space of Lebesgue measurable $\mathbb{C}^6$-valued functions defined on R that are square integrable. The total energy stored in the electromagnetic field within volume R defines a semi-norm $\rho_{Y_R}(\cdot)$ on the space Y_R. Let $\widetilde{Y}_R$ be the space of equivalence classes of functions similar to the source current density case above and let $\|\cdot\|_{\widetilde{Y}_R}$ be the norm induced by the semi-norm ρ_{Y_R} on this space. I assume that the electromagnetic field is modeled by an element of the space $\widetilde{Y}_R$ with a finite norm.

3. **Constraint:** Physical quantities, such as source current densities and electromagnetic fields must be continuous functions.
   Approach taken to model this constraint: An indirect approach is taken to model this constraint. Obviously, the current density on the antenna cannot be an equivalence class of square integrable functions. However, as shown in section 2.3, for all $\widetilde{x} \in \widetilde{X}_T$ and $\epsilon > 0$, there exists an $\widetilde{x}_\epsilon \in \widetilde{X}_T$ such that $\|\widetilde{x} - \widetilde{x}_\epsilon\|_{\widetilde{X}_T} \leq \epsilon$ and at least one member of the equivalence class $\widetilde{x}_\epsilon$ is a continuous function. Physically, we can interpret this statement as follows: for any given $\widetilde{\mathbf{J}} \in \widetilde{X}_T$, there is a current density $\mathbf{J}$ that is continuous and the electromagnetic field generated by $\mathbf{J}$ is approximately equal to the field generated by $\widetilde{\mathbf{J}}$. We can similarly interpret the elements of $\widetilde{Y}_R$ as electromagnetic fields.

4. **Constraint:** Receiver sensitivity cannot be arbitrarily high.
   Approach taken to model this constraint: This constraint is not directly modeled in the definition of SWCs. However, the definition gives enough structure to an SWC, so that this structure can be used in studying the properties of SWCs that

[2] See section 2.3.

depend on the sensitivity of the receiver. At any instant in time, the energy radiated by the transmitter is stored in the electromagnetic field. For the receiver to be able to measure the field, the energy stored in the field in volume R cannot be too small. As explained earlier, the energy stored in the received field induces a norm on the space $\widetilde{Y}_R$. So, for a receiver with finite sensitivity to be able to measure the electromagnetic field we assume that only those fields with norm greater than some small positive constant ϵ are measurable. A simple way of incorporating this idea is to say that two signals $[\mathbf{E}_1\ \mathbf{H}_1], [\mathbf{E}_2\ \mathbf{H}_2] \in \widetilde{Y}_R$ are physically indistinguishable at some level $\epsilon > 0$ if the energy of the field $[\mathbf{E}\ \mathbf{H}] = [\mathbf{E}_2\ \mathbf{H}_2] - [\mathbf{E}_1\ \mathbf{H}_1]$ is less than ϵ. Or equivalently, if $\|[\mathbf{E}\ \mathbf{H}]\|_{\widetilde{Y}_R} < \epsilon$. This concept is not novel and has been used for instance in Bucci *et. al.* [25] and Landau *et. al.* [27].

Finite power case: The finite power case is very similar to the finite energy case. Now the space $\widetilde{X}_T$ consists of equivalence classes of $\mathbb{C}^3$-valued functions defined on some compact set T. Again, it can be shown that the total power lost and radiated defines a norm on the space $\widetilde{X}_T$. One significant difference is that we can now look at the power received by a volume $R \subset \mathbb{R}^3$ or the energy stored in the electromagnetic field in the volume $R \subset \mathbb{R}^3$. If we assume that the receiving volume consists of empty space then the total power received would be zero. If on the other hand we assume that it consists of a conductor with finite resistance then it would be difficult to determine the total power received by it. So I use the total energy stored in the receiving volume to determine the number of degrees of freedom of the channel.

The above discussion motivates the definition of an SWC given in the following section. In sections 2.3 and 2.4, I show how the energy/power can be used to define norms on the space of current densities and electromagnetic fields.

2.1 Definition of Spatial Waveform Channels

At the beginning of this chapter, I explained that one could use the energy/power to define a norm on the space of source current densities. However, in defining an SWC channel, I allow any physical constraint to be used to define a norm on the space of source current densities. This ensures that an SWC is sufficiently general to accommodate several models used in the literature.

Let $T \subset \mathbb{R}^n$ be compact and Lebesgue measurable and let $\mathcal{F}_T$ be some linear vector space of $\mathbb{C}^m$ valued functions defined on T. Here, $m \in \mathbb{Z}^+$. For instance, $\mathcal{F}_T$ can be the space of all Lebesgue measurable, square integrable functions defined on T. Now suppose that $\rho_{\mathcal{F}_T} : \mathcal{F}_T \to [0, \infty]$ is some semi-norm defined on the space $\mathcal{F}_T$. Then we can define an equivalence relation $\sim$ on $\mathcal{F}_T$ as $f \sim g$ if $f, g \in \mathcal{F}_T$ and $\rho_{\mathcal{F}_T}(f - g) = 0$. It follows from the definition of a semi-norm that this is in fact an equivalence relation. Let $\widetilde{\mathcal{F}}_T$ be the set of all equivalence classes in $\mathcal{F}_T$. Then, $\rho_{\mathcal{F}_T}$ induces a norm on the

elements of $\widetilde{\mathcal{F}}_T$ as follows

$$\|\widetilde{f}\|_{\widetilde{\mathcal{F}}_T} = \rho_{\mathcal{F}_T}(f).$$

Here $\widetilde{f} \in \widetilde{\mathcal{F}_T}$ is an equivalence class of functions in $\mathcal{F}_T$ and $f \in \mathcal{F}_T$ is some member of the equivalence class $\widetilde{f}$. From the definition of the equivalence classes it is apparent that the above norm is well defined. Note that if the semi-norm $\rho_{\mathcal{F}_T}$ is in fact a norm on the space $\mathcal{F}_T$ then $\mathcal{F}_T$ is isomorphic to $\widetilde{\mathcal{F}}_T$.

Definition 2.1. *A **spatial waveform channel (SWC)** consists of a triple $(\widetilde{\mathcal{F}}_T, \widetilde{\mathcal{F}}_R, \Gamma)$. Here, $T \subset \mathbb{R}^n$ and $R \subset \mathbb{R}^n$ are compact and Lebesgue measurable*[3]*. Also, $\widetilde{\mathcal{F}}_T$ is a normed space of equivalence classes of $\mathbb{C}^m$-valued functions defined on T with norm $\|\cdot\|_{\widetilde{\mathcal{F}}_T}$ and similarly $\widetilde{\mathcal{F}}_R$ is a normed space of equivalence classes of $\mathbb{C}^p$-valued functions defined on R with norm $\|\cdot\|_{\widetilde{\mathcal{F}}_R}$. Here m, n and p are positive integers. Further, $\Gamma : \widetilde{\mathcal{F}}_T \rightarrow \widetilde{\mathcal{F}}_R$ is a bounded linear operator.*

In the following example I will show how the special case of communication using scalar waves between finite volumes can be written as an SWC. This channel was studied by Miller [26]. This example also demonstrates how a general channel, that does not necessarily constrain the total power/energy used, can be written as an SWC.

Example 2.1. *Consider communication using scalar waves between a transmitting volume $T \subset \mathbb{R}^3$ and a receiving volume $R \subset \mathbb{R}^3$ which are compact and measurable such that $T \cap R = \emptyset$. Suppose that $\mathcal{F}_T$ and $\mathcal{F}_R$ are the spaces of all complex valued, Lebesgue measurable, square integrable functions defined on T and R, respectively. Then we can define semi-norms on these spaces as*

$$\begin{aligned} \rho_{\mathcal{F}_T}(\cdot) &= \left[\int_T |\cdot|^2 d\mathbf{r}\right]^{1/2} \\ \rho_{\mathcal{F}_R}(\cdot) &= \left[\int_R |\cdot|^2 d\mathbf{r}\right]^{1/2} \end{aligned}$$

Therefore, we have $\widetilde{\mathcal{F}}_T = \mathcal{L}^2(T, \mathbb{C})$ and $\widetilde{\mathcal{F}}_R = \mathcal{L}^2(R, \mathbb{C})$.

Now suppose that there are sources $\psi \in \mathcal{L}^2(T, \mathbb{C})$ in the transmitting volume T that generate waves $\phi(\mathbf{r})$ according to the scalar Helmholtz equation

$$\nabla^2\phi(\mathbf{r}) + k^2\phi(\mathbf{r}) = -\psi(\mathbf{r}) \; \forall \mathbf{r} \in \mathbb{R}^3.$$

Here k is the wave number. Then it can be shown (see eg. [26]) that for all $\mathbf{r} \in R$

$$\phi(\mathbf{r}) = \int_T G(\mathbf{r}, \mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}'.$$

[3]We can think of T and R as subsets of $\mathbb{R}^3$ to which the transmitting and receiving antennas are constrained.

Here,

$$G(\mathbf{r},\mathbf{r}') = \frac{\exp\{-ik|\mathbf{r}-\mathbf{r}'|\}}{4\pi|\mathbf{r}-\mathbf{r}'|}.$$

We can now define an operator $\Gamma : \mathcal{L}^2(T,\mathbb{C}) \to \mathcal{L}^2(R,\mathbb{C})$ *as* $\psi \mapsto \phi$*. It is obvious that* Γ *is a linear map. To prove that it is bounded, note that because* T *and* R *are compact and have empty intersection, there exists an* $r > 0$ *such that for all* $\mathbf{r} \in R$ *and* $\mathbf{r}' \in T$*,* $r < |\mathbf{r}-\mathbf{r}'|$*. Therefore* $|G(\mathbf{r},\mathbf{r}')|$ *is bounded for* $\mathbf{r} \in R, \mathbf{r}' \in T$*. Suppose* b *is some upper bound for* $|G(\mathbf{r},\mathbf{r}')|$ *for all* $\mathbf{r} \in T$ *and* $\mathbf{r}' \in R$*. Then*

$$\begin{aligned}
\|\phi\|^2_{\mathcal{L}^2(R,\mathbb{C})} &= \int_R |\phi(\mathbf{r})|^2 d\mathbf{r} \\
&\leq b^2 \int_R \int_T |\psi(\mathbf{r}')|^2 d\mathbf{r}' d^3\mathbf{r} \\
&\leq b^2 \mu(R) \|\psi\|^2_{\mathcal{L}^2(T,\mathbb{C})}
\end{aligned}$$

Here, $\mu(R)$*, the volume of* R *is finite because* R *is compact. This proves that* Γ *is bounded. Therefore* $(\mathcal{L}^2(T,\mathbb{C}), \mathcal{L}^2(R,\mathbb{C}), \Gamma)$ *is an SWC. It describes communication through free space between a finite volume transmitter and a finite volume receiver using scalar waves.*

In the following section I will describe how one can calculate the total energy/power radiated by a given source current density. This radiated energy/power is then used to define semi-norms on the spaces of transmitter and receiver functions.

2.2 Energy and Power for Electromagnetic Waves

Suppose T is a compact subset of $\mathbb{R}^3$ and let $\mathbf{J}(\mathbf{x},t)$, with $\mathbf{x} \in T$ and $t \in (-\infty,\infty)$, be the source current density that generates an electric field $\mathbf{E}(\mathbf{x},t)$ and a magnetic field $\mathbf{H}(\mathbf{x},t)$ defined for all $\mathbf{x} \in \mathbb{R}^3$ and $t \in (-\infty,\infty)$.

Let us now consider the finite energy and finite power cases separately. In the finite power case we assume a monochromatic time-dependence. That is, we assume that the source current density is of the form

$$\mathbf{J}(\mathbf{r},t) = \Re\{\mathbf{J}'(\mathbf{r})e^{j\omega t}\} \; \forall t \in (-\infty,\infty).$$

Here $\Re\{\cdot\}$ denotes the real part of a complex number and $\mathbf{J}' : T \to \mathbb{C}^3$. Because Maxwell's equations are linear and time-invariant, the electric and magnetic fields are of the form

$$\begin{aligned}
\mathbf{E}(\mathbf{r},t) &= \Re\{\mathbf{E}'(\mathbf{r})e^{j\omega t}\} \; \forall t \in (-\infty,\infty) \text{ and} \\
\mathbf{H}(\mathbf{r},t) &= \Re\{\mathbf{H}'(\mathbf{r})e^{j\omega t}\} \; \forall t \in (-\infty,\infty),
\end{aligned}$$

with $\mathbf{E}', \mathbf{H}' : \mathbb{R}^3 \to \mathbb{C}^3$.

Now suppose that Ω is some smooth surface the interior of which contains T. Then the time-averaged, total power radiated by the current of density $\Re\{\mathbf{J}'(\mathbf{r})e^{j\omega t}\}$ can be calculated using Poynting's theorem [44, section 6.9]:

$$P_{rad}(\mathbf{J}') = \frac{1}{2}\int_{\Omega} \Re\{(\mathbf{E}'(\mathbf{r}) \times \mathbf{H}'^{*}(\mathbf{r})) \cdot \mathbf{n}\} da. \tag{2.1}$$

Here, $\cdot \times \cdot$ denotes the vector-product in $\mathbb{R}^3$ and da is the surface area element. Note that the energy stored in the near-field of the transmitting volume does not contribute to the above integral. Therefore, this stored energy plays no part in calculating the norm on the space of transmitter functions. Also, I define the time-averaged power lost (as heat) in the transmitting volume to be

$$P_{lost}(\mathbf{J}') = \frac{1}{2} R_{loss} \int_{T} |\mathbf{J}'(\mathbf{r})|^2 d\mathbf{r}. \tag{2.2}$$

This definition is motivated by the analogous definition for power lost in radiating antenna. So the total power, lost as heat is proportional to the square of the current and a loss resistance R_{loss}. The loss resistance is a constant that depends on the material used to construct the antenna and the antenna configuration. We study the behavior of an SWC for varying loss resistances.

Let $R \subset \mathbb{R}^3$ be some compact three dimensional receiving volume. Then the time-averaged power entering the volume R is given by

$$P_{rec}(\mathbf{J}') = -\frac{1}{2}\int_{\partial R} \Re\{(\mathbf{E}'(\mathbf{r}) \times \mathbf{H}'^{*}(\mathbf{r})) \cdot \mathbf{n}\} da. \tag{2.3}$$

Here, ∂R is the boundary of R, $\mathbf{n}$ is the unit normal vector to the surface pointing outwards and da is the surface area element. We get the negative sign because we use the convention that $\mathbf{n}$ is pointing out from the surface. If we assume that $R \cap T = \emptyset$, then the time-averaged power received by the receiving volume is in fact zero because we assume a monochromatic time-dependance. However, in a real antenna system, due to the presence of conductors that have finite conductivity in the receiving volume, some of the incident electromagnetic power is absorbed by the receiver. So, we can calculate the time-averaged incident power on the receiving volume R as follows:

$$P_{inc}(\mathbf{J}') = -\frac{1}{2}\int_{\partial R} \min\{\Re\{(\mathbf{E}'(\mathbf{r}) \times \mathbf{H}'^{*}(\mathbf{r})) \cdot \mathbf{n}\}, 0\} da. \tag{2.4}$$

We assume that the power received by the receiving volume is equal to $R_{rad}P_{inc}(\mathbf{J}')$. Here, R_{rad} is the radiation resistance of the receiving antenna. This approach is however not ideal because the radiation resistance of an antenna system depends on the antenna configuration. I therefore do not take this approach in this thesis.

Instead, we look at the energy stored in the electromagnetic field within the receiving volume R. At any instant in time the total power radiated by the current density $\mathbf{J}(\mathbf{x}, t)$ is stored in the electromagnetic field throughout $\mathbb{R}^3$. The time-averaged energy stored in the field within the volume R is [44, section 6.7]

$$E_{rec}(\mathbf{J}') = \frac{1}{4}\int_R \epsilon_0 |\mathbf{E}'(\mathbf{r})|^2 + \frac{1}{\mu_0}|\mathbf{H}'(\mathbf{r})|^2 d\mathbf{r}. \tag{2.5}$$

Here, ϵ_0 and μ_0 are the permittivity and permeability of free space. The motivation behind using this energy is that it simplifies calculations considerably. We do not specifically calculate the total energy re-radiated from the receiver. We assume that the total energy lost as heat and also lost due to re-radiation is proportional to the incident energy. We absorb this constant of proportionality into R_{loss} the loss resistance of the transmitting antenna and therefore do not need to state it explicitly.

Now consider the finite energy case wherein we assume that the transmitting volume has current flowing in it only during a finite time interval $[0, t_1]$. I again assume a monochromatic time-dependance for the source current density, i.e.

$$\mathbf{J}(\mathbf{r}, t) = \Re\{\mathbf{J}'(\mathbf{r})e^{j\omega t}\xi_{t_1}(t)\}.$$

Here,

$$\xi_{t_1}(t) = \begin{cases} 1 & 0 \leq t \leq t_1 \\ 0 & \text{otherwise} \end{cases}$$

Note that because of the indicator function ξ_{t_1}, we cannot assume that the electric and magnetic fields are monochromatic.

In this case we need to calculate the total energy radiated by the volume T during the time interval $[0, t_1]$. One needs to be careful in calculating this total radiated energy, because the electromagnetic field is not well defined in the transmitting volume T. I assume that the total radiated energy is the energy that is stored in the field outside a closed surface Ω the interior of which contains T. By making the surface Ω arbitrarily close to the volume T, we can ensure that the total radiated energy is very close to the actual radiated energy. This is not a problem when modeling a real antenna system because it would have finite conductivity and therefore the energy stored in the field in the small volume $\Omega_{int} \setminus T$ will be very small and goes to zero as Ω gets closer to the volume T. Here, Ω_{int} is the interior of the closed surface Ω.

In this case we can write the total energy radiated by the transmitting volume as

$$[E_{rad}(\mathbf{J}')](t_1) = \int_{\Omega_{ext}} \epsilon_0 |\mathbf{E}(\mathbf{r}, t_1)|^2 + \frac{1}{\mu_0}|\mathbf{H}(\mathbf{r}, t_1)|^2 d\mathbf{r} \tag{2.6}$$

where Ω_{ext} is the exterior of the surface Ω. Again, I define the total energy lost (as heat) to be

$$[E_{lost}(\mathbf{J}')](t_1) = R_{loss}\int_0^{t_1}\int_T |\mathbf{J}(\mathbf{r}, t)|^2 d\mathbf{r}dt. \tag{2.7}$$

Finally, if $R \subset \mathbb{R}^3$ is some compact receiving volume, then the total energy stored in this volume at time instant t_1 is given by

$$[E_{rec}(\mathbf{J}')](t_1) = \int_R \epsilon_0 |\mathbf{E}(\mathbf{r}, t_1)|^2 + \frac{1}{\mu_0} |\mathbf{H}(\mathbf{r}, t_1)|^2 d\mathbf{r} \tag{2.8}$$

Note that for the remainder of this document I will refer to both $\mathbf{J}(\mathbf{r}, t)$ and $\mathbf{J}'(\mathbf{r})$ as the source current density in both the finite energy and finite power cases. The reader should interpret the current density $\mathbf{J}'(\mathbf{r})$ to be $\Re\{\mathbf{J}'(\mathbf{r}) e^{j\omega t} \xi_{t_1}(t)\}$ in the finite energy case and $\Re\{\mathbf{J}'(\mathbf{r}) e^{j\omega t}\}$ in the finite power case.

2.3 Finite Energy Case

In this section I will describe how a MIMO system with antennas that are constrained to finite volumes and radiate for a finite time duration can be modeled as an SWC. I assume that the transmitting antennas have a finite amount of energy available for transmission and that the energy is radiated in the form of electromagnetic fields and lost as heat in the transmitting volume.

Suppose $T \subset \mathbb{R}^3$ is a compact set and $0 < t_0 < \infty$ is some instant in time and the antennas are constrained to volume T and can only radiate energy during the interval $[0, t_0]$. Let $\mathcal{L}^2(T, \mathbb{C}^3)$ be the space of equivalence classes of $\mathbb{C}^3$-valued functions defined on T with finite L^2 norm. In this section, I will define a different norm on this space of functions that depends on the total energy lost and radiated.

Also suppose $\mathbf{J} \in \mathcal{L}^2(T, \mathbb{C}^3)$ is some function, such that $\Re\{\mathbf{J}(\mathbf{r}) e^{j\omega t} \xi_{t_0}(t)\}$, $\mathbf{r} \in T$, $t \in (-\infty, \infty)$ is the current density. Then we can calculate the total energy radiated $[E_{rad}(\mathbf{J})](t_0)$ from equation (2.6) and the total energy lost as heat, $[E_{lost}(\mathbf{J})](t_0)$ from equation (2.7).

Theorem 2.1. *Let $T \subset \mathbb{R}^3$ be some compact set and let $t_0 < \infty$ be some instant in time and let Ω be a closed surface the interior of which contains T. Then we can define an inner product on $\mathcal{L}^2(T, \mathbb{C}^3)$ as follows. For all $\mathbf{J}_1, \mathbf{J}_2 \in \mathcal{L}^2(T, \mathbb{C}^3)$,*

$$\begin{aligned} \langle \mathbf{J}_1, \mathbf{J}_2 \rangle_{\widetilde{X}_{T,t_0}} &= I_1 + I_2 && \textit{where} \\ I_1 &= R_{loss} \int_0^{t_0} \int_T \mathbf{J}_1'^*(\mathbf{r}, t) \mathbf{J}_2'(\mathbf{r}, t) d\mathbf{r} dt && \textit{and} \\ I_2 &= \epsilon_0 \int_{\Omega_{ext}} \mathbf{E}_1^*(\mathbf{r}, t_0) \mathbf{E}_2(\mathbf{r}, t_0) d\mathbf{r} + \tfrac{1}{\mu_0} \int_{\Omega_{ext}} \mathbf{H}_1^*(\mathbf{r}, t_0) \mathbf{H}_2(\mathbf{r}, t_0) d\mathbf{r}. \end{aligned} \tag{2.9}$$

Here, for $i = 1, 2$, $\mathbf{J}_i'(\mathbf{r}, t)$ generates the electric field $\mathbf{E}_i(\mathbf{r}, t)$ and magnetic field $\mathbf{H}_i(\mathbf{r}, t)$ and $\mathbf{J}_i'(\mathbf{r}, t) = \mathbf{J}_i(\mathbf{r}) e^{j\omega t} \xi_{t_0}(t)$.

Proof. Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{L}^2(T, \mathbb{C}^3), \alpha \in \mathbb{C}$ and let $\mathbf{E}_x, \mathbf{E}_y, \mathbf{E}_z$ and $\mathbf{H}_x, \mathbf{H}_y, \mathbf{H}_z$ be the corresponding electric and magnetic fields generated by $\mathbf{x}, \mathbf{y}, \mathbf{z}$. From the linearity of

Maxwell's equations, we know that the electric and magnetic fields generated by $\mathbf{x}+\mathbf{y}$ are $\mathbf{E}_x+\mathbf{E}_y$ and $\mathbf{H}_x+\mathbf{H}_y$. Therefore, we get the following after some simple algebraic manipulations

$$\langle \mathbf{x}+\mathbf{y},\mathbf{z}\rangle_{\widetilde{X}_{T,t_0}} = \langle \mathbf{x},\mathbf{z}\rangle_{\widetilde{X}_{T,t_0}} + \langle \mathbf{y},\mathbf{z}\rangle_{\widetilde{X}_{T,t_0}},$$
$$\langle \alpha\mathbf{x},\mathbf{z}\rangle = \alpha^*\langle \mathbf{x},\mathbf{z}\rangle_{\widetilde{X}_{T,t_0}}.$$

It is easy to show that $\langle \mathbf{x},\mathbf{y}\rangle_{\widetilde{X}_{T,t_0}}$ is the complex conjugate of $\langle \mathbf{y},\mathbf{x}\rangle_{\widetilde{X}_{T,t_0}}$. Finally, note that if $\mathbf{J}_1=\mathbf{J}_2$ then I_2 is always non-negative and I_1 is proportional to the L^2 norm. Therefore, $\langle \mathbf{J}_1,\mathbf{J}_1\rangle_{\widetilde{X}_{T,t_0}}$ is always non-negative and if it is zero, then $\mathbf{J}_1$ must be the equivalence class of functions that are zero almost everywhere.

Consequently, $\langle\cdot,\cdot\rangle_{X_{T,t_0}}$ is an inner product on $\mathcal{L}^2(T,\mathbb{C}^3)$. □

Let $\|\cdot\|_{\widetilde{X}_{T,t_0}}$ denote the norm induced by this inner product and let $\widetilde{X}_{T,t_0}$ denote the subset of $\mathcal{L}^2(T,\mathbb{C}^3)$ functions with finite $\|\cdot\|_{\widetilde{X}_{T,t_0}}$ norm[4]. From the definition of the inner product in equation (2.9), we can tell that the norm on $\widetilde{X}_{T,t_0}$ is equal to the total energy lost and radiated by the source current of density $\mathbf{J}$.

Similarly, if $R\subset\mathbb{R}^3$ is compact, then we can define an inner product on the space $\mathcal{L}^2(R,\mathbb{C}^6)$, that gives the total energy stored in the field in volume R. The inner product is

$$\langle[\mathbf{E}_1\ \mathbf{H}_1],[\mathbf{E}_2\ \mathbf{H}_2]\rangle_{\widetilde{Y}_R} = \int_R \epsilon_0\mathbf{E}_1^*(\mathbf{r})\mathbf{E}_2(\mathbf{r}) + \frac{1}{\mu_0}\mathbf{H}_1^*(\mathbf{r})\mathbf{H}_2(\mathbf{r})d\mathbf{r}. \tag{2.10}$$

It is trivial to show that $\langle\cdot,\cdot\rangle_{\widetilde{Y}_R}$ is an inner product. Let $\|\cdot\|_{\widetilde{Y}_R}$ be the norm induced by this inner product and let $\widetilde{Y}_R$ denote the set of functions whose $\|\cdot\|_{\widetilde{Y}_R}$ norm is finite. The following theorem shows that the case of communication using electromagnetic waves, with an energy constraint on the transmitter can be written as an SWC.

Theorem 2.2. *Let, $\widetilde{X}_{T,t_0}$ and $\widetilde{Y}_R$ be as described above. Then, the operator $\Gamma:\widetilde{X}_{T,t_0}\to\widetilde{Y}_R$ which maps any current density in $\widetilde{X}_{T,t_0}$ to the corresponding electric and magnetic fields in $\widetilde{Y}_R$ at time t_0 is linear and bounded.*

Proof. Because each current density in the space $\mathbf{J}\in\widetilde{X}_{T,t_0}$ generates a unique field in R, there exists an operator Γ that maps $\mathbf{J}$ to some field $[\mathbf{E}\ \mathbf{H}]$. The linearity of the operator Γ follows from the linearity of Maxwell's equations. To show that Γ is bounded, note that if the total energy radiated is bounded (i.e. if $\|\mathbf{J}\|_{\widetilde{X}_{T,t_0}}<b<\infty$) then the energy stored within volume R is less than the same bound (i.e. $\|[\mathbf{E}\ \mathbf{H}]\|_{\widetilde{Y}_R}<b$). Therefore Γ is bounded and $\|\Gamma\|\le 1$. □

[4]It will be shown in theorem 2.3 that all elements of $\mathcal{L}^2(T,\mathbb{C}^3)$ are also elements of $\widetilde{X}_{T,t_0}$.

Therefore the triple $(\widetilde{X}_{T,t_0}, \widetilde{Y}_R, \Gamma)$ is a spatial waveform channel. The electromagnetic field for a given current density depends on the position of the scatterers. Therefore, the operator Γ, in the above theorem depends on the position of the scatterers. Note that it is not important to have a closed form description of the operator Γ. One only needs to calculate the electromagnetic field for a given source current density in order to study the properties of the SWC. This calculation is done numerically in section 4.3.

Finally note that if $R \cap T = \emptyset$ then the operator Γ is compact. This is a direct consequence of the following theorem.

Theorem 2.3. *Let T be a compact set and let $\widetilde{X}_{T,t_0}$ be the normed space described above. If $\mathbf{x} \in \mathcal{L}^2(T, \mathbb{C}^3)$ then $\mathbf{x} \in \widetilde{X}_{T,t_0}$. Moreover, if $\mathbf{x} \in \mathcal{L}^2(T, \mathbb{C}^3)$ and $\{\mathbf{x}_1, \mathbf{x}_2, \ldots\}$ is a sequence in $\mathcal{L}^2(T, \mathbb{C}^3)$, then*

$$\lim_{n\to\infty} \|\mathbf{x}_n - \mathbf{x}\|_{\widetilde{X}_{T,t_0}} = 0 \Leftrightarrow \lim_{n\to\infty} \|\mathbf{x}_n - \mathbf{x}\|_{\mathcal{L}^2(T,\mathbb{C}^3)} = 0.$$

Proof. Let $\mathbf{x} \in \mathcal{L}^2(T, \mathbb{C}^3)$. Then to prove that $\mathbf{x} \in \widetilde{X}_{T,t_0}$, we need to show that $\|\mathbf{x}\|_{\widetilde{X}_{T,t_0}}$ is finite. But,

$$\|\mathbf{x}\|^2_{\widetilde{X}_{T,t_0}} = R_{loss}\|\mathbf{x}\|^2_{\mathcal{L}^2(T,\mathbb{C}^3)} + [E_{rad}(\mathbf{x})](t_0)$$

Then as shown in appendix A.2, there exists a constant $b < \infty$ such that

$$[E_{rad}(\mathbf{x})](t_0) \leq b\|\mathbf{x}\|^2_{\mathcal{L}^2(T,\mathbb{C}^3)}.$$

Therefore,

$$\|\mathbf{x}\|^2_{\widetilde{X}_{T,t_0}} \leq (R_{loss} + b)\|\mathbf{x}\|^2_{\mathcal{L}^2(T,\mathbb{C}^3)} < \infty \tag{2.11}$$

and $\mathbf{x} \in \widetilde{X}_{T,t_0}$.

Now suppose $\{\mathbf{x}_i\}_{i=1}^{\infty}$ is a sequence in $\mathcal{L}^2(T, \mathbb{C}^3)$. Then it is also a sequence in $\widetilde{X}_{T,t_0}$. Also, suppose that $\lim_{i\to\infty} \|\mathbf{x} - \mathbf{x}_i\|_{\mathcal{L}^2(T,\mathbb{C}^3)} = 0$ and let $\delta > 0$ be given. Then there exists an N such that for all $n > N$,

$$\|\mathbf{x}_n - \mathbf{x}\|_{\mathcal{L}^2(T,\mathbb{C}^3)} \leq \frac{\delta}{b + R_{loss}}.$$

Therefore, from equation (2.11) we know that for all $n > N$,

$$\|\mathbf{x}_n - \mathbf{x}\|_{\widetilde{X}_{T,t_0}} \leq \delta.$$

Hence, $\lim_{i\to\infty} \|\mathbf{x} - \mathbf{x}_i\|_{\widetilde{X}_{T,t_0}} = 0$.

The converse statement follows very easily from,

$$\|\mathbf{x}\|_{\mathcal{L}^2(T,\mathbb{C}^3)} \leq \frac{1}{R_{loss}}\|\mathbf{x}\|_{\widetilde{X}_{T,t_0}}. \tag{2.12}$$

□

The above theorem is very useful because it shows that the $\|\cdot\|_{\widetilde{X}_{T,t_0}}$ norm induces the same topology on the space $\mathcal{L}^2(T,\mathbb{C}^3)$ as $\|\cdot\|_{\mathcal{L}^2(T,\mathbb{C}^3)}$ norm. This is useful because it allows us to use several well established properties of $\mathcal{L}^2(T,\mathbb{C}^3)$. The following corollaries are simple consequences of the above theorem.

Corollary 2.3.1. *Let T be some compact set and let $\widetilde{X}_{T,t_0}$ be as described above. Then the space of $\mathbb{C}^3$-valued continuous functions defined on T is dense in $\widetilde{X}_{T,t_0}$.*

Proof. This statement follows form theorem 2.3 and the fact that $\mathbb{C}^3$-valued continuous functions defined on T are dense in $\mathcal{L}^2(T,\mathbb{C}^3)$. □

Corollary 2.3.2. *Let T be some compact set and let $\widetilde{X}_{T,t_0}$ be as described above. If $\{\mathbf{x}_i\}_{i=1}^{\infty}$ is a complete Schauder basis for $\mathcal{L}^2(T,\mathbb{C}^3)$ then it is a complete Schauder basis for $\widetilde{X}_{T,t_0}$.*

Corollary 2.3.3. *Let $\widetilde{X}_{T,t_0}$, $\widetilde{Y}_R$ and Γ be as described in theorem 2.2. Then, the operator Γ is compact.*

Proof. The integral kernel that describes the operator Γ is continuous on $T \times R$. Therefore the operator is compact if it is defined on $\mathcal{L}^2(T,\mathbb{C}^3)$. Hence the operator Γ is compact from theorem 2.2. □

The first corollary above is used to ensure that constraint 3 specified in the introduction to this chapter is satisfied. Because the set of continuous functions is dense in $\widetilde{X}_{T,t_0}$, for any $\mathbf{J} \in \widetilde{X}_{T,t_0}$ and any $\delta > 0$, there exists a $\mathbf{J}_1 \in \widetilde{X}_{T,t_0}$ such that $\|\mathbf{J} - \mathbf{J}_1\|_{\widetilde{X}_{T,t_0}} \leq \delta$ and a continuous function is a member of the equivalence class $\mathbf{J}_1$. The second corollary is useful in the numerical computation of degrees of freedom. As explained in the following two chapters, the existence of a complete Schauder basis is essential for the numerical computation of degrees of freedom. Because there are several well known Schauder bases for $\mathcal{L}^2(T,\mathbb{C}^3)$, the corollary ensures that we have an adequate supply of Schauder bases for $\widetilde{X}_{T,t_0}$.

2.4 Finite Power Case

The finite power case is almost identical to the finite energy case. The only difference is that the total lost and radiated power is used to define a norm on the space of source current densities. The proofs of the following theorems are almost identical to the finite energy case and are omitted here.

Suppose $\mathbf{J} \in \mathcal{L}^2(T,\mathbb{C}^3)$ is some function, such that $\Re\{\mathbf{J}(\mathbf{r})e^{j\omega t}\}$, $\mathbf{r} \in T$, $t \in \mathbb{R}$ is the current density. Then we can calculate the total energy radiated $P_{rad}(\mathbf{J})$ from equation (2.1) and the total energy lost as heat, $P_{loss}(\mathbf{J})$ from equation (2.2).

Theorem 2.4. *Let $T \subset \mathbb{R}^3$ be some compact set. Then we can define a norm on $\mathcal{L}^2(T, \mathbb{C}^3)$ as follows. For all $\mathbf{J} \in \mathcal{L}^2(T, \mathbb{C}^3)$,*

$$\|\mathbf{J}\|_{\widetilde{X}_T} = [P_{loss}(\mathbf{J}) + P_{rad}(\mathbf{J})]^{1/2} .$$

Here, P_{loss} and P_{rad} are as defined in equations (2.1) *and* (2.2) *respectively.*

The major difference between the finite energy case and the finite power case is that there is no obvious way of defining an inner product on the space of source functions $\widetilde{X}_T$ in the finite power case. This is because the radiated power in equation (2.1) is determined using a vector product and not a scalar product. Therefore the space of source current densities is not an inner product space. In the following section I develop a novel theory to cope with compact operators on normed spaces without an inner product structure.

Now, let R be some compact subset of $\mathbb{R}^3$. Then the space of receiver functions $\widetilde{Y}_R$ is similar to the space of receiver functions in the finite energy case. Instead of using the energy stored in the receiving volume R at time instant t_0, we use the time-averaged energy stored in the receiving volume as defined in equation (2.5) to induce a norm on the space $\widetilde{Y}_R$. We can again prove the following theorem which shows that the finite power case can be written as an SWC.

Theorem 2.5. *Let $\widetilde{X}_T$ and $\widetilde{Y}_R$ be as described above. Then, the operator $\Gamma : \widetilde{X}_T \to \widetilde{Y}_R$ which maps any current density in $\widetilde{X}_T$ to the corresponding electric and magnetic fields in $\widetilde{Y}_R$ is linear and bounded.*

We can also prove the following theorem and corollaries, but the proofs are omitted here because they are identical to the finite energy case.

Theorem 2.6. *Let T be a compact set and let $\widetilde{X}_T$ be the normed space described above. If $\mathbf{x} \in \mathcal{L}^2(T, \mathbb{C}^3)$ then $\mathbf{x} \in \widetilde{X}_T$. Moreover, if $\mathbf{x} \in \mathcal{L}^2(T, \mathbb{C}^3)$ and $\{\mathbf{x}_1, \mathbf{x}_2, \ldots\}$ is a sequence in $\mathcal{L}^2(T, \mathbb{C}^3)$, then*

$$\lim_{n\to\infty} \|\mathbf{x}_n - \mathbf{x}\|_{\widetilde{X}_T} = 0 \Leftrightarrow \lim_{n\to\infty} \|\mathbf{x}_n - \mathbf{x}\|_{\mathcal{L}^2(T,\mathbb{C}^3)} = 0.$$

Corollary 2.6.1. *Let T be some compact set and let $\widetilde{X}_T$ be as described above. Then the space of $\mathbb{C}^3$ valued continuous functions defined on T is dense in $\widetilde{X}_T$.*

Corollary 2.6.2. *Let T be some compact set and let $\widetilde{X}_T$ be as described above. If $\mathbf{x}_i$ is a Schauder basis for $\mathcal{L}^2(T, \mathbb{C}^3)$ then it is a Schauder basis for $\widetilde{X}_T$.*

Corollary 2.6.3. *Let $\widetilde{X}_T$, $\widetilde{Y}_R$ and Γ be as described in theorem 2.5. Then, the operator Γ is compact.*

2.5 Discussion

The abstract concept of an SWC was introduced in definition 2.1. In this section I explain how the channels discussed in several papers can be written as SWCs.

As discussed in section 1.4, Bucci *et. al.* [25,33] study the degrees of freedom of electromagnetic fields. Suppose that all sources and scatterers are restricted to be within a sphere $B \subset \mathbb{R}^3$ of radius a and the field is measured over an observation curve $C \subset \mathbb{R}^3$ of arc length $2S$. Suppose that $\mathbf{J}(\mathbf{r}'), \mathbf{r}' \in B$ gives the total current density of the sources and that induced on the scatterers. The total current (source and induced) density is assumed to be constrained so that

$$\int_B |\mathbf{J}(\mathbf{r}')|^2 d^3\mathbf{r}' \leq Ia. \tag{2.13}$$

where I is some positive constant. Then the electric field at a point $\mathbf{r}$ is given by

$$\mathbf{E}(\mathbf{r}) = \int_B \mathbf{J}(\mathbf{r}')\mathbf{G}(\mathbf{r}, \mathbf{r}')d\mathbf{r}' \quad \forall \mathbf{r} \in C. \tag{2.14}$$

Here, the dyadic Green's function [25]

$$\mathbf{G}(\mathbf{r}, \mathbf{r}') = -i\tfrac{\omega\mu_0}{4\pi R} \exp[i\beta(\mathbf{r} - R)]\mathbf{N}(R), \tag{2.15}$$

$$\mathbf{N}(R) = \mathbf{I} + R\exp(j\beta R)\tfrac{\nabla\nabla}{\beta^2}\tfrac{\exp(-i\beta\mathbf{r})}{R} \tag{2.16}$$

with, $\mathbf{r} \in C, \mathbf{r}' \in B$ and $R = |\mathbf{r} - \mathbf{r}'|$. Also, ω is the angular frequency of the electromagnetic wave, β is the wave number and μ_0 is the permeability of free space. One can now define the operator $\Gamma : \mathbf{J} \mapsto \mathbf{E}$. Therefore, $(\mathcal{L}^2(\overline{B}_{a,\mathbb{R}^3}(0), \mathbb{C}^3), \mathcal{L}^2(C, \mathbb{C}^3), \Gamma)$ is the SWC studied in [33]. There is a subtle difference between the physical situation modeled by this channel and the corresponding SWC. In an SWC, we assume that any source current density in the transmitter space of functions, the norm of which is less than some pre-specified bound, can be generated on a real antenna. However, Bucci *et. al.* [25] assume that the scatterers and sources are within the volume $\overline{B}_{a,\mathbb{R}^3}(0)$. Therefore, only certain source/induced current densities are possible that depend on the scatterer distribution within this volume. However, in all subsequent analysis in [25,33] it is assumed that any element of $\mathcal{L}^2(\overline{B}_{a,\mathbb{R}^3}(0), \mathbb{C}^3)$ that satisfies the constraint in equation (2.13) can be a source/induced current density. So, we can assume that the channel used in [25,33] can be written as an SWC.

The work of Bucci *et. al.* [25] is similar to that of Miller [26] where scalar waves are studied. Miller [26] analyses communication between volumes using scalar waves. The main assumption in Miller [26] is that the transmitter is constrained to be in a volume $V \subset \mathbb{R}^3$ and the receiver is constrained to a volume $W \subset \mathbb{R}^3$. We can assume that V and W are compact. Miller [26] assumes that the transmitter wave function $\psi \in \mathcal{L}^2(V, \mathbb{C})$

generates a wave function $\Gamma\psi \in \mathcal{L}^2(W, \mathbb{C})$ according to

$$(\Gamma\psi)(\mathbf{y}) = \int_V G(\mathbf{x}, \mathbf{y})\psi(\mathbf{x})d\mathbf{x} \quad \forall \mathbf{y} \in W.$$

Here, the operator $\Gamma : \mathcal{L}^2(V, \mathbb{C}) \to \mathcal{L}^2(W, \mathbb{C})$ and [26]

$$G(\mathbf{x}, \mathbf{y}) = \frac{\exp\{ik|\mathbf{x} - \mathbf{y} + \mathbf{e}_3 r_0|\}}{4\pi|\mathbf{x} - \mathbf{y} + \mathbf{e}_3 r_0|}.$$

Here $\mathbf{x} \in V, \mathbf{y} \in W$, $r_0 \in \mathbb{R}$ is the distance between the centers of the transmitter and receiver and $\mathbf{e}_3$ is a unit vector in the direction of the vector connecting these centers. It was shown in example 2.1 that this channel can be written as an SWC. Piestun *et. al.* [34] study communication using vector waves and their work is very similar to that of [26] which studies communication using scalar waves. Therefore, the channel studied in [34] can similarly be written as an SWC.

Hanlen *et. al.* [36] generalise the work in [26] by including the effect of scatterers. The effect of scatterers is incorporated into the channel operator Γ in an SWC. Therefore, the channel studied in [36] can be written as an SWC. Similarly, the channel studied in Xu *et. al.* [37], which generalises the work of [34] to include the effect of scatterers can be written as an SWC by changing the operator Γ.

Poon *et. al.* [28] also consider communication between finite volume transmitters and receivers in the presence of scatterers. But instead of considering individual scatterers, they consider a cluster of scatterers that are constrained to be within some finite volume. The scatterers reflect waves from the transmitter to the receiver provided the waves from the transmitter emerge from some solid angle Ω_t [28]. The Green's function, $G(\mathbf{r}_R, \mathbf{r}_T)$ is split up into three components

$$G(\mathbf{r}_R, \mathbf{r}_T) = \int\int A_R(\mathbf{r}_R, \hat{\mathbf{K}})H(\hat{\mathbf{K}}, \hat{\mathbf{k}})A_T(\hat{\mathbf{k}}, \mathbf{r}_T)d\hat{\mathbf{K}}d\hat{\mathbf{k}}. \tag{2.17}$$

Here A_T models the transmitting volume, A_R models the receiving volume and H models the scatterers. Also, $\hat{\mathbf{K}}$ and $\hat{\mathbf{k}}$ are unit vectors in the transmitting and receiving volumes and the integration is performed over the unit spheres. In order to model the scattering environment, [28] assumes that

$$H(\hat{\mathbf{K}}, \hat{\mathbf{k}}) \neq 0 \text{ if and only if } \mathbf{k} \in \Omega_t \text{ and } \mathbf{K} \in \Omega_r.$$

Here, Ω_t and Ω_r are solid angles over which the scatterers are assumed to be visible from the transmitting and receiving volumes [28]. The channel studied by Poon *et. al.* [28] can be written as an SWC similar to the channel studied by Miller [26]. The SWC is simply $(\mathcal{L}^2(A_T, \mathbb{C}^3), \mathcal{L}^2(A_R, \mathbb{C}^3), \Gamma)$. Here the channel operator Γ is determined by the Green's function in equation (2.17).

2.6 Chapter Conclusion

Spatial waveform channels have previously been used to calculate performance bounds for MIMO channels [26,28,36,37,45]. However, SWCs have not been clearly defined in any of these papers. In this chapter I give a novel definition for SWCs that is sufficiently general to include the works of [25,26,28,36,37,45] as special cases. This definition gives SWCs sufficient structure to model the different physical constraints imposed on MIMO systems.

According to definition 2.1 an SWC is a triple $(\widetilde{\mathcal{F}}_T, \widetilde{\mathcal{F}}_R, \Gamma)$. Here $\widetilde{\mathcal{F}}_T$ is the space of all the transmitter current densities, $\widetilde{\mathcal{F}}_R$ is the space of electromagnetic fields in the receiving volume and $\Gamma : \widetilde{\mathcal{F}}_T \rightarrow \widetilde{\mathcal{F}}_R$ is a bounded linear operator that determines the electromagnetic field in the receiving volume given the current density in the transmitting volume. The spaces $\widetilde{\mathcal{F}}_T$ and $\widetilde{\mathcal{F}}_R$ are normed spaces and the norms on these spaces can be physically interpreted as energy or power.

The main contributions of this chapter are as follows:

1. Novel definition for spatial waveform channels and a physical interpretation of the definition.

2. Demonstrated that communication using electromagnetic waves in MIMO systems can be written as an SWC. The constraints imposed on a MIMO system can be incorporated into the corresponding SWC model.

3. Established several properties in sections 2.3 and 2.4 of SWCs that can be used to model communication using electromagnetic waves in: *finite energy* and *finite power* cases.

Chapter 3

Degrees of Freedom and Essential Dimension for Compact Operators

The concepts of degrees of freedom and essential dimension are very general. In this chapter I motivate the definition for degrees of freedom at level ϵ for a compact operator on an arbitrary normed space. I then explain how the notion of degrees of freedom can be used to define generalised singular values. As the name suggests, singular values of compact operators on Hilbert spaces are special cases of generalised singular values. These generalised singular values play the same role that the well known singular values of compact operators on Hilbert spaces do. I also distinguish between the terms essential dimension and degrees of freedom and show that one can associate a unique essential dimension with any compact operators on normed spaces.

3.1 Motivation

I motivate the definition of degrees of freedom at level ϵ for compact operators on normed spaces by considering linear operators on finite dimensional spaces. I will discuss several possible definitions for degrees of freedom. I will explain the physical intuition behind each definition using a simple example.

Consider a communication channel that uses n transmitting antennas and m receiving antennas which can be mathematically modeled as follows. Let the current on the n transmitting antennas be given by $\mathbf{x} \in \mathbb{C}^n$. This current on the transmitting antennas generates a current $\mathbf{y} \in \mathbb{C}^m$ in the m receiving antennas according to the equation

$$\mathbf{y} = \mathbf{H}\mathbf{x}.$$

Here, $\mathbf{H} \in \mathbb{C}^{m \times n}$ is the channel matrix. We can define the operator $H : \mathbb{C}^n \to \mathbb{C}^m$ as $\mathbf{x} \mapsto \mathbf{y}$.

Note that $\mathbf{y}$ is not assumed to be the current on the receiving antenna but is the current generated by the transmitting antenna in the receiving antenna. To complete the model let the current on the receiving antenna be given by

$$\mathbf{y}_r = \mathbf{H}\mathbf{x} + \mathbf{n}.$$

Here $\mathbf{n}$ is the additive noise present at the receiver. The fundamental problem in communication is to determine the signal $\mathbf{x}$ that was transmitted given that some signal $\mathbf{y}_r$ was received. Generally this problem is ill-conditioned.

Firstly, H need not be injective. Therefore, there might exist a $\mathbf{y} \in H\mathbb{C}^n$ and $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}^n$ such that $\mathbf{x}_1 \neq \mathbf{x}_2$ and

$$H\mathbf{x}_1 = H\mathbf{x}_2 = \mathbf{y}.$$

Even if H is injective the problem could be ill-conditioned owing to the presence of noise. Suppose the noise in the receiver can be modeled as follows:

$$\text{Prob}(\mathbf{n}) = \begin{cases} \frac{1}{N(\epsilon)} & |\mathbf{n}| \leq \epsilon, \\ 0 & |\mathbf{n}| > \epsilon. \end{cases}$$

Here, $\epsilon > 0$ is some constant and $N(\epsilon)$ is a normalisation constant. So, if $\mathbf{y}$ is the received signal then we can only tell for certain that the transmitted signal $\mathbf{x}$ is in the set $H^{-1}(\overline{B}_{\epsilon,\mathbb{C}^m}(\mathbf{y}) \cap H\mathbb{C}^n)$. Here, $\overline{B}_{\epsilon,\mathbb{C}^m}(\mathbf{y})$ is the closed ball of radius ϵ centered at $\mathbf{y}$. Therefore, the noise in the receiver fundamentally determines whether or not one can calculate the transmitted signal given the received signal.

It is possible to overcome the above problem by using the following method. For all sets $Q \subset \mathbb{C}^m$ let $\overline{B}_{\epsilon,\mathbb{C}^m}(Q) = \cup_{\mathbf{y}\in Q}\overline{B}_{\epsilon,\mathbb{C}^m}(\mathbf{y})$. Also, let the set $S \subset \mathbb{C}^n$ be such that

$$\forall \mathbf{y} \in \overline{B}_{\epsilon,\mathbb{C}^m}(HS) : \; H^{-1}(\overline{B}_{\epsilon,\mathbb{C}^m}(\mathbf{y}) \cap HS) \cap S \text{ contains exactly one element.} \quad (3.1)$$

If we ensure that the transmitted signal is an element of S, then for each received signal, one can uniquely determine the transmitted signal. Generally, there will exist sets that satisfy the above condition and have infinitely many elements.

However, along with non-zero receiver noise, in most physical channels only a finite amount of power $P > 0$ is available for transmission. That is,

$$|\mathbf{x}| \leq P.$$

Without loss of generality, we can assume that $P = 1$ because if the available power is different from 1, then we can always scale ϵ accordingly. This leads us to

Possible Definition for Degrees of Freedom 1. *Suppose $H : \mathbb{C}^n \to \mathbb{C}^m$ is a linear operator and let $\epsilon > 0$. Then the number of degrees of freedom at level-ϵ is the largest number N such that there exists a set $S \subset \overline{B}_{1,\mathbb{C}^n}(0) \subset \mathbb{C}^n$ with N elements that satisfies property* (3.1).

One can think of the number of degrees of freedom as the maximal number of different signals that a transmitter can send so that, even in the presence of noise and a power constraint on the transmitted signal, the receiver can decode the received signal to accurately determine the transmitted signal.

An alternative way of defining the degrees of freedom is in terms of finite ϵ-nets[1]. If $\mathbf{y}$ and $\mathbf{y}'$ are two received signals then I will call the signals $\mathbf{y}$ and $\mathbf{y}'$ physically indistinguishable at level-ϵ if $|\mathbf{y}-\mathbf{y}'| \leq \epsilon$ (see eg. [25]). The intuition behind calling two such functions physically indistinguishable is fairly obvious. If the distance between the vectors $\mathbf{y}$ and $\mathbf{y}'$ is less than ϵ, the noise at the receiver can make $\mathbf{y}$ look like $\mathbf{y}'$ so a real physical receiver cannot tell the difference between $\mathbf{y}$ and $\mathbf{y}'$. Also we will call a set S physically indistinguishable from a set S' at level-ϵ if for all $\mathbf{x} \in S$ there exists an $\mathbf{x}' \in S'$ such that $|\mathbf{x} - \mathbf{x}'| \leq \epsilon$. That is, S' is an ϵ-net for S.

The notion of physical indistinguishability can be used to define the degrees of freedom because the set $H\overline{B}_{1,\mathbb{C}^n}(0)$ is physically indistinguishable from some set with a finite number of elements. To see this note that $H\overline{B}_{1,\mathbb{C}^n}(0)$ is a bounded subset of a finite dimensional space and is therefore totally bounded. It therefore has a finite ϵ-net.

Possible Definition for Degrees of Freedom 2. *Suppose $H : \mathbb{C}^n \to \mathbb{C}^m$ is a linear operator and let $\epsilon > 0$ be given. Then the number of degrees of freedom at level-ϵ of H is the smallest number N such that $H\overline{B}_{1,\mathbb{C}^n}(0)$ has an ϵ-net with N elements.*

One can think of the degrees of freedom N as the minimum number of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathbb{C}^n$ so that for all possible transmitter functions $\mathbf{x} \in \overline{B}_{1,\mathbb{C}^n}(0)$, there exists some $\mathbf{x}_i, 1 \leq i \leq N$ such that $H\mathbf{x}$ and $H\mathbf{x}_i$ are physically indistinguishable at the receiver.

The following example will make the two definitions clearer.

Example 3.1. *Let $m = n = 2$, $P = 1$, $\epsilon = 0.25$ and*

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Now suppose that N_1 and N_2 are respectively the degrees of freedom of $\mathbf{H}$ if degrees of freedom are defined as in the possible definitions 1 and 2

The set $S = \{(x, y) : x \in \{0, 0.5\}, y \in \{0, 0.25, 0.5\}\}$ satisfies (3.1)*. Therefore, the number of elements of S, 6 is a lower bound for N_1.*

The set $S_\epsilon = \{(x, y) : x = 0.1n, y = 0.1m, n = 1, \ldots, 10$ and $m = 1, \ldots, 20\}$ is an ϵ-net for $\mathbf{H}\overline{B}_{P,\mathbb{C}^2}(0)$. Therefore the number of elements in S_ϵ, 200 is an upper bound for N_2.

[1] If S is any subset of a metric space then S_ϵ is an ϵ-net for S if for all $x \in S$, there exists an $x_\epsilon \in S_\epsilon$ such that $\mathrm{dist}(x, x_\epsilon) < \epsilon$.

The above example illustrates the physical interpretation of the degrees of freedom. But, the problem with using either of these definitions for degrees of freedom is that it is very difficult to calculate the exact number of degrees of freedom. This fact is obvious from the very simple example considered above. This problem is commonly know as the sphere packing problem. Though we cannot easily solve this problem, we can look at the dimensionality of the sphere packing problem. That is, instead of looking at specific subsets S and S_ϵ of $\mathbb{C}^n$ and $\mathbb{C}^m$, we can look at linear subspaces.

The singular value decomposition theorem tells us that there exists a set of basis functions in $\mathbb{C}^n$ and $\mathbb{C}^m$ such that the matrix representation for H in these basis function is diagonal. Let $\mathbf{H}$ be such a matrix with the basis functions ordered such that the diagonal elements are in decreasing order. A simple examination of the diagonal matrix proves that[2]

1. for all $\epsilon > 0$ there exist a number N_1 and a set of mutually orthogonal vectors, $\{\mathbf{v}_1, \ldots, \mathbf{v}_{N_1}\} \subset \mathbb{C}^n$ such that if any vector $\mathbf{v}$ is orthogonal to $\mathbf{v}_i, i = 1, \ldots, N_1$ and if $|\mathbf{v}| \leq 1$, then $|\mathbf{H}\mathbf{v}| < \epsilon$. For a given ϵ, call the smallest number that satisfies the above condition $N_1(\epsilon)$. Note that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{N_1}$ span the space of all linear combinations of the right singular vectors of H whose corresponding singular values are greater than ϵ.

2. for all $\epsilon > 0$ there exist a number N_2 and a set of linearly independent vectors, $\{\mathbf{u}_1, \ldots, \mathbf{u}_{N_2}\} \subset \mathbb{C}^m$ such that for all $\mathbf{v} \in \overline{B}_{1,\mathbb{C}^n}(0)$,

$$\inf_{a_i} |\mathbf{H}\mathbf{v} - \sum_{i=1}^{N_2} a_i \mathbf{u}_i| < \epsilon.$$

   For a given ϵ, call the smallest number that satisfies the above condition $N_2(\epsilon)$. Note that the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_{N_1}$ span the space of all linear combinations of the left singular vectors of H whose corresponding singular values are greater than ϵ.

A simple examination of the diagonal matrix tells us that both $N_1(\epsilon)$ and $N_2(\epsilon)$ are equal to the number of singular values of $\mathbf{H}$ that are greater than ϵ. This leads us to our definition for degrees of freedom in finite dimensional spaces.

Definition 3.1. *Let $H : \mathbb{C}^n \to \mathbb{C}^m$ be a linear operator and let $\epsilon > 0$ be given. Then the number of degrees of freedom at level-ϵ for H is the smallest number N such that there exists a set of vectors $\mathbf{u}_1, \ldots, \mathbf{u}_N \in \mathbb{C}^m$ such that for all $\mathbf{v} \in \overline{B}_{1,\mathbb{C}^n}(0)$,*

$$\inf_{a_1,\ldots,a_N} |H\mathbf{v} - \sum_{i=1}^{N} a_i \mathbf{u}_i| < \epsilon.$$

[2] I will not go through a detailed proof here because this is just a special case of theorem 3.4.

This particular characterisation for the number of degrees of freedom is the one that is most easily generalised to compact operators on normed spaces. Moreover, it is fairly straightforward to compute the singular values of matrices and therefore one can calculate the degrees of freedom for a given matrix. The following example illustrates the use of the above definition for degrees of freedom.

Example 3.2. *Let* $n = 5$, $m = 4$ *and* [3]

$$\mathbf{H} = \begin{bmatrix} 0.2 & 0.2 & 0.5 & 1.0 & 0.8 \\ 0.8 & 0.6 & 0.9 & 0.2 & 0.1 \\ 0.1 & 0.3 & 0.6 & 0.8 & 0.6 \\ 0.5 & 0.1 & 0.0 & 0.6 & 0.1 \end{bmatrix}$$

The singular values of $\mathbf{H}$ *are* $4.6, 1.0, 0.30, 0.01$ *and* 0. *Therefore the number of degrees of freedom at level-*0.1 *is 3 and the number of degrees of freedom at level-*0.001 *is 4.*

I will look at one final example to show how the ideas in the previous example can be generalised to infinite dimensional function spaces. The main tool used in this example is the singular value decomposition of compact operators in Hilbert Spaces.

Example 3.3. *Consider communication using scalar waves in free space using finite volume transmitting and receiving antennas (see example 2.1). Let* $V, W \subset \mathbb{R}^3$ *be measurable, compact and* $V \bigcap W = \emptyset$ *and* $G : V \times W \to \mathbb{C}$ *be defined by*

$$G(\mathbf{x}, \mathbf{y}) = \frac{e^{jk|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}.$$

Then,

1. $\Gamma : \mathcal{L}^2(V, \mathbb{C}) \to \mathcal{L}^2(W, \mathbb{C})$ *defined by*

$$(\Gamma\phi)(\mathbf{z}) = \int_V G(\mathbf{x}, \mathbf{z})\phi(\mathbf{x})d\mathbf{x}$$

   is compact [46].

2. $K : V \times V \to \mathbb{C}$ *defined by*

$$K(\mathbf{x}, \mathbf{y}) = \int_W G^*(\mathbf{x}, \mathbf{z})G(\mathbf{y}, \mathbf{z})d\mathbf{z}$$

   is continuous and Hermitian.

[3] This matrix was generated using the rand function in Matlab.

3. $\Psi : \mathcal{L}^2(V, \mathbb{C}) \to \mathcal{L}^2(W, \mathbb{C})$ *defined by*

$$(\Psi\phi)(\mathbf{y}) = \int_V k(\mathbf{x}, \mathbf{y})\phi(\mathbf{x})d\mathbf{x}$$

is compact, self-adjoint and non-negative.

4. *We now have the spectral theorem [47, pp. 261]*

$$\begin{aligned} \Psi &= \Gamma^*\Gamma \\ &= \sum_{j=1}^{\infty} \alpha_j^2 \langle \cdot, \phi_j \rangle_{\mathcal{L}^2(V,\mathbb{C})} \phi_j. \end{aligned}$$

Here, ϕ_j are the set of orthonormal eigenfunctions of Ψ that we can get via Gram-Schmidt orthogonalisation and α_j are the singular values of Γ. Note that we can have $\alpha_j = \alpha_i$ even if $i \neq j$.

5. *We also have the singular value decomposition (SVD) theorem [47, pp. 261]*

$$\Gamma = \sum_j \langle \cdot, \phi_j \rangle_{\mathcal{L}^2(V,\mathbb{C})} (\Gamma\phi_j) \tag{3.2}$$

6. *The eigenvalues $\{\alpha_j^2\}$ of Ψ form a countable set with 0 being the only possible point of accumulation [48, thm 8.3-1].*

Now suppose $\epsilon > 0$ is some small positive number. Then the number of mutually orthogonal functions $\phi \in \mathcal{L}^2(V, \mathbb{C})$ such that $\|\phi\|_{\mathcal{L}^2(V,\mathbb{C})} \leq 1$ and $\|\Gamma\phi\|_{\mathcal{L}^2(W,\mathbb{C})} > \epsilon$ is finite. To see this note that if ϕ_i is an eigenfunction of Ψ then

$$\begin{aligned} \|\Gamma\phi_i\|^2_{\mathcal{L}^2(W,\mathbb{C})} &= \int_W \left\{ \int_V G(\mathbf{x}, \mathbf{z})\phi_i(\mathbf{x})d\mathbf{x} \right\}^* \int_V G(\mathbf{y}, \mathbf{z})\phi_i(\mathbf{y})d\mathbf{y}d\mathbf{z} \\ &= \int_V \phi_i^*(\mathbf{x})(\Psi\phi_i)(\mathbf{x})d\mathbf{x} \\ &= \alpha_i^2. \end{aligned}$$

From point 6 above, we know that if $\{\lambda_j = \alpha_j^2\}$ is the set of all eigenvalues of Ψ then there is a finite subset $\{\lambda_{j_i}\}_{i=1}^N$ of the set of eigenvalues such that $\lambda_j > \epsilon$ if and only if $j \in \{j_1, \ldots, j_N\}$. The set $\{\phi_{j_1}, \ldots, \phi_{j_N}\}$ has the required property. Now, if $\phi \in \mathcal{L}^2(V, \mathbb{C})$ with $\|\phi\|_{\mathcal{L}^2(V,\mathbb{C})} \leq 1$, then we can write $\phi = \sum_j a_j\phi_j + \phi_r$. Here, $a_j = \langle \phi, \phi_j \rangle_{\mathcal{L}^2(V,\mathbb{C})}$ and ϕ_r is the remainder term. From Parseval's theorem we have

$$\begin{aligned} \sum_j |a_j|^2 + \|\phi_r\|^2_{\mathcal{L}^2(V,\mathbb{C})} &\leq \|\phi\|^2_{\mathcal{L}^2(V,\mathbb{C})} \\ &\leq 1. \end{aligned}$$

Finally, if $\phi \in \mathcal{L}^2(V, \mathbb{C})$ is orthogonal to all the $\phi_{j_i}, i = 1, \ldots, N$, the SVD theorem (point 5) gives

$$\begin{aligned}
\|\Gamma\phi\|^2_{\mathcal{L}^2(W,\mathbb{C})} &= \sum_j |a_j|^2 \lambda_j \\
&= \sum_{\substack{j \\ j \neq j_1, \ldots, j_N}} |a_j|^2 \lambda_j \\
&\leq \sum_{\substack{j \\ j \neq j_1, \ldots, j_N}} \epsilon |a_j|^2 \\
&\leq \epsilon.
\end{aligned}$$

Equally, for all $\epsilon > 0$ there exists a number N and a set of functions $\psi_1, \ldots \psi_N$ such that for all $\phi \in \mathcal{L}^2(V, \mathbb{C}), \|\phi\|_{\mathcal{L}^2(V,\mathbb{C})} \leq 1$

$$\inf_{a_1,\ldots,a_N} \|\Gamma\phi - \sum_{i=1}^{N} a_i \psi_i\|_{\mathcal{L}^2(W,\mathbb{C})} \leq \epsilon$$

The proof follows from choosing $\psi_i = \Gamma\phi_{j_i}$.

The singular value decomposition theorem was used in the previous example to prove the required result. However, the compactness of the operator Γ is essential to ensure that the eigenvalues of $\Gamma^*\Gamma$ form a countable set with zero being the only possible point of accumulation. This leads us to the idea that the compactness of Γ is sufficient to ensure that we can generalise the notion of degrees of freedom to compact operators in normed spaces. In fact, the definition for degrees of freedom can then be used to define singular values for compact operators on arbitrary normed spaces. This will be extremely useful in numerically calculating the degrees of freedom for compact operators.

3.2 Degrees of Freedom for Compact Operators

In this section I show how the definition of degrees of freedom in finite dimensional spaces can be generalised to compact operators on arbitrary normed spaces. The following theorem shows that it makes sense to talk about the number of degrees of freedom for compact operators.

Theorem 3.1. *Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively and $T : X \to Y$ is a compact operator. Then for all $\epsilon > 0$ there exists $N \in \mathbb{Z}_0^+$ and $\{\psi_i\}_{i=1}^N \subset Y$ such that*

$$\inf_{a_1,\ldots,a_N} \|Tx - \sum_{i=1}^{N} a_i\psi_i\|_Y \leq \epsilon \; \forall x \in \overline{B}_{1,X}(0) \tag{3.3}$$

Proof. The proof is by contradiction. Let $\epsilon > 0$ be given. Suppose no such N exists.

Let $x_1 \in \overline{B}_{1,X}(0)$ be any vector. Choose $\psi_1 = Tx_1$. Then,

$$\inf_{a_1} \|Tx_1 - \sum_{i=1}^{1} a_i\psi_i\|_Y = 0 \leq \epsilon.$$

Suppose that $\{\psi_1, \ldots, \psi_N\}$ and $\{x_1, \ldots, x_N\}$ have been chosen. Then, by our assumption, there exists an $x_{N+1} \in \overline{B}_{1,X}(0)$ such that

$$\inf_{a_1,\ldots,a_N} \|Tx_{N+1} - \sum_{i=1}^{N} a_i\psi_i\|_Y > \epsilon. \tag{3.4}$$

Choose, $\psi_{N+1} = Tx_{N+1}$. Also, if $M \leq N$, by choosing $a_i = 0$ if $i \leq N, i \neq M$ and $a_i = 1$, $i = M$ in the inequality (3.4), we have

$$\|Tx_{N+1} - Tx_M\|_Y > \epsilon.$$

Therefore, using the Cauchy criterion, the sequence $\{Tx_n\}_{n=1}^{\infty}$ chosen by induction cannot have a convergent subsequence. This is the required contradiction because $\{x_n\}_{n=1}^{\infty}$ is a bounded sequence and T is compact. □

So we can use the following definition for the number of degrees of freedom at level-ϵ for compact operators on normed spaces.

Definition 3.2 (Number of degrees of freedom at level-ϵ)**.** *Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively and $T : X \to Y$ is a compact operator. Then the number of degrees of freedom at level-ϵ is the smallest $N \in \mathbb{Z}^+$ such that there exists a set of vectors $\{\psi_1, \ldots, \psi_N\} \subset Y$ such that*

$$\inf_{a_1,\ldots,a_N} \|Tx - \sum_{i=1}^{N} a_i\psi_i\|_Y \leq \epsilon \; \forall x \in \overline{B}_{1,X}(0).$$

This definition is a descriptive one and can not be used to calculate the number of degrees of freedom for a given compact operator. In the finite dimensional case we can calculate the degrees of freedom by calculating the singular values. However, as far as

I am aware, there is no known description of singular values for compact operators on arbitrary normed spaces. In the following I will give a novel definition for generalised singular values of compact operators[4]. In fact we use the degrees of freedom to define singular values.

I will first establish some very simple properties of degrees of freedom at level-ϵ of compact operators.

Theorem 3.2. *Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively and $T : X \to Y$ is a compact operator. Let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom T at level-ϵ. Then*

1. *There exists an $\epsilon_0 \leq \|T\| < \infty$ such that for all $\epsilon > \epsilon_0$, $\mathcal{N}(\epsilon) = 0$.*
2. *$\mathcal{N}(\epsilon)$ is a non-increasing function of ϵ.*
3. *In any finite interval, $(\epsilon_1, \epsilon_2) \subset \mathbb{R}$, with $\epsilon_2 > \epsilon_1 > 0$, $\mathcal{N}(\epsilon)$ has only finitely many discontinuities (different values).*
4. *Unless T is identically zero, there exists an ϵ_0 such that for all $\epsilon < \epsilon_0$, $\mathcal{N}(\epsilon) \geq 1$.*

Proof.

1. Because T is compact, it is bounded and therefore $\|T\| < \infty$. Suppose $\epsilon > \|T\|$ then from the definition of $\|T\|$, if $\phi \in X$ and $\|\phi\|_X \leq 1$, then $\|T\phi\| \leq \|T\| < \epsilon$. Therefore $\mathcal{N}(\epsilon) = 0$.

2. Suppose $\epsilon_1 < \epsilon_2$. Then there exist functions $\psi_1, \ldots, \psi_{\mathcal{N}(\epsilon_1)}$ such that for all $\phi \in \overline{B}_{1,X}(0)$,

$$\inf_{a_1,\ldots,a_{\mathcal{N}(\epsilon_1)}} \|T\phi - \sum_{i=1}^{\mathcal{N}(\epsilon_1)} a_i\psi_i\|_Y < \epsilon_1 < \epsilon_2$$

   Therefore $\mathcal{N}(\epsilon_2) \leq \mathcal{N}(\epsilon_1)$ from the definition of degrees of freedom at level-ϵ.

3. This follows from part 2 and the definition of degrees of freedom at level-ϵ.

4. Because $\|T\| > 0$, there exists a function $\phi \in X$, $\|\phi\|_X \leq 1$ such that $\|T\phi\|_Y > 0$. Then for all $\epsilon < \|T\phi\|$, $\mathcal{N}(\epsilon) \geq 1$.

□

The following examples show that as ϵ goes to zero, $\mathcal{N}(\epsilon)$ can both be bounded or go to infinity.

[4] In chapter 4 I will explain how one can use numerical methods to calculate these singular values.

Example 3.4. *Let l^1 be the Banach space of all real-valued sequences with finite l^1 norm and let $(e_1, e_2, \ldots)$ be the standard Schauder basis for l^1. Then define the operator $T : l^1 \to l^1$ as $e_n \mapsto e_1 \; \forall n$. This operator is compact.*

To see this let $x_n = \sum_{i=1}^{\infty} \alpha_{in} e_i$, $n = 1, 2, \ldots$ be a sequence in the unit ball in l^1. Then

$$\sum_{i=1}^{\infty} |\alpha_{in}| \leq 1.$$

So

$$Tx_n = \left[\sum_{i=1}^{\infty} \alpha_{in} \right] e_1$$

is well defined and bounded. Because the dimension of the range of T is one, it is compact [48] and for all $\epsilon > 0$ the $\mathcal{N}(\epsilon) \leq 1$.

Example 3.5. *Let l^1 and $(e_1, e_2, \ldots)$ be as defined in the previous example. Define $T : l^1 \to l^1$ as $e_n \mapsto \frac{1}{n} e_n \; \forall n$. It is fairly trivial to show that T is compact and also that $\lim_{\epsilon \to 0} \mathcal{N}(\epsilon) = \infty$.*

Figure 3.1 shows a typical example of degrees of freedom at level-ϵ for some compact operator that satisfies all the properties in the above theorem. We identify the discontinuities in the number of degrees of freedom of T at level-ϵ with the singular values of T.

Definition 3.3 (Generalised Singular Values)**.** *Suppose X and Y are normed spaces and $T : X \to Y$ is a compact operator. Let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom of T at level ϵ. Then ϵ_m is the m^{th} generalised singular value of T if*

$$\begin{aligned} &\sup_{\epsilon > \epsilon_m} \mathcal{N}(\epsilon) = m - 1 \quad \textit{and} \\ &\inf_{\epsilon < \epsilon_m} \mathcal{N}(\epsilon) = M \geq m. \end{aligned}$$

Further, if $m < M$ then for all $m < n \leq M$, $\epsilon_n := \epsilon_m$ is the n^{th} generalised singular value of T.

Note that we used the behavior of degrees of freedom to identify the generalised singular values. Let the degrees of freedom of the operator Γ be as shown in figure 3.1. Then the generalised singular values, ϵ_m, of Γ identify the jumps in the degrees of freedom. So, $\epsilon_1 = 0.9, \epsilon_2 = \epsilon_3 = \epsilon_4 = 0.8, \epsilon_5 = 0.6 \ldots$.

The intuition behind the definition for generalised singular values needs further clarification. In the finite dimensional case, if ϵ_p is the p^{th} singular value of some operator $\Gamma : \mathbb{C}^n \to \mathbb{C}^m$, then there exist corresponding left and right singular functions $v_p \in \mathbb{C}^n$ and $u_p \in \mathbb{C}^m$ such that v_p is of unit norm, $\Gamma v_p = u_p$ and the norm of u_p is ϵ_p. This is not necessarily true for arbitrary compact operators on normed spaces as the following example proves.

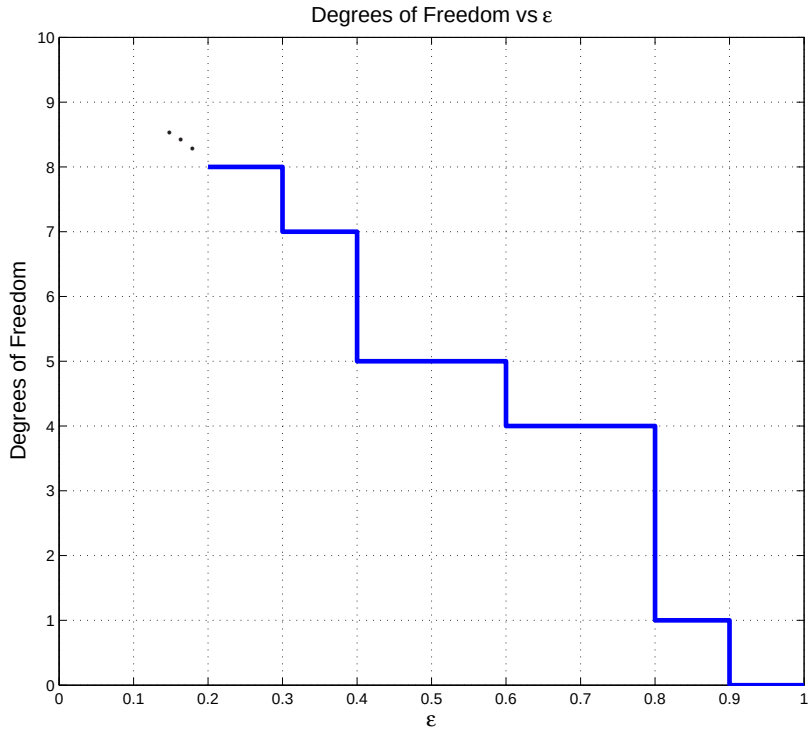

Figure 3.1: Degrees of Freedom of a Compact Operator

Example 3.6. *Let l^1 be the sequence space of real numbers with the standard Schauder basis $(e_1, e_2, \ldots)$. Define the operator $\Gamma : l^1 \to l^1$ as $\Gamma e_n = (1 - \frac{1}{n})e_1 \; \forall n$. Then Γ is bounded and because the range of Γ is finite dimensional it is compact. Also, the number of degrees of freedom of Γ at level-ϵ is*

$$\mathcal{N}(\epsilon) = \begin{cases} 0 & \text{if } \epsilon \geq 1, \\ 1 & \text{if } \epsilon < 1. \end{cases}$$

So $\epsilon_1 = 1$. However, for any vector $\mathbf{u}$ in the unit ball in l^1, $\|\Gamma u\|_{l^1} < 1$.

The above example motivates the following theorem which explains the intuition behind the definition of generalised singular values.

Theorem 3.3. *Suppose X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively and $\Gamma : X \to Y$ is a compact operator. Let ϵ_m be a singular value of the operator Γ. Then for all $\theta > 0$, there exists a function $\psi \in \overline{B}_{1,X}(0)$ such that*

$$\epsilon_m + \theta \geq \|\Gamma\psi\|_Y \geq \epsilon_m - \theta.$$

Proof. The proof is by contradiction. Assume that there exists a $\theta > 0$ such that for all $\psi \in X$, if $\|\psi\|_X = 1$ then $\|\Gamma\psi\|_Y \notin [\epsilon_m - \theta, \epsilon_m + \theta]$. Let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom at level ϵ of the operator Γ. From the definition of degrees of freedom at level-ϵ we have

$$\mathcal{N}(\epsilon_m + \theta) \leq m - 1, \tag{3.5}$$
$$\mathcal{N}(\epsilon_m - \theta) \geq m. \tag{3.6}$$

So there exist functions $\phi_1, \ldots, \phi_{m-1}$ such that for all $\psi \in \overline{B}_{1,X}(0)$

$$\inf_{a_1,\ldots,a_{m-1}} \|\Gamma\psi - \sum_{i=1}^{m-1} a_i\phi_i\| \leq \epsilon_m + \theta.$$

Because for all $\psi \in X$, if $\|\psi\|_X = 1$ then $\|\Gamma\psi\|_Y \notin [\epsilon_m - \theta, \epsilon_m + \theta]$,

$$\inf_{a_1,\ldots,a_{m-1}} \|\Gamma\psi - \sum_{i=1}^{m-1} a_i\phi_i\| \leq \epsilon_m - \theta.$$

if $\|\psi\|_X = 1$. So $\mathcal{N}(\epsilon_m - \theta) \leq m - 1$. This contradicts equation (3.6). Therefore there exists a $\psi \in \overline{B}_{1,X}(0)$ that satisfies the conditions of the theorem. □

The above theorem shows how the generalised singular values are related to the traditionally accepted notion of singular values of compact operators on Hilbert spaces. However, we still need to prove that in the special case of Hilbert spaces, the new definition for generalised singular values agrees with the traditionally accepted definition.

Recall that if $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces with inner products $\langle\cdot,\cdot\rangle_{\mathcal{H}_1}$ and $\langle\cdot,\cdot\rangle_{\mathcal{H}_1}$ respectively and if $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator then the Hilbert adjoint operator for T is defined as the operator $T^* : \mathcal{H}_2 \to \mathcal{H}_1$ that satisfies [48, Sec. 3.9]

$$\langle Tx, y\rangle_{\mathcal{H}_2} = \langle x, T^*y\rangle_{\mathcal{H}_1} \; \forall x \in \mathcal{H}_1 \text{ and } y \in \mathcal{H}_2.$$

The singular values of T are defined to be the square roots of the eigenvalues of the operator $T^*T : \mathcal{H}_1 \to \mathcal{H}_1$. I will refer to these as Hilbert space singular values to distinguish them from generalised singular values.

In order to prove that the generalised singular values are equal to Hilbert space singular values we need the following theorem. This theorem is important in its own right because it shows that there are two other equivalent ways of calculating the degrees of freedom of a Hilbert space operator.

Theorem 3.4. *Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator. Then for all $\epsilon > 0$ there exists an $N \in \mathbb{Z}^+$ and a set of N mutually orthogonal functions $\{\phi_i\}_{i=1}^N$ such that if*

$$x \in \mathcal{H}_1,\ \|x\|_{\mathcal{H}_1} \leq 1 \text{ and } \langle x, \phi_i\rangle_{\mathcal{H}_1} = 0$$

then

$$\|Tx\|_{\mathcal{H}_2} \leq \epsilon.$$

Moreover the smallest N that satisfies the above condition for a given ϵ is equal to the number of degrees of freedom of T at level-ϵ.

Proof. I first prove that such an N exists and then prove that the smallest such N is in fact equal to the number of degrees of freedom.

Let $\epsilon > 0$ be given. Because T is compact, we can use the singular value decomposition theorem which says [47, pp. 261]

$$T\cdot = \sum_i \sigma_i \langle\cdot, \phi_i\rangle_{\mathcal{H}_1} \psi_i. \tag{3.7}$$

Here, σ_i, ϕ_i and ψ_i are the Hilbert space singular values and left and right singular functions of T, respectively. Because T is compact, the set of singular values is at most countable [48, thm 8.3-1] and we can reorder the Hilbert space singular values so that $\sigma_i \geq \sigma_j$ if $i < j$. Also, because 0 is the only possible point of accumulation for the Hilbert space singular values [48, thm 8.3-1], there exists a number $N_1 \in \mathbb{Z}^+$ such that $\sigma_n > \epsilon$ if and only if $n \leq N_1$.

Now, if x is orthogonal to $\phi_i, i = 1, \ldots, N_1$ and if $\|x\|_{\mathcal{H}_1} \leq 1$ then from equation (3.7)

$$\begin{aligned}
\|Tx\|^2_{\mathcal{H}_2} &= \sum_{i=1}^{\infty} \sigma_i^2 |\langle x, \phi_i \rangle_{\mathcal{H}_1}|^2 \|\psi_i\|^2_{\mathcal{H}_2} \\
&< \epsilon^2 \sum_{i=N_1+1}^{\infty} |\langle x, \phi_i \rangle_{\mathcal{H}_1}|^2 \\
&\leq \epsilon^2.
\end{aligned}$$

It is fairly trivial to prove that N_1 is the smallest number that satisfies the conditions of this theorem because otherwise some element in the span of $\{\psi_1, \ldots, \psi_{N_1}\}$ will be orthogonal to the span of any chosen set of elements $\{\psi'_1, \ldots, \psi'_M\}$ if $M < N_1$.

To prove the second part of the theorem let N_2 denote the number of degrees of freedom of T at level-ϵ. Then to prove that $N_2 \geq N_1$ note that if x is in the unit ball in $\mathcal{H}_1$ then we can write $x = \sum_{i=1}^{\infty} \langle x, \phi_i \rangle_{\mathcal{H}_1} \phi_i + x_r$. Here x_r is the remainder term that is orthogonal to all the ϕ_i. So, from the singular value decomposition theorem (equation (3.7))

$$\|Tx - \sum_{i=1}^{N_1} \sigma_i \langle x, \phi_i \rangle_{\mathcal{H}_1} \psi_i \|_{\mathcal{H}_2} \leq \epsilon.$$

To prove that $N_1 \geq N_2$ assume that $N_2 > N_1$ for a contradiction. Then there exists a set $\{\psi'_i\}_{i=1}^{N_1}$ such that

$$\inf_{a_1, \ldots, a_{N_1}} \|Tx - \sum_{i=1}^{N_1} a_i \psi'_i\|_{\mathcal{H}_2} \leq \epsilon \; \forall x \in \mathcal{H}_1, \|x\|_{\mathcal{H}_1} \leq 1.$$

Because we assume $N_2 > N_1$, there exists a $y \in \text{span}\{\psi_1, \ldots, \psi_{N_2}\}$ which is orthogonal to all ψ'_i. Let $y = \sum_{i=1}^{N_1} b_i \psi_i$. Then $y = Tx$ where $x = \sum_{i=1}^{N_1} \frac{b_i}{\sigma_i} \phi_i$ from the singular value decomposition theorem. We can assume, w.l.o.g. that the b_i are normalised so that $\|x\|_{\mathcal{H}_1} = 1$. Then

$$\begin{aligned}
\inf_{a_1, \ldots, a_N} \|Tx - \sum_{i=1}^{N_2} a_i \psi'_i\|^2_{\mathcal{H}_2} &= \|y\|^2_{\mathcal{H}_2} && (3.8) \\
&= \sum_{i=1}^{N_1} b_i^2 && (3.9) \\
&> \sum_{i=1}^{N_1} \frac{b_i^2}{\sigma_i^2} \epsilon^2 && (3.10) \\
&= \epsilon^2. && (3.11)
\end{aligned}$$

In the above we get equation (3.8) from the fact that y is orthogonal to ψ_i', inequality (3.10) from the definition of N_1 and equation (3.11) from $\|x\|_{\mathcal{H}_1} = 1$.

The inequality (3.11) is the required contradiction. This proves that $N_2 = N_1$. □

Corollary 3.4.1. *Suppose that $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator. Then the number of degrees of freedom at level-ϵ is equal to the number of Hilbert space singular values of T that are greater than or equal to ϵ.*

We can now easily prove that the new definition for generalised singular values agrees with the traditionally accepted definition for singular values on Hilbert spaces.

Theorem 3.5. *Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator. Suppose $\{\epsilon_m\}$ are the generalised singular values of T and $\{\lambda_m\}$ are the possibly repeated eigenvalues written in non-increasing order of the operator $T^*T : \mathcal{H}_1 \to \mathcal{H}_1$. Then*

$$\lambda_m = \epsilon_m^2.$$

Proof. From corollary 3.4.1 we know that $\mathcal{N}(\epsilon)$ is equal to the number of Hilbert space singular values of T that are greater than or equal to ϵ. Therefore if ϵ_m is the m^{th} generalised singular value of T then ϵ_m must also be the m^{th} Hilbert space singular value of T which is $\sqrt{\lambda_m}$. □

In Hilbert spaces we have three characterizations for degrees of freedom: 1) As in Definition 3.3, 2) As in Corollary 3.4.1 in terms of singular values and 3) As in Theorem 3.4 in terms of mutually orthogonal functions in the domain.

The first two characterisations can be generalised to normed spaces. However, the final characterisation is more difficult to generalise. It would be extremely useful to generalise the final characterisation because, for the Hilbert space case, the functions ϕ_i in theorem 3.4 are the best functions to transmit. One could possibly replace the mutual orthogonality by using Riesz's lemma which states (see eg. [48, pp. 78])

Lemma 3.6 (Riesz's lemma)**.** *Let Y and Z be subspaces of a normed space X and suppose that Y is closed and is a proper subspace of Z. Then for all $\theta \in (0, 1)$ there is a $z \in Z$ such that*

$$\|z\| = 1 \text{ and } \|y - z\| > \theta \forall y \in Y. \tag{3.12}$$

The following conjecture is still an open question.

Conjecture 3.1. *Let X and Y be reflexive Banach spaces and let $T : X \to Y$ be compact. Given any $\epsilon > 0$ and some $\theta \in (0, 1)$, there exists a finite set of vectors $\{\phi_i\}_{i=1}^N$ such that for all $x \in X$*

$$\|x\|_X \leq 1 \text{ and } \inf_{a_1,\ldots,a_N} \|x - \sum_{i=1}^{N} a_i \phi_i\|_X \geq \theta, \tag{3.13}$$

implies

$$\|Tx\|_Y < \epsilon.$$

Comparing with theorem 3.4, equation (3.13) is analogous to requiring that x be orthogonal to ϕ_i. The conjecture is definitely not true if the requirement of reflexivity is removed as the next example proves.

Example 3.7. *Consider the sequence space l^1 and let $\{e_n\}$ be the standard Schauder basis for l^1. Then define the operator $T : l^1 \to l^1$ as $e_n \mapsto e_1$ for all n. It was proven in example 3.4 that T is compact.*

Now if $\epsilon < 1$, for any function $x \in l^1$ if $\|x\| = 1$, $\|Tx\| = \|x\| = 1 > \epsilon$. So no finite set of functions can satisfy the conditions in the conjecture.

3.3 Essential Dimension for Compact Operators

The definition for degrees of freedom given in the previous section depends on the arbitrarily chosen number ϵ and therefore this definition does not give a unique number for a given channel. The physical intuition behind choosing this arbitrary small number ϵ is best explained in [37]. In this paper $\epsilon = \sigma^2$ is the noise at the receiver, and Xu and Janaswamy [37] claim that the number of degrees of freedom fundamentally depends on this noise at the receiver.

However, in several important cases the number of degrees of a channel is essentially independent of this arbitrarily chosen positive number [25–28,31,45]. This is due to the fact that in these cases the singular values of the channel operator show a step like behavior. Therefore, for a big range of values ϵ, the number of degrees of freedom at level-ϵ is constant. This leads us to the concept of essential dimensionality[5] which is only a function of the channel and not the arbitrarily chosen positive number ϵ. Some of the properties that one might require from the essential dimension of an operator are:

1. It must be uniquely defined for a given operator Γ.

2. The definition must be applicable to a general class of operators under consideration so that comparisons can be made between different operators. (c.f. the essential dimension definition in [49] that is only applicable to the time-bandwidth problem.)

3. It must in some sense *represent* the number of degrees of freedom at level-ϵ.

[5]Note that the term essential dimension has been used instead of degrees of freedom in several papers. As far as I am aware, this is the first time a distinction has been made between the two terms.

The final requirement in defining the essential dimension needs further clarification. Obviously the essential dimension of Γ can not in general be equal to the number of degrees of freedom at level-ϵ because the latter is a function of ϵ. However, if the singular values of Γ (in decreasing order) change suddenly from being large to small then the 'knee' is at the essential dimension of Γ. The following definition for the essential dimension tries to identify this 'knee' in the set of generalised singular values.

Each level-ϵ defines a unique number of degrees of freedom $N_\Gamma(\epsilon)$ for a given compact operator Γ. So for each positive integer $n \in \mathbb{Z}^+$ we can calculate $E(n) = \mu(\{\epsilon : n = N_\Gamma(\epsilon)\})$. Here $\mu(\cdot)$ is the Lebesgue measure. This function is well defined because of the properties of generalised singular values discussed in theorem 3.2. We can now define the essential dimension of Γ as follows.

Definition 3.4. *The essential dimension of an operator Γ is*

$$\mathrm{EssDim}(\Gamma) = \mathrm{argmax}\{E(n) : n \in \mathbb{Z}^+\}$$

where $E(n)$ is as defined above. If argmax *above is not uniquely defined then one can choose the smallest n of all the n that achieve* $\max\{E(n)\}$ *as the essential dimension.*

In this definition we are simply calculating the maximum range of values of the arbitrarily chosen ϵ over which the number of degrees of freedom of an operator does not change.

This definition uniquely determines the essential dimension of all compact operators. Further, it is equal to the number of degrees of freedom at level-ϵ for the maximum range of ϵ. Choosing this value for the number of degrees of freedom in order to model MIMO communication has the big advantage that it is largely independent of the noise level at the receiver. Further, if for a given noise level the number of degrees of freedom is greater than the essential dimension then one can be sure that even if the noise level varies by a significant amount the number of degrees of freedom is always greater than the essential dimension.

The essential dimension of Γ is the number of generalised singular values of Γ after which the change in two consecutive singular values is at a maximum. However, one can also look at how the generalised singular values are changing gradually. The above definition is a special case of the following definition of essential dimension of order-n with $n = 1$.

Definition 3.5. *Let X, Y be normed spaces and let $\Gamma : X \to Y$ be a compact operator. Let λ_i be the generalised singular values of Γ written in descending order. For n even, define the essential dimension of Γ of order-n to be N if*

$$\lambda_{N-n/2} - \lambda_{N+n/2} \geq \lambda_{M-n/2} - \lambda_{M+n/2}$$

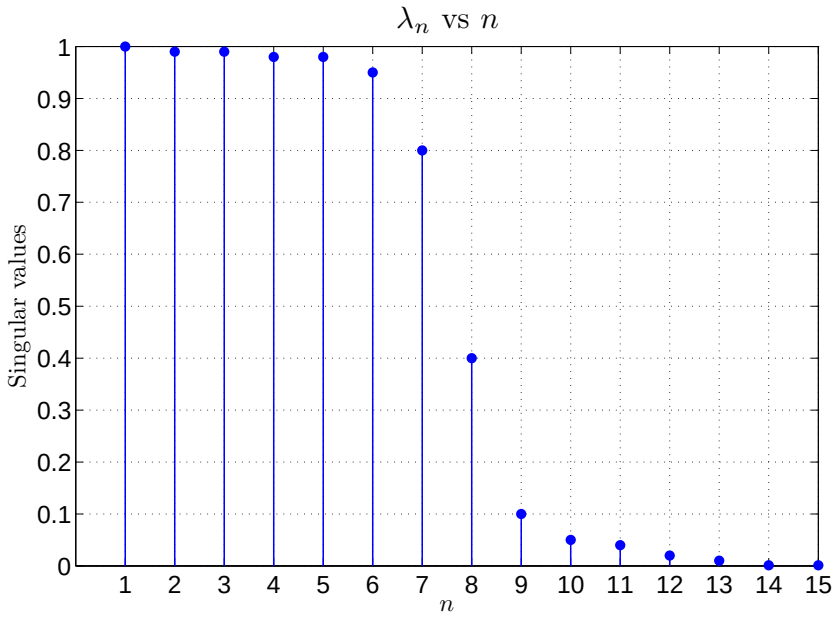

Figure 3.2: Singular values of an Operator

for all $M \neq N$. If there are several N that satisfy the above condition, then choose the smallest N that satisfies the above condition. If n is odd then we require N to satisfy

$$\lambda_{N-(n-1)/2} - \lambda_{N+(n+1)/2} \geq \lambda_{M-(n-1)/2} - \lambda_{M+(n+1)/2}$$

for all $M \neq N$.

A simple example illustrates the concepts of essential dimensionality and degrees of freedom.

Example 3.8. *Figure 3.2 shows the singular values of some operator Γ. For this operator the number of degrees of freedom at level-0.75 is 7 and at level-0.05 is 9. So the number of degrees of freedom do not change much even though the number ϵ changed by a great amount. One can identify the location of the 'knee' in the singular values with the essential dimension of the channel.*

The essential dimension of the channel is 7. This is because for $\epsilon \in (0.8, 0.4)$, $N_\Gamma(\epsilon) = 7$. Therefore $E(7) = 0.4$ which is greater than $E(n)$ for all $n \neq 7$. The essential dimension of order-2 is 8 because $\lambda_7 - \lambda_9 = 0.7$ which is greater than $\lambda_{M-1} - \lambda_{M+1}$ for all $M \neq 8$.

3.4 Discussion

There are several other known definitions for essential dimension and degrees of freedom as discussed in section 1.4. In this section I compare my definition with others that are commonly used in the literature. Because no clear distinction has been made between degrees of freedom and essential dimension in the literature, I compare definitions that depend on some arbitrary constant with my definition for degrees of freedom at level-ϵ and those that determine a unique number for a given channel to that of essential dimension.

3.4.1 Degrees of Freedom

There are several definitions for degrees of freedom at level-ϵ that are used in the literature. The definition of Bucci *et. al.* [25] is a special case of definition 3.2 for degrees of freedom at level-ϵ. They study the case of communication using electromagnetic waves. Suppose the current, both source and induced are elements of $\overline{B}_{Ia,\mathcal{L}^2(S,\mathbb{C}^3)}(0)$ and the electric field is an element of $\mathcal{L}^2(C,\mathbb{C}^3)$. Here, S is a sphere of radius a, C is some compact observation curve and I provides a finite bound on the source and induced currents. Also, suppose that the operator $\Gamma : \mathcal{L}^2(S,\mathbb{C}^3) \to \mathcal{L}^2(C,\mathbb{C}^3)$ determines the electric field for a given source current.

Now let $\mathscr{E} = \{g = \Gamma f : f \in \mathcal{L}^2(B,\mathbb{C}^3), \|f\|_{\mathcal{L}^2(B,\mathbb{C}^3)} \leq Ia\}$ and given any two sets $A, B \subset \mathcal{L}^2(C)$ let

$$\delta(A,B) = \sup_{a\in A}\inf_{b\in B}\|a-b\|_C.$$

Bucci *et. al.* [25] assert that the operator Γ has a finite number of degrees of freedom if the set $\mathscr{E}$ is physically indistinguishable at level ϵ from another set that has a finite number of functions. Bucci *et. al.* [25] defines $\mathscr{E}$ to be physically indistinguishable from a set $\mathscr{L} \subset \mathcal{L}^2(C,\mathbb{C}^3)$ at level ϵ if $\mathscr{L}$ is an ϵ-net of the set $\mathscr{E}$. That is, if

$$\delta(\mathscr{E},\mathscr{L}) \leq \epsilon.$$

One could define the number of elements in the smallest ϵ net of $\mathscr{E}$ to be the number of degrees of freedom of the channel. However, as Bucci *et. al.* [25] point out, this definition is not the easiest one to use because it is extremely difficult to determine the smallest ϵ-net in many important cases. They suggest an alternative definition that is easier to use in practice. The number of degrees of freedom of the operator Γ at level-ϵ is defined to be the dimension of the smallest subspace $\mathscr{L}_n \subset \mathcal{L}^2(C,\mathbb{C}^3)$ such that

$$\delta(\mathscr{E},\mathscr{L}_n) \leq \epsilon.$$

This definition is a special case of my definition for degrees of freedom at level-ϵ.

Miller [26] analyses communication between volumes using scalar waves. The main assumption in [26] is that the transmitter is constrained to be in a volume $V \subset \mathbb{R}^3$ and the receiver is constrained to a volume $W \subset \mathbb{R}^3$. We can assume that V and W are compact. Miller [26] assumes that the transmitter wave function $\psi \in \mathcal{L}^2(V, \mathbb{C})$ generates a wave function $\Gamma\psi \in \mathcal{L}^2(W, \mathbb{C})$ according to

$$(\Gamma\psi)(\mathbf{y}) = \int_V G(\mathbf{x}, \mathbf{y})\psi(\mathbf{x})d\mathbf{x} \quad \forall \mathbf{y} \in W.$$

Here, the operator $\Gamma : \mathcal{L}^2(V, \mathbb{C}) \rightarrow \mathcal{L}^2(W, \mathbb{C})$ and

$$G(\mathbf{x}, \mathbf{y}) = \frac{\exp\{ik|\mathbf{x} - \mathbf{y} + \mathbf{e}_3 r_0|\}}{4\pi|\mathbf{x} - \mathbf{y} + \mathbf{e}_3 r_0|}.$$

Here $\mathbf{x} \in V, \mathbf{y} \in W$, $r_0 \in \mathbb{R}$ is the distance between the centers of the transmitter and receiver and $\mathbf{e}_3$ is a unit vector in the direction of the vector connecting these centers.

Miller [26] asserts that in order to find the best *communication modes*, we need to find the source wave function ψ that maximises the norm of the received wave function ϕ which is given by

$$\begin{aligned} \|\phi\|^2_{\mathcal{L}^2(V,\mathbb{C})} &= \int_W \phi^*(\mathbf{y})\phi(\mathbf{y})d\mathbf{z} \\ &= \int_V \psi^*(\mathbf{x}_1) \int_V \int_W G^*(\mathbf{y}, \mathbf{x}_1)G(\mathbf{y}, \mathbf{x}_2)d\mathbf{y}\psi(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2 \\ &= \int_V \psi^*(\mathbf{x}_1) \int_V K(\mathbf{x}_1, \mathbf{x}_2)\psi(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2. \qquad (3.14) \\ & \qquad (3.15) \end{aligned}$$

Here

$$K(\mathbf{x}_1, \mathbf{x}_2) = \int_W G^*(\mathbf{y}, \mathbf{x}_1)G(\mathbf{y}, \mathbf{x}_2)d\mathbf{y}.$$

It is well known that the function ψ that achieves the maximum value in equation (3.14) is the eigenfunction with the highest eigenvalue of the integral equation

$$\lambda\psi(\mathbf{x}_1) = \int_V K(\mathbf{x}_1, \mathbf{x}_2)\psi(\mathbf{x}_2)d\mathbf{x}_2. \qquad (3.16)$$

Let $\{\lambda_i\}$ and $\{\psi_i\}$ be the sets of eigenvalues and their corresponding normalised eigenfunctions of equation (3.16), respectively, written in descending order of eigenvalues. Because $K(\mathbf{x}_1, \mathbf{x}_2)$ is self-adjoint the eigenfunctions $\{\psi_i\}$ form an orthonormal set [6].

[6] For repeated eigenvalues one can use the Gram-Schmidt method to ensure orthogonality.

The eigenfunctions $\psi_1, \psi_2, \ldots$ are the best source functions to transmit in the sense that out of all functions $\psi \in \mathcal{L}^2(V, \mathbb{C})$ of unit norm, ψ_1 maximises $\|\Gamma\psi\|_{\mathcal{L}^2(W,\mathbb{C})}$. Also, given functions $\{\psi_i\}_1^{n-1}$, the function orthogonal to $\psi_1, \ldots, \psi_{n-1}$ and of unit norm that maximises $\|\Gamma\psi\|_{\mathcal{L}^2(W,\mathbb{C})}$ is ψ_n [26]. Moreover, if $\phi_i = \Gamma\psi_i$ then

$$\begin{aligned} \|\phi_i\|^2_{\mathcal{L}^2(V,\mathbb{C})} &= \int_V \psi_i^*(\mathbf{x}_1) \int_V K(\mathbf{x}_1, \mathbf{x}_2)\psi_i(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \\ &= \lambda_i. \end{aligned}$$

Therefore Miller [26] asserts that the number of *significant* eigenvalues λ_i determine the number of physically distinguishable signals ϕ_i at the receiver. Hence the number of *significant* eigenvalues of equation (3.16) gives the number of modes of communication. Piestun and Miller [34] similarly analyse the case of vector waves. If we can assume that an eigenvalue is significant if it is greater than some pre-specified level ϵ, then we know from corollary 3.4.1 that the number of modes of communication is equal to the number of degrees of freedom of the channel operator at level-$\sqrt{\epsilon}$ using my definition.

Xu *et. al.* [37] build on the results of [34] for vector waves to include the effect of scatterers. They numerically evaluate the Green's function for vector waves in the presence of reflective scatterers in two dimensions using the finite moments method (FMM). Xu *et. al.* [37] then numerically calculate the singular values of the operator defined using the Green's function. Suppose λ_n are the singular values in descending order and also suppose that there is transmit power constraint $P_T > 0$ and receiver noise level $Q > 0$. Xu *et. al.* [37] define the number of degrees of freedom of Γ to be[7]

$$\begin{aligned} N_{dof}(P_T, Q) = \max\{N \quad : \quad & P_n = |a_n|^2|\lambda_n| \geq Q, n = 1, 2, \ldots, N \\ & \text{and } \sum_{n=1}^{N} |a_n|^2 = P_t\}. \end{aligned}$$

The definition of Xu and Ganaswamy [37] is equivalent to definition 3.2 for the special case of operators on Hilbert spaces. This is a direct consequence of corollary 3.4.1.

Newsam and Barakat [41] approach the problem of finding the degrees of freedom of an operator differently. The primary interest of [41] is in the inverse problem: Suppose $g(t) = \int k(s-t)f(t)dt$. Here $g(t), f(t) \in \mathcal{L}^2(\mathbb{R}, \mathbb{C})$. Then given g can we determine f? This inverse problem is ill-posed in the sense that it is extremely sensitive to noise [41]. However, some components of $f(t)$ can be accurately determined and the number of such components is defined to be the number of degrees of freedom of the operator [41].

[7] [37] use Q_n instead of Q in their definition. However they make $Q_n = Q$ subsequently and use that in all their calculations. I do not think using Q_n instead of Q makes sense because then N_{dof} will depend on how Q_n is chosen which in turn depends on ψ_n, the n^{th} eigenfunction of Γ. Therefore, the definition for DOF is not unique anymore.

Let X and Y be Hilbert spaces and let $\Gamma : X \rightarrow Y$ be a compact operator. If $g' \in \Gamma(X)$ is a perturbation of $g \in \Gamma(X)$ such that $\|g - g'\|_Y < \epsilon$ and if $\Gamma f = g$ and $\Gamma f' = g'$, then for some tolerance δ, [41] define the number of degrees of freedom to be

$$N_{dof} = \max\{dim(V) : V \text{ subspace of } X \text{ and } \|P_V(f - f')\| < \delta \; \forall \|g - g'\| < \epsilon\}$$

Here P_V is the projection operator.

Newsam and Barakat [41] prove that if σ_n^2 is the n^{th} eigenvalue of $\Gamma^*\Gamma$ then the number of degrees of freedom is N_1 if $\sigma_{N_1+1} \leq \epsilon/\delta \leq \sigma_{N_1}$. The definition of Newsam and Barakat [41] is equivalent to definition 3.2 for the special case of operators on Hilbert spaces. This is a direct consequence of corollary 3.4.1.

3.4.2 Essential Dimension

Several articles do not give an explicit definition for essential dimension but argue that the "knee" in the singular values of the channel operator corresponds to the essential dimension of the channel [25,26,28,45]. It is therefore difficult to compare the definition of essential dimension with these works. However, the definition 3.4 for essential dimension tries to identify the knee in the singular values and gives a unique number for a given channel. Further, one can use the definition 3.5 for essential dimension of order-n to ensure that outliers do not effect the identification of the knee. In summary, the definition of essential dimension enables one to uniquely quantify the results of [25,26,28,45].

A step like behavior of eigenvalues is also found in [31,38,39]. The approach of Kennedy *et. al.* [31] is substantially different from those discussed earlier in this section wherein the transmitting volume was assumed to be constrained. In contrast, Kennedy *et. al.* [31] consider the possible wavefields within a volume constrained to a radius R provided all sources are in the far field (outside a ball of radius R_S). Most of their results are for wavefields in two dimensions. However, they emphasise that their results can be extended to three dimensions. The channel model used in these papers is considerably different from that considered in this thesis. However, their definition of degrees of freedom is comparable to my definitions of essential dimension because they evaluate the number of significant eigenvalues of a specific operator to calculate the degrees of freedom.

Let $B_R(0) \subset \mathbb{R}^2$ be an open ball of radius R centered at the origin. A wave field in $B_R(0)$ is a function $F : B_R(0) \rightarrow \mathbb{C}$ that satisfies the scalar Helmholtz equation [31].

$$\nabla^2 F(\mathbf{x}) + k^2 F(\mathbf{x}) = 0, \;\; x \in B_R(0).$$

Here $k = 2\pi/\lambda$ is the wave number and λ is the wavelength. Kennedy *et. al.* [31] use the following model: if all the sources are in the far field, then $F(x)$ can be written as

$$F(\mathbf{x}) = \int_0^{2\pi} a(\phi) e^{ik\langle \mathbf{x}, \hat{\mathbf{y}}(\phi)\rangle} d\phi \tag{3.17}$$

using polar coordinates ($\mathbf{x} = [\|x\|, \arg(\mathbf{x})] = [r_x, \phi(x)]$). Here, $\hat{\mathbf{y}} = (1, \phi)$ is a unit vector with azimuth angle ϕ. Physically, the wave field $F(\mathbf{x})$ can be interpreted as the superposition of planar waves arriving from each azimuth angle ϕ with amplitude $a(\phi)$. This model is suitable for far field sources because the waves arriving from these sources can be approximated by planar waves.

Using an orthonormal series expansion for the planar waves, equation (3.17) can be rewritten as [31]

$$F(\mathbf{x}) = \sum_{n=-\infty}^{\infty} \sqrt{2\pi \mathcal{J}(R)} \alpha_n \Phi_n(\mathbf{x}). \tag{3.18}$$

Here, the set $\{\Phi_n(\mathbf{x})\}$ are a set of functions orthonormal over the unit circle and can be written as

$$\Phi_n(\mathbf{x}) = i^n \frac{J_n(k\|\mathbf{x}\|)}{\sqrt{\mathcal{J}_n(R)}} \frac{\exp\{in\phi(x)\}}{\sqrt{2\pi}},$$

where $J_n(\cdot)$ is the n^{th} order Bessel function of the first kind and

$$\mathcal{J}_n(R) = \int_0^R J_n^2(r) r dr$$

is a normalization term. Also, α_n are the Fourier expansion coefficients of $a(\phi)$:

$$\alpha_n = \int_0^{2\pi} a(\phi) e^{-in\phi} d\phi.$$

Even though the series representation of the wave field requires infinitely many terms, it is shown that $F(\mathbf{x})$ can be approximated with arbitrary accuracy using only finitely many terms. Specifically if

$$F_N(\mathbf{x}) = \sum_{n=-N}^{N} \sqrt{2\pi \mathcal{J}(R)} \alpha_n \Phi_n(\mathbf{x})$$

and

$$A = \int_0^{2\pi} |a(\phi)| d\phi$$

then the relative truncation error for the wave field within $B_R(0)$ and for a truncated series of length $2N + 1$ can be defined as

$$\epsilon_N(R) = \frac{1}{\pi R^2} \int_{B_R(0)} \frac{|F(\mathbf{x}) - F_N(\mathbf{x})|}{A} d\mathbf{x}.$$

Kennedy *et. al.* prove that [31]

$$\epsilon_N < \eta e^{-\Delta}$$

if

$$N = \lceil e\pi R/\lambda \rceil + \Delta.$$

Because the normalised truncation error decreases exponentially with increasing N for all $N > \lceil e\pi R/\lambda \rceil$, [31] asserts that $\lceil e\pi R/\lambda \rceil$ is the number of degrees of freedom of the multipath field. A similar definition for degrees of freedom is also found in Dickins and Hanlen [40]. Let X be some normed space and Y be a Hilbert spaces. Also, let $\Gamma : X \to Y$ be some operator. Then if for some choice of N_0 and a set of functions $\{\phi_i\}_{i=1}^{\infty}$ and for any $x \in X$ with $\|x\|_X \leq 1$[8]

$$\left\| \Gamma x - \sum_{n=1}^{N_0} \langle \Gamma x, \phi_n \rangle_Y \phi_n \right\|_Y \leq \epsilon < \infty$$

and if $\forall n > N_0$

$$\left\| \Gamma x - \sum_{n=1}^{n} \langle \Gamma x, \phi_n \rangle_Y \phi_n \right\|_Y \leq \epsilon e^{-\alpha(n-N_0)}$$

then Dickins and Hanlen [40] assert that N_0 is the number of degrees of freedom of the operator Γ.

However, this definition does not give a unique number N_0 for a given channel because if the above condition is satisfied by some finite N then it is also satisfied by all $N \in \mathbb{Z}^+$. So it does not satisfy all the conditions that the essential dimension of an operator is required to satisfy as explained in the previous section. Another commonly used definition is that of Landau and Pollak [27][9]. This definition however is only applicable to time-bandwidth problems and cannot be used for arbitrary SWCs.

3.5 Chapter Conclusion

In this chapter I prove that for a compact operator Γ on some normed space, for any given number $\epsilon > 0$ there is a unique number $N_{dof}(\epsilon)$ which is the number of degrees of freedom of the operator Γ at level ϵ. Physically we can interpret this number as the maximum number of linearly independent functions a receiver that has noise of variance

[8] The definition given here is different from that in [40], which does not assume anything about $\|x\|_X$ but using that definition the spaces that have degrees of freedom N_0 would actually be finite dimensional.

[9] Also see [49].

$\sigma^2 = \epsilon$ can measure. This definition has been used previously for specific operators on Hilbert spaces (see eg. [25]). However, as far as I am aware, this is the first definition that is applicable to an arbitrary compact operator on normed spaces.

I prove that one can use this definition for degrees of freedom to define generalised singular values which are generalisations of commonly accepted singular values defined for Hilbert space operators. A direct consequence of these definitions is that the number of degrees of freedom of compact operators can be characterised in terms of their generalised singular values. The advantage of this characterisation is that it lends itself to numerical computations as shown in the next chapter.

In this chapter, I also distinguish between the terms "degrees of freedom" and "essential dimension" though they have been used interchangeably in the literature. In the situation where the singular values of an operator change rapidly from being large to small, the position of the "knee" in the singular values is unique for a given channel. In such channels, the number of degrees of freedom at level ϵ depends very little on the actual value of ϵ. I define the essential dimension of an operator as the smallest number of singular values after which the difference between two consecutive singular values is at a maximum.

The main contributions of this chapter are:

1. Novel definition of degrees of freedom at level-ϵ for arbitrary compact operators on normed spaces. I prove that the definition gives a unique number for a given level ϵ for any compact operator.

2. Novel definition for normed space singular values. I give a physical interpretation for this definition and show that it is a generalisation of singular values of compact operators on Hilbert spaces.

3. Novel definitions for essential dimension and essential dimension of order-n for any normed space compact operator.

Chapter 4

Computation of Generalised Singular Values

In this chapter I develop techniques for the calculation of generalised singular values of normed space operators. In some special cases it is possible to analytically calculate bounds on the generalised singular values using perturbation theory. This is discussed in the following section. However, in a majority of cases analytical computations are nigh on impossible and one has to resort to numerical techniques. Several numerical techniques are known for the computation of singular values of integral operators. The one that can be most easily generalised to normed space operators is the Galerkin method (see eg. [32]). The essential idea behind this method is to use some complete Hilbert basis and use finite dimensional approximations of the integral operator. I will prove in section 4.2 that in the special case of normed spaces with complete Schauder bases it is possible to use the same techniques to calculate the generalised singular values. The results of the numerical computation of generalised singular values in some special cases is given in section 4.3

4.1 Perturbation Theory Applied to Scalar Wave Communication

In several cases it is difficult to evaluate the singular values and the left and right singular functions of the operator Γ for a given spatial waveform channel $(\widetilde{X}_T, \widetilde{Y}_R, \Gamma)$. However, in some cases it is possible to calculate the singular values and functions for another operator $\Gamma' : \widetilde{X}_T \to \widetilde{Y}_R$ that closely approximates Γ. Therefore, one can use perturbation theory to show that the the singular values of Γ can in some sense be approximated by the singular values of Γ'.

We can use the following perturbation theory result [47, ch V, th. 4.10]: If Ψ :

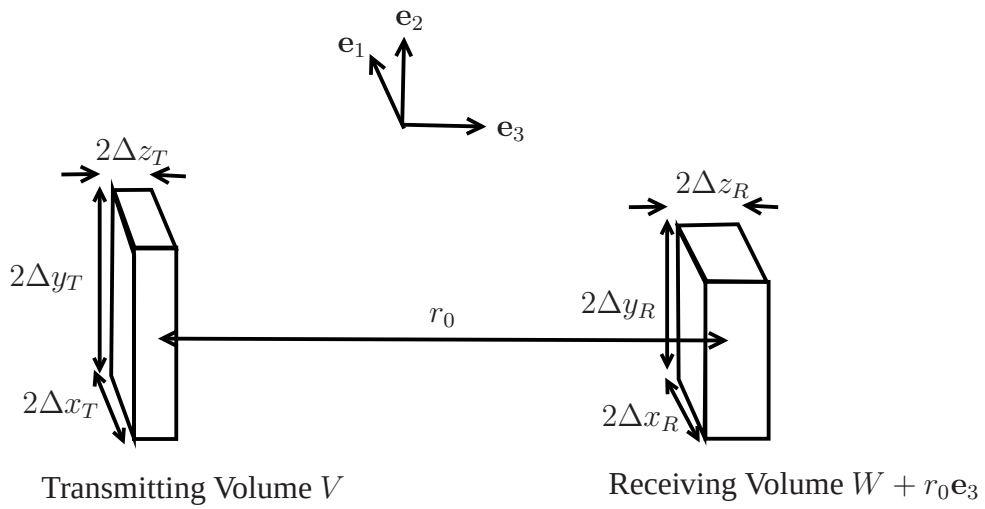

Figure 4.1: Communication Using Scalar Waves Between Rectangular Prisms

$\mathcal{L}^2(T) \to \mathcal{L}^2(T)$ and $\Psi_\Delta : \mathcal{L}^2(T) \to \mathcal{L}^2(T)$ are self-adjoint, compact then

$$\operatorname{dist}(\sigma(\Psi), \sigma(\Psi + \Psi_\Delta)) \leq \|\Psi_\Delta\| \tag{4.1}$$

Here $\sigma(\Psi) = \{\sigma_i(\Psi)\}_{i=1}^N$ is the spectrum of Ψ indexed in decreasing order and $N \in \mathbb{Z}^+ \cup \{\infty\}$ and

$$\operatorname{dist}(\sigma(\Psi), \sigma(\Psi + \Psi_\Delta)) = \sup_i \inf_j |\sigma_i(\Psi) - \sigma_j(\Psi + \Psi_\Delta)|.$$

Also if we can write

$$(\Psi_\Delta \phi)(y) = \int_V K_\Delta(x, y)\phi(y)dy \tag{4.2}$$

then

$$\|\Psi_\Delta\| \leq \sqrt{\int_{V\times V} |K_\Delta(x, y)|^2 dxdy} \tag{4.3}$$

Now consider communication between rectangular prisms using scalar waves. This problem was studied by Miller [26]. I use slightly different notation to that used in [26]. Let $V, W \subset \mathbb{R}^3$ be two compact and measurable sets, let $0 \in V, W$ and let $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ be unit vectors in $\mathbb{R}^3$. Also, let $r_0 > 0$ such that $V \cap (W + r_0\mathbf{e}_3) = \emptyset$ (see figure 4.1). We can define the linear operator $\Gamma : \mathcal{L}^2(V) \to \mathcal{L}^2(W)$ as (c.f. example 2.1)

$$(\Gamma f)(\mathbf{z}) = \int_V G(\mathbf{x}, \mathbf{z}) f(\mathbf{x}) d\mathbf{x} \;\; \forall \mathbf{x} \in W$$

where

$$G(\mathbf{x},\mathbf{z}) = \frac{\exp\{jk|\mathbf{z}+r_0\mathbf{e}_3-\mathbf{x}|\}}{4\pi|\mathbf{z}+r_0\mathbf{e}_3-\mathbf{x}|}.$$

Because $G|_{V\times W}$ is continuous, Γ is compact and $(\mathcal{L}^2(V),\mathcal{L}^2(W),\Gamma)$ is an SWC. Now we can define the operator $\Psi : \mathcal{L}^2(V) \to \mathcal{L}^2(V)$ as

$$(\Psi f)(\mathbf{y}) = \int_V k(\mathbf{x},\mathbf{y}) f(\mathbf{x}) d\mathbf{x}.$$

Here

$$k(\mathbf{x},\mathbf{y}) = \int_W G^*(\mathbf{z},\mathbf{x}) G(\mathbf{z},\mathbf{y}) d\mathbf{z}.$$

We wish to calculate the eigenvalues of the operator Ψ. Miller [26] uses the paraxial approximation to find these eigenvalues. The paraxial approximation consists of using Taylor series expansions of the terms under the modulus signs in the denominator and in the exponent of the Green's function $G(\mathbf{x},\mathbf{y})$, respectively. Note that this approximation is commonly used in antenna theory and is also used in calculating the singular values in other papers on waveform channels [25,33].

The approximating operator $\Psi_T : \mathcal{L}^2(V) \to \mathcal{L}^2(V)$ is defined as

$$(\Psi_T f)(\mathbf{y}) = \int_V k_T(\mathbf{x},\mathbf{y}) f(\mathbf{x}) d\mathbf{x}.$$

Here

$$k_T(\mathbf{x},\mathbf{y}) = \int_W G_T^*(\mathbf{z},\mathbf{x}) G_T(\mathbf{z},\mathbf{y}) dz$$

and

$$G_T(\mathbf{z},\mathbf{x}) = \frac{\exp\{jk f_{T_2}(\mathbf{z},\mathbf{x})\}}{f_{T_0}(\mathbf{z},\mathbf{x})},$$

where f_{T_2} and f_{T_0} are the second and zeroth order Taylor Series expansions of $|\mathbf{z}+r_0\mathbf{e}_3-\mathbf{x}|$ with respect to $\mathbf{x}$ and $\mathbf{z}$ about the respective origins.

We can write $\Psi = \Psi_T + \Psi_\Delta$ where $\Psi_\Delta : \mathcal{L}^2(V) \to \mathcal{L}^2(V)$ is defined by

$$\begin{aligned}
[\Psi_\Delta f](\mathbf{y}) &= [(\Psi - \Psi_T)(f)](\mathbf{y}) \\
&= \int_V (K(\mathbf{x},\mathbf{y}) - k_T(\mathbf{x},\mathbf{y})) f(\mathbf{x}) d\mathbf{x} \\
&= \int_V K_\Delta(\mathbf{x},\mathbf{y}) f(\mathbf{x}) d\mathbf{x}.
\end{aligned}$$

Let x_i be the i^{th} Cartesian coordinate of $\mathbf{x}$. If [1]

$$\begin{aligned} b_i &= \sup_{\mathbf{x}\in V-W} |x_i|\,,\ i=1,2,3, \\ b &= \max_i b_i, \\ l &= \inf_{\mathbf{x}\in V-W} |r_0\mathbf{e}_3+\mathbf{x}| \end{aligned}$$

then it can be shown that (see appendix A.1)

$$|K_\Delta(x,y)| \le \frac{b\mu(W)}{l^3}\left[e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b}{l}c_1 + c_3 + c_3 e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1\right].$$

Here, c_1, c_3 are constants and $\mu(W)$ is the Lebesgue measure of W. Therefore,

$$\begin{aligned} \|\Psi_\Delta\| &\le \left[\int_V\int_V |K_\Delta(x,y)|^2 dxdy\right]^{1/2} \\ &\le \frac{b\mu(W)\mu(V)}{l^3}\times \\ &\quad \left[e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b}{l}c_1 + c_3 + c_3 e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1\right] \end{aligned} \tag{4.4}$$

Equation (4.4) provides a bound for $\|\Psi_\Delta\|$. All the terms in the square brackets in this equation, except for the constant c_3 can be made arbitrarily small as $b/l \to 0$. Therefore, for small b/l, $|k_\Delta(x,y)|$ is proportional to $b\mu(W)\mu(V)/l^3$. It is very interesting to note that the norm of $\|\Psi_\Delta\|$ does not actually go to zero if b/l goes to zero. We not only need b to be small compared to l but also to be small compared to λ for this approximation to be valid.

Miller [26] proves that the eigenvalues of the operator Ψ_T show a step like behavior. The eigenvalues, indexed in decreasing order of magnitude are close to 1 until they reach a critical value N_{max} and then decrease rapidly to zero. Here,

$$N_{max} = \frac{\mu(W)\mu(V)}{r^2}\frac{1}{\lambda^2\Delta z_T\Delta z_R}$$

and $\Delta z_T, \Delta z_R$ and r are as shown in figure 4.1. Therefore, if the right hand side of equation (4.4) is small, then the eigenvalues of Ψ show a similar behavior and if this is the case, then N_{max} is the essential dimension of the channel.

[1] Specifically, referring to figure 4.1, $b_1 = 2\Delta x_T + 2\Delta x_R$, $b_2 = 2\Delta y_T + 2\Delta y_R$ and $b_3 = 2\Delta z_T + 2\Delta z_R$ and $l = r_0 + 2\Delta z_T + 2\Delta z_R$.

4.2 Computing DOF for Compact Operators

In chapter 3 it was proved that in order to calculate the degrees of freedom of a compact operator on a normed space, one needs to calculate its generalised singular values. However, no known method exists for computing these singular values. In this section, I will develop a numerical method based on finite dimensional approximations that could be used to calculate the generalised singular values.

Suppose X and Y are normed spaces and $T : X \to Y$ is a compact operator. Also suppose that X has a complete Schauder basis $\{\psi_1, \psi_2, \ldots\}$. Let $S_n = \text{span}\{\psi_1, \ldots, \psi_n\}$. Then we can define the operator $T_n = T|_{S_n} : S_n \to Y$. Let ϵ_m and $\epsilon_{m,n}$ be the m^{th} singular values of T and T_n respectively. Then, I will prove in theorem 4.1 that

1. As $n \to \infty$, $\epsilon_{m,n} \to \epsilon_m$.

2. For all n, if $\epsilon_{m,n}$ exists, then it is a lower bound for ϵ_m.

The crux of the argument used to prove the theorem is as follows. Assume $\epsilon > 0$ is given and let $\mathcal{N}(\epsilon)$ denote the number of degrees of freedom at level ϵ for the operator T. This is the case if and only if there exist functions $\{\phi_1, \ldots, \phi_{\mathcal{N}(\epsilon)}\} \subset Y$ such that for all $\psi \in X$, $T\psi$ can be approximated to level-ϵ by a linear combination of the ϕ_i and further, no set of functions, $\{\phi'_1, \ldots, \phi'_N\} \subset Y$ can approximate all $\psi \in X$ if $N < \mathcal{N}(\epsilon)$. Equivalently, there is a function in the unit ball in X whose image under T can be approximated by a function in the $\text{span}\{\phi_1, \ldots, \phi_{\mathcal{N}(\epsilon)}\}$ that cannot be approximated by $\text{span}\{\phi'_1, \ldots, \phi'_N\}$.

So we take the inverse image of an ϵ-net of points in $\text{span}\{\phi_1, \ldots, \phi_{\mathcal{N}(\epsilon)}\}$ and choose n large enough so that all the inverse images are close to S_n. We can do this because the ψ_i form a complete Schauder basis for X. I then show that there exists a function in S_n such that its image under T cannot be approximated by a linear combination of $\phi'_1, \ldots, \phi'_N$ for $N < \mathcal{N}(\epsilon)$. This will prove that the number of degrees of freedom at level-ϵ of T_n approaches that of T and consequently so do the singular values. The details are as follows.

Theorem 4.1. *Suppose X and Y are normed spaces and $T : X \to Y$ is a compact operator. Suppose that X has a complete Schauder basis $\{\psi_1, \psi_2, \ldots\}$ and let $S_n = \text{span}\{\psi_1, \ldots, \psi_n\}$. Let $T_n = T|_{S_n} : S_n \to Y$. If ϵ_m, the m^{th} singular value of T exists then for n large enough, $\epsilon_{m,n}$, the m^{th} singular value of T_n will exist and*

$$\lim_{n\to\infty} \epsilon_{m,n} = \epsilon_m.$$

Furthermore, if $\epsilon_{m,n}$ exists then it is a lower bound for ϵ_m.

Proof. I will prove this theorem in two parts. In part a) I will prove that if $\epsilon_{m,n}$ exists for some $n = N \in \mathbb{N}$ then $\epsilon_{m,n}$ exists for all $n > N$ and is a non-decreasing sequence

that is bounded from above by ϵ_m. In part b) I prove using a contradiction argument that $\epsilon_{m,n}$ exists for some $n \in \mathbb{N}$ and that $\epsilon_{m,n}$ must converge to ϵ_m.

I will use the following notation in the proof:

$$\text{span}_\epsilon\{\phi_1, \ldots, \phi_N\} = \{\phi \in Y : \inf_{a_1,\ldots,a_N} \|\phi - \textstyle\sum_{i=1}^N a_i\phi_i\| < \epsilon\} \quad \text{and}$$
$$B_r = \{\psi \in X : \|\psi\|_X \leq r\}.$$

Part a Let T and T_n be as described in the theorem and let $\mathcal{N}(\epsilon)$ and $\mathcal{N}_n(\epsilon)$ be the numbers of degrees of freedom at level ϵ of T and T_n, respectively. Assume that ϵ_{m,n_1} exists and let $n_2 > n_1$.

Then for all $\{\phi_1, \ldots, \phi_{\mathcal{N}_{n_1}(\epsilon)-1}\} \subset Y$ there is a $\psi \in S_{n_1} \cap B_1$ such that

$$T\psi_{n_1} = T\psi \notin \text{span}_\epsilon\{\phi_1, \ldots, \phi_{\mathcal{N}_{n_1}(\epsilon)-1}\}.$$

Because $S_{n_1} \subset S_{n_2}$, for all $\{\phi_1, \ldots, \phi_{\mathcal{N}_{n_1}(\epsilon)-1}\} \subset Y$ we have $\psi \in S_{n_2} \cap B_1$ and

$$T_{n_2}\psi = T\psi \notin \text{span}_\epsilon\{\phi_1, \ldots, \phi_{\mathcal{N}_{n_1}(\epsilon)-1}\}.$$

Therefore,

$$\mathcal{N}_{n_2}(\epsilon) \geq \mathcal{N}_{n_1}(\epsilon) \ \ \forall \epsilon > 0. \tag{4.5}$$

Because

$$\inf_{\epsilon < \epsilon_{m,n_1}} \mathcal{N}_{n_1}(\epsilon) \geq m \tag{4.6}$$

we have $\mathcal{N}_{n_2}(\epsilon) \geq \mathcal{N}_{n_1}(\epsilon) \geq m$ for $\epsilon < \epsilon_{m,n_1}$. Hence ϵ_{m,n_2} must exist.

From the definition of generalised singular values, we have equation (4.6) and

$$\sup_{\epsilon > \epsilon_{m,n_2}} \mathcal{N}_{n_2}(\epsilon) \leq m - 1$$

If $\epsilon_{m,n_1} > \epsilon_{m,n_2}$ then there exists an ϵ_1 such that $\epsilon_{m,n_1} > \epsilon_1 > \epsilon_{m,n_2}$. Therefore,

$$\mathcal{N}_{n_1}(\epsilon_1) \geq m > m - 1 \geq \mathcal{N}_{n_2}(\epsilon_1).$$

This contradicts equation (4.5). Therefore $\epsilon_{m,n_1} \leq \epsilon_{m,n_2}$.

The same argument used above shows that if $\epsilon_{m,n}$ exists then $\epsilon_m \geq \epsilon_{m,n}$. Therefore, if $\epsilon_{m,n}$ exists for $n = n_1 \in \mathbb{N}$ then $\epsilon_{m,n}$ is a non-decreasing sequence in $n \geq n_1$ that is bounded from above by ϵ_m.

Part b By part a, if $\epsilon_{m,n}$ exists for $n \geq n_1$ then because $\epsilon_{m,n}$ is a bounded monotonic sequence in n it must converge to some $\epsilon'_m \leq \epsilon_m$.

Now there are two situations to consider. Firstly, $\epsilon_{m,n}$ might not exist for any $n \in \mathbb{N}$. Secondly, $\epsilon_{m,n}$ might exist for some n but the limit ϵ'_m might be strictly less than ϵ_m. We consider the two situations separately and arrive at the same set of equations. We then show a contradiction to this set of equations.

Situation 1: Assume, to arrive at a contradiction, that $\epsilon_{m,n}$ does not exist for any $n \in \mathbb{N}$. Then

$$\mathcal{N}_n(\epsilon) \leq m-1 \ \ \forall n \in \mathbb{N} \text{ and } \forall \epsilon > 0. \tag{4.7}$$

From the definition of degrees of freedom, there exist constants $\alpha < \beta < \epsilon_m$ such that

$$\mathcal{N}_n(\alpha) \leq m-1 \quad \forall n \in \mathbb{N} \text{ and} \tag{4.8}$$
$$\mathcal{N}(\beta) \geq m. \tag{4.9}$$

Situation 2: Assume, to arrive at a contradiction, that $\epsilon'_m < \epsilon_m$. From the definition of generalised singular values we know,

$$\sup_{\epsilon > \epsilon_{m,n}} \mathcal{N}_n(\epsilon) \leq m-1 \quad \text{and}$$
$$\inf_{\epsilon < \epsilon_m} \mathcal{N}(\epsilon) \geq m.$$

Because $\epsilon_{m,n} \leq \epsilon'_m$, we know that there exist numbers α and β, $\epsilon'_m < \alpha < \beta < \epsilon_m$ such that

$$\mathcal{N}_n(\alpha) \leq m-1 \quad \forall n \in \mathbb{N} \text{ and} \tag{4.10}$$
$$\mathcal{N}(\beta) \geq m. \tag{4.11}$$

Therefore, in both situation 1 and situation 2, we need to prove that equations (4.11) and (4.10) cannot be simultaneously true.

Because T is compact, TB_1 is totally bounded [48, ch. 8]. Therefore, for all ϵ TB_1 has a finite ϵ-net. Hence there exists a set of vectors $\{\xi_1, \ldots, \xi_P\} \subset B_1$ such that for all $\phi \in TB_1$, there exists a p, $1 \leq p \leq P$ with

$$\|\phi - T\xi_p\|_Y < \frac{\beta - \alpha}{2}. \tag{4.12}$$

Now, because $\{\psi_1, \psi_2, \ldots\}$ is a complete Schauder basis for X and because $P < \infty$, there exists a number N such that for all $n > N$ and for all p, $1 \leq p \leq P$, there exists a $\xi_{p,n} \in S_n \cap B_1$ such that

$$\|\xi_p - \xi_{p,n}\|_X < \frac{\beta - \alpha}{2\|T\|}. \tag{4.13}$$

Therefore, for all $\phi \in TB_1$ and for all $n > N$ there exists a $p, 1 \leq p \leq P$ and a $\xi_{p,n} \in S_n \cap B_1$ such that

$$\begin{aligned}
\|T\xi_{p,n} - \phi\|_Y &= \|T\xi_{p,n} - T\xi_p + T\xi_p - \phi\|_Y \\
&\leq \|T\xi_{p,n} - T\xi_p\|_Y + \|T\xi_p - \phi\|_Y \text{ (Triangle Inequality)} \\
&< \|T(\xi_{p,n} - \xi_p)\|_Y + \frac{\beta - \alpha}{2} \text{ (equation (4.12))} \\
&< \|T\|\frac{\beta - \alpha}{2\|T\|} + \frac{\beta - \alpha}{2} \text{ (equation (4.13))} \\
&= \beta - \alpha.
\end{aligned} \tag{4.14}$$

From equation (4.10) and the definition of degrees of freedom, we know that for all $n \in \mathbb{N}$ there exist vectors $\{\phi_{1,n}, \ldots, \phi_{m-1,n}\} \subset Y$ such that

$$\xi \in \text{span}_\alpha\{\phi_{1,n}, \ldots, \phi_{m-1,n}\} \ \forall \xi \in T(B_1 \cap S_n). \tag{4.15}$$

But, from the definition of degrees of freedom and equation (4.11) we know that for all n and all vectors $\{\phi_{1,n}, \ldots, \phi_{m-1,n}\}$ there exists a vector $\phi \in TB_1$ such that

$$\phi \notin \text{span}_\beta\{\phi_{1,n}, \ldots, \phi_{m-1,n}\}.$$

From equation (4.14) we know that for all $n > N$ there exists a $\xi_{p,n} \in S_n \cap B_1$ such that

$$\|\phi - T\xi_{p,n}\| < \beta - \alpha.$$

Therefore, for all $n > N$ there exists a $\xi_{p,n} \in S_n \cap B_1$ such that

$$T\xi_{p,n} \notin \text{span}_\alpha\{\phi_{1,n}, \ldots, \phi_{m-1,n}\}. \tag{4.16}$$

This directly contradicts equation (4.15). Therefore, if ϵ_m exists then $\epsilon_{m,n}$ exists for n large enough and

$$\lim_{n\to\infty} \epsilon_{m,n} = \epsilon_m.$$

□

The theorem shows that if the domain of the operator has some complete Schauder basis then we can calculate the generalised singular values of the operator restricted to finite dimensional subspaces and as the subspaces get bigger we will approach the generalised singular values of the original operator. Moreover, the theorem also proves that the generalised singular values of the finite dimensional operators provide lower bounds for the original generalised singular values. We, however, need a practical method of calculating the generalise singular values of linear operators defined on finite dimensional normed spaces.

Let X, Y be two finite dimensional Banach spaces and let $T : X \to Y$ be a linear operator. Suppose $\epsilon_1, \ldots, \epsilon_n$ are the generalised singular values of T. Denote $B_1 = \{x \in X : \|x\|_X \leq 1\}$ and let

$$\epsilon_1' = \sup_{x \in B_1} \|Tx\|_Y.$$

Then for all $\epsilon > \epsilon_1'$

$$\epsilon > \sup_{x \in B_1} \|Tx\|_Y$$

and for all $\epsilon < \epsilon_1'$, there exists an $x \in B_1$ such that $\|Tx\|_Y > \epsilon$. Therefore $\epsilon_1' = \epsilon_1$, the first generalised singular value of T. Suppose for ease of argument that $\lim_{\epsilon \to \epsilon_p^-} \mathcal{N}(\epsilon) = p$ (i.e. ϵ_p is not a repeated singular value). Therefore, for all $\epsilon \in [\epsilon_{p+1}, \epsilon_p)$, $\mathcal{N}(\epsilon) = p$. Now for each $\epsilon \in [\epsilon_{p+1}, \epsilon_p)$ there exists a $\{\psi_i\}_{i=1}^p \in Y$ such that

$$\sup_{x \in B_1} \inf_{\{a_i\}_{i=1}^p} \|Tx - \sum_{i=1}^p a_i \psi_i\|_Y \leq \epsilon.$$

Let $\Psi_{p,\epsilon}$ denote the set of all sets $\{\psi_i : \|\psi_i\|_Y \leq 1\}_{i=1}^p$ that satisfy the above equation for a given $\epsilon \in [\epsilon_{p+1}, \epsilon_p)$ and let

$$\Psi_p = \bigcup_{\epsilon \in [\epsilon_{p+1}, \epsilon_p)} \Psi_{p,\epsilon}.$$

Suppose the sets $\Psi_1, \ldots, \Psi_p$ have been chosen as above. Then

$$\inf_{\{\psi_i\}_{i=1}^p \in \Psi_p} \sup_{x \in B_1} \inf_{\{a_i\}_{i=1}^p} \left\{ \left\| x - \sum_{i=1}^p a_i \psi_i \right\| \right\} = \epsilon_{p+1}.$$

To prove this let ϵ_{p+1}' denote the left hand side of the above equation. We show by contradiction that $\epsilon_{p+1} \leq \epsilon_{p+1}'$. Otherwise, let $\epsilon \in (\epsilon_{p+1}', \epsilon_{p+1})$. Then because $\epsilon > \epsilon_{p+1}'$, we know that there exists a $\{\psi_i\}_{i=1}^p$ such that

$$\sup_{x \in B_1} \inf_{\{a_i\}_{i=1}^p} \left\{ \left\| x - \sum_{i=1}^p a_i \psi_i \right\| \right\} < \epsilon.$$

If this is the case then $\mathcal{N}(\epsilon) \leq p$. But we know from the definition of generalised singular values that if $\epsilon < \epsilon_{p+1}$ then $\mathcal{N}(\epsilon) \geq p+1$. This is a contradiction and proves that $\epsilon_{p+1} \leq \epsilon'_{p+1}$.

To prove the converse, suppose, to arrive at a contradiction, that $\epsilon'_{p+1} > \epsilon_{p+1}$. Let $\epsilon \in (\epsilon_{p+1}, \epsilon'_{p+1})$. Then, $\mathcal{N}(\epsilon) = p$. Therefore there exists a set of functions $\{\phi_i\}_{i=1}^p$ such that

$$\sup_{x \in B_1} \inf_{\{a_i\}_{i=1}^p} \left\{ \left\| x - \sum_{i=1}^p a_i \phi_i \right\| \right\} < \epsilon.$$

Therefore $\{\phi_i\}_{i=1}^p \in \Psi_{p,\epsilon} \subset \Psi_p$. Hence,

$$\inf_{\{\psi_i\}_{i=1}^p \in \Psi_p} \sup_{x \in B_1} \inf_{\{a_i\}_{i=1}^p} \left\{ \left\| x - \sum_{i=1}^p a_i \psi_i \right\| \right\} < \epsilon.$$

By definition, the left hand side of the above equation is ϵ'_{p+1} and this contradicts the assumption that $\epsilon \in (\epsilon_{p+1}, \epsilon'_{p+1})$. The following theorem summarises the above result.

Theorem 4.2. *Let X, Y be two finite dimensional Banach spaces and let $T : X \to Y$ be a linear operator. Also let B_1 be the closed unit ball in X and suppose Ψ_p is defined as explained above. Then,*

$$\sup_{x \in B_1} \|Tx\|_Y = \epsilon_1$$

and if ϵ_p is not a repeated generalised singular value then

$$\inf_{\{\psi_i\}_{i=1}^p \in \Psi_p} \sup_{x \in B_1} \inf_{\{a_i\}_{i=1}^p} \left\{ \left\| x - \sum_{i=1}^p a_i \psi_i \right\| \right\} = \epsilon_{p+1}.$$

The above theorem characterises the singular values in terms of a maximisation problem over a finite dimensional domain and one can use several well-know maximisation methods to calculate the generalised singular values. However, it is difficult to check whether a given set of functions $\{\psi_i\}_{i=1}^p$ is an element of Ψ_p. I therefore use the following algorithm to calculate bounds on the generalised singular values.

Suppose $X, Y, T : X \to Y$ and $\epsilon_1, \ldots, \epsilon_n$ are as defined above. Let

$$\epsilon'_1 = \sup_{x \in B_1} \|Tx\|_Y.$$

Because $B_1 \subset X$ is a compact set and $\|\cdot\|_Y$ and T are continuous, there exists an $x_1 \in B_1$ such that $\|x_1\|_Y = \epsilon'_1$. Choose $\psi_1 = Tx_1$.

Now suppose $\psi_1, \ldots, \psi_p$ have been chosen. Then let

$$\sup_{x \in B_1} \inf_{\{a_i\}_{i=1}^{p}} \left\{ \left\| x - \sum_{i=1}^{p} a_i \psi_i \right\| \right\} = \epsilon'_{p+1}. \tag{4.17}$$

Again, because $B_1 \subset X$ is a compact set and $\| \cdot \|_Y$ and T are continuous, there exists an $x_{p+1} \in B_1$ such that x_{p+1} attains the maximum in the above equation. Choose $\psi_{p+1} = T x_{p+1}$. Comparing with theorem 4.2 we note that ϵ'_p is an upper bound for ϵ_p. It is an open conjecture as to whether $\epsilon'_p = \epsilon_p$.

In this algorithm, instead of searching over all possible sets $\{\psi_i\} \in \Psi_p$ we select a special set that is in some sense (it is the image of the $x \in B_1$ that maximises equation (4.17)) the best possible set to use. This choice is essential because otherwise the calculation of generalised singular values becomes too cumbersome (one needs to find the set Ψ_p before calculating ϵ_{p+1}.) Note however, that the above algorithm gives the right value for ϵ_1. In the next section, I use this algorithm to calculate approximations for the generalised singular values of some specific spatial waveform channels.

4.3 Software design and techniques used for numerical simulations

In this chapter I present the results of numerical simulations. I describe the software used for the simulations and several techniques that are used to reduce the running time of the simulations are also discussed. All simulations were done in two dimensions because the number of grid points required to do simulations in three dimensions makes it infeasible to do computations over distances greater than a few wavelengths (c.f. eg. [37] and [31], where simulations were also done in two dimensions). In this thesis, Finite-Difference Time-Domain (FDTD) simulations were used to calculate the electromagnetic field in the presence of reflective scatterers.

In order to calculate the degrees of freedom of SWCs, one needs to calculate the generalised singular values of the compact operators in SWCs. This is a direct consequence of the definition of generalised singular values. In order to calculate these generalised singular values we use several theorems proved in earlier chapters. Suppose T is the transmitting volume, R is the receiving volume, t_0 is some instant in time and $(\widetilde{X}_{T,t_0}, \widetilde{Y}_R, \Gamma)$ is an SWC. The essential idea involved in the calculation of the generalised singular values is:

1. Let $\{\mathbf{x}_i\}_{i=1}^{\infty}$ be a Schauder basis for $\mathcal{L}^2(T, \mathbb{C}^3)$. Then we know from the second corollary to theorem 2.3 that $\{\mathbf{x}_i\}_{i=1}^{\infty}$ is also a Schauder basis for $\widetilde{X}_{T,t_0}$.

2. Let $\Gamma_n = \Gamma|_{\mathrm{span}\{\mathbf{x}_i\}_{i=1}^{N}}$. Then we know from theorem 4.1 that the generalised singular values of Γ_n approach those of Γ as $n \to \infty$.

3. We can use any convenient, finite subset of a Schauder basis of $\mathcal{L}^2(T, \mathbb{C}^3)$ to calculate approximations of the generalised singular values of Γ.

In this thesis, I assume that the source and receiver antennas are restricted to square areas of size $1\lambda \times 1\lambda$ or $2\lambda \times 2\lambda$, where λ is the wavelength (see eg. [37] where a similar assumption was made). I divide the transmitting and receiving squares into several smaller squares and use functions that are non-zero on one of the smaller squares to form a subset of a Schauder basis for the space of transmitting and receiving functions.

The task of finding the generalised singular values can be split into three components. Firstly, one needs to find the orthonormal source functions in $\widetilde{X}_{T,t_0}$ from a finite subset of the chosen Schauder basis using the Gram-Schmidt method. This step is only required for the finite energy case. It is easy to use existing `Matlab` routines to calculate singular values of compact operators defined on finite dimensional spaces once the channel operator has been specified as a matrix expressed in terms of orthonormal bases. In the finite power case, the space of source functions does not necessarily carry an inner product structure and therefore we need to develop alternative algorithms to calculate the singular values as explained in the previous section. Secondly, one needs to calculate the received field for each one of the (orthonormal) source functions in the presence of randomly placed scatterers. Finally one needs to calculate the transfer matrix in terms of the Schauder bases for the source and receiver functions and calculate the singular values of the operator Γ_n. The first two tasks were completed using programs written in `C++` and the last one was undertaken using `Matlab` in the finite energy case and `C++` in the finite power case. The source code used for these programs can be found on the attached compact disk or downloaded from `http://users.rsise.anu.edu.au/~somaraju`.

The interactive `C++` program consist of five files and allows the user to calculate the orthonormal source functions and the received field. Several settings such as loss resistance and simulation domain size can be set interactively before starting the simulations. The functionality of different files in the program is briefly explained below.

`main.cpp`: The program execution starts in the function `main()` within this file.

`field.cpp`,`field.hpp`: These files are used to define the `field` object. This object is used to calculate the electromagnetic field for a given set of sources and scatterers possibly in the presence of a perfectly matched layer (PML) (see subsection 4.3.2). The field object allows the user to specify the simulation region, step sizes, the number of randomly placed scatterers in the simulation region and the position of all the sources and their magnitudes. One can also specify if the simulation domain has a PML. Once all the above have been specified, the simulation can start and the field object can be used to calculate the electromagnetic field after n time steps by calling the function `timeStep(int n)`.

`field_util.cpp,field_util.hpp`: These files contain several utility functions that can be used on objects of the `field` class.

`gramSchmidt.cpp,gramSchmidt.hpp`: These files are used to calculate the orthonormal source functions. The basis expansion of the orthonormal source functions in terms of the subset of the Schauder basis $\{\mathbf{x}_i\}_{i=1}^{\infty}$ is stored in a file. See subsection 4.3.1 for a more detailed description of the algorithms used in this file.

`receiverField.cpp,receiverField.hpp`: These files are used to calculate the field at several randomly placed receivers in the presence of randomly placed scatterers and the calculated fields are stored in a file. See subsection 4.3.2 for a more detailed description.

`svalClass.cpp,svalClass.hpp` Objects of this class are used to calculate the generalised singular values in the finite power case as explained in subsection 4.3.3

`genSval.cpp,genSval.hpp` The functions in this file are used to perform the maximisation routines for calculating the generalised singular values. Functions from the the GNU scientific library are used to perform the maximisation (see `http://www.gnu.org/software/gsl/`).

`sval.m` The matlab file `sval.m` uses the orthonormal basis expansion and receiver fields calculated using the `C++` program to evaluate the singular values of SWCs in the finite energy case.

4.3.1 Calculating Orthonormal Source Functions for $\widetilde{X}_{T,t_o}$

Suppose $\{\mathbf{x}_i\}_{i=1}^{\infty}$ is an orthonormal basis in $\mathcal{L}^2(T,\mathbb{C}^3)$. Then the set $\{\mathbf{x}_i\}_{i=1}^{\infty}$ is a complete Schauder basis for $\widetilde{X}_{T,t_0}$ with respect to the norm $\|\cdot\|_{\widetilde{X}_{T,t_0}}$ (see theorem 2.3). However, the functions are not orthogonal with respect to the inner product $\langle\cdot,\cdot\rangle_{\widetilde{X}_{T,t_0}}$ where,

$$\begin{aligned}\langle \mathbf{J}_1,\mathbf{J}_2\rangle_{\widetilde{X}_{T,t_0}} &= I_1+I_2 && \text{with}\\ I_1 &= R_{loss}\int_0^{t_0}\int_T \mathbf{J}_1'^{*}(\mathbf{r},t)\mathbf{J}_2'(\mathbf{r},t)d\mathbf{r}dt && \text{and}\\ I_2 &= \epsilon_0\int_{\Omega_{ext}}\mathbf{E}_1^{*}(\mathbf{r},t_0)\mathbf{E}_2(\mathbf{r},t_0)d\mathbf{r}+\tfrac{1}{\mu_0}\int_{\Omega_{ext}}\mathbf{H}_1^{*}(\mathbf{r},t_0)\mathbf{H}_2(\mathbf{r},t_0)d\mathbf{r}.\end{aligned}$$

The source functions $\{\mathbf{x}_i\}_{i=1}^{N}$ need to be orthonormalised with respect to the inner product $\langle\cdot,\cdot\rangle_{\widetilde{X}_{T,t_0}}$ using the Gram-Schmidt process.

In order to calculate the inner product $\langle\mathbf{x}_1,\mathbf{x}_2\rangle_{\widetilde{X}_{T,t_0}}$ for a given time instant t_0, one needs to calculate both I_1 and I_2 in the above equation. However, to calculate I_2, one needs to calculate the fields generated by the source functions at time instant t_0 in free space. Therefore the simulation region needs to be as big as the distance travelled by

light in the time t_0. This makes the simulation region extremely large and the numerical computations take a great amount of time. We can however assume that the sources are contained in a cell made of a perfect electric conductor. Suppose Ω' is such a cell and Ω'_{int} is the interior of this cell. If I'_2 is defined by

$$I'_2 = \epsilon_0 \int_{\Omega_{ext} \cap \Omega'_{int}} \mathbf{E}'^{*}_1(\mathbf{r}, t_0)\mathbf{E}'_2(\mathbf{r}, t_0)d\mathbf{r} + \frac{1}{\mu_0} \int_{\Omega_{ext} \cap \Omega'_{int}} \mathbf{H}'^{*}_1(\mathbf{r}, t_0)\mathbf{H}'_2(\mathbf{r}, t_0)d\mathbf{r},$$

then $I_1 + I'_2$ defines an inner product on the space of source current densities. Here, $\mathbf{E}'_i$ and $\mathbf{H}'_i$ are the electric and magnetic fields generated by the sources $\mathbf{J}_i$ for $i = 1, 2$ that are contained within the cell Ω'. Now, the norm defined by this inner product gives the total energy radiated by the sources. Because the perfect electrical conductors absorb no energy, this norm is actually equal to the norm induced by the inner product $\langle \cdot, \cdot \rangle_{\widetilde{X}_{T,t_0}}$. Therefore we can tell from the parallelogram law for inner products that $I_2 = I'_2$. Because the perfect conductors are also perfect reflectors of electromagnetic radiation, the field outside the conductors is zero. Therefore, we can reduce the total simulation region and speed up the calculation of the orthonormal source functions by using I'_2 instead of I_2.

4.3.2 Calculating the Receiver Field

FDTD simulations using the Yee scheme [50] are used to calculate the field at different receiver locations for a given source function (see eg. [51–54]). Sixteen points per wavelength were chosen in each cartesian direction to minimise the effects of numerical dispersion (see eg [52]). In order to maximise the distance between the transmitting and receiving volumes, it was decided that Perfectly Matched Layers (PMLs) be used[2] [56]. PMLs are absorbing layers that do not reflect any of the electromagnetic waves that are incident upon them and the magnitude of electromagnetic waves reduces exponentially as they travel in the PML. Also, reflective scatterers were placed at random locations in the simulation region. The reflective scatterers were also placed within the PML. This enables one to model the effect of distant scatterers because the PML reduces the magnitude of the electromagnetic waves and the total energy reflected is very small. The received field is calculated at random positions within the simulation region. Figure 4.2 on page 68 shows an example of a typical simulation region.

4.3.3 Calculating the Generalised Singular Values

In the finite energy case, the spaces of source and receiver functions are Hilbert spaces. Therefore, as was proven in theorem 3.5, the generalised singular values of the channel

[2] Also see [52,55]

operator are equal to the traditionally accepted notion of the singular values of compact operators defined on Hilbert spaces. There are several well known methods that can be used to calculate these singular values and the `Matlab` function `SVD` is used to calculate these singular values.

In the finite power case, the space of source functions is not an inner product space. Therefore a new program was written in `C++` in order to calculate the generalised singular values. Suppose $(\widetilde{X}_T, \widetilde{X}_R, \Gamma)$ is an SWC and $\{\mathbf{x}_i\}_{i=1}^{\infty}$ is a complete Schauder basis for $\widetilde{X}_T$ and $\Gamma_n = \Gamma|_{\text{span}\{\mathbf{x}_i\}_{i=1}^n}$ is the restriction of Γ to the span of the first n vectors in the Schauder basis. Then we know from theorem 4.1 that the singular values of Γ_n are approximately equal to those of Γ for n large enough. Therefore, we only need to calculate the singular values of the operator Γ_n. The algorithm described after theorem 4.2 is used to calculate these generalised singular values and the results of these calculations are presented in the following section.

4.4 Results and Discussion

The singular values of various SWCs were calculated. The source resistance, number of scatterers and the receiver and scatterer locations were varied and the results of these simulations are presented in this subsection. The simulations were performed under the following conditions (see figure 4.2 on page 68):

1. Because we are performing the simulations in 2D, the sources and scatterers were assumed to be invariant in the z-direction and only transverse magnetic fields were considered[3]. The sources and receivers were assumed to be restricted to within squares of sizes $1\lambda \times 1\lambda$ or $2\lambda \times 2\lambda$.

2. The total size of the simulation region is $16\lambda \times 256\lambda$ (c.f. Xu and Janaswamy [37] where the simulation region is of a similar size). It was decided that 16 grid points be chosen per wavelength in order to minimise the effect of dispersion (see eg [52]). The simulation region was surrounded by a PML of thickness 2λ which proved to be sufficient to absorb all the incident electromagnetic waves.

3. The center of the source was always located at $8\lambda \times 8\lambda$. Ten receivers were placed at random locations in the half of the simulation region that did not contain the source. A variable number of scatterers were placed both within the simulation region and in the PML. The scatterers were assumed to be perfectly conducting materials of variable length between 1λ and 4λ. The exact form of the scatterers is not important because we are using them to create a rich scattering environment.

[3]Reciprocity between the electric and magnetic fields shows that exactly the same behavior can be observed for the transverse electric case.

Figure 4.3 on page 69 shows the first twenty singular values of a typical SWC in the finite energy case. In this figure, the field was calculated at 10 randomly placed receivers. Hundred scatterers were randomly placed in the simulation region and the loss resistance was assumed to be 5Ω. The error bars indicate the variability of the singular values for different receivers. The number of degrees of freedom for this SWC at level-10^{-4} is about 6.

This behavior is typical of all the SWCs that were simulated. No 'knee' type behavior is detected for any of the simulations performed. The essential dimension of all the simulated channels is 1. This is similar to the behavior predicted by Xu *et. al.* [37] and other numerical results [34,36]. Notice that the 'knee' type behavior is only observed in analytic evaluations of singular values of channels with specific assumptions made about the sources and/or scatterers [25,26,28,31]. For instance, in Bucci *et. al.* [25] it was assumed that all sources and scatterers were constrained to a ball of radius a. In Miller [26] it was assumed that the sources and receivers were rectangular prisms. In Poon *et. al.* it was assumed that the scatterers coupled the electromagnetic field from the source to the receiver only if the source radiation was within some solid angle (see section 1.4).

Figure 4.4 on page 70 shows how the singular values change with varying loss resistance R_{loss} in the finite energy case. As expected, as the loss resistance increases, the generalised singular values decrease. This is because most of the energy is lost as heat in the transmitting antenna. The energy lost as heat is comparable to the energy radiated if the loss resistance is about 1Ω. Therefore, below this value of loss resistance, the total energy radiated dominates and there is little change in the singular values as the loss resistance changes.

Figure 4.5 on page 71 shows the behavior of the singular values with increasing number of scatterers in the finite energy case. As the number of scatterers increases, the singular values reduce in magnitude. This is explained if one examines the physical situation modeled by the simulation region. Reflective scatterers that are randomly placed between the source and the receiver ensure that a large fraction of the energy is reflected. Therefore, the total energy in the receiver is reduced. However, the amount of reduction in the magnitude of the singular values ϵ_n reduces with increasing n. This indicates that there are a greater number of reflected signals in the path with increasing number of scatterers. Therefore, the smaller singular values are effected less by the increasing number of scatterers due to the presence of multiple reflective paths.

Figure 4.6 on page 72 shows the behavior of the singular values for zero loss resistance in the finite energy case. In this case the singular values show a similar behavior to the case where the loss resistance is not zero. This is contrary to what is predicted by Wallace and Jensen [43]. If the norm on the space of source functions is given by the total power/energy radiated the resulting operator of the SWC need not be compact. Therefore, the singular values should not go to zero. However, because

the channel operator is not compact in this case, the theory developed in this thesis can not be used. Specifically, we can not prove that the singular values of the operator Γ_n approach those of Γ. This is to be expected because the operator Γ_n defined on a finite dimensional space is compact but the operator Γ is not. Therefore we cannot have $\lim_{n\to\infty} \|\Gamma_n - \Gamma\| = 0$ because the uniform operator limit of a compact operator must be compact.

Figure 4.7 on page 73 shows the values[4] ϵ'_n for various values of loss resistance calculated for the finite power case. If ϵ_m is the m^{th} singular value of the channel operator Γ and $\epsilon_{m,n}$ is the m^{th} singular value of Γ_n then ϵ'_n is an upper bound for $\epsilon_{m,n}$. The legend shows the loss resistance in Ohms for different plots shown in the figure. The values ϵ'_n were calculated by maximising the function in equation (4.17). The maximisation was performed by choosing a random initial vector and this procedure was repeated several times. The error bars in the graph indicate the maximum range of ϵ'_n for different initial vectors. As expected as the loss resistance increases ϵ'_n gets smaller. It should be noted that the loss resistance chosen was always greater than 5Ω. This is because of the following: it is possible to get high concentrations of energy in a small receiving volume R for arbitrarily small amount of radiated power (e.g. think of a source between two perfect reflectors, so that all the energy ever radiated by the source is restricted to be between the reflectors).

Therefore it is possible to have current configurations for which the radiated power is close to zero. The total power radiated is calculated using Poynting's theorem. In order to use Poynting's theorem one needs to calculate the vector product of the electric and magnetic fields on a surface enclosing the source. However, because the Yee scheme is used for FTDT simulations, the electric field is calculated on integer grid points and the magnetic field is calculated on half-integer grid points. Therefore if the surface is on the integer grid points then one can only approximate the magnetic field on the surface. Therefore there is a small, but inevitable numerical error that enters the calculation of the total power radiated. This error is not significant when the loss resistance is large because the functions which radiate very small amount of energy but still concentrate high amounts of energy in the receiving volume tend to have high amplitudes. Therefore the total power lost as heat is high and this factor dominates the power used by the source functions. However, for small loss resistances, it is possible that a numerical calculation of the total power used to set up a current might be negative and this could result in some absurdly large values for singular values. So different types of simulations need to be performed to study the performance of the channel for smaller loss resistances.

[4] see section 4.2 for a definition of ϵ'_n

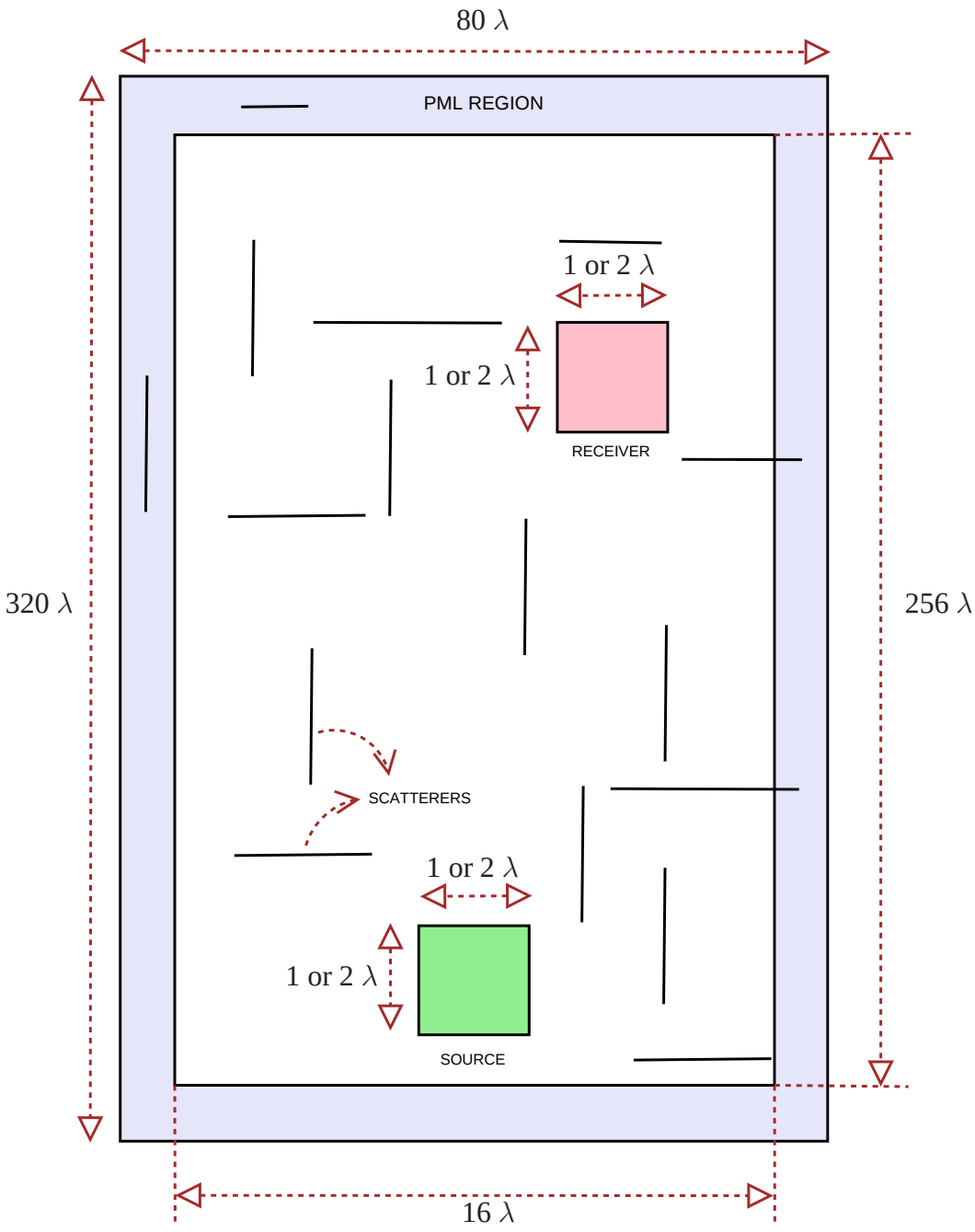

Figure 4.2: Typical Simulation Region for FDTD Simulations

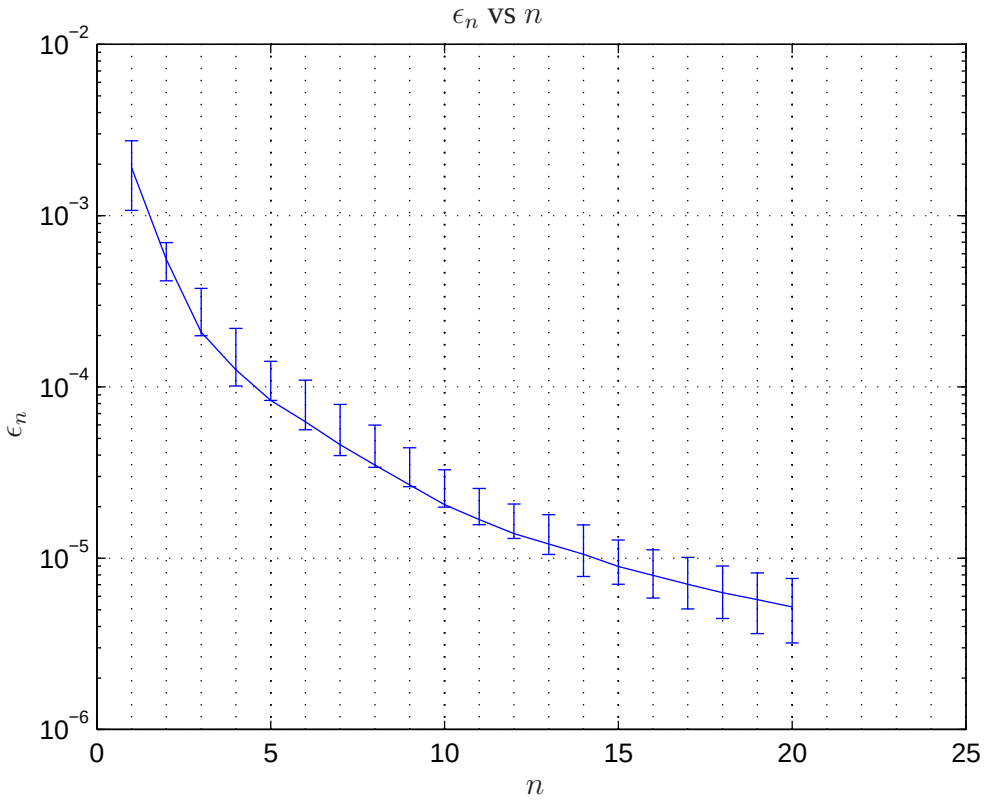

Figure 4.3: Singular Values for Finite Energy Case

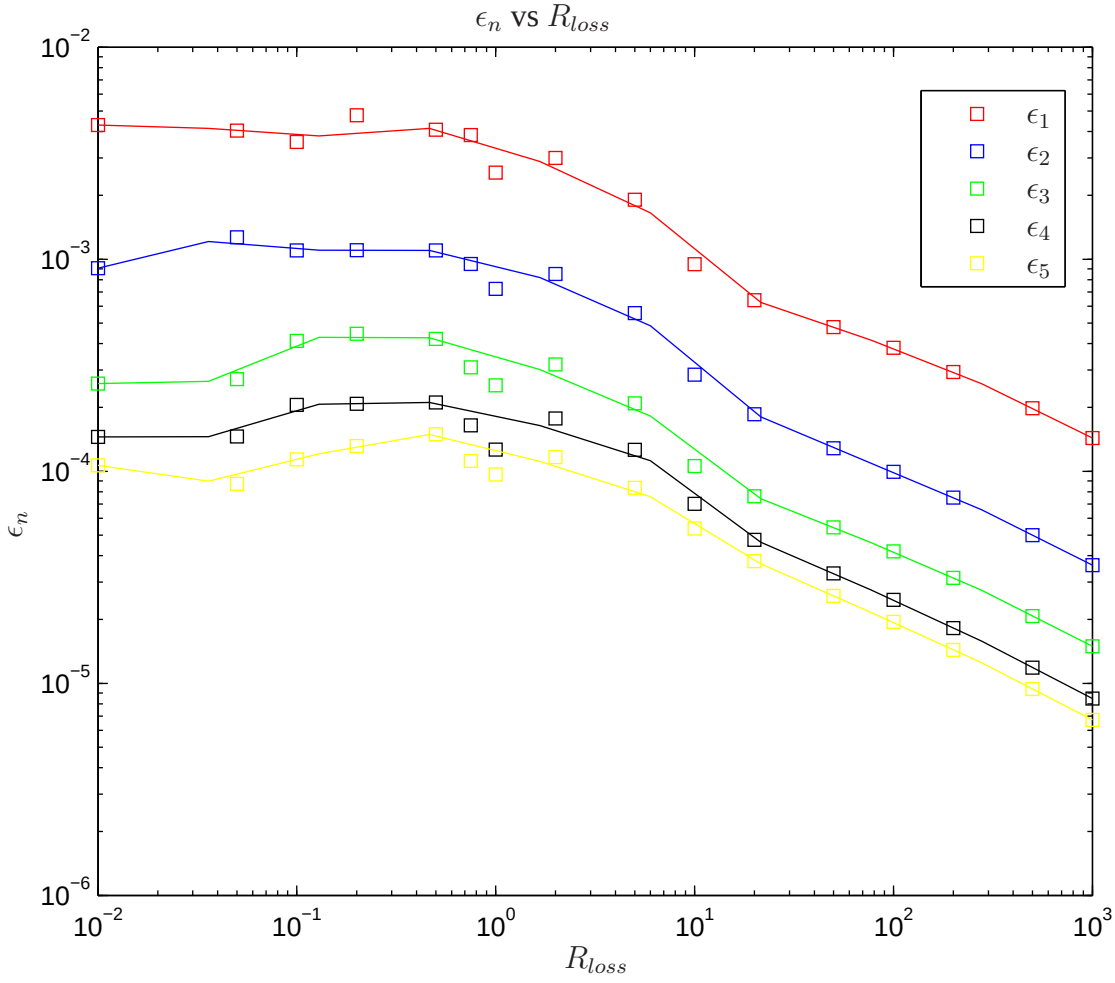

Figure 4.4: Variation of Singular Values with Loss Resistance

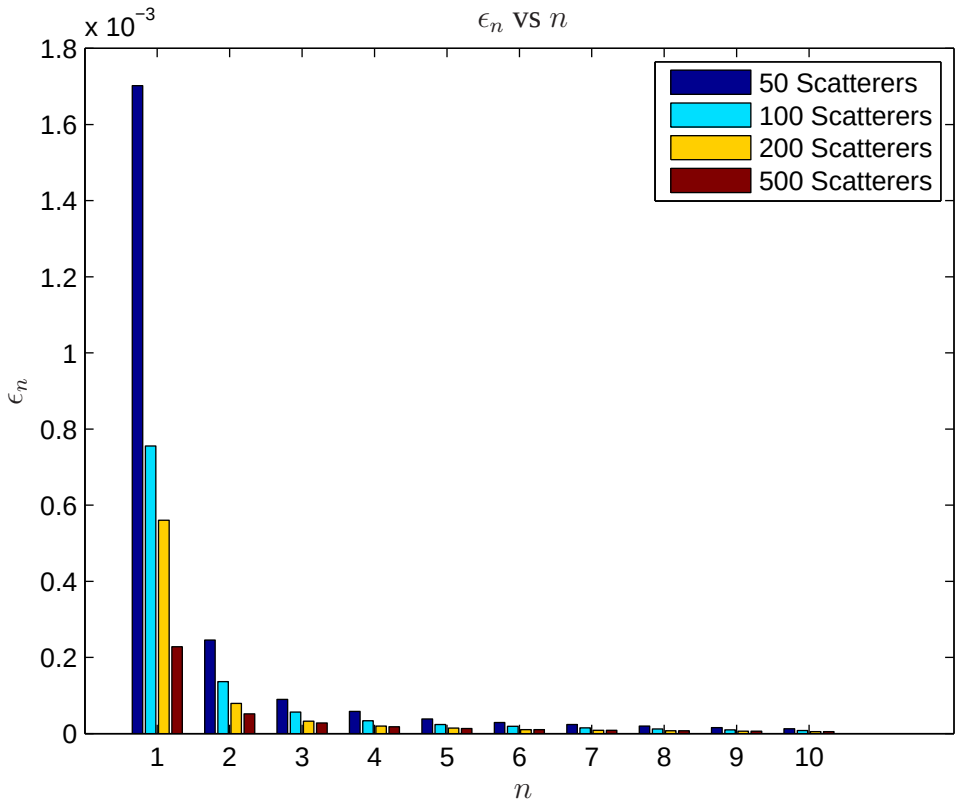

Figure 4.5: Variation of Singular Values with Number of Scatterers

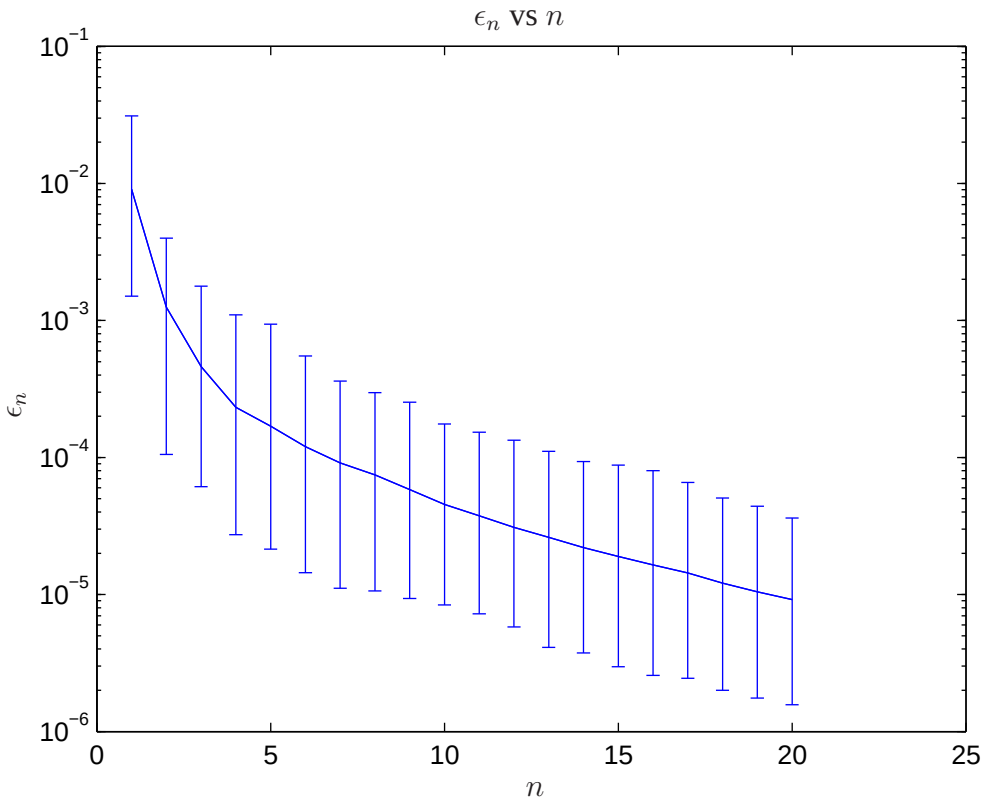

Figure 4.6: Singular Values for Zero Loss Resistance

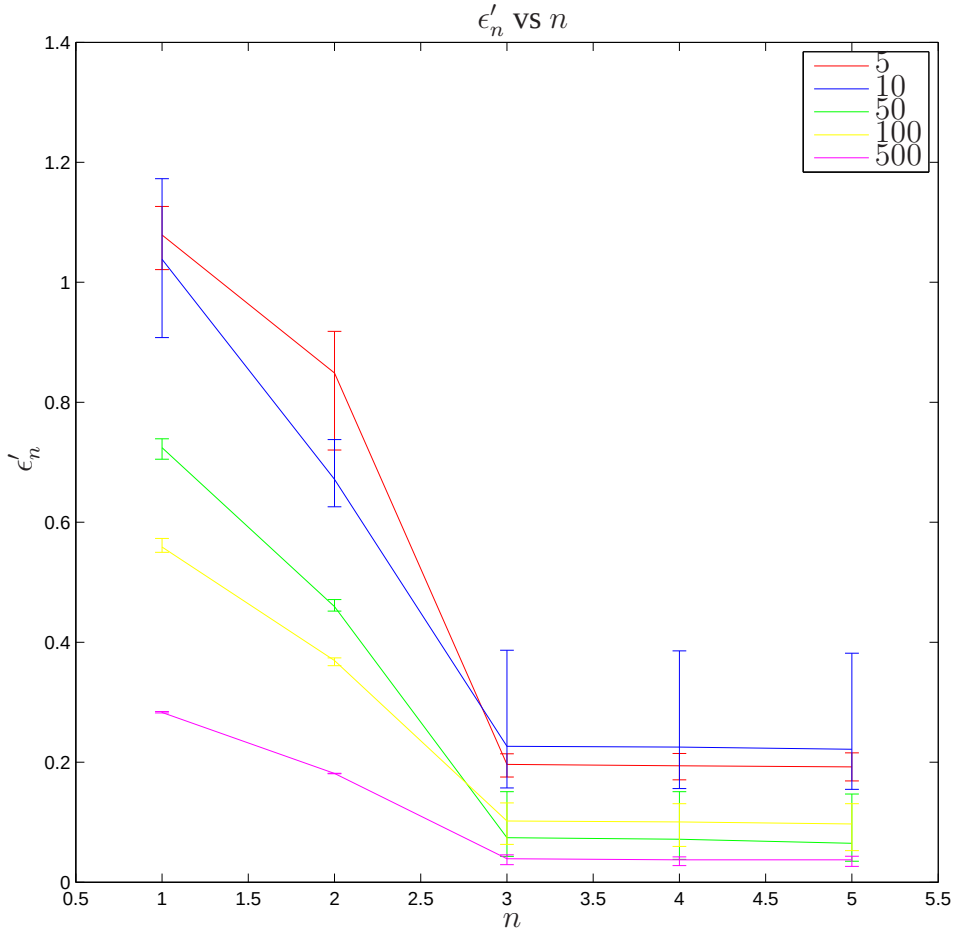

Figure 4.7: Upper Bound on Singular Values for Finite Power Case

4.5 Chapter Conclusion

In this chapter I showed that it is possible to use perturbation theory to calculate the essential dimension of the scalar waveform channel studied by Miller [26]. But, as it is generally not possible to analytically calculate the degrees of freedom, numerical techniques are required. In this chapter I prove that if the normed space on which the compact operator is defined has a complete Schauder basis then we can use numerical techniques similar to Galerkin's method. I proved in theorem 4.1 that the singular values of finite dimensional approximations of an operator approach those of the original operator if the domain of the operator has a complete Schauder basis. I also developed an alternative formulation of singular values in theorem 4.2 that is more conducive to numerical calculations. Numerical simulations were performed in both the finite energy and finite power cases using the finite-difference time-domain (FDTD) method and the simulation results were presented in section 4.4 The main contributions of this chapter are:

1. Used perturbation theory to analytically calculate the essential dimension of a scalar waveform channel.

2. Developed a numerical technique to calculate normed space generalised singular values.

3. Numerical results of simulations were presented.

Chapter 5

Uncertainty Principles for Energy Concentrations

Uncertainty Principles (UPs) have gained great popularity since Heisenberg [57]. The famous original example is Heisenberg's Uncertainty Principle: it is impossible to exactly measure the location *and* the momentum of a particle simultaneously. This is a special case of a more general framework, which may be formulated (via various Hilbert space techniques) to apply to a wide range of scenarios.

In a communication theory setting, a similar uncertainty principle has been well known: that a signal cannot be arbitrarily confined in both time and frequency. The reader is directed to Slepian [49] for a discussion. The works of Landau, Slepian and Pollak [27,30,58] has formalised this result, although without explicit reference to the operator theoretic nature of the problem[1]. Similar to the classical UP, one can generalise this UP to arbitrary operators on Hilbert spaces as well. This generalisation is particularly pertinent to SWCs. We shall pose the following question: *Given a signal (source current) which has energy in one volume V_A, how well can we constrain a function (electromagnetic field) of that signal to have energy in another volume V_B?*

The remainder of this chapter is arranged as follows: In Section 5.1 I collate classic results on uncertainty and formulate UPs in terms of operator theoretic terminology. Section 5.2 develops an Uncertainty Principle for communication between volumes with a particular form of operator channel. In section 5.3 I prove a second uncertainty principle and explain its interpretation in a communication theory setting.

5.1 A Review of the Uncertainty Principle

Before explaining the Uncertainty Principle, we develop some relevant notation.

[1] Landau provides some work in this regard [27].

5.1.1 Notation

Following Selig [29], let $\mathcal{H}$ be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \| \equiv \langle \cdot, \cdot \rangle^{1/2}$. Further, let A be a linear operator with domain $\mathcal{D}(A) \subseteq \mathcal{H}$ and range in $\mathcal{H}$. We then define the normalised expectation value, $\tau_A(f)$ and standard deviation (uncertainty), $\sigma_A(f)$ of the operator A with respect to $f \in \mathcal{D}(A)$ to be [29]

$$\tau_A(f) \equiv \frac{\langle Af, f \rangle}{\langle f, f \rangle} \tag{5.1}$$

$$\sigma_A(f) \equiv \|(A - \tau_A(f))f\|. \tag{5.2}$$

If the domain of some operator A is dense in $\mathcal{H}$ then we can define its adjoint$A^\dagger$ using the following equation [48]

$$\langle Ax, y \rangle = \langle x, A^\dagger y \rangle \ \ \forall x \in \mathcal{D}(A), y \in \mathcal{D}(A^\dagger). \tag{5.3}$$

The domain of $\mathcal{D}(A^\dagger)$ consists of vectors $x \in H$ such that the function

$$y \mapsto \langle x, Ay \rangle$$

(which is a linear map defined on a dense subset of $\mathcal{H}$) is a continuous linear functional. Furthermore, A is said to be Hermitian or self-adjoint if $A = A^\dagger$. Therefore, for a self-adjoint operator

$$\langle Ax, y \rangle = \langle x, Ay \rangle \ \ \forall x, y \in \mathcal{D}(A). \tag{5.4}$$

Any operator that obeys equation (5.4) is said to be symmetric. Note that it is possible that an operator A is symmetric but not self-adjoint if its domain $\mathcal{D}(A)$ is not dense in $\mathcal{H}$.

Given two linear operators A and B with domains $\mathcal{D}(A)$ and $\mathcal{D}(B)$ respectively, the commutator is defined as

$$[A, B] \equiv AB - BA$$

and the anti-commutator is defined as

$$[A, B]_+ \equiv AB + BA$$

with domains $\mathcal{D}(AB) \cap \mathcal{D}(BA)$ for either one. The operators A and B are said to *commute* with each other if $[A, B] = 0$. Otherwise they are called non-commutative operators.

Also, let $L^2[\mathbb{R}]$ be the set of all square integrable real-valued functions defined on the real line with norm $\|f\|_2 = \langle f, f \rangle^{1/2}$. Here the inner product $\langle \cdot, \cdot \rangle$ is defined as

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(t)\overline{g(t)}dt. \tag{5.5}$$

We can now think of $\|f\|_2^2$ as the energy of the function $f \in L^2[\mathbb{R}]$. The angle between two non-zero functions f and g is defined as

$$\theta(f,g) = \cos^{-1} \frac{\Re\{\langle f, g\rangle\}}{\|f\|_2 \|g\|_2}. \tag{5.6}$$

We also define

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx \tag{5.7}$$

to be the fourier transform of $f(x)$ whenever this integral exists.

5.1.2 The Classical Uncertainty Principle

The classical uncertainty principle states that the concurrent values of two non-commuting observables such as position and momentum cannot be precisely determined in any quantum state. That is, the standard deviation of two non-commuting operators cannot be made arbitrarily small simultaneously.

Theorem 5.1. *If A and B are two self-adjoint operators on a Hilbert space $\mathcal{H}$, then*

$$\|(A-a)f\|\|(B-b)f\| \geq \frac{1}{2}|\langle [A,B]f, f\rangle|$$

for all $f \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$ and all $a, b \in \mathbb{R}$. Equality holds precisely when $(A-a)f$ and $(B-b)f$ are purely imaginary scalar multiples of one another.

Selig realizes that the only property of self-adjoint operators required in the proof of the above theorem is given by equation (5.4). He therefore reformulates theorem 5.1 in terms of symmetric operators.

Theorem 5.2. *[29, theorem 3.4] If A and B are two symmetric operators on a Hilbert space $\mathcal{H}$, then*

$$\begin{aligned}\|(A-a)f\|\|(B-b)f\| \geq \frac{1}{2}\{&|\langle [A,B]f, f\rangle|^2 \\ &+ |\langle [A-aI, B-bI]_+ f, f\rangle|^2\}^{1/2}\end{aligned}$$

for all $f \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$ and all $a, b \in \mathbb{R}$. Equality holds precisely when $(A-a)f$ and $(B-b)f$ are purely imaginary scalar multiples of one another.

Proof. Let A, B, a, b and f be as stated in the theorem. From the Cauchy-Schwarz inequality we have

$$\begin{aligned}2\|(B-b)f\|\|(A-a)f\| &= 2|\langle (B-b)f, (A-a)f\rangle| \\ &= 2[\Im\{\langle (B-b)f, (A-a)f\rangle\}^2 + \\ &\quad \Re\{\langle (B-b)f, (A-a)f\rangle\}^2]^{1/2}\end{aligned} \tag{5.8}$$

Here, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a complex number. We can evaluate the real and imaginary parts of $\langle (B-b)f, (A-a)f\rangle$.

$$\begin{aligned} 2i\Im\{\langle (B-b)f,(A-a)f\rangle\} &= \langle (B-b)f,(A-a)f\rangle \\ &\quad - \langle (A-a)f,(B-b)f\rangle \\ &= \langle (A-a)(B-b)f, f\rangle \\ &\quad - \langle (B-b)(A-a)f, f\rangle \\ &= \langle [A-aI, B-bI]f, f\rangle \\ &= \langle [A,B]f, f\rangle \end{aligned} \tag{5.9}$$

The symmetry of the operators is used in the second step and the fact that scalar multiplication commutes with all linear operators is used in the last. Also,

$$\begin{aligned} 2\Re\langle (B-b)f,(A-a)f\rangle &= \langle (B-b)f,(A-a)f\rangle \\ &\quad + \langle (A-a)f,(B-b)f\rangle \\ &= \langle [A-aI, B-bI]_+ f, f\rangle \end{aligned} \tag{5.10}$$

Substituting equations (5.10) and (5.9) into equation (5.8) proves the theorem. □

Theorem 5.2 is valid for arbitrary values of a and b. However, the left hand side of the inequality in this theorem is minimized if a and b are the magnitudes of the projection vectors of Af and Bf onto f, respectively. That is

$$\begin{aligned} \min_a \|Af - af\| &= \|Af - \frac{\langle Af, f\rangle}{\langle f, f\rangle} f\| \\ &= \|Af - \tau_A(f) f\| \\ &= \sigma_A(f) \end{aligned}$$

Note that $\sigma_A(f)$ is the uncertainty (standard deviation) of the operator A. Therefore we obtain the following special case of the UP.

Corollary 5.2.1. *[29, Corollary 3.6] If A and B are two symmetric operators on a Hilbert space $\mathcal{H}$, then*

$$\sigma_A(f)\sigma_B(f) \geq \frac{1}{2}\sqrt{|\langle [A,B]f, f\rangle|^2 + 4\mathrm{COV}_{AB}^2(f)}$$

for all $f \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$ and all $a, b \in \mathbb{R}$. Equality holds precisely when $(A - \tau_A(f))f$ and $(B - \tau_B(f))f$ are scalar multiples of one another.

Here, the covariance of the operators A and B is defined as

$$\mathrm{COV}_{AB}(f) = \frac{1}{2}\langle [A - \tau_A(f)I, B - \tau_B(f)I]_+ f, f\rangle$$

Heisenberg's Uncertainty Principle

The first UP discovered by Heisenberg is a special case of Corollary 5.2.1. We can derive Heisenberg's UP by substituting $Af = \cdot f$, the position operator and $Bf = -if'$, the momentum operator into corollary 5.2.1. Here $i = \sqrt{-1}$. Making this substitution[2] and noting that $[A, B]f = if$ we get

$$\begin{aligned}
\sigma_A(f)\sigma_B(f) &\geq \frac{1}{2}\sqrt{|\langle [A,B]f, f\rangle|^2 + 4\text{COV}^2_{AB}(f)} \\
&\geq \frac{1}{2}|\langle [A,B]f, f\rangle| \\
&= \frac{1}{2}|\langle if, f\rangle| \\
&= \frac{1}{2}\|f\|^2
\end{aligned}$$

Here, $\sigma_A(f)$ and $\sigma_B(f)$ are the uncertainties in position and momentum respectively. Note that equality is attained iff $(A - \tau_A(f))f$ and $(B - \tau_B(f))f$ are purely imaginary scalar multiples of each other (see Theorem 5.1). That is

$$-if'(x) - \tau_B(f)f(x) = ir(x - \tau_A(f))f(x)$$

and $r \in \mathbb{R}$. Therefore, f(x) is a complex gaussian function:

$$f(x) = ce^{ibx}e^{-r(x-a)^2/2}$$

for some $c \in \mathbb{C}$ and $r > 0$. Here $a = \tau_A(f)$ and $b = \tau_B(f)$ are the expectation values of the position and momentum operators respectively.

Heisenberg's principle shows that a quantum system described by a wave function f cannot have precise values for position and momentum at the same time. This principle is often stated in terms of a function and its fourier transform as explained in the following theorem .

Theorem 5.3. *[29, theorem 6.1]] Let $f \in L^2(\mathbb{R}), \|f\| = 1$ and set*

$$\begin{aligned}
x_o &= \int x|f(x)|^2 dx \\
\omega_o &= \int \omega|\hat{f}(\omega)|^2 d\omega \\
\Delta x &= \int (x - x_o)^2|f(x)|^2 dx \\
\Delta\omega &= \int (\omega - \omega_o)^2|\hat{f}(\omega)|^2 d\omega
\end{aligned}$$

[2]We can make this substitution because both operators are symmetric.

whenever these integrals exist. Then $\Delta x \Delta \omega \geq \pi/2$*, where equality is attained iff* $f(x) = (r/\pi)^{1/4} e^{i\omega_o x} e^{-r(x-x_o)^2/2}$ *for any* $r > 0$*.*

The theorem is proved trivially by noting that $\sigma_A(f) = \Delta x$ and $\sigma_B(f) = \Delta\omega/2\pi$, where $Af = \cdot f$ and $Bf = -if'$ [29].

This theorem gives valuable insight into how localized a function can be in both time and frequency. If one defines Δx and $\Delta\omega$ to be the root mean square (RMS) measure of approximate time duration and bandwidth of the signal, respectively, then the theorem says that the product of time duration and bandwidth of a function is bounded from below by $\pi/2$. Therefore, if the time-spread gets very small, the frequency-spread must be large and vice-versa. It therefore validates our intuition that a function cannot simultaneously be both time and frequency limited.

Though a very good qualitative tool, this theorem is inadequate for the purposes of signal processing. The theorem does not for instance answer the question, given a bandlimited function (i.e. $\hat{f}(\omega) = 0$ for $\omega \notin [-\Omega, \Omega]$), how much of the energy of f is 'concentrated' in any finite duration of time. This would be useful in answering the question: "given a bandlimited channel, how much of the transmitted signal can a receiver measure over a finite period of time?"

5.1.3 An Uncertainty Principle for energy concentrations

Landau and Pollak propose that for the purposes of signal processing, a more relevant uncertainty principle should use sharper measures of concentrations in time and frequency than the one used in Heisenberg's principle [30]. To help derive their uncertainty principle [30] define $\mathcal{D} = \{f : f \in L^2[\mathbb{R}], f(t) = 0 \; \forall |t| > T/2\}$ to be the class of all time-limited functions and $\mathcal{B} = \{f : f \in L^2[\mathbb{R}], \hat{f}(\omega) = 0 \; \forall |\omega| > \Omega\}$ to be the class of all band-limited functions. Here, $\hat{f}(\omega)$ is the fourier transform of $f(t)$ as defined in equation (5.7). Also T, the time duration and $W = \Omega/2\pi$, the bandwidth are fixed for the remainder of this chapter. It is easy to prove [58] that $\mathcal{D}$ and $\mathcal{B}$ are complete subspaces of $L^2[\mathbb{R}]$.

We can therefore define the projection operators $B : L^2[\mathbb{R}] \rightarrow L^2[\mathbb{R}]$ and $D : L^2[\mathbb{R}] \rightarrow L^2[\mathbb{R}]$ as follows

$$Df(t) = \begin{cases} f(t), & |t| \leq T/2 \\ 0, & |t| > T/2 \end{cases} \tag{5.11}$$

$$Bf(t) = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{j\omega t} d\omega \tag{5.12}$$

Note that the ranges of operators B and D are $\mathcal{B}$ and $\mathcal{D}$, respectively. Using these operators we can calculate the fraction of energy, α^2 of any function $f \in L^2[\mathbb{R}]$ in the

finite duration of time $[-T/2, T/2]$.

$$\alpha^2 = \frac{\|Df\|_2^2}{\|f\|_2^2}. \tag{5.13}$$

Similarly, we can calculate, β^2, the fraction of energy of a function in a finite bandwidth $[-\Omega, \Omega]$.

$$\beta^2 = \frac{\|Bf\|_2^2}{\|f\|_2^2}. \tag{5.14}$$

Note that $\alpha, \beta \leq 1$.

Prolate Spheroidal Wave Functions

Slepian and Pollak [58] show that the prolate spheroidal functions are eigenfunctions of the finite fourier transform and discuss several interesting properties for these functions. In the following I motivate the usefulness of these functions and list some of their key properties.

A special case of the UP for energy concentrations arises when we choose the functions to be time-limited, i.e. $f \in \mathcal{D}$ or equivalently, $\alpha = 1$. The UP then puts constraints on the values β can take as explained in the following theorem.

Theorem 5.4. *Let $f(t) \in \mathcal{D}$. Then, $\beta^2 = \|Bf\|_2^2/\|f\|_2^2 \leq \lambda_0$. Equality is achieved if and only if $f(t) = cD\psi_0(t)$, where $\psi_0(t)$ is an eigenfunction corresponding to the largest eigenvalue λ_0 of the integral equation*

$$\lambda f(t) = \int_{-T/2}^{T/2} \frac{\sin \Omega(t-s)}{\pi(t-s)} f(s)dt \tag{5.15}$$

and c is an arbitrary multiplicative constant.

Proof. Let $f(t)$ and β be as stated in the theorem. Then,

$$\begin{aligned} Bf(t) &= \frac{1}{2\pi}\int_{-\Omega}^{\Omega} e^{i\omega t}\hat{f}(\omega)d\omega \\ &= \frac{1}{2\pi}\int_{-\Omega}^{\Omega} e^{i\omega t}\int_{-T/2}^{T/2} f(s)e^{-i\omega s}dsd\omega \\ &= \int_{-T/2}^{T/2} \frac{\sin \Omega(t-s)}{\pi(t-s)} f(s)ds \end{aligned}$$

Because,

$$\int_{-\infty}^{\infty} \frac{\sin \Omega(t-u)}{\pi(t-u)} \frac{\sin \Omega(u-s)}{\pi(u-s)} du = \frac{\sin \Omega(t-s)}{\pi(t-s)}$$

we have

$$\begin{aligned}
\|Bf(t)\|_2^2 &= \int_{-\infty}^{\infty} Bf(t)\overline{Bf(t)}dt \\
&= \int_{-\infty}^{\infty}\int_{-T/2}^{T/2}\int_{-T/2}^{T/2} \frac{\sin\Omega(u-t)}{\pi(u-t)}\frac{\sin\Omega(t-s)}{\pi(t-s)} f(u)\overline{f(s)}dudsdt \\
&= \int_{-T/2}^{T/2}\int_{-T/2}^{T/2} \frac{\sin\Omega(u-s)}{\pi(u-s)} f(u)\overline{f(s)}duds
\end{aligned}$$

It is well known that the maximum value of the right hand side is λ_0, the largest eigenvalue of the integral equation (5.15) [59]. The maximum is attained when $f = cD\psi_0$ is the time-limited version of the corresponding eigenfunction. Here, c is an arbitrary multiplicative constant. Note that f is proportional to $D\psi_0$ and not to ψ_0 because ψ_0 is not time-limited. □

We can conclude from this theorem that the time-limited function that has the greatest fraction of energy in the finite bandwidth W is $cD\psi_0$. From the symmetry of the fourier transform, we can also deduce that the bandlimited function that has maximum energy in the finite time interval $[-T/2, T/2]$ is $g = c\psi_0$.[3] We therefore have the following corollary.

Corollary 5.4.1. *Let $f(t) \in \mathcal{B}$. Then, $\alpha^2 = \|Df\|_2^2/\|f\|_2^2 \leq \lambda_0$. Equality is achieved if and only if $f(t) = c\psi_0(t)$, where $\psi_0(t)$ is an eigenfunction corresponding to the largest eigenvalue λ_0 of the integral equation*

$$\lambda f(t) = \int_{-T/2}^{T/2} \frac{\sin\Omega(t-s)}{\pi(t-s)} f(s)dt \tag{5.16}$$

and c is an arbitrary multiplicative constant.

The integral equation (5.15) is the defining equation for the prolate spheroidal wave functions. Therefore they play an important role in the analysis of time/frequency limited functions. Slepian and Pollak [58] prove that there exist a countably infinite set of functions $\psi_0(t)$, $\psi_1(t)$, $\psi_2(t), \ldots$ and a set of positive real numbers $\lambda_0 > \lambda_1 > \lambda_2 > \ldots$ that satisfy the integral equation (5.15), that is

$$\lambda_i\psi_i(t) = \int_{-T/2}^{T/2} \frac{\sin\Omega(t-s)}{\pi(t-s)} \psi_i(s)dt \tag{5.17}$$

Equivalently, we can write this equation in terms of our operators B and D

$$\lambda_i\psi_i(t) = BD\psi_i(t)$$

[3] In fact the symmetry argument tells us that $g = cB\psi_0$. But, it turns out that ψ_0 is band limited, i.e. $B\psi_0 = \psi_0$

Hence, λ_i and ψ_i are eigenvalues and eigenfunctions of the operator BD, respectively. The functions ψ_i are called the prolate spheroidal wave functions. They have several interesting properties[4]:

1. The set of functions $\{\psi_0(t), \psi_1(t), \ldots\}$ is bandlimited, complete in $\mathcal{B}$ and orthonormal in $L^2[\mathbb{R}]$:

   $$\langle \psi_i, \psi_j \rangle = \delta_{ij},$$

   where δ_{ij} is the Kronecker-Delta.

2. The set of functions $\{D\psi_0(t), D\psi_1(t), \ldots\}$ is time-limited, complete in $\mathcal{D}$ and orthogonal in $L^2[\mathbb{R}]$:

   $$\langle D\psi_i, D\psi_j \rangle = \lambda_i \delta_{ij}$$

   This implies that the energy of ψ_i in the time interval $[-T/2, T/2]$ is λ_i.

3. The dependance of ψ_i and λ_i on T and Ω has been suppressed in the notation used. In fact, $\psi_i(t) = \psi_1(t, T, c)$ and $\lambda_i = \lambda_i(c)$, where $c = \Omega T/2$. The fact that λ_i only depends on ΩT and not on Ω and T separately becomes significant when one discusses the dimensionality of the time-bandwidth product.

4. For all i, $\lambda_i < 1$. Also, the eigenvalues fall off rapidly to zero once $i > 2c/\pi = WT$. Therefore, for $i > WT$, ψ_i has little energy in the time-interval $[-T/2, T/2]$. Further, for fixed i, λ_i increases with increasing values of c (see [58] for tabulated values of λ_i). That is, the greater the time-bandwidth product, the greater the concentration of energy of $\psi_i(t)$ in $[-T/2, T/2]$.

For the remainder of this chapter let ψ_i denote the prolate spheroidal wave functions and let λ_i denote the corresponding eigenvalues.

The Uncertainty Principle

Let $f \in L^2[\mathbb{R}]$ be a nonzero function. We already know that $\alpha = \frac{\|Df\|}{\|f\|}$ and $\beta = \frac{\|Bf\|}{\|f\|}$ cannot simultaneously be equal to 1. Landau and Pollok [30] show that α and β cannot simultaneously be arbitrarily close to 1. Specifically, they prove

Theorem 5.5. *[30, theorem 2] Let $0 \leq \alpha, \beta \leq 1$. Then there exists a function $f \in L^2[\mathbb{R}]$, $\|f\|_2 = 1$ with $\|Df\|_2 = \alpha$ and $\|Bf\|_2 = \beta$ if and only if $(\alpha, \beta) \neq (0, 1)$ and $(\alpha, \beta) \neq (1, 0)$ and*

$$\cos^{-1}\alpha + \cos^{-1}\beta \geq \cos^{-1}\sqrt{\lambda_0},$$

[4] See [30,58] for more details

where λ_0 is the largest eigenvalue of the equation

$$\lambda\psi = BD\psi$$

This theorem constrains the possible values of α and β because $\lambda_0 < 1$ [58]. Therefore, we can conclude that any function cannot have arbitrarily large fractions of energy in both a finite time duration and a finite frequency bandwidth.

Alternatively, if $\epsilon_\omega^2 = 1 - \beta^2$ then f is said to be ϵ_ω-concentrated to the frequency interval $[-W, W]$. Similarly, if $\epsilon_t^2 = 1 - \beta^2$ then f is said to be ϵ_t-concentrated to the time interval $[-T/2, T/2]$. Also, f is said to have 'fraction out of band energy' (FOBE) bandwidth W and FOBE time duration T. The UP then implies [5]

$$\sin^{-1}\epsilon_t + \sin^{-1}\epsilon_\omega \geq \cos^{-1}\sqrt{\lambda_0(\Omega T)}$$

The principle constrains the range of values the product ΩT can take exactly like in the classical principle.

We will show in the next section that this theorem is a special case of a more general theorem just like Heisenberg's UP is a special case of the classical uncertainty principle as alluded to by [30]. We show that this more general theory can be used to understand certain fundamental limits on communication through arbitrary channels.

5.2 A General Uncertainty Principle

Though the UP derived in the previous section is for a function defined on the real line and its fourier transform, the principle can be extended to include arbitrary transforms defined on $\mathbb{R}^n$. To derive results similar to those in the previous two sections for general transformations we need to ask what are the essential properties of the operators B and D. It turns out that the key property is that the subspaces $\mathcal{B}$ and $\mathcal{D}$ form a nonzero minimum angle. Before deriving these results, we explain the physical model.

5.2.1 Physical Problem and Notation

Let $\mathscr{X}$ and $\mathscr{Y}$ be two Hilbert spaces with inner-products $\langle\cdot,\cdot\rangle_X$ and $\langle\cdot,\cdot\rangle_Y$ and norms $\|\cdot\|_X = \langle\cdot,\cdot\rangle_X^{1/2}$ and $\|\cdot\|_Y = \langle\cdot,\cdot\rangle_Y^{1/2}$, respectively. Let $\Gamma : \mathscr{X} \to \mathscr{Y}$ be a linear operator and let

$$\hat{f} = \Gamma f. \tag{5.18}$$

Let $\mathcal{D} \subset \mathscr{Y}$ and $\mathcal{A} \subset \mathscr{X}$. Also let $\mathcal{B} = \Gamma(\mathcal{A})$ be the image of the set $\mathcal{A}$ under Γ. Physically, we can interpret $\mathscr{X}$ to be a set of possible transmitter functions and $\mathscr{Y}$ to

[5] See property 3 of the prolate spheroidal functions in section 5.1.3

be a set of possible receiver functions and Γ to be some operator which determines what receiver function each transmitter function generates. Also, because of physical constraints (e.g. transmitted signals must be bandlimited) the transmitter might not be able to generate all the functions in the transmitter function space $\mathscr{X}$. So one can interpret $\mathcal{A}$ to be the set of transmitter functions that a physical transmitter can generate. Therefore, $\mathcal{B}$ is the set of functions that the transmitter can generate at the receiver and will also be referred to as the space of transmitter functions. Similarly, the receiver might not be able to measure all functions that are in the receiver space due to physical constraints (e.g. received signals can only be measured over a finite time-duration). One can think of the set $\mathcal{D}$ as the set of functions that the receiver can measure and it will be referred to as the space of receiver functions.

Let us also assume that the sets $\mathcal{D}$ and $\mathcal{B}$ are complete subspaces of $\mathscr{Y}$. We can then define projection operators $D : \mathscr{Y} \to \mathscr{Y}$ and $B : \mathscr{Y} \to \mathscr{Y}$ such that $D(\mathscr{Y}) = \mathcal{D}$ and $B(\mathscr{Y}) = \mathcal{B}$. We can also define an angle between these two subspaces as follows:

$$\theta(\mathcal{B}, \mathcal{D}) = \inf_{\substack{f \in \mathcal{B}, g \in \mathcal{D} \\ f \neq 0, g \neq 0}} \theta(f, g). \tag{5.19}$$

Finally, given any linear operator $L : \mathscr{Y} \to \mathscr{Y}$ we can define the operator norm

$$\|L\|_Y = \sup_{y \in \mathscr{Y}} \frac{\|Ly\|_Y}{\|y\|_Y}.$$

5.2.2 An Uncertainty Principle for arbitrary subspaces

One can think of $\|f\|_Y^2$ as the energy of a function $f \in \mathscr{Y}$. Then $\alpha^2 = \|Df\|_Y^2 / \|f\|_Y^2$ is the fraction of energy of f in the space of receiver functions and $\beta^2 = \|Bf\|_Y^2 / \|f\|_Y^2$ is the fraction of the energy of f in the space of transmitter functions. In order to prove the Uncertainty Principle, we need the following lemma.

Lemma 5.6. *Let $f, g, h \in \mathscr{Y}$. Then,*

$$\theta(f, g) \leq \theta(f, h) + \theta(g, h). \tag{5.20}$$

Proof. Let $\hat{f} = f / \|f\|_Y$ and $\hat{g} = g / \|g\|_Y$. Then

$$\begin{aligned} \theta(f, g) &= \cos^{-1} \frac{\Re\{\langle \|f\|_Y \hat{f}, \|g\|_Y \hat{g} \rangle\}}{\|f\|_Y \|g\|_Y} \\ &= \cos^{-1} \Re\{\langle \hat{f}, \hat{g} \rangle\} \\ &= \theta(\hat{f}, \hat{g}) \end{aligned} \tag{5.21}$$

Let $\theta(f,g) \neq 0$. Otherwise, there is nothing to prove. Also, let $S = span\{f,g\}$ be the space of all functions spanned by f and g. Then this space is complete and we can write $\hat{h} = h/\|h\|_Y$ as [48]

$$\hat{h} = h^{\|} + h^{\perp},$$

where $h^{\|} \in S$ and $h^{\perp}$ is orthogonal to both f and g. Because $\|h^{\|}\|_Y \leq \|\hat{h}\|_Y = 1$ we have

$$\begin{aligned}
\theta(f,h) &= \theta(\hat{f},\hat{h}) \\
&= \cos^{-1}\Re\{\langle \hat{f},\hat{h}\rangle\} \\
&= \cos^{-1}\Re\{\langle \hat{f},h^{\|}\rangle\} \\
&\geq \cos^{-1}\Re\left\{\frac{\langle \hat{f},h^{\|}\rangle}{\|h^{\|}\|_Y}\right\} \\
&= \theta(\hat{f},h^{\|}) \qquad (5.22)
\end{aligned}$$

Similarly,

$$\theta(g,h) \geq \theta(\hat{g},h^{\|}) \qquad (5.23)$$

Now, if $\theta(\hat{g},h^{\|}) = 0$ then the proof is complete because $\theta(f,g) = \theta(\hat{f},\hat{g}) = \theta(\hat{f},h^{\|}) \leq \theta(f,h) \leq \theta(f,h) + \theta(g,h)$.

If $\theta(\hat{g},h^{\|}) \neq 0$, let

$$\hat{h}_1 = h^{\|}/\|h^{\|}\|_Y, \qquad (5.24)$$

$$\hat{h}_2 = \frac{g - \hat{h}_1\langle g,\hat{h}_1\rangle}{\|(g - \hat{h}_1\langle g,\hat{h}_1\rangle\|_Y} \qquad (5.25)$$

be two unit vectors that are orthogonal to each other and whose span is S. We can therefore write,

$$\hat{f} = a_1\hat{h}_1 + a_2\hat{h}_2 \qquad (5.26)$$

$$\hat{g} = b_1\hat{h}_1 + b_2\hat{h}_2 \qquad (5.27)$$

where $a_1 = a_1' + ia_1'', a_2 = a_2' + ia_2'', b_1 = b_1' + ib_1''$ and $b_2 = b_2' + ib_2''$ are complex numbers. From the orthogonality of $\hat{h}_1$ and $\hat{h}_2$ we have

$$\cos\theta(\hat{f},h^{\|}) = a_1' \qquad (5.28)$$

$$\cos\theta(\hat{g},h^{\|}) = b_1' \qquad (5.29)$$

$$\cos\theta(\hat{f},\hat{g}) = \Re\{a_1^*b_1 + a_2^*b_2\} \qquad (5.30)$$

$$= a_1'b_1' + a_2'b_2' + a_1''b_1'' + a_2''b_2'' \qquad (5.31)$$

Here, a_i^* is the complex conjugate of a_i. From the orthogonality of $\hat{h}_1$ and $\hat{h}_2$ and the fact that $\hat{f}$ and $\hat{g}$ have unit norm, we have

$$a_1'^2 + a_1''^2 + a_2'^2 + a_2''^2 = 1 \tag{5.32}$$
$$b_1'^2 + b_1''^2 + b_2'^2 + b_2''^2 = 1 \tag{5.33}$$

We can think of (a_1'', a_2', a_2'') and (b_1'', b_2', b_2'') as two three dimensional vectors. Then the Cauchy-Schwarz inequality gives us

$$\sqrt{a_1''^2 + a_2''^2 + a_2'^2}\sqrt{b_1''^2 + b_2''^2 + b_2'^2} \geq |a_1''b_1'' + a_2''b_2'' + a_2'b_2'|$$

Here, $|\cdot|$ denotes the absolute value of a real number. Therefore,

$$a_1''b_1'' + a_2''b_2'' + a_2'b_2' \geq -\sqrt{a_1''^2 + a_2''^2 + a_2'^2}\sqrt{b_1''^2 + b_2''^2 + b_2'^2} \tag{5.34}$$

Now,

$$\cos(\theta(\hat{f}, h^{\parallel}) + \theta(\hat{g}, h^{\parallel})) \tag{5.35}$$
$$= \cos(\theta(\hat{f}, h^{\parallel}))\cos(\theta(\hat{g}, h^{\parallel})) - \sin(\theta(\hat{f}, h^{\parallel}))\sin(\theta(\hat{g}, h^{\parallel})) \tag{5.36}$$
$$= a_1'b_1' - \sqrt{1 - a_1'^2}\sqrt{1 - b_1'^2} \tag{5.37}$$
$$= a_1'b_1' - \sqrt{a_1''^2 + a_2''^2 + a_2'^2}\sqrt{b_1''^2 + b_2''^2 + b_2'^2} \tag{5.38}$$
$$\leq a_1'b_1' + a_1''b_1'' + a_2''b_2'' + a_2'b_2' \tag{5.39}$$
$$= \cos(\theta(\hat{f}, \hat{g})) \tag{5.40}$$

We get equation (5.37) from equations (5.28) and (5.29). Equations (5.32) and (5.33) are used to get (5.38) and finally we use inequality (5.34) to get (5.39). Now, from the monotonicity of $\cos$, we have

$$\theta(\hat{f}, \hat{g}) \leq \theta(\hat{f}, h^{\parallel}) + \theta(\hat{g}, h^{\parallel}). \tag{5.41}$$

Substituting inequalities (5.22) and (5.23) into the above and using equation (5.21) proves the lemma. □

The uncertainty principle constrains the range of values that α and β can take provided the subspaces $\mathcal{B}$ and $\mathcal{D}$ form a non-zero minimum angle.

Theorem 5.7. *[The Uncertainty Principle] If the two subspaces $\mathcal{B}$ and $\mathcal{D}$ form a non-zero minimum angle θ_0 then*

$$\cos^{-1}\alpha + \cos^{-1}\beta \geq \theta_0. \tag{5.42}$$

Proof. From the definition of $\theta(f, g)$ we have

$$\begin{aligned}\cos(\theta(f, Df)) &= \frac{\Re\{\langle f, Df\rangle\}}{\|Df\|_Y \|f\|_Y} \\ &= \frac{\Re\{\langle Df, Df\rangle\}}{\|Df\|_Y \|f\|_Y} \\ &= \frac{\|Df\|_Y}{\|f\|_Y} \\ &= \alpha.\end{aligned}$$

We get the second step in the above derivation from the fact that $f = Df + f - Df$ and $\langle f - Df, Df\rangle = 0$. Similarly we have, $\beta = \cos(\theta(f, Bf))$. Therefore from lemma 5.6 we get

$$\begin{aligned}\cos^{-1}\alpha + \cos^{-1}\beta &= \theta(f, Df) + \theta(f, Bf) \\ &\geq \theta(Df, Bf) \\ &\geq \theta_0.\end{aligned}$$

We get the last step from the fact that $Df \in \mathcal{D}$ and $Bf \in \mathcal{B}$ and these two subspaces have the minimum angle θ_0. □

We can calculate the minimum angle between the two subspaces by calculating the norm of the operator BD and this is the subject of our next theorem.

Theorem 5.8. *The angle between two complete subspaces $\mathcal{B}$ and $\mathcal{D}$ with projection operators B and D is*

$$\theta(\mathcal{B}, \mathcal{D}) = \cos^{-1}\|BD\|_Y \tag{5.43}$$

Proof. The angle between the two subspaces is given by

$$\theta(\mathcal{B}, \mathcal{D}) = \inf_{\substack{f\in\mathcal{B}, g\in\mathcal{D} \\ f\neq 0, g\neq 0}} \theta(f, g) \tag{5.44}$$

Using $\cos\theta(f,g)$ from the proof of the last theorem, we can write

$$\begin{aligned}
\cos\theta(\mathcal{B},\mathcal{D}) &= \sup_{\substack{f\in\mathcal{B},g\in\mathcal{D}\\ \|f\|_Y=1,\|g\|_Y=1}} \Re\{\langle f,g\rangle\} \\
&= \sup_{\substack{f\in\mathscr{Y},g\in\mathscr{Y}\\ \|f\|_Y=1,\|g\|_Y=1}} \Re\{\langle Bf,Dg\rangle\} \\
&= \sup_{\substack{f\in\mathscr{Y},g\in\mathscr{Y}\\ \|f\|_Y=1,\|g\|_Y=1}} \Re\{\langle f,BDg\rangle\} \\
&= \|BD\|_Y
\end{aligned}$$

Here, we get the second to last step from the fact that B is self adjoint and the last step from the definition of the operator norm. □

5.2.3 Discussion

The UP in theorem 5.7 has a very simple physical interpretation based on a very simple geometric idea. If the space of all the functions that a transmitter can generate and the space of all the functions a receiver can receive form a non-zero minimum angle then there exist no functions that can have arbitrarily large fractions of energy in these two spaces of functions. Also, one can find the minimum angle between these two subspaces using projection operators as explained in theorem 5.8.

Note that because B and D are projection operators we have, for $\mathcal{B},\mathcal{D}\neq\{0\}$, $\|B\|_Y = \|D\|_Y = 1$. Therefore, $\|BD\|_Y \leq \|B\|_Y\|D\|_Y = 1$. If $\|BD\| = 1$ then the uncertainty principle is not very useful because for all α and β we have $\cos^{-1}\alpha + \cos^{-1}\beta \geq \cos^{-1}1 = 0$. So the theorem is only useful if we can prove that $\|BD\|_Y < 1$ which is the case for the time-bandwidth problem considered by Landau and Pollak [27].

In order to find $\|BD\|_Y$ one still needs to evaluate the projection operators B and D. It is fairly easy to do this for the projection operator D if the subspace $\mathscr{D}$ is known. For instance if $\mathscr{Y}$ is $L^2[\mathbb{R}]$ and if the receiver can only measure functions for a time-duration of T seconds, then the projection operator can be defined as in equation (5.11).

There seems to be no general way in which one can define the projection operator B. One exception to this is the case where Γ is compact. In this case one can use the singular value decomposition (SVD) for the operator Γ to define the operator B. We can write [47]

$$\begin{aligned}
\Gamma &= \sum_j \langle\cdot,\phi_j\rangle_X \Gamma(\phi_j) \\
&= \alpha_j \sum_j \langle\cdot,\phi_j\rangle_X \psi_j
\end{aligned} \tag{5.45}$$

Here, α_j is the j^{th} singular value of Γ and ϕ_j and ψ_j are the j^{th} left and right singular functions of Γ, respectively. From equation (5.45) it is obvious that if a function belongs to $\mathcal{B}$ then it must be in the span of the set of right singular functions $\{\psi_j\}$ of Γ. Moreover, because the functions ψ_j are orthonormal in $\mathscr{Y}$ we can define the projection operator $B : \mathscr{Y} \to \mathscr{Y}$ as

$$Bf = \sum_j \langle f, \psi_j \rangle_Y \psi_j.$$

5.3 A Second General Uncertainty principle

In this section I prove a second general UP that can be used to constrain the total amount of energy that can be concentrated in the receiving volume. We prove a slightly modified version of the general uncertainty theorem proved in Donoho an Stark [60]. In the following $L^1[V]$, $L^2[V]$ and $L^\infty[V]$ are the spaces of real or complex valued functions defined on $V \subset \mathbb{R}^3$ with finite L_1 ($\|\cdot\|_1 = \int_V |\cdot|$), L_2 ($\|\cdot\|_1 = \int_V |\cdot|^2$) and L_∞ ($\|\cdot\|_\infty = \sup_V |\cdot|$) norms, respectively.

Theorem 5.9. *Suppose $f \in L^1[\mathbb{R}^3] \bigcap L^2[\mathbb{R}^3]$ and $\Gamma : f \mapsto \hat{f}$, where $\hat{f} \in L^2[\mathbb{R}^3] \bigcap L^\infty[\mathbb{R}^3]$ and satisfies*

1. $\|f\|_2 = \alpha \|\hat{f}\|_2$
2. $\|\hat{f}\|_\infty \leq \beta \|f\|_1$.

Let V_T and V_R be two compact subsets of $\mathbb{R}^3$ with empty intersection. Suppose f is ϵ_T-concentrated to V_T in the L_1 norm and $\hat{f}$ is ϵ_R-concentrated to V_R in the L_2 norm. Then,

$$|V_T||V_R|\alpha^2\beta^2 \geq (1-\epsilon_T)^2(1-\epsilon_R^2)$$

Proof.

$$\begin{aligned}
\|f\|_2^2 &= \alpha^2 \|\hat{f}\|_2^2 \\
&\leq \alpha^2 (1-\epsilon_R^2)^{-1} \int_{V_R} \hat{f}^2 \\
&\leq \alpha^2 (1-\epsilon_R^2)^{-1} |V_R| \|\hat{f}\|_\infty^2 \\
&\leq \alpha^2 (1-\epsilon_R^2)^{-1} |V_R| \beta^2 \|f\|_1^2 \\
&\leq \alpha^2 (1-\epsilon_R^2)^{-1} |V_R| \beta^2 (1-\epsilon_T)^{-2} \int_{V_T} |f| \\
&\leq \alpha^2 (1-\epsilon_R^2)^{-1} |V_R| \beta^2 (1-\epsilon_T)^{-2} |V_T| \|f\|_2^2
\end{aligned}$$

We get the last step using the Cauchy-Schwarz inequality. By rearranging the above inequality we get the required result. □

The theorem has a very simple physical interpretation for communication between finite volumes. Firstly, by requiring $\|f\|_2 = \alpha\|\hat{f}\|_2$, we ensure that the energy of the received signal is proportional to that of the transmitting signal. So α determines the attenuation in the signal and we expect it to be greater than 1. Secondly, $\|\hat{f}\|_\infty \leq \beta\|f\|_1$ can be thought of as a stability condition (i.e. bounded input gives bounded output).

Also, if the transmitting volume is finite, then f must be perfectly concentrated in V_T and so $\epsilon_T = 0$. The theorem then implies

$$|V_T||V_R|\alpha^2\beta^2 \geq (1 - \epsilon_R^2)$$

That is, the maximum fraction of energy that can be inside the receiving volume is bounded from above. Note that this bound increases if the transmitting and receiving volumes get bigger. This statement is valid for arbitrary channels provided the two conditions in the above theorem are satisfied.

5.4 Chapter Conclusion

In this chapter I reviewed the classical UP and examined its formulation in operator theoretic terminology. I then reviewed the time-bandwidth problem studied by Slepian, Landau and Pollak in a series of papers [27,30,58]. I generalised their results to arbitrary operators on Hilbert spaces and gave a physical interpretation of the resulting generalised Uncertainty Principle for SWCs. I also developed a second generalised UP. The main contributions of this chapter are:

1. Review of the classical and time-bandwidth UP.

2. Two new generalised UPs in Theorems 5.7 and 5.9

Chapter 6

Conclusion

In this chapter, I present the conclusions drawn from my study of SWCs. In each chapter, a summary of the work and the contributions made was presented. Here a summary of the thesis is followed by possible future research directions.

6.1 Thesis Summary

1. In chapter 2 I give a novel definition for SWCs that has sufficient structure to model the different physical constraints imposed on MIMO systems. According to definition 2.1 an SWC is a triple (X_T, Y_R, Γ). Here, X_T models the space of all the transmitter current densities and Y_R models the space of electromagnetic fields at the receiver and $\Gamma : X_T \rightarrow Y_R$ is a bounded linear operator that determines the electromagnetic field in the receiving volume given the current density in the transmitter volume. The spaces X_T and Y_R are normed spaces and the norms on these spaces can be physically interpreted as energy or power. I also establish several properties of SWCs for finite power and finite energy channels.

2. In chapter 3 I prove that for a compact operator Γ on some normed space, for any given number $\epsilon > 0$ there is a unique number $N_{dof}(\epsilon)$ which is the number of degrees of freedom of the operator Γ at level ϵ. Physically we can interpret this number as the maximum number of linearly independent functions a receiver that has noise level proportional to ϵ can measure. I prove that one can use this definition for degrees of freedom to define generalised singular values which are generalisations of the commonly accepted singular values defined for Hilbert space operators. These generalised singular values can be used to numerically compute the degrees of freedom of SWCs. In this chapter, I also distinguish between the terms “degrees of freedom” and “essential dimension” though they have been used interchangeably in the literature. In the situation where the singular values of an operator change rapidly from being large to small, the position of the “knee”

in the singular values is unique for a given channel. In such channels, the number of degrees of freedom at level ϵ depends very little on the actual value of ϵ. I define the essential dimension of an operator as the smallest number of singular values after which the difference between two consecutive singular values is at a maximum.

3. In chapter 4 I showed that it is possible to use perturbation theory to calculate the essential dimension of the scalar waveform channel studied by Miller [26]. In most situations, it is not possible to use such techniques to analytically compute the singular values. I therefore developed numerical techniques to compute generalised singular values of compact operators defined on a normed space that has a complete Schauder basis similar to Galerkin's method. I proved theorem 4.1 that shows that the singular values of finite dimensional approximations of an operator approach those of the original operator if the domain of the operator has a complete Schauder basis. I used this numerical technique to compute the singular values, degrees of freedom and essential dimension of several SWCs.

4. In chapter 5 I reviewed the classical Uncertainty Principle (UP) and examined its formulation in operator theoretic terminology. I then reviewed the time-bandwidth problem studied by Slepian, Landau and Pollak in a series of papers [27,30,58]. I generalised their results to arbitrary operators on Hilbert spaces and gave a physical interpretation of the resulting generalised UP for SWCs. I also developed a second generalised UP.

6.2 Future Directions

In this thesis I mainly concentrated on deterministic channels. However, several practical communication systems are random in nature and are also time varying. It would therefore be useful to extend the definition of SWCs to include random operators that depend on the positions of arbitrarily placed scatterers. Similarly, the definitons of essential dimension and degrees of freedom could be generalised to random variables instead of the deterministic ones studied in this thesis.

A second important extension would be to assume arbitrary time dependence instead of exponential time-dependance for the finite energy case. If this is done then one can study the number of linearly independent signals available in a MIMO system that can radiate any waveform.

Finally, properties of generalised singular values would need to be studied to develop more numerical techniques to calculate them. For instance, if A is an $n \times m$ matrix, then the singular values of A can be calculated from the stationary points of $\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$. If generalised singular values satisfy a similar condition, then this condition could be used to compute the generalised singular values.

Appendix A

Proofs Used In Theorems

A.1 Upper Bound on $\|\triangle\|$ used in section 4.1

The original Green's function is:

$$G(x,y) = \frac{\exp\{|\mathbf{x}-\mathbf{y}+\mathbf{e}_3 r_0|\}}{|\mathbf{x}-\mathbf{y}+\mathbf{e}_3 r_0|} \tag{A.1}$$

Because $G(x,y)$ is only a function of $\Delta = x - y$, we can rewrite it as

$$G(x,y) = \frac{\exp\{|\Delta + r_0 e_3|\}}{|\Delta + r_0 e_3|} \tag{A.2}$$

Now let

$$f(\Delta) = \sqrt{\Delta_1^2 + \Delta_2^2 + (\Delta_3 + r_0)^2} \tag{A.3}$$

In the following I will calculate all the partial derivatives of up to order 3 and for ease of notation I will use f instead of $f(\Delta)$. I show brief working for the calculation of derivatives w.r.t Δ_1 and then we can use symmetry to calculate derivatives w.r.t Δ_2 and substitution/symmetry to calculate derivatives w.r.t Δ_3.

$$\begin{aligned}
\frac{\partial f}{\partial \Delta_1} &= \frac{\Delta_1}{f} \\
\frac{\partial f}{\partial \Delta_2} &= \frac{\Delta_2}{f} \\
\frac{\partial f}{\partial \Delta_3} &= \frac{(\Delta_3 + r_0)}{f}
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 f}{\partial \Delta_1^2} &= \frac{f - \Delta_1 \frac{\partial f}{\partial \Delta_1}}{f^2} \\
&= \frac{1}{f}\left[1 - \frac{\Delta_1^2}{f^2}\right] \\
\frac{\partial^2 f}{\partial \Delta_2^2} &= \frac{1}{f}\left[1 - \frac{\Delta_2^2}{f^2}\right] \\
\frac{\partial^2 f}{\partial \Delta_3^2} &= \frac{1}{f}\left[1 - \frac{(\Delta_3 + r_0)^2}{f^2}\right] \\
\frac{\partial^2 f}{\partial \Delta_2 \partial \Delta_1} &= \frac{\partial}{\partial \Delta_2}\left\{\frac{\Delta_1}{f}\right\} \\
&= \frac{f\frac{\partial \Delta_1}{\partial \Delta_2} - \Delta_1 \frac{\partial f}{\partial \Delta_2}}{f^2} \\
&= \frac{-\Delta_1 \Delta_2}{f^3} \\
\frac{\partial^2 f}{\partial \Delta_3 \partial \Delta_1} &= \frac{-\Delta_1 (\Delta_3 + r_0)}{f^3} \\
\frac{\partial^2 f}{\partial \Delta_3 \partial \Delta_2} &= \frac{-\Delta_2 (\Delta_3 + r_0)}{f^3} \\
\frac{\partial^3 f}{\partial \Delta_1^3} &= \frac{\partial}{\partial \Delta_1}\left\{\frac{1}{f} - \frac{\Delta_1^2}{f^3}\right\} \\
&= \frac{-1}{f^2}\frac{\partial f}{\partial \Delta_1} - \left[\frac{f^3 \frac{\partial \Delta_1^2}{\partial \Delta_1} - \Delta_1^2 \frac{\partial f^3}{\partial \Delta_1}}{f^6}\right] \\
&= \frac{-\Delta_1}{f^3} - \frac{f^3 2\Delta_1 - \Delta_1^2 3 f^2 \frac{\Delta_1}{f}}{f^6} \\
&= \frac{3\Delta_1}{f^3}\left[\frac{\Delta_1^2}{f^2} - 1\right] \\
\frac{\partial^3 f}{\partial \Delta_2^3} &= \frac{3\Delta_2}{f^3}\left[\frac{\Delta_2^2}{f^2} - 1\right] \\
\frac{\partial^3 f}{\partial \Delta_3^3} &= \frac{3(\Delta_3 + r_0)}{f^3}\left[\frac{(\Delta_3 + r_0)^2}{f^2} - 1\right]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^3 f}{\partial\Delta_2\partial\Delta_1^2} &= \frac{\partial}{\partial\Delta_2}\left\{\frac{1}{f}-\frac{\Delta_1^2}{f^3}\right\} \\
&= \frac{-\Delta_2}{f^3}-\left[\frac{f^3\frac{\partial\Delta_1^2}{\partial\Delta_2}-\Delta_1^2\frac{\partial f^3}{\partial\Delta_2}}{f^6}\right] \\
&= \frac{\Delta_2}{f^3}\left[\frac{3\Delta_1^2}{f^2}-1\right] \\
\frac{\partial^3 f}{\partial\Delta_2^2\partial\Delta_1} &= \frac{\Delta_1}{f^3}\left[\frac{3\Delta_2^2}{f^2}-1\right] \\
\frac{\partial^3 f}{\partial\Delta_3\partial\Delta_1^2} &= \frac{(\Delta_3+r_0)}{f^3}\left[\frac{3\Delta_1^2}{f^2}-1\right] \\
\frac{\partial^3 f}{\partial\Delta_3^2\partial\Delta_1} &= \frac{\Delta_1}{f^3}\left[\frac{3\left(\Delta_3+r_0\right)^2}{f^2}-1\right] \\
\frac{\partial^3 f}{\partial\Delta_3\partial\Delta_2^2} &= \frac{(\Delta_3+r_0)}{f^3}\left[\frac{3\Delta_2^2}{f^2}-1\right] \\
\frac{\partial^3 f}{\partial\Delta_3^2\partial\Delta_2} &= \frac{\Delta_2}{f^3}\left[\frac{3\left(\Delta_3+r_0\right)^2}{f^2}-1\right] \\
\frac{\partial^3 f}{\partial\Delta_3\partial\Delta_2\partial\Delta_1} &= \frac{\partial}{\partial\Delta_3}\left\{\frac{-\Delta_2\Delta_1}{f^3}\right\} \\
&= \frac{f^3\frac{\partial-\Delta_1\Delta_2}{\partial\Delta_3}+\Delta_2\Delta_1\frac{\partial f^3}{\partial\Delta_3}}{f^6} \\
&= \frac{3\Delta_1\Delta_2\left(\Delta_3+r_0\right)}{f^5}
\end{aligned}
$$

If $a \in V - W$, using Taylor's theorem we can write

$$f(a) = \sum_{n=0}^{N-1} \frac{1}{n!}(a\cdot\nabla)^n f(0) + R_N(a) \tag{A.4}$$

and the remainder $R_N(a)$ satisfies

$$R_N(a) \leq \sup_{\alpha\in V-W} \frac{1}{N!}(a\cdot\nabla)^N f(\alpha). \tag{A.5}$$

To calculate a bound on this remainder, define

$$\begin{aligned} b_i &= \sup_{x \in V-W} |x_i|\,,\ i = 1,2,3 \\ b &= \max_i b_i \\ l &= \inf_{x \in V-W} f(x) \\ u &= \sup_{x \in V-W} r_0 + x_3 = r_0 + b_3 \leq r_0 + b. \end{aligned}$$

Now, if $x \in V - W$ then we have the following inequalities if we assume $b \leq l$ (Note that in these inequalities, $i = 1, 2$ or 3):

$$\begin{aligned} \left|\frac{x_i}{f(x)}\right| &\leq \frac{b}{l} \\ \left|\frac{x_3 + r_0}{f(x)}\right| &\leq \frac{u}{l} \\ &\leq \frac{r_0 - b + 2b}{l} \\ &\leq 1 + \frac{2b}{l} \\ \left|\frac{3x_i^2}{f^2} - 1\right| &\leq \left|\frac{3b^2}{f^2} - 1\right| \\ &\leq 2 \\ \left|\frac{3(x_3 + r_0)^2}{f^2} - 1\right| &\leq \left|\frac{3u^2}{f^2} - 1\right| \\ &\leq \left|3\left(1 + \frac{2b}{l}\right)^2 - 1\right| \\ &\leq 24 \end{aligned}$$

Now if we further assume that b/l is much smaller than 1, for example if

$$b/l \leq \frac{2 - \sqrt{3}}{2\sqrt{3}}, \tag{A.6}$$

then we can get tighter bounds for the last two inequalities:

$$\begin{aligned} \left|\frac{3x_i^2}{f^2} - 1\right| &\leq 1 \\ \left|\frac{3(x_3 + r_0)^2}{f^2} - 1\right| &\leq 3 \end{aligned}$$

Assuming (A.6) we prove the bounds for each term in the remainder $R_3(\Delta)$.

$$
\begin{aligned}
\left|\Delta_1^3 \frac{\partial^3 f}{\partial \Delta_1^3}\Big|_{\Delta=\alpha}\right| &= \left|\frac{\Delta_1^3 \cdot 3\alpha_1}{f^3}\left[\frac{\alpha_1^2}{f^2}-1\right]\right| \\
&\leq \frac{3b^4}{l^3} \\
\left|\Delta_2^3 \frac{\partial^3 f}{\partial \Delta_2^3}\Big|_{\Delta=\alpha}\right| &\leq \frac{3b^4}{l^3} \\
\left|\Delta_3^3 \frac{\partial^3 f}{\partial \Delta_3^3}\Big|_{\Delta=\alpha}\right| &= \left|\frac{\Delta_3^3 \cdot 3(\alpha_3+r_0)}{f^3}\left[\frac{(\alpha_3+r_0)^2}{f^2}-1\right]\right| \\
&\leq \frac{3b^3}{l^2}\cdot\frac{u}{l}\cdot 3 \\
\left|3\Delta_1^2\Delta_2 \frac{\partial^3 f}{\partial \Delta_2 \partial \Delta_1^2}\Big|_{\Delta=\alpha}\right| &= \left|\frac{3\Delta_1^2\Delta_2 \cdot \alpha_2}{f^3}\left[\frac{3\alpha_1^2}{f^2}-1\right]\right| \\
&\leq \frac{3b^4}{l^3} \\
\left|3\Delta_2^2\Delta_1 \frac{\partial^3 f}{\partial \Delta_2^2 \partial \Delta_1}\Big|_{\Delta=\alpha}\right| &\leq \frac{3b^4}{l^3} \\
\left|3\Delta_1^2\Delta_3 \frac{\partial^3 f}{\partial \Delta_3 \partial \Delta_1^2}\Big|_{\Delta=\alpha}\right| &= \left|\frac{3\Delta_1^2\Delta_3 \cdot (\alpha_3+r_0)}{f^3}\left[\frac{3\alpha_1^2}{f^2}-1\right]\right| \\
&\leq \frac{3b^3}{l^2}\cdot\frac{u}{l} \\
\left|3\Delta_2^2\Delta_3 \frac{\partial^3 f}{\partial \Delta_3 \partial \Delta_2^2}\Big|_{\Delta=\alpha}\right| &\leq \frac{3b^3}{l^2}\cdot\frac{u}{l} \\
\left|3\Delta_3^2\Delta_1 \frac{\partial^3 f}{\partial \Delta_3^2 \partial \Delta_1}\Big|_{\Delta=\alpha}\right| &= \left|\frac{3\Delta_3^2\Delta_1 \cdot \alpha_1}{f^3}\left[\frac{3(\alpha_3+r_0)^2}{f^2}-1\right]\right| \\
&\leq \frac{3b^4}{l^3}\cdot 3 \\
\left|3\Delta_3^2\Delta_2 \frac{\partial^3 f}{\partial \Delta_3^2 \partial \Delta_2}\Big|_{\Delta=\alpha}\right| &\leq \frac{3b^4}{l^3}\cdot 3 \\
\left|6\Delta_3\Delta_2\Delta_1 \frac{\partial^3 f}{\partial \Delta_3 \partial \Delta_2 \partial \Delta_1}\Big|_{\Delta=\alpha}\right| &= \left|\frac{6\Delta_3\Delta_2\Delta_1 \cdot 3(\alpha_3+r_0)\alpha_2\alpha_1}{f^5}\right| \\
&\leq \frac{18b^5}{l^4}\cdot\frac{u}{l}
\end{aligned}
$$

Therefore,

$$\begin{aligned}
|R_3(\Delta)| &\leq \frac{1}{3!}\left[10\frac{3b^4}{l^3} + 5\frac{3b^3}{l^2}\cdot\frac{u}{l} + \frac{18b^5}{l^4}\cdot\frac{u}{l}\right] \\
&= \frac{b}{2}\cdot\frac{b^2}{l^2}\left[10\frac{b}{l} + 5\frac{u}{l} + \frac{6b^2}{l^2}\right] \\
&\leq \frac{b}{2}\cdot\frac{b^2}{l^2}\left[5 + 20\frac{b}{l} + \frac{6b^2}{l^2}\right] \qquad \text{(A.7)} \\
&\leq \frac{b}{2}\cdot\frac{b^2}{l^2}c_1 \qquad \text{(A.8)}
\end{aligned}$$

Here, c_1 is constant which, assuming the validity of equation (A.6) is 5.58. If a different condition is imposed on b/l then we will get a different value for c_1. Now we will calculate the error in the approximate Green's function $G_T(x, y)$:

$$G_T(x, y) = \frac{\exp\{ikf_{T_2}(x-y)\}}{f_{T_0}(x-y)} \qquad \text{(A.9)}$$

Here, f_{T_i} is the i^{th} order Taylor series expansion of f. From Taylor's theorem, we have

$$f_{T_2}(\Delta) = f(\Delta) - R_3(\Delta) \qquad \text{(A.10)}$$

For ease of notation, I will use h_{xy} and δ_{xy} instead of $f(x-y)$ and $R_3(x-y)$ respectively.

$$\begin{aligned}
e^{-ikf_T(z-x)}e^{ikf_T(y-z)} &= e^{ik[h_{yz}-h_{zx}]}\cdot e^{ik[\delta_{zx}-\delta yz]} \\
&= e^{ik[h_{yz}-h_{zx}]}\cdot[1 + e^{\alpha}ik(\delta_{zx}-\delta_{yz})] \qquad \text{(A.11)} \\
&= e^{ik[h_{yz}-h_{zx}]} + R_e \qquad \text{(A.12)}
\end{aligned}$$

Here, we get equation (A.11) from the Taylor's theorem for exponential function. $\alpha \in \mathbb{C}$ satisfies $|\alpha| \leq k|\delta_{zx} - \delta yz|$ and R_e is the remainder term and is a function of x, y and z. From equations (A.12) and (A.8) we get

$$\begin{aligned}
|R_e| &= \left|e^{ik[h_{yz}-h_{zx}]}\right|\cdot|e^{\alpha}ik(\delta_{zx}-\delta_{yz})| \\
&\leq |e^{\alpha}|\,k(|\delta_{zx}| + |\delta_{yz}|) \\
&\leq e^{k\frac{b}{2}\frac{b^2}{l^2}c_1}\cdot kb\frac{b^2}{l^2}c_1 \\
&= e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1 \qquad \text{(A.13)}
\end{aligned}$$

Here, $\lambda = 2\pi/k$ is the wavelength under consideration.

Similarly we can calculate the error term in the denominator. Let

$$g(\Delta) = \frac{1}{f(\Delta)}. \qquad \text{(A.14)}$$

Then the first order derivatives of $g = g(\Delta)$ are

$$\begin{aligned}
\frac{\partial g}{\partial \Delta_1} &= \frac{-\Delta_1}{f^3} \\
\frac{\partial g}{\partial \Delta_2} &= \frac{-\Delta_2}{f^3} \\
\frac{\partial g}{\partial \Delta_3} &= \frac{-\left(\Delta_3 + r_0\right)}{f^3}
\end{aligned}$$

From Taylor's theorem, we can write

$$g(\Delta) = \frac{1}{r_0} + R_{d\Delta}$$

and the remainder $R_{d\Delta}$ is bounded. Shown below are bounds for each term in the remainder.

$$\begin{aligned}
\left|\Delta_1 \frac{\partial g}{\partial \Delta_1}\right|_{\Delta=\alpha}\Bigg| &= \left|\Delta_1 \cdot \frac{-\alpha_1}{f^3}\right| \\
&\leq \frac{b^2}{l^3} \\
\left|\Delta_2 \frac{\partial g}{\partial \Delta_2}\right|_{\Delta=\alpha}\Bigg| &\leq \frac{b^2}{l^3} \\
\left|\Delta_3 \frac{\partial g}{\partial \Delta_3}\right|_{\Delta=\alpha}\Bigg| &= \left|\Delta_2 \cdot \frac{-\alpha_3}{f^3}\right| \\
&\leq \frac{b}{l^2} \cdot \frac{u}{l}
\end{aligned}$$

Therefore,

$$\begin{aligned}
|R_{d\Delta}| &\leq \frac{b^2}{l^3} + \frac{b^2}{l^3} + \frac{b}{l^2} \cdot \frac{u}{l} \\
&\leq \frac{b}{l^2} \cdot \left[1 + \frac{4b}{l}\right] \\
&= \frac{b}{l^2} c_2
\end{aligned} \tag{A.15}$$

Here, c_2 is a constant and if $b < l$ then $1 \leq c_2 \leq 5$.

Now let $R_{dyz} = R_{d(y-z)}$ and $R_{dzx} = R_{d(z-x)}$. Then we have

$$\begin{aligned}
\frac{1}{f_T(y-z) f_T(z-x)} &= \frac{1}{h_{yz} h_{zx}} + \frac{R_{dzx}}{h_{yz}} + \frac{R_{dyz}}{h_{zx}} + R_{dyz} R_{dzx} \\
&= \frac{1}{h_{yz} h_{zx}} + R_d
\end{aligned}$$

The remainder R_d, which is a function of x, y and z is bounded:

$$\begin{aligned} |R_d| &\leq \left|\frac{R_{dzx}}{h_{yz}}\right| + \left|\frac{R_{dyz}}{h_{zx}}\right| + |R_{dyz}R_{dzx}| \\ &\leq \frac{2c_2 b}{l^3}\left[1 + \frac{c_2 b}{2l}\right] \qquad \text{(A.16)} \\ &= \frac{c_3 b}{l^3} \qquad \text{(A.17)} \end{aligned}$$

Here, c_3 is a constant and $2c_2 \leq c_3 \leq 1 + c_2^2 b/l$. Therefore, we have

$$\begin{aligned} &|G^*(z,x)G(y,z) - G_T^*(z,x)G_T(y,z)| \\ &= \left|\frac{R_e}{f_{zx}f_{yz}} + e^{ik[f_{zx}+f_{yz}]}R_d + R_e R_d\right| \\ &\leq e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1 \cdot \frac{1}{l^2} + \frac{c_3 b}{l^3} + \frac{c_3 b}{l^3}e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1 \\ &= \frac{b}{l^3}\left[e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b}{l}c_1 + c_3 + c_3 e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1\right] \qquad \text{(A.18)} \end{aligned}$$

Therefore we have

$$\begin{aligned} |k_\Delta(x,y)| &\leq \int_W |G^*(z,x)G(z,x) - G_T^*(z,x)G_T(z,x)|dz \\ &\leq \frac{b}{l^3}\left[e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b}{l}c_1 + c_3 + c_3 e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1\right]\int_W dz \\ &\leq \frac{b\mu(W)}{l^3}\left[e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b}{l}c_1 + c_3 + c_3 e^{\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1}\cdot 2\pi\frac{b}{\lambda}\frac{b^2}{l^2}c_1\right] \end{aligned}$$

Here $\mu(W)$ is the Lebesgue measure of the volume W.

A.2 Bound on the Energy Stored in the Electromagnetic Field

In the finite energy case, we assume that the source current has exponential time dependance and is zero outside some time interval $[0, t_0] \subset \mathbb{R}^3$. Let

$$\xi_{t_0}(t) = \begin{cases} 1 & 0 \leq t \leq t_0 \\ 0 & \text{Otherwise} \end{cases}$$

Let T be some compact subset of $\mathbb{R}^3$ and let $\mathbf{J} \in \mathcal{L}^2(T, \mathbb{C}^3)$ be such that $\Re\{\mathbf{J}(\mathbf{r})e^{j\omega t}\xi_{t_0}(t)\}$ is the source current. Here, ω is the angular frequency of oscillation of the source current. Then the magnetic field is the real part of [44, pp. 247]

$$\mathbf{H}(\mathbf{r}, t_0) = \frac{e^{j\omega t_0}}{4\pi}\int_T \left[\frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|^2} + \frac{ke^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\right]\xi_{t_0}\left(t_0 - \frac{|\mathbf{r}-\mathbf{r}'|}{c}\right)\mathbf{J}(\mathbf{r}') \times \hat{\mathbf{R}}d\mathbf{r}.$$

Here, c is the speed of light, $k = \frac{\omega}{c}$ is the wave number and $\hat{\mathbf{R}}$ is a unit vector in the direction of $\mathbf{r} - \mathbf{r}'$. Because $c < \infty$ and $\xi_{t_0}(t) = 0$ for $t < 0$, the magnetic field $\mathbf{H}$ is zero outside the set $(T + \overline{B}_{t_0 c, \mathbb{R}^3}(0))$ which has finite measure . Also if there exists an ϵ such that

$$\inf_{\mathbf{r}' \in T} |\mathbf{r} - \mathbf{r}'| \geq \epsilon \tag{A.19}$$

then there exists a constant $c(\epsilon) < \infty$ such that

$$\left| \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|^2} + \frac{ke^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \right| \leq c.$$

Therefore, a simple application of the Hölder's inequality proves that there is a constant $c_1(\epsilon)$ such that

$$|\mathbf{H}(\mathbf{r}, t_0)| \leq c_1(\epsilon) \|\mathbf{J}\|_{\mathcal{L}^2(T, \mathbb{C}^3)}.$$

Now, if Ω is some closed surface the interior of which contains T, then because T is compact, there exists an $\epsilon > 0$ such that for all $\mathbf{r}' \in T$ and $\mathbf{r} \in \Omega_{ext}$,

$$|\mathbf{r} - \mathbf{r}'| > \epsilon.$$

Therefore for all $\mathbf{r} \in \Omega_{ext}$,

$$|\mathbf{H}(\mathbf{r}, t_0)| \leq c_1(\epsilon) \|\mathbf{J}\|_{\mathcal{L}^2(T, \mathbb{C}^3)}.$$

Also, because $\mathbf{H}$ is zero outside a set of finite measure,

$$\int_{\Omega_{ext}} \frac{1}{\mu_0} |\mathbf{H}(\mathbf{r}, t_0)|^2 d\mathbf{r} \leq c_2(\epsilon) \|\mathbf{J}\|^2_{\mathcal{L}^2(T, \mathbb{C}^3)}.$$

Here $c_1(\epsilon), c_2(\epsilon) < \infty \forall \epsilon > 0$. Therefore, the energy stored in the magnetic field is less than a bound proportional to $\|\mathbf{J}\|^2_{\mathcal{L}^2(T, \mathbb{C}^3)}$. Because in free space, with exponential time dependence, the electric field is proportional to the magnetic field, we have a similar condition for the energy stored in the electric field. Therefore, there exists some constant $b = b(\epsilon) < \infty$ such that for all $\mathbf{J} \in \mathcal{L}^2(T, \mathbb{C}^3)$

$$[E_{rad}(\mathbf{J})](t_0) \leq b(\epsilon) \|\mathbf{J}\|^2_{\mathcal{L}^2(T, \mathbb{C}^3)}.$$

Here E_{rad} is as defined in equation (2.6) and b is a constant that only depends on T and Ω and not on $\mathbf{J}$.

Bibliography

[1] Simon Haykin. *Communication Systems*. Wiley; 4th edition, 2000.

[2] Bernard Sklar. Rayleigh fading channels in mobile digital communication systems, part i: Characterization. *IEEE Communications Magazine*, 35(7):90–100, July 1997.

[3] Theodore S. Rappaport. *Wireless Communications*. Prentice-Hall India, New Delhi-110001, 2 edition, 2005.

[4] Bernard Sklar. Rayleigh fading channels in mobile digital communication systems, part ii: Mitigation. *IEEE Communications Magazine*, 35(10):148–155, September 1997.

[5] John G. Proakis and Masoud Salehi. *Fundamentals of Communication Systems*. Prentice Hall, 2004.

[6] Jack H. Winters. On the capacity of radio communication systems with diversity in a Rayleigh fading environment. *IEEE Journal on Selected Areas in Communication*, 5(5):871–878, 1987.

[7] I. E. Telatar. Capacity of multi-antenna gaussian channels. Technical Report BL0112170-950615-07TM, AT & T Bell Laboratories, 1995.

[8] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6:311–335, 1998.

[9] Ezio Biglieri, John Proakis, and Shlomo Shamai. Fading channels: Information-theoretic and communications aspects. *IEEE Transactions on Information Theory*, 44:2619–2692, 1998.

[10] Thomas L. Marzetta and Bertrand M. Hochwald. Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading. *IEEE Transactions on Information Theory*, 45(1):139–157, January 1999.

[11] Gregory G. Raleigh and John M. Cioffi. Spatio-temporal coding for wireless communication. *IEEE Transactions on Communication*, 46(3):357–366, March 1998.

[12] Gregory W. Wornell and Mitchell D. Trott. Efficient signal processing techniques for exploiting transmit antenna diversity on fading channels. *IEEE Transactions on Signal Processing*, 45(1):191–205, 1997.

[13] G.D. Golden, C.J. Foschini, R.A. Valenzuela, and P.W. Wolniansky. Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture. *Electronics Letters*, 35(1):14–16, January 1999.

[14] P. W. Wolniansky, G. D. Golden, G. J. Foschini, and R. A. Valenzuela. V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel. In *URSI Int. Symp. on Signals, Systems, and Electronics, ISSSE-98.*, pages 295–300, 1998.

[15] David Gesbert, Mansoor Shafi, Da shan Shiu, Peter J. Smith, and Ayman Naguib. From theory to practice: An overview of MIMO spacetime coded wireless systems. *IEEE Journal on Selected Areas in Communication*, 21(3):281–302, April 2003.

[16] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H Bolcskei. An overview of MIMO communication - A key to gigabit wireless. *Proceedings of the IEEE*, 92(2):198–218, 2004.

[17] H. Bölcskei, D. Gesbert, and A. Paulraj. On the capacity of OFDM-based spatial multiplexing systems. *IEEE Transactions on Communication*, 50(2):225–234, February 2002.

[18] B. Clerckx, C. Craeye, D. Vanhoenacker-Janvier, and C. Oestges. Impact of antenna coupling on 2×2 MIMO communications. *IEEE Transactions on Vehicular Technology*, 56(3):1009–1018, 2007.

[19] Chen-Nee Chuah, David N. C. Tse, Joseph M. Kahn, and Reinaldo A. Valenzuela. Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Transactions on Information Theory*, 48(3):637–650, 2002.

[20] D. Gesbert, H. Bölcskei, D.A. Gore, and A.J. Paulraj. Outdoor MIMO wireless channels: Models and performance prediction. *IEEE Transactions on Communications*, 50(12):1926–1934, December 2002.

[21] B. N. Getu and R. Janaswamy. The effect of mutual coupling on the capacity of the MIMO cube. *IEEE Antennas and Wireless Propagation Letters*, 4:240–244, 2005.

[22] S. Loyka and A. Kouki. New compound upper bound on MIMO channel capacity. *IEEE Communications Letters*, 6(3):96–98, 2002.

[23] Mehmet Kemal Ozdemir, Huseyin Arslan, and Ercument Arvas. A mutual coupling model for MIMO systems. In *IEEE Topical Conference on Wireless Communication Technology*, pages 306–307, 2003.

[24] Jon W. Wallace and Michael A. Jensen. Mutual coupling in MIMO wireless systems: A rigorous network theory analysis. *IEEE Transactions on Wireless Communications*, 3(4):1317–1325, 2004.

[25] Ovidio M. Bucci and Giorgio Franceschetti. On the degrees of freedom of scattered fields. *IEEE Transactions on Antennas Propagation*, 37(7):318–326, July 1989.

[26] David A.B. Miller. Communicating with waves between volumes: evaluating orthogonal spatial channels and limits on coupling strengths. *Applied Optics*, 39(11):1681–1699, April 2000.

[27] H.J. Landau and H.O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty - iii: The dimension of the space of essentially time- and band-limited signals. *The Bell System Technical Journal*, 41:1295–1336, Jul 1962.

[28] Ada S. Y. Poon, Robert W. Brodersen, and David N. C. Tse. Degrees of freedom in multiple-antenna channels: A signal space approach. *IEEE Transactions on Information Theory*, 51(2):523–536, February 2005.

[29] Kathi K. Selig. Uncertainty principles revisited. Technical Report TUM M0101(1), Techn. Univ. Mnchen, Fak. f. Math., 2001.

[30] H.J. Landau and H.O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty - ii. *The Bell System Technical Journal*, 40:65–84, Jan 1961.

[31] Rodney A. Kennedy, Parastoo Sadeghi, Thushara D Abhayapala, and Haley M. Jones. Intrinsic limits of dimensionality and richness in random multipath fields. *IEEE Transactions on Signal Processing*, To appear, 2007.

[32] Christopher T. H. Baker. *The Numerical Treatment of Integral Equations*. Clarendon Press, Oxford, 1977.

[33] Ovidio M. Bucci and Giorgio Franceschetti. On spatial bandwidth of scattered fields. *IEEE Transactions on Antennas and Propagation*, 35(12):1445–1455, December 1987.

[34] R. Piestun and D. A. B. Miller. Electromagnetic degrees of freedom of an optical system. *Journal of the Optical Society of America A*, 17(5):892–902, 2000.

[35] L. Hanlen and M. Fu. Capacity of MIMO channels: A volumetric approach. In *International Conference on Communications (ICC03)*, pages 2673–2677, Anchorage, Alaska, May 2003.

[36] Leif Hanlen and Minyue Fu. Wireless communication systems with spatial diversity: A volumetric model. *IEEE Transactions on Wireless Communications*, 5(1):133–142, January 2006.

[37] Jie Xu and Ramakrishna Janaswamy. Electromagnetic degrees of freedom in 2-d scattering environments. *IEEE Transactions on Antennas and Propagation*, 54(12):3882–3894, December 2006.

[38] H. M. Jones, R. A. Kennedy, and T. D. Abahayapala. On dimensionality of multipath fileds: spatial extent and richness. In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, page 28372840, 2002.

[39] Parastoo Sadeghi, Thushara D. Abhayapala, and Rodney A. Kennedy. Intrinsic finite dimensionality of random multipath fields. In *ICASSP*, pages 17–20, Toulouse, France, May 2006.

[40] M. W. G. Dickins and L. W. Hanlen. On the dimensionality of spatial fields with restricted angle of arrival. In *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, pages 1033–1037, Adelaide, Australia, Sept 2005.

[41] G. Newsam and R. Barakat. Essential dimension as a well-defined number of degrees of freedom of finite-convolution operators appearing in optics. *Journal of the Optical Society of America A*, 2(11):2040–2045, November 1985.

[42] M. A. Jensen and J. W. Wallace. Antenna-independent capacity bound of electromagnetic channels. In *Antennas and Propagation Society International Symposium*, volume 2A, pages 317–320, July 2005.

[43] J.W. Wallace and M.A. Jensen. Intrinsic capacity of the MIMO wireless channel. In *Vehicular Technology Conference, 2002. Proceedings. VTC 2002-Fall. 2002 IEEE 56th*, volume 2, pages 701– 705, 2002.

[44] John David Jackson. *Classical Electrodynamics*. John Wiley & Sons, 1999.

[45] Marco Donald Migliore. On the role of the number of degrees of freedom of the field in MIMO channels. *IEEE Transactions on Antennas Propagation*, 54(2):620–628, February 2006.

[46] Van G. Graham and Ivan H. Sloan. On the compactness of certain integral operators. *Journal of Mathematical Analysis and Applications*, 68:580–594, 1979.

[47] Kato. *Perturbation theory for linear operators*. Springer, 1966.

[48] Erwin Kreyszig. *Introductory functional analysis with applications*. John Wiley & Sons, 1989.

[49] David Slepian. On bandwidth. *Proc. IEEE*, 64(3):292–300, March 1976.

[50] Kane S. Yee. Numerical solution of initial boundary value problems involving maxwells equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):802–807, 1966.

[51] Jeffrey L. Young and Ronald O. Nelson. A summary and systematic analysis of FDTD algorithms for linearly dispersive media. *IEEE Antennas and Propagation Magazine*, 43(1):61–77, February 2001.

[52] Anders Bondeson, Thomas Rylander, and Pär Ingelström. *Computational electromagnetics*. Springer, 2005.

[53] Allen Taflove. Review of the formulation and applications of the finite-difference time-domain method for the numerical modeling of electromagnetic wave interactions with arbitrary structures. *Wave Motion*, 10:547–582, 1988.

[54] Peter G. Petropoulos. Analysis of exponential time-differencing for FDTD in lossy dielectrics. *IEEE Transactions on Antennas and Propagation*, 45(6):1054–1057, June 1997.

[55] Jonathon C. Veihl and Raj Mittra. An efficient implementation of Berenger's perfectly matched layer (PML) for finite-difference time-domain mesh truncation. *IEEE Microwave and Guided Wave Letters*, 6(2):94–96, 1996.

[56] Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200, 1994.

[57] Werner Heisenberg. *The Physical Principles of Quantum Theory*. Dover Pub, 1949. First German edition in Leipzig, 1930.

[58] D. Slepian and H.O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty - i. *The Bell System Technical Journal*, 40:43–63, Jan 1961.

[59] George Arfken. *Mathematical methods for physicists*, chapter 15. Academic Press Inc., 1966.

[60] David L. Donoho and Phillip B. Stark. Uncertainty principle and signal recovery. *SIAM J. App, Math.*, 49(3):906–931, Jun 1989.